Heliyon

ELSEVIER

Received:
13 February 2018
Revised:
18 May 2018
Accepted:
26 June 2018

Cite as: Angela Collins, Edward K. Wakeland, Prithvi Raj, Min S. Kim, Jiwoong Kim, Naureen G. Tareen, Lawson A. B. Copley. The impact of *Staphylococcus aureus* genomic variation on clinical phenotype of children with acute hematogenous osteomyelitis. Heliyon 4 (2018) e00674. doi: 10.1016/j.heliyon.2018.e00674

The impact of *Staphylococcus aureus* genomic variation on clinical phenotype of children with acute hematogenous osteomyelitis

Angela Collins [a], **Edward K. Wakeland** [b], **Prithvi Raj** [b], **Min S. Kim** [c], **Jiwoong Kim** [c], **Naureen G. Tareen** [d,*], **Lawson A. B. Copley** [e,*]

[a] *McLaren Medical Center, Department of Orthopaedic Surgery, Flint, Michigan, USA*

[b] *University of Texas Southwestern, Department of Immunology, Dallas, Texas, USA*

[c] *University of Texas Southwestern, Department of Clinical Science, Dallas, Texas, USA*

[d] *Children's Health System of Texas, Department of Orthopaedic Surgery, Dallas, Texas, USA*

[e] *University of Texas Southwestern, Department of Orthopaedic Surgery, Dallas, Texas, USA*

[*] Corresponding authors.

E-mail addresses: Naureen.tareen@childrens.com (N.G. Tareen), Lawson.copley@childrens.com (L.A.B. Copley).

Abstract

Background: Children with acute hematogenous osteomyelitis (AHO) have a broad spectrum of illness ranging from mild to severe. The purpose of this study is to evaluate the impact of genomic variation of *Staphylococcus aureus* on clinical phenotype of affected children and determine which virulence genes correlate with severity of illness.

Methods: De novo whole genome sequencing was conducted for a strain of Community Acquired Methicillin Resistant *Staphylococcus aureus* (CA-MRSA), using PacBio Hierarchical Genome Assembly Process (HGAP) from 6 Single Molecule Real Time (SMRT) Cells, as a reference for DNA library assembly of 71 *Staphylococcus aureus* isolates from children with AHO. Virulence gene annotation was based on exhaustive literature review and genomic data in NCBI for *Staphylococcus aureus*. Clinical phenotype was assessed using a validated

https://doi.org/10.1016/j.heliyon.2018.e00674

severity score. Kruskal-Wallis rank sum test determined association between clinical severity and virulence gene presence using False Discovery Rate (FDR), significance <0.01.

Results: PacBio produced an assembled genome of 2,898,306 bp and 2054 Open Reading Frames (ORFs). Annotation confirmed 201 virulence genes. Statistical analysis of gene presence by clinical severity found 40 genes significantly associated with severity of illness (FDR ≤ 0.009). MRSA isolates encoded a significantly greater number of virulence genes than did MSSA ($p < 0.0001$). Phylogenetic analysis by maximum likelihood (PAML) demonstrated the relatedness of genomic distance to clinical phenotype.

Conclusions: The *Staphylococcus aureus* genome contains virulence genes which are significantly associated with severity of illness in children with osteomyelitis. This study introduces a novel reference strain and detailed annotation of *Staphylococcus aureus* virulence genes. While this study does not address bacterial gene expression, a platform is created for future transcriptome investigations to elucidate the complex mechanisms involved in childhood osteomyelitis.

Keywords: Microbiology, Pediatrics, Pathology, Infectious disease, Genetics, Immunology

1. Introduction

Community acquired *Staphylococcus aureus* is the leading cause of acute hematogenous osteomyelitis (AHO) among children worldwide [1, 2, 3]. There is substantial variation in clinical presentation and response to treatment of affected children ranging in spectrum from mild to severe [3, 4, 5, 6, 7]. Some children require only a short course of antibiotics and brief hospitalization without the need for surgery, whereas others require intensive care, undergo multiple surgical procedures, experience prolonged hospitalization, and receive long-term antibiotic therapy [7, 8]. Differences in bacterial virulence behavior may contribute to this variation in clinical presentation among these otherwise healthy, immunologically competent children [9, 10, 11, 12].

