Infection (2019) 47:729–737
https://doi.org/10.1007/s15010-019-01282-w

ORIGINAL PAPER

Genomic epidemiology of MRSA infection and colonization isolates among military trainees with skin and soft tissue infection

Eugene V. Millar[1,2] · Gregory K. Rice[3,4] · Carey D. Schlett[1,2] · Emad M. Elassal[1,2] · Regina Z. Cer[3,4] · Kenneth G. Frey[4] · Theron Hamilton[4] · Michael W. Ellis[5] · David R. Tribble[1] · Kimberly A. Bishop-Lilly[4] · Jason W. Bennett[6,7]

Received: 15 November 2018 / Accepted: 12 February 2019 / Published online: 22 February 2019

Abstract

Purpose Individuals with methicillin-resistant *Staphylococcus aureus* (MRSA) skin and soft tissue infection (SSTI) can be simultaneously colonized with MRSA on multiple body sites. Using whole genome sequencing (WGS), the intrahost relatedness of MRSA colonization and infection isolates was investigated.

Methods In the context of a prospective case–control study of SSTI, we analyzed colonization and infection isolates from US Army Infantry trainees with purulent infection due to MRSA. At the time of clinical presentation for SSTI, culture swabs were obtained from the infection site, as well as from the patient's nasal, oral, inguinal, and perianal regions. *S. aureus* culture and susceptibility was performed by standard methods. DNA from MRSA isolates was extracted and libraries were produced. Sequences were generated on an Illumina MiSeq, sequence reads were assembled, and single nucleotide variant (SNV) data were analyzed.

Results Of 74 trainees with MRSA SSTI, 19 (25.7%) were colonized with MRSA. Ten (52.6%) were colonized on more than one body site. Colonization frequency by anatomic site was as follows: inguinal region (33%), nasal region (30%), perianal region (22%), and oral region (14%). A total of 36 MRSA colonization isolates were characterized. The intrahost median number of SNVs between infection and colonization isolates was 17. Among trainees with recurrent MRSA SSTI, limited intrahost diversity suggests that persistent colonization is a major contributor to recurrence risk.

Conclusions Among military trainees with MRSA SSTI, genomic characterization of infection and colonization isolates revealed a high degree of strain relatedness. Single acquisition events may account for MRSA colonization and infection in this population.

Keywords Methicillin-resistant *Staphylococcus aureus* (MRSA) · Skin and soft tissue infection · Colonization · Whole genome sequencing · Genomics · Military

Introduction

Staphylococcus aureus causes invasive (e.g., bacteremia, pneumonia) and non-invasive (e.g., abscess, cellulitis) disease in humans. Colonization plays a role in the pathogenesis

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s15010-019-01282-w) contains supplementary material, which is available to authorized users.

✉ Eugene V. Millar
emillar@idcrp.org

1 Infectious Disease Clinical Research Program, Department of Preventive Medicine and Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

2 Henry M. Jackson Foundation for the Advancement of Military Medicine, Rockville, MD, USA

3 Leidos, Reston, VA, USA

4 Naval Medical Research Center, Biological Defense Research Directorate, Fort Detrick, MD, USA

5 University of Toledo College of Medicine and Life Sciences, Toledo, OH, USA

6 Walter Reed Army Institute of Research, Silver Spring, MD, USA

7 Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

of *S. aureus*, although the relationship between colonization and subsequent infection events is much less clear for skin and soft tissue infection (SSTI) [1–5] than it is for invasive disease [6, 7]. There is variability in how colonization is defined; the anterior nares, often considered the primary reservoir of *S. aureus*, are sampled most frequently. However, *S. aureus* colonization of non-nasal sites (e.g., oral, inguinal, and perianal regions) is common and may also contribute to infection risk [3, 8–10].

Establishing the connection between colonization and disease in individuals with *S. aureus* and methicillin-resistant *S. aureus* (MRSA) infection requires demonstration of strain concordance [1, 3, 5, 11]. Two important limitations of past studies were that only the nares were sampled in the determination of colonization status and strains were characterized by pulsed-field type [1, 5]. To date, only one study has utilized highly discriminatory whole genome sequencing (WGS) methods to assess intrahost strain relatedness among individuals with MRSA colonization and infection [11]. Improved understanding of the association between *S. aureus*/MRSA colonization and SSTI warrants: (1) inclusion of non-nasal sites in the colonization sampling strategy, and (2) utilization of WGS in the determination of strain relatedness.

We previously conducted an epidemiologic investigation of MRSA SSTI among US Army Infantry trainees, a population known to be at increased risk for MRSA colonization and infection [12]. Specifically, we identified clusters of MRSA SSTI in several military training classes and characterized clinical isolates by WGS to deduce intraclass patterns of MRSA transmission. As part of this same investigation, we swabbed each SSTI case on multiple body sites to assess his MRSA colonization status at the time of clinical presentation. Through the use of WGS, we are now able to describe the genomic epidemiology of MRSA colonization isolates and the intrahost relatedness of MRSA infection and colonization isolates among US Army Infantry trainees with MRSA SSTI.