Previous studies have attempted to elucidate the pathogenetic mechanisms of *Staphylococcus aureus* virulence based on a variety of candidate genes, in isolation or combination, which may play a role in disease pathogenesis [13, 14, 15, 16, 17, 18, 19]. The advent of next generation sequencing allows for a greater discriminatory power to aid in this investigation, however, there are current limitations posed by the lack of a suitable reference strain for DNA library assembly. The majority of clinical isolates for which whole genome sequencing data is publicly available in the GenBank database of the National Center for Biotechnology Information (NCBI)

https://doi.org/10.1016/j.heliyon.2018.e00674

have been isolated from skin and soft tissue infections or from adults [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. Of the two reference strains which are commonly used for USA300 methicillin-resistant *Staphylococcus aureus* (MRSA) research, TCH1516 and FPR3757, neither are ideal to study AHO [32, 33]. TCH1516 was obtained from an adolescent describe as having "severe sepsis syndrome" without osteomyelitis [32]. FPR3757 was isolated from the wrist abscess of a 36-year-old HIV-positive white male with a history of injection drug use from San Francisco, California [33]. In a recent study of *Staphylococcus aureus* isolates obtained from children with AHO within a single community, phylogenetic analysis utilizing TCH1516 and FPR3757 as reference strains demonstrated substantial genetic distance of the all study isolates from either reference strain [9].

The purpose of this study is to evaluate the impact of genomic variation of *Staphylococcus aureus* on clinical phenotype of affected children and determine which virulence genes correlate with severity of illness. A secondary purpose is to carefully annotate candidate virulence genes within the *Staphylococcus aureus* genome following de novo assembly of the whole genome of a novel isolate of MRSA which caused severe AHO in a child in the Southwestern United States.

2. Materials and methods

2.1. Ethics statement

This study was conducted according to the principles expressed in the Declaration of Helsinki. The study was approved by the Institutional Review Board of the University of Texas Southwestern Medical Center Dallas and Children's Medical Center of Dallas (IRB# STU 032013-022).

2.2. Clinical phenotype of the child from whom the reference strain was isolated

UTSW55 was isolated from an otherwise healthy 12 year old female with osteomyelitis of the right proximal femur, septic arthritis of the right hip, and adjacent pyomyositis. She had sustained a recent minor fall which was followed by fever and pain in the hip region. She was hospitalized in septic shock with disseminated infection (bacteremia, deep venous thrombosis of the right iliac vein, septic pulmonary emboli, pneumonia and a urinary tract infection). She had eight consecutive days of positive blood cultures. Cultures were positive for clindamycin-sensitive MRSA. She was treated with five days of inotropic support, two weeks of mechanical ventilation and five surgical procedures. Her initial hospitalization was 51 days, 19 of which were intensive care. A total hip replacement was ultimately performed due to avascular necrosis, chondrolysis, and a pathologic femoral neck fracture. Severity of illness scored (Table 1) as 10 out of 10 (average 2.9, standard deviation

https://doi.org/10.1016/j.heliyon.2018.e00674

Table 1. Modified severity of illness scoring system for AHO [5, 6].

Scoring parameter	Criteria	Points	Total
Initial CRP (mg/dL)	>15	2	0−2
	10 to 15	1	
	<10	0	
CRP hospital day 4−5 (mg/dL)	>10	2	0−2
	5 to 10	1	
	<5	0	
CRP hospital day 2−3 (mg/dL)	>10	2	0−2
	5 to 10	1	
	<5	0	
Band percentage of WBC	≥1.5%	1	0−1
	<1.5%	0	
Febrile days on antibiotic	≥2	1	0−1
	<2	0	
ICU admission	Yes	1	0−1
	No	0	
Disseminated disease[a]	Yes	1	0−1
	No	0	
Total severity of illness score			0−10

[a] DVT, septic pulmonary embolism, pneumonia, endocarditis, multi-focal infection.

3.2) [5, 6]. Health Information Management (HIM) scoring methods determined a Medicare Severity Diagnosis Related Group (MS DRG) score of 5.4 (average 1.7, range 0.7−5.5, standard deviation 0.9); All Patient Refined Diagnosis Related Groups Case Mix Index (APR-DRG CMI) of 5.8 (average 1.7, range 0.6−8.3, standard deviation 1.4); Severity of Illness (SOI) score of 4.0 (average 2.3, range 1.0−4.0, standard deviation 0.7); and Risk of Mortality (ROM) score of 4.0 (average 1.3, range 1.0−4.0, standard deviation 0.8).

2.3. Study population

Previously healthy children birth to 18 years who were admitted and confirmed to have AHO due to methicillin-sensitive or methicillin-resistant *Staphylococcus aureus* (MSSA or MRSA, respectively) between January 1, 2009 and March 1, 2014 were retrospectively studied. AHO was defined as infection diagnosed within 2 weeks of the onset of symptoms. Diagnosis was confirmed by magnetic resonance imaging (MRI), positive cultures, and histopathology of bone obtained from the site of infection. Bacterial isolates from infected bone were obtained under sterile conditions in the operating room, plated on sheep blood agar, and incubated at 37°. Bacterial isolates were then catalogued and stored at −80 °C in the microbiology lab. Excluded from the study were children with underlying medical conditions associated with immunocompromise, children with AHO caused by any other bacterial organism, culture-negative cases, and children with history of surgery or penetrating inoculation as the cause of infection.

https://doi.org/10.1016/j.heliyon.2018.e00674

2.4. Clinical data

Clinical and laboratory data were collected to calculate severity of illness for each child using the previously validated method based on CRP values (mg/dL) on admission, at 48 hours, and at 96 hours; febrile days on antibiotics; band percentage of the white blood cell count (WBC); ICU admission; and evidence of disseminated disease (multi-focal infection, deep venous thrombosis, septic pulmonary embolism, pneumonia, or endocarditis) (see Table 1) [5, 6]. Tables 2a, 2b and 3 provide relevant details of the clinical disease of the affected children, aggregated into groups according to severity of illness and antibiotic resistance of the causative organism.

2.5. Specimen processing and PacBio of UTSW55

A frozen specimen of UTSW55 was thawed and inoculated on blood-agar. Isolated colonies were grown overnight in Brain Heart Infusion (BHI) medium at 37 °C in nephlo flasks. Genomic DNA was extracted with DNeasy blood and tissue kits (Qiagen, Germantown, MD) with lysostaphin replacing lysozyme. Quality of the genomic DNA was assessed on 1% agarose gels and measured with the use of Picogreen. A total of 4.4 μg DNA was sent to the sequencing service at the University of Washington (https://pacbio.gs.washington.edu/). De novo whole genome sequencing and assembly was conducted for this isolate using the PacBio Hierarchical Genome Assembly Process (HGAP) from 6 Single Molecule Real Time Sequencing (SMRT) Cells.

For all other isolates, frozen specimens were thawed and inoculated on blood-agar at 37° for 24 hours. Isolated colonies were transferred to enriched Mueller Hinton Broth Media overnight. DNeasy Blood and tissue kits (Qiagen) were used for DNA extraction, with lysostaphin replacing lysozyme.

2.6. Next generation sequencing of clinical isolates

A range of 2−5 ug of isolated genomic DNA, quantified by Picogreen, was submitted to the UT Southwestern Genomics Core (see http://genomics.swmed.edu/) from each study isolate for sequence analysis. Sequencing libraries were produced using TruSeq adaptors (Illumina, Inc., San Diego, CA) with inserts of 300−500 bp in length. DNA quality was analyzed with a Bioanalyzer (Agilent technologies). Sequencing was performed on an Illumina HiSeq 2000 using a paired-end protocol in which 100 bp sequences were obtained from each end of the sequencing library fragments. An average of 6.9 million reads was obtained for the study samples, for an average of 243-fold coverage of the *Staphylococcus aureus* genome. Sequencing protocols and basic information concerning the sequence generation are available from the UT Southwestern Genomics Core website.

https://doi.org/10.1016/j.heliyon.2018.e00674

Table 2a. Comparisons of clinical isolates (Grouped by severity of illness and antibiotic resistance) according to clinical parameters of the affected children.