Methods

In the context of a prospective case–control study of SSTI [12], we analyzed colonization and infection isolates from US Army Infantry trainees with purulent infection due to MRSA. A case was defined as a trainee who presented to the clinic or was admitted to the hospital with cellulitis, abscess, folliculitis, impetigo, paronychia, infected blister, or pilonidal cyst. A *S. aureus* SSTI case was defined as a SSTI with a positive culture from the corresponding clinical site. A recurrent episode was defined as a subsequent SSTI at a different anatomical site ≥ 30 days after the first episode. Information on SSTI risk factors and training class (i.e., platoon, company, battalion, etc.) was collected at enrollment.

At the time of clinical presentation for SSTI, participants were also swabbed at several anatomical sites (nasal, oropharyngeal, inguinal, and perianal regions) to assess *S. aureus* colonization status. For individuals with recurrent SSTI, colonization swabs were collected from the first episode only.

Cultures were processed according to standard protocols. All *S. aureus* isolates underwent identification and susceptibility testing using Microscan® Walk-Away-96 (Dade Behring Inc.; Deerfield, Illinois), according to Clinical Laboratory Standards Institute (CLSI) methods. *S. aureus* isolates underwent typing with pulsed-field gel electrophoresis (PFGE). Control strains of known pulsed-field types (PFT) were obtained from the Biodefense and Emerging Infections Research Resources Repository [13]. PFGE findings were resolved and analyzed using BioNumerics (Applied Math; Austin, TX).

Genomic characterization of colonization isolates was restricted to MRSA only. To describe the genomic epidemiology of MRSA colonization among high-risk military trainees with SSTI, we selected MRSA colonization isolates from those individuals whose MRSA SSTI isolates were previously characterized by WGS [12].

DNA extraction was performed using the Wizard Kit (Promega; Madison, WI) and libraries produced using the Nextera XT DNA Library Preparation Kit (Illumina, Inc.; San Diego, CA) according to the manufacturer's instructions. Libraries were multiplexed and sequenced using an Illumina MiSeq 600-cycle kit and 2 × 300-bp read lengths. Sequence read quality was analyzed with FastQC (version 0.11.5) [14] and low-quality bases were trimmed with sickle (version 1.3.3) [15]. In order to determine the closest reference for single nucleotide variant (SNV) analysis, sequence reads were assembled using SPAdes (version 3.11.0) [16] and the longest contig from each assembly was aligned against the NCBI nt database (accessed on November 30, 2017) using BLAST [17]. *S. aureus* USA300 strain TCH1516 (NCBI accession no. CP000730) was determined to be the closest reference genome based on the percentage of total pairwise identity.

SNV data were analyzed by the Bacterial and Archaeal Genome Analyser (BAGA; version 0.2.1) [18], a wrapper for proven third party bioinformatics tools. Sequence reads were mapped to the reference using bwa (version 0.7.17-r1188) [19], and variant calls and filtering were performed with GATK (version 3.6.0) [20]. Genomics regions containing insertions/deletions (indels), potential chromosomal rearrangements and sequence repeats, known to increase the likelihood of false-positive variant calls, were excluded from the SNV set. A multiple sequence alignment (MSA) was created from nucleotide substitutions, small deletions

called by GATK and putative large deletions detected in the bwa sequence alignments where no reads mapped. A maximum-likelihood tree was constructed from the nucleic acid MSA using PhyML (version 3.3.20170530) [21] tree search with the GTR substitution model, 100 bootstrap replicates, and the BEST tree topology search operation option. Sequence data were submitted to NCBI under BioProject PRJNA356758.

For this paper, isolate numbers are Subject ID numbers from the primary manuscript describing the genomic epidemiology of MRSA SSTI isolates [12], followed by a letter denoting the anatomic site from which the MRSA colonization isolate was obtained (N: nasal; O: oral: I: inguinal; P: perianal region).

Statistical analysis was performed in SAS (SAS Version 9.3; SAS Institute; Cary, NC).

Each of the trainees provided written informed consent to participate in the study. This study was approved by the Uniformed Services University Infectious Disease Institutional Review Board (IDCRP-074).

Results

Study population and case characteristics

Among the 74 individuals whose MRSA SSTI isolates were previously characterized [12], 19 (25.7%) were colonized with MRSA. The demographic and clinical characteristics of the 19 cases are presented in Table 1. The median (range) age was 21 (17–29) years. All were male. Sixty-three percent were White. The most frequent clinical diagnoses were purulent cellulitis (73.7%) and abscess (57.9), the majority (63.2%) of which occurred on the lower extremities. SSTI cases generally occurred after 6–7 weeks of training (median [range] number of days after the start of training, 45 [10–84]).

With respect to SSTI risk factors, 23.5% reported antibiotic use in the past 6 months, 17.6% reported a known/suspected SSTI in the past year, and 5.9% reported having a preceding medically attended SSTI while at Fort Benning. There were no differences in demographic, clinical, nor risk factor characteristics when comparing those with versus those without MRSA-positive colonization swabs (data not shown).

Characteristics of *S. aureus* colonization

For all participants, infection and colonization swabs were collected during the same visit. Of 296 colonization swabs, 36 (12.1%) were MRSA-positive and 97 (33%) were methicillin-susceptible *S. aureus* (MSSA) positive. The most frequently MRSA colonized site was the inguinal region (33.3%; 12/36), followed by the nasal region (27.8%; 10/36), perianal region (25%; 9/36), and oral region (13.9%; 5/36) (Table 2). Ten (52.6%) individuals were colonized with MRSA at more than one anatomic site. Five (50%) were colonized at two sites, three (30%) were colonized at three sites, and two (20%) were colonized at all four sites.