Variable	Summary statistics							
	1 (MSSA 0−5)		2 (MSSA 6−10)		3 (MRSA 0−5)		4 (MRSA 6−10)	
	N	Mean ± std, med (range)	N	Mean ± std, med (range)	N	Mean ± std, med (range)	N	Mean ± std, med (range)
LOS (days)	20	5.72 ± 3.22, 4.85 (2, 16.54)	7	14.16 ± 7.18, 10.08 (8, 27)	13	6.34 ± 2.03, 7 (2.43, 8.64)	30	22.63 ± 11.14, 21.03 (7.75, 51)
TLOS	20	5.87 ± 3.11, 5 (2.69, 16.54)	7	14.3 ± 7.35, 10.08 (8, 27)	13	6.34 ± 2.03, 7 (2.43, 8.64)	30	24.39 ± 14.22, 21.03 (7.75, 69)
ICU LOS	20	0.15 ± 0.67, 0 (0, 3)	7	1.14 ± 1.57, 0 (0, 4)	13	0 ± 0, 0 (0, 0)	30	5.8 ± 7.12, 1.5 (0, 22)
Days of persistent bacteremia	20	0.65 ± 0.99, 0 (0, 4)	7	3.14 ± 1.86, 4 (1, 6)	13	1.08 ± 1.04, 1 (0, 3)	30	3.83 ± 3.24, 3 (0, 14)
CRP initial (mg/dL)	20	6.42 ± 5.48, 4.58 (0.68, 21.9)	7	25.87 ± 8.08, 24.2 (14.5, 38.2)	13	8.13 ± 5.53, 7 (0.1, 16.2)	30	25.38 ± 10.17, 23 (9.8, 55.6)
CRP 48 hours (mg/dL)	20	5 ± 2.84, 4.74 (0.88, 13)	7	22.63 ± 7.32, 21.9 (10.5, 33)	13	5.82 ± 3.26, 5.1 (0.1, 10.2)	30	23.63 ± 10.76, 20.55 (11.8, 67.7)
CRP 96 hours (mg/dL)	20	2.05 ± 1.69, 2.1 (0, 7.2)	7	16.49 ± 8.28, 16 (6.3, 32.9)	13	2.91 ± 2.14, 2.6 (0, 7)	30	19.63 ± 9.39, 19.9 (2.5, 39.7)
Band % of white blood cell count	20	1.42 ± 2.83, 0 (0, 12)	7	14.06 ± 7.56, 12.2 (2.5, 27)	13	2.64 ± 5.17, 0 (0, 17.4)	30	14.38 ± 15.73, 10.6 (0, 64)
Febrile days on Antibiotic	20	1.0 ± 1.34, 0.5 (0, 5)	7	2.29 ± 0.76, 2 (1, 3)	13	0.54 ± 0.88, 0 (0, 3)	30	7.93 ± 7.64, 6 (0, 41)
Number of surgeries	20	0.85 ± 0.59, 1 (0, 2)	7	1.43 ± 0.79, 2 (0, 2)	13	1 ± 0.71, 1 (0, 2)	30	2.37 ± 1.43, 2 (0, 6)
Days until CRP <2.0 mg/dL	20	5.25 ± 3.08, 5 (0, 11)	7	14.0 ± 5, 14 (7, 20)	13	6.31 ± 3.15, 6 (0, 14)	30	24.93 ± 24.99, 17.5 (8, 141)
Antibiotic duration (days)	20	38.35 ± 13.76, 34.5 (20, 74)	7	63.29 ± 40.65, 48 (28, 138)	13	53 ± 33.11, 42 (28, 151)	30	73.03 ± 52.73, 48.5 (21, 265)
SIS	20	1.45 ± 1.47, 1 (0, 5)	7	8.0 ± 1.15, 8 (6, 9)	13	1.77 ± 1.48, 2 (0, 4)	30	8.33 ± 1.49, 8 (4, 10)

One-way ANOVA was used to compare 4 groups, followed by post-hoc Tukey (which accounts for multiple comparisons). Likewise, a nonparametric Kruskal Walis test was used to compare 4 groups.
LOS (length of stay); TLOS (total length of stay including readmission days); ICU LOS (length of intensive care stay); CRP (C-Reactive Protein); SIS (severity of illness score).

https://doi.org/10.1016/j.heliyon.2018.e00674

Table 2b. Statistical significance of the differences between the cohorts comparing the variables in Table 2a.

ANOVA P-value	Posthoc Tukey test						Kruskal P-value
	1 vs.			2 vs.		3 vs.	
	2	3	4	3	4	4	
0.0000	0.0829	0.9962	0.0000	0.1636	0.0626	0.0000	0.000
0.0000	0.2191	0.9991	0.0000	0.3212	0.0799	0.0000	0.000
0.0002	0.9642	0.9997	0.0006	0.9558	0.1008	0.0027	0.000
0.0000	0.0790	0.9551	0.0001	0.2396	0.8937	0.0036	0.000
0.0000	0.0000	0.9336	0.0000	0.0001	0.9989	0.0000	0.000
0.0000	0.0000	0.9908	0.0000	0.0001	0.9897	0.0000	0.000
0.0000	0.0001	0.9847	0.0000	0.0004	0.6914	0.0000	0.000
0.0002	0.0524	0.9894	0.0007	0.1301	0.9999	0.0107	0.000
0.0000	0.9405	0.9943	0.0001	0.8863	0.0518	0.0003	0.000
0.0000	0.6077	0.9790	0.0000	0.8272	0.1660	0.0015	0.000
0.0004	0.6364	0.9980	0.0007	0.7621	0.4121	0.0073	0.000
0.0335	0.4988	0.7384	0.0205	0.9478	0.9389	0.4455	0.012
0.0000	0.0000	0.9268	0.0000	0.0000	0.9475	0.0000	0.000

2.7. Genome assembly and gene prediction

Genome assembly for the isolates was performed using the A5-miseq microbial genomic assembly pipeline specifically designed for Illumina sequencing data [34]. The assembled isolates had an average contig count of 48 with an N50 length of 607,387. Assembled genomes were then run through the gene prediction tool Glimmer3 (https://ccb.jhu.edu/software/glimmer/) before annotation.

2.8. Gene presence and annotation

To assess the presence of genes across *Staphylococcus aureus* isolates, reference sequences for genes were retrieved from the NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene) using the taxonomy filter for *Staphylococcus aureus* (ID: 1280). Exhaustive literature review was performed to identify a list of candidate virulence genes. The predicted genes were then aligned to the reference database using the blastp functionality within the DIAMOND program [35]. Due to various levels of sequence similarity between genes, only sequence alignment results with similarity identity scores greater than 0.8 were selected as positive matches. No other filtering criterion was applied to determine presence or absence of gene for a particular isolate. Despite the presence of multiple alternate gene names within the NCBI Gene database, the detailed annotation effort resulted in a unique list of 201 genes, identified within the *Staphylococcus aureus* species, to which predicted protein sequences could be unambiguously assigned.

https://doi.org/10.1016/j.heliyon.2018.e00674

Table 3. Comparison of clinical severity parameters between cohorts of antibiotic resistance and high versus low severity scores.

Variable	MSSA 0–5 (n = 20)	MSSA 6–10 (n = 7)	MRSA 0–5 (n-13)	MRSA 6–10 (n = 30)	*p-value
Readmission (rate)	1 (5%)	1 (14.3%)	0 (0%)	8 (26.7%)	0.060
ICU admission (rate)	1 (5%)	3 (42.9%)	0 (0%)	17 (56.7%)	0.000
Bacteremia (rate)	9 (45%)	7 (100%)	8 (61.5%)	26 (86.7%)	0.001
Surgery (rate)	15 (75%)	6 (85.7%)	10 (76.9%)	27 (90%)	0.071
DVT (rate)	0 (0%)	0 (0%)	0 (0%)	29 (96.7%)	0.001
SPE or Pneumonia (rate)	0 (0%)	0 (0%)	0 (0%)	11 (36.7%)	0.011

A Fisher's exact test (*) was conducted to test if there is association among groups and variables of interest.
ICU (intensive care unit); DVT (deep vein thrombosis); SPE (septic pulmonary embolism).