Among those individuals colonized with MRSA, the overall prevalence of MSSA co-colonization was 26%.

Table 1 Demographic and clinical characteristics of US Army Infantry Trainees with MRSA skin and soft tissue infection included in genomic analysis of infection and colonization isolates

Subject characteristics	Number (%)
Median (range) age, years	21 (17–29)
Race/ethnicity	
White, non-Hispanic	12 (63.2)
Hispanic	6 (31.6)
Black, non-Hispanic	1 (5.3)
Median (range) no. of days from training start to presentation for SSTI	45 (10–84)
Clinical diagnosis	
Cellulitis	14 (73.7)
Abscess	11 (57.9)
Folliculitis	2 (10.5)
Infected blister	1 (5.3)
Site of infection	
Lower extremity	12 (63.2)
Upper extremity	6 (31.6)
Head	1 (5.3)
Thorax	1 (5.3)
Groin/inguinal/perineal	2 (10.5)
SSTI risk factors	
Known or suspected SSTI infection in past year	3 (17.6)
Prior medically attended SSTI at Fort Benning	1 (5.9)
Self-reported antibiotic use in prior 6 months	4 (23.5)

MRSA methicillin-resistant *S. aureus, SSTI* skin and soft tissue infection

Table 2 Distribution of MRSA colonization status by anatomic site and number of sites colonized among US Army Infantry Trainees with MRSA skin and soft tissue infection

Number (%) of MRSA-positive cultures, $n=36$	
Inguinal	12 (33.3)
Nasal	10 (27.8)
Perianal	9 (25)
Oral	5 (13.9)
Number (%) of anatomic sites colonized with MRSA, $n=19$	
One	9 (47.4)
Two	5 (26.3)
Three	3 (15.8)
Four	2 (10.5)

Seven MSSA colonization isolates from five individuals were obtained. By anatomic site, the proportion of MSSA-positive isolates was as follows: 43% (oral), 29% (perianal), 14% (nasal), and 14% (inguinal).

Distribution of SSTI cases by training class

The 19 cases stemmed from seven different training classes, each composed of ~200 soldiers who were segregated from other soldiers from other classes for the duration of the 14-week training period. Five cases each were from classes A (3025) and B (4043); two cases each were from classes C (3006), D (4030), G (3104) and H (2042); one case was from class F (3049).

Recurrent SSTI

Two of the subjects (1098 and 1564) had recurrent MRSA SSTI. Subject 1098 first presented with purulent cellulitis (knee) and then with purulent cellulitis (forearm) 67 days later [12]. Two of four swabs (nasal and perianal), collected at the time of the first SSTI, were positive for MRSA. Subject 1564 first presented with abscess and folliculitis (abdomen and leg) and then abscess (forearm) 49 days later [12]. Two of four swabs (nasal and inguinal), also collected at the time of the first SSTI, were positive for MRSA.

Molecular and genomic characteristics

All but one of the colonization isolates was pulsed-field type (PFT) USA300 (2758.O was PFT USA100). MRSA sequence data revealed the following sequence types (ST): 35 (97.2%) of the isolates were ST-8 and 1 (2.8%), isolate 2758.O, was ST-5. Among the USA300 MRSA colonization isolates, a total of 1411 SNVs were identified. The overall median (range) SNV difference between the USA300 MRSA isolates ($n = 35$) was 123 (83–545). The phylogenetic tree of colonization isolates is presented in Fig. 1.