2.9. Identification of genes associated with severity of illness

The Kruskal-Wallis rank sum test was used to measure the significance of association between the severity of illness scores of affected children and presence or absence of genes within each genome of the associated causative organism [36]. Calculated p-values were further adjusted using False Discovery Rate (FDR) and only genes with adjusted p-values less than 0.01 were considered to be significant [37].

2.10. Phylogenetic tree analysis

Using the single contig of UTSW55 as reference, SNP calling was performed using MUMmer [38]. The haplotypes for each of the study samples were generated with confident homozygous single nucleotide variant (SNV) sites for which the read fractions supporting major alleles were greater than 0.9 for every sample. These haplotypes were used as the nucleotide input for MEGA 6.06 (Molecular Evolutionary Genetics Analysis http://www.megasoftware.net/mega.php) for phylogenetic analysis. Maximum Likelihood (ML) phylogenetic reconstruction was performed with the Jukes-Cantor model using the ML Heuristic Method of Nearest-Neighbor-Interchange (NNI). The initial tree was assembled using maximum parsimony to estimate the evolutionary distances between sequences by computing the proportion of nucleotide differences between each pair of sequences (Fig. 4).

3. Results

3.1. Patient population and reference strain

The clinical severity of illness scores of the study population ranged from 0 to 10 (mean 5.3 ± 3.7). Denovo assembly of UTSW55 using Pacbio HGAP on sequencing reads from 6 SMRT Cells yielded a prefilter read count of 318,329 and mean read length of 5,750. The final assembly of the *Staphylococcus aureus* chromosome

https://doi.org/10.1016/j.heliyon.2018.e00674

had a contig count of 1 with N50 contig length of 2,898,306, plus 4 non-overlapping contiguous sequences with an average length of 35,372 bp, most likely representing plasmids. There were 2,054 assigned ORF's within the UTSW55 reference genome.

3.2. Determination of gene presence and annotation

Extensive review of the literature identified a list of potential virulence genes of *Staphylococcus aureus* (supplementary table S1). These putative virulence genes were then divided into the following categories: adhesins, immune evasion proteins, proteases, regulatory proteins, toxins, and other proteins. This resulted in the identification of 201 unique virulence genes. UTSW55 was found to contain 148 of the 211 putative virulence genes (Table 4), including 22 involved in global virulence regulation and 36 used to evade the immune response. The presence and absence of virulence genes was assessed for each of the study isolates as well as the GenBank strains for which whole genome sequence data is publicly reported (supplementary table S2). Fig. 1 depicts a heat map of the alphabetized list of these *Staphylococcus aureus* virulence genes and their presence or absence among the GenBank and study isolates.

Table 4. Virulence genes of UTSW55 by category.

Category	Gene name
Adhesions	clfA, clfB, ebpS, fbpA, fnbA, fnbB, icaA, icaB, icaC, pls, sdrC, sdrE, srtA
Antibiotic resistance	dltA, ileS
Immune evasion	adsA, aur, capA, capB, capC, capD, capE, capF, capG, capH, capl, capJ, capK, capL, capM, capN, capO, capP, coa, ebhB, hlgA, hlgB, hlgC, lukE, lukF-PV, lukG, lukH, lukS-PV, map, mprF, oatA, sbi, scpA, ssl11nm, ssl6nm, ssl7nm
Other	arlR, arlS, arsB, atl, epiB, epiC, fhuB, fhuD, fmt, geh, graR, graS, htrA, hysA, isaA, isdA, lip, lpi11, lpl2, lpl3, lpl4, lpl7, lpl8, lytR, lytS, nuc, phoU, putP, rot, rsbU, sceD, scrA, sigB
Protease	opp3A, opp3C, opp3D opp3F, sak, splA, splB, splC, splF, sspA, sspB
Regulatory	agrA, agrB, agrC, arcA, arcB, arc, arcD, arcR, clpC, clpP, clpX, codY, cshA, essC, saeQ, saeR, saeS, sarS, sarU, sarZ, srrA, srrB
Stress resistance	ahpC, kata, ksgA, plc
Toxin	ear, eta, hla, hlb, sek, seq, set6, set7, set9, set10, set12, set13, tst