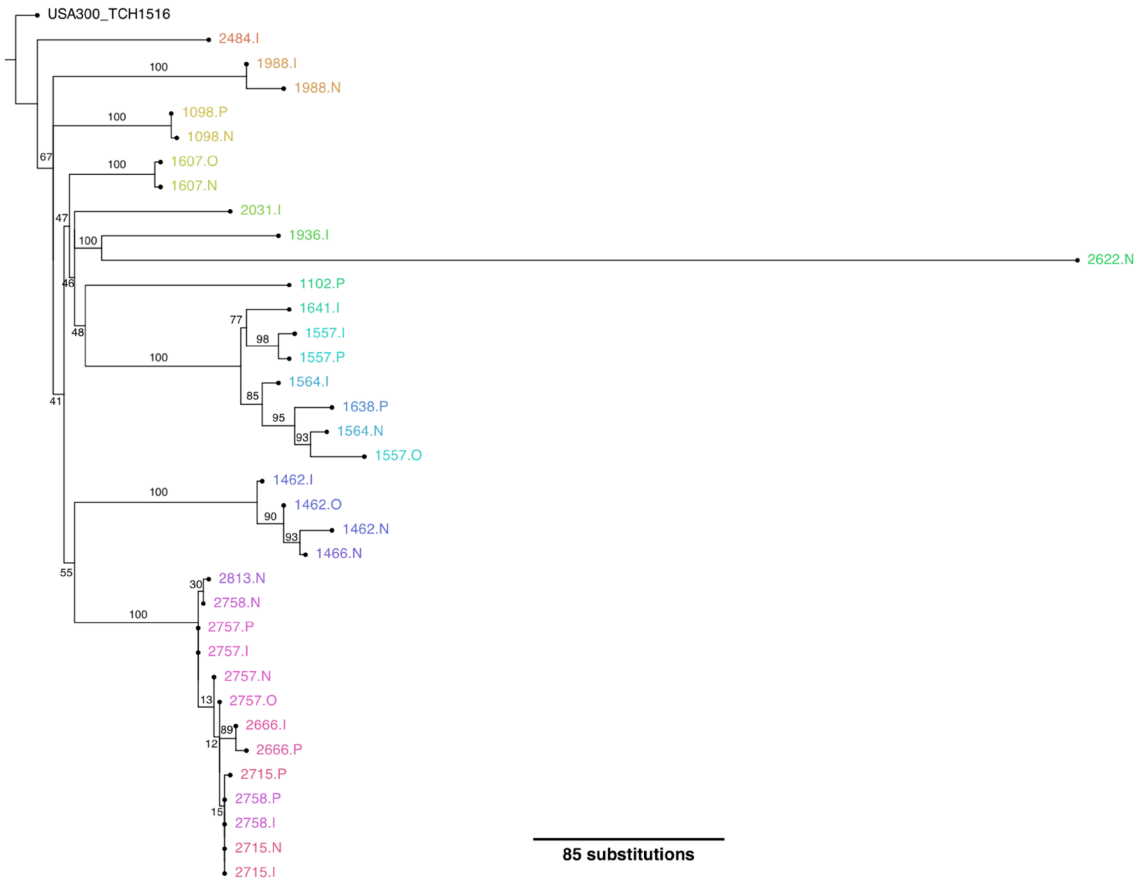

Fig. 1 Distance scaled maximum-likelihood (PhyML) phylogenetic tree inferred from nucleotide sequences of 35 MRSA USA300 colonization isolates from 19 Military Trainees with MRSA SSTI. Isolate numbers are Subject ID numbers from the primary manuscript describing the genomic epidemiology of MRSA SSTI isolates [12], followed by a letter denoting the anatomic site from which the MRSA colonization isolate was obtained (*N* nasal, *O* oral, *I* inguinal, *P* perianal region). Bootstrap proportion values are indicated on the branches of the tree. A phylogenetic tree with MRSA USA300 colonization and infection isolates combined is presented in Supplemental Fig. 1

Intrahost relatedness of MRSA infection and colonization isolates

Excluding the one colonization isolate that was clearly divergent from the others (i.e., 2758.O), a low diversity of isolates was observed (Fig. 2). The overall median (range) intrahost SNV difference between infection and colonization isolates was 16.5 (1–19,396). There were no associations between particular anatomic sites and the median number of SNV differences between colonization and infection isolates (Fig. 3). The phylogenetic tree of colonization isolates with the accompanying infection isolates, described previously [12], is presented in Supplemental Fig. 1.

Intrahost relatedness of MRSA colonization isolates

Among the ten individuals who were colonized with MRSA USA300 at more than one anatomic site, the median (range) intrahost SNV difference between colonization isolates was 18 (2–67). There were no associations between number of anatomic sites colonized and median number of SNV differences observed, either among colonization pairs, or infection–colonization pairs (Table 3).

Recurrent MRSA SSTI

Among the two trainees with recurrent MRSA SSTI, infection and colonization isolates were highly related (Table 4). For subject 1098, the median (range) number of SNVs between isolates obtained from the first SSTI episode, the second SSTI episode, and the two colonization isolates (1 nasal, 1 perianal) was 1.5 (0–3). Notably, for this subject, there were only two SNVs between his nasal MRSA isolate and the MRSA isolate associated with the second SSTI episode, identified 67 days after the first. Moreover, the perianal isolate and the isolate associated with the second SSTI episode were identical. For subject 1564, the median (range) number of SNVs between isolates obtained from the first SSTI episode, the second SSTI episode, and the two colonization isolates (1 nasal, 1 inguinal) was 21 (19–28). There were ~20 SNVs each between his nasal and inguinal MRSA isolates and the MRSA isolate associated with the second SSTI episode, identified 49 days after the first. Both subjects were prescribed co-trimoxazole at the time of the first infection for the treatment of SSTI.

Intraclass relatedness of MRSA colonization isolates

We stratified the analysis of colonization isolates by military training class in order to assess the relatedness of MRSA strains obtained from a congregate setting (i.e., strain circulation). Five cases each were from classes A (Subject IDs 1557, 1564, 1607, 1638, and 1641; 9 colonization isolates) and B (Subject IDs 2666, 2715, 2757, 2758, and 2813; 14 colonization isolates). Among trainees in class A who were positive for MRSA colonization, the first and last cases of MRSA SSTI in this class were separated by 41 days. The intraclass median (range) number

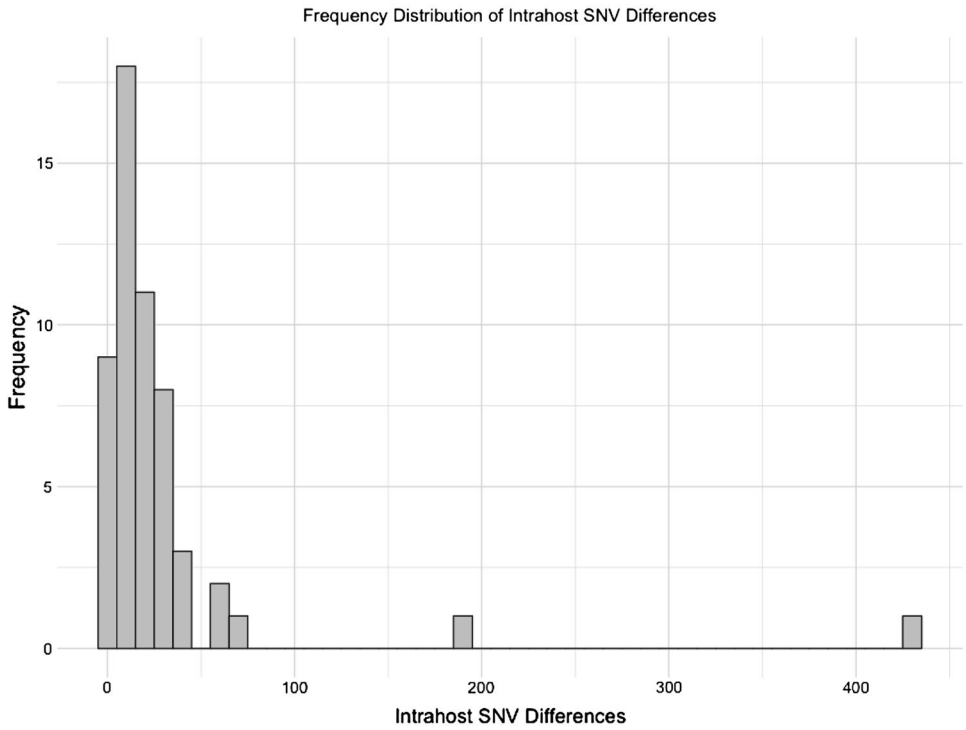

Fig. 2 Frequency distribution of the number of intrahost single nucleotide variant differences among MRSA USA300 colonization isolates obtained from US Army Infantry Trainees with MRSA skin and soft tissue infection

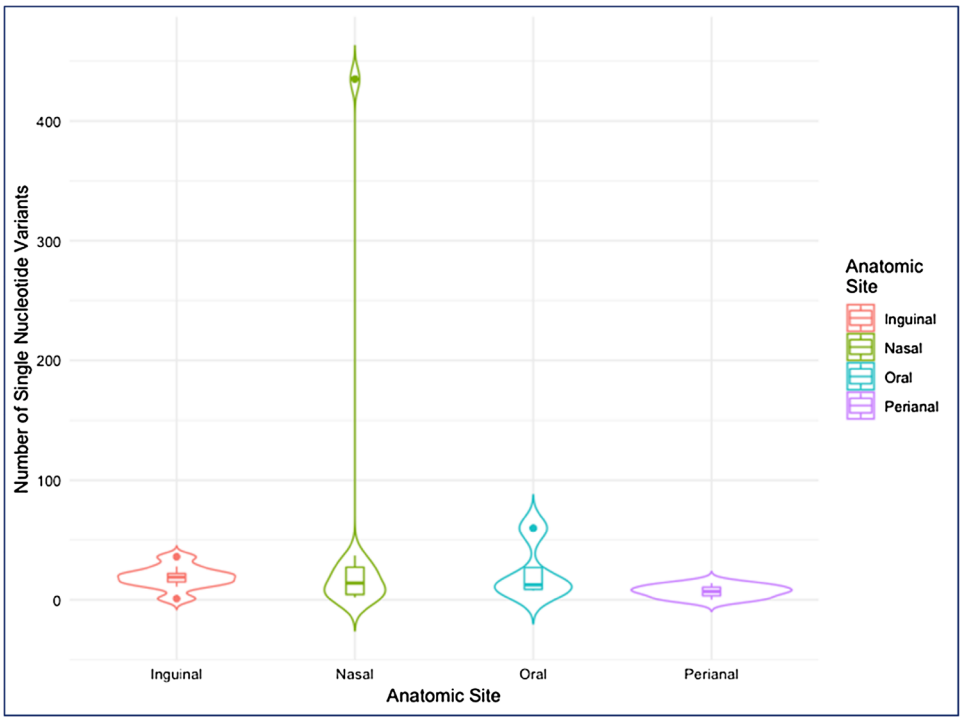

Fig. 3 Violin plot displaying the number of intrahost single nucleotide variants between MRSA colonization and infection isolates by anatomic site among US Army Infantry Trainees with purulent skin and soft tissue infection due to MRSA

Table 3 Median (range) number of single nucleotide variants between colonization pairs and infection–colonization pairs by the number of anatomic sites colonized among US Army Infantry Trainees with purulent skin and soft tissue infection due to MRSA

No. of anatomic sites colonized	Median (range) number of SNVs between:	
	Colonization pairs	Infection and colonization pairs
One, $n=9$	N/A	25 (6–435)
Two, $n=5$	14 (2–29)	12 (1–28)
Three, $n=3$	28 (6–67)	16 (2–60)
Four, $n=2$	19 (3–19,396)	11.5 (1–19,386)

SNV single nucleotide variants, *N/A* not applicable

of SNVs between MRSA colonization isolates obtained from class A was 66.5 (7–186). Among trainees in class B who were positive for MRSA colonization, the first and last cases of MRSA SSTI in this class were separated by 40 days. The intraclass median (range) number of SNVs between MRSA colonization isolates obtained from class B was 17 (2–19,466). Although the case numbers per class were relatively small, the relatively low diversity of MRSA colonization isolates in each suggests that person-to-person transmission plays an important role in the persistence of MRSA in military training settings. Small sample size precluded similar analyses of MRSA colonization isolates from the other five training classes.

Discussion

In the current study, we evaluated the epidemiology of MRSA colonization among military trainees with MRSA SSTI, finding that 25% were colonized at the time of clinical presentation and that, among those colonized, ~50% were colonized at more than one anatomic site. Other studies of community-associated MRSA infection have reported MRSA colonization prevalence estimates of 34–64% at the time of clinical presentation [3, 5, 9, 22–24], with 6–64% of colonized subjects being MRSA-positive at more than one site [3, 22, 24]. Of the four anatomic sites that were sampled (nasal, oral, inguinal, and perianal regions), the frequency of MRSA colonization was highest not in the nares (28%), but rather, in the groin (33%). Several other studies have demonstrated that non-nasal sites are important reservoirs for MRSA [3, 8, 9, 22, 25].

It is interesting to note that all but one of the MRSA colonization isolates were pulsed-field type (PFT) USA300, whereas a greater diversity of PFTs among MRSA colonization isolates was expected. This observation may be an artifact of the study design, in which we restricted our selection and subsequent genomic characterization of MRSA colonization isolates to those individuals who had purulent SSTI due to MRSA, all of which were due to the USA 300 strain.

Previously, we used whole genome sequencing (WGS) in the investigation of clusters of MRSA SSTI, in order to describe the intra- and interclass relatedness of strains associated with infection [12]. Herein, we used WGS to examine

Table 4 Numbers of single-nucleotide variants between infection and colonization isolates by anatomic site for two US Army Infantry Trainees with recurrent purulent skin and soft tissue infection due to MRSA

Subject 1098	Clinical isolate	
	First infection	Second infection, 67 days later
Clinical isolate: first infection	–	1
Colonization isolate[§]: nasal	3	2
Colonization isolate[§]: perianal	1	0
Subject 1564	Clinical isolate	
	First infection	Second infection, 49 days later
Clinical isolate: first infection	–	9
Colonization isolate[§]: nasal	28	21
Colonization isolate[§]: inguinal	15	22

[§]Colonization swabs were collected only at the time of the first infection

the intrahost relatedness of strains among trainees who were both infected with and colonized with MRSA. Whether a trainee was colonized at one or multiple anatomic sites, a high degree of intrahost strain relatedness was observed. With the exception of one individual who was colonized with at least two highly divergent strains of MRSA, the maximum intrahost SNV difference was low (median: 18 SNVs). In the comparison of intrahost pairwise differences, there were no indications that colonization isolates from any particular body site were more closely related than others to the isolate associated with infection.

Similar findings have been reported in other populations at increased risk for MRSA colonization and infection. Frazee et al. [23] reported 100% concordance of MRSA isolates obtained from nares and wounds of patients presenting with SSTI to an urban emergency department. Popovich et al. [26] conducted a MRSA colonization study in a Chicago public hospital, sampling participants at four anatomic sites and characterizing MRSA isolates by WGS. Intrahost diversity of MRSA was low, as indicated by median pairwise differences of 2–3 SNVs and a maximum intrahost SNV difference of 41. Several other studies conducted in hospitals or long-term care facilities yielded similar results: MRSA can be isolated from multiple anatomic sites, and the intrahost diversity of colonization strains is low [27–30].

Our genomic characterization of MRSA infection and colonization isolates among the two trainees with recurrent SSTI provides additional insight on the pathogenesis of recurrence. Despite antibiotic treatment and resolution of the initial infections, these individuals remained at risk for SSTI and suffered recurrent infections < 2 months later. The high degree of relatedness between the colonization strains and the strains associated with subsequent infections (0 SNVs for one subject and ≤ 28 SNVs for the other) strongly supports that a pre-existing reservoir of MRSA (i.e., nasal, inguinal, or perianal colonization) seeded the subject's second infection. Similar findings of strain persistence were reported in a longitudinal study of recurrent MRSA SSTI; among two patients evaluated for 225 and 750 days, a maximum of 5 and 11 SNVs, respectively, differentiated the strains associated with recurrent SSTI [11].

It is well recognized that training cycles are associated with an increased risk of MRSA acquisition, transmission and SSTI outbreaks for military recruits [4, 12, 31]. Crowded living conditions and recent contact with an individual with SSTI are well-established risk factors for infection [31]. Crowding also increases the likelihood of MRSA colonization: two-thirds of household contacts of MRSA index patients acquire MRSA, and the number of household members and the frequency of interaction are positively correlated with risk of acquisition [32].

In an Infantry training setting, up to 50 soldiers may occupy a single ("open bay") barrack where, for 14 weeks, the sleeping/living, restroom and laundry facilities are shared. Factors such as crowding, inadequate personal hygiene, and infrequent laundering of clothing, towels, etc., likely contribute to the ongoing transmission of MRSA in the trainee population. Ultimately, in high-risk settings such as the military training environment, where frequent, prolonged exposures to MRSA are commonplace, multiple personal and environmental hygiene-based strategies (e.g., administration of intranasal mupirocin, use of chlorhexidine-based body wash, routine disinfection of high-touch common surfaces, etc.) may be needed in order to reduce the burden of MRSA colonization and, in turn, reduce the risk of MRSA SSTIs.

There are limitations to this study. First, the cases were enrolled and colonization swabs collected at the time of clinical presentation for SSTI. As a result, we were not able to discern whether colonization episodes preceded infection events or vice versa. Second, determination of a trainee's MRSA colonization status was based on the selection of a single, purified colony from the laboratory culture of a colonization swab. It is possible that some individuals who were colonized with MRSA were inaccurately classified as not colonized, either because of a low density of colonization at a given body site or because of an insufficient sampling of colonies from the cultured specimen. Third, our genomic characterization of these isolates was limited to the core genome; we did not attempt to isolate and characterize extrachromosomal genetic elements (e.g., plasmids) that encode toxins, other virulence factors, and mechanisms of antibiotic resistance for MRSA. Ultimately, elucidating the association between colonization and infection in the study of MRSA SSTI pathogenesis will require longitudinal,

observational studies of acquisition, colonization, and subsequent infection, paired with the use of WGS in order to accurately characterize the strains involved in the transition from an asymptomatic carrier state to a clinically apparent infection.

WGS has emerged as an invaluable tool for the study of MRSA SSTI epidemiology, whether it be for the characterization of MRSA transmission dynamics in congregate populations at high-risk for infection, or to determine the intrahost relatedness of strains among patients both colonized and infected with MRSA. As the role of *S. aureus* colonization in the pathogenesis of SSTI continues to be examined, the identification of important human and non-human reservoirs for incident and recurrent infection may provide insight into new opportunities for host decolonization and environmental disinfection. Ultimately, these strategies may be a critical component for the prevention and control of SSTI in high-risk settings.

Acknowledgements We are indebted to the study team of clinical research coordinators, laboratory personnel, and data management staff for their dedication to the project.

Funding This work was supported by a US Department of Defense Program Project Grant [HT9404-12-1-0019 to M.W.E.]. Additional support for this work was provided by the Department of Defense Global Emerging Infections Surveillance and Response System [GEIS;HU0001-10-1-0018 to M.W.E] and the Military Infectious Diseases Research Program [MIDRP; HT9404-12-1-0012 to M.W.E]. The protocol was conducted by the Infectious Disease Clinical Research Program (IDCRP), a Department of Defense (DoD) program executed through the Uniformed Services UNIVERSITY of the Health Sciences through a cooperative agreement with The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), under Inter-Agency Agreement Y1-AI-5072.

Compliance with ethical standards

Conflict of interest All authors: no conflicts.

Disclaimer The contents of this publication are the sole responsibility of the author(s) and do not necessarily reflect the views, opinions or policies of Uniformed Services University of the Health Sciences (USUHS), the Department of Defense (DoD), the Departments of the Army, Navy, or Air Force, or the Henry M. Jackson Foundation for the Advancement of Military Medicine. Mention of trade names, commercial products, or organizations does not imply endorsement by the US Government. Drs. Bennett, Bishop-Lilly, Hamilton, and Tribble are service members or employees of the U.S. Government. This work was prepared as part of their official duties. Title 17 U.S.C. §105 provides that 'Copyright protection under this title is not available for any work of the United States Government.' Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties.

References

1. Chen AE, Cantey JB, Carroll KC, Ross T, Speser S, Siberry GK. Discordance between *Staphylococcus aureus* nasal colonization and skin infections in children. Pediatr Infect Dis J. 2009;28:244–6. https://doi.org/10.1097/INF.0b013e31818cb0c4.
2. Fritz SA, Epplin EK, Garbutt J, Storch GA. Skin infection in children colonized with community-associated methicillin-resistant Staphylococcus aureus. J Infect. 2009;59:394–401. https://doi.org/10.1016/j.jinf.2009.09.001.
3. Miller LG, Eells SJ, Taylor AR, David MZ, Ortiz N, Zychowski D, et al. *Staphylococcus aureus* colonization among household contacts of patients with skin infections: risk factors, strain discordance, and complex ecology. Clin Infect Dis. 2012;54:1523–35. https://doi.org/10.1093/cid/cis213.
4. Ellis MW, Hospenthal DR, Dooley DP, Gray PJ, Murray CK. Natural history of community-acquired methicillin-resistant Staphylococcus aureus colonization and infection in soldiers. Clin Infect Dis. 2004;39:971–9. https://doi.org/10.1086/423965.
5. Ellis MW, Schlett CD, Millar EV, Crawford KB, Cui T, Lanier JB, et al. Prevalence of nasal colonization and strain concordance in patients with community-associated Staphylococcus aureus skin and soft-tissue infections. Infect Control Hosp Epidemiol. 2014;35:1251–6. https://doi.org/10.1086/678060.
6. Wertheim HF, Vos MC, Ott A, van Belkum A, Voss A, Kluytmans JA, et al. Risk and outcome of nosocomial *Staphylococcus aureus* bacteraemia in nasal carriers versus non-carriers. Lancet. 2004;364:703–5. https://doi.org/10.1016/s0140-6736(04)16897-9.
7. von Eiff C, Becker K, Machka K, Stammer H, Peters G. Nasal carriage as a source of *Staphylococcus aureus* bacteremia. Study Group. N Engl J Med. 2001;344:11–6. https://doi.org/10.1056/nejm200101043440102.
8. Mermel LA, Cartony JM, Covington P, Maxey G, Morse D. Methicillin-resistant *Staphylococcus aureus* colonization at different body sites: a prospective, quantitative analysis. J Clin Microbiol. 2011;49:1119–21. https://doi.org/10.1128/jcm.02601-10.
9. Yang ES, Tan J, Eells S, Rieg G, Tagudar G, Miller LG. Body site colonization in patients with community-associated methicillin-resistant *Staphylococcus aureus* and other types of *S. aureus* skin infections. Clin Microbiol Infect. 2010;16:425–31. https://doi.org/10.1111/j.1469-0691.2009.02836.x.
10. Singh J, Johnson RC, Schlett CD, Elassal EM, Crawford KB, Mor D, et al. Multi-body-site microbiome and culture profiling of military trainees suffering from skin and soft tissue infections at fort benning, Georgia. mSphere. 2016;1. https://doi.org/10.1128/mSphere.00232-16.
11. Azarian T, Daum RS, Petty LA, Steinbeck JL, Yin Z, Nolan D, et al. Intrahost evolution of methicillin-resistant *Staphylococcus aureus* USA300 among individuals with reoccurring skin and soft-tissue infections. J Infect Dis. 2016;214:895–905. https://doi.org/10.1093/infdis/jiw242.
12. Millar EV, Rice GK, Elassal EM, Schlett CD, Bennett JW, Redden CL, et al. Genomic characterization of USA300 MRSA to evaluate intraclass transmission and recurrence of SSTI among high risk military trainees. Clin Infect Dis. 2017. https://doi.org/10.1093/cid/cix327.
13. Biodefense and Emerging Infections Research Resources Repository, Manassas, VA. beiresources.org.
14. FastQC. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 2 Nov 2017.
15. Joshi NA. Sickle. A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). Available at https://github.com/najoshi/sickle. 2011. Accessed 2 Nov 2017.

16. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. https://doi.org/10.1089/cmb.2012.0021.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10. https://doi.org/10.1016/s0022-2836(05)80360-2.
18. BAGA. https://github.com/daveuu/baga. Accessed 2 Nov 2017.
19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60. https://doi.org/10.1093/bioinformatics/btp324.
20. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303. https://doi.org/10.1101/gr.107524.110.
21. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21. https://doi.org/10.1093/sysbio/syq010.
22. Albrecht VS, Limbago BM, Moran GJ, Krishnadasan A, Gorwitz RJ, McDougal LK, et al. *Staphylococcus aureus* colonization and strain type at various body sites among patients with a closed abscess and uninfected controls at US emergency departments. J Clin Microbiol. 2015;53:3478–84. https://doi.org/10.1128/jcm.01371-15.
23. Frazee BW, Lynn J, Charlebois ED, Lambert L, Lowery D, Perdreau-Remington F. High prevalence of methicillin-resistant *Staphylococcus aureus* in emergency department skin and soft tissue infections. Ann Emerg Med. 2005;45:311–20. https://doi.org/10.1016/j.annemergmed.2004.10.011.
24. Kumar N, David MZ, Boyle-Vavra S, Sieth J, Daum RS. High *Staphylococcus aureus* colonization prevalence among patients with skin and soft tissue infections and controls in an urban emergency department. J Clin Microbiol. 2015;53:810–5. https://doi.org/10.1128/jcm.03221-14.
25. Faden H, Lesse AJ, Trask J, Hill JA, Hess DJ, Dryja D, et al. Importance of colonization site in the current epidemic of Staphylococcal skin abscesses. Pediatrics. 2010;125:e618-24. https://doi.org/10.1542/peds.2009-1523.
26. Popovich KJ, Snitkin E, Green SJ, Aroutcheva A, Hayden MK, Hota B, et al. Genomic epidemiology of USA300 methicillin-resistant *Staphylococcus aureus* in an urban community. Clin Infect Dis. 2016;62:37–44. https://doi.org/10.1093/cid/civ794.
27. Golubchik T, Batty EM, Miller RR, Farr H, Young BC, Larner-Svensson H, et al. Within-host evolution of *Staphylococcus aureus* during asymptomatic carriage. PloS One. 2013;8:e61319. https://doi.org/10.1371/journal.pone.0061319.
28. Lemmens N, van Wamel W, Snijders S, Lesse AJ, Faden H, van Belkum A. Genomic comparisons of USA300 *Staphylococcus aureus* colonizating the nose and rectum of children with skin abscesses. Microb Pathog. 2011;50:192–9. https://doi.org/10.1016/j.micpath.2010.12.006.
29. Price JR, Golubchik T, Cole K, Wilson DJ, Crook DW, Thwaites GE, et al. Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of *Staphylococcus aureus* in an intensive care unit. Clin Infect Dis. 2014;58:609–18. https://doi.org/10.1093/cid/cit807.
30. Stine OC, Burrowes S, David S, Johnson JK, Roghmann MC. Transmission clusters of methicillin-resistant *Staphylococcus Aureus* in long-term care facilities based on whole-genome sequencing. Infect Control Hosp Epidemiol. 2016;37:685–91. https://doi.org/10.1017/ice.2016.41.
31. Campbell KM, Vaughn AF, Russell KL, Smith B, Jimenez DL, Barrozo CP, et al. Risk factors for community-associated methicillin-resistant *Staphylococcus aureus* infections in an outbreak of disease among military trainees in San Diego, California, in 2002. J Clin Microbiol. 2004;42:4050–3. https://doi.org/10.1128/jcm.42.9.4050-4053.2004.
32. Mollema FP, Richardus JH, Behrendt M, Vaessen N, Lodder W, Hendriks W, et al. Transmission of methicillin-resistant *Staphylococcus aureus* to household contacts. J Clin Microbiol. 2010;48:202–7. https://doi.org/10.1128/jcm.01499-09.