https://doi.org/10.1016/j.heliyon.2018.e00674

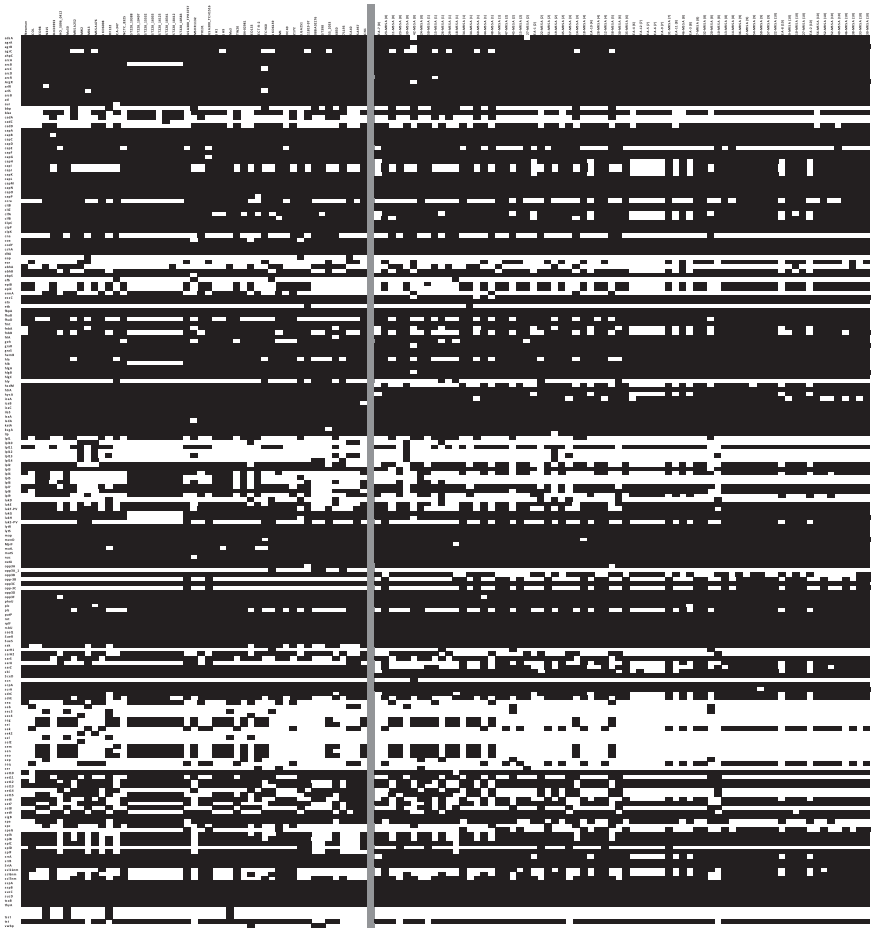

Fig. 1. Heatmap of virulence genes (alphabetical) listed vertically and *Staphylococcus aureus* isolates of GenBank in chronologic order from 1952 to 2013 (left of the vertical black line) or UTSW isolates in order of severity from 0 to 10 (right of the vertical black line) listed horizontally. White = absent, Black = present.

3.3. Identification of genes associated with severity of illness

Table 5 lists the 40 genes which were significantly associated with severity of illness of the affected children ($p < 0.01$). MRSA isolates were found to encode a significantly greater number of virulence genes ($p < 0.0001$) than did MSSA isolates (Fig. 2A). However, no significant difference in the number of virulence genes was identified when comparison was made between MRSA or MSSA isolates causing mild versus severe illness (Fig. 2B). MRSA isolates demonstrated greater numbers of adhesion, immune evasion, proteases, regulatory, and other protein virulence genes (Fig. 3). There was no difference in the number of toxin genes present between MSSA and MRSA isolates. For each virulence gene class, the genes which were present demonstrated greater variation when comparison was made between MRSA and MSSA than when comparison was made between isolates which had caused mild versus severe illness phenotypes.

https://doi.org/10.1016/j.heliyon.2018.e00674

Table 5. Virulence genes significantly associated with severity of illness among 71 children with acute osteomyelitis.

Gene	p-value	Adjusted p-value	Gene	p-value	Adjusted p-value	Gene	p-value	Adjusted p-value	Gene	p-value	Adjusted p-value
capE	0.000	0.000	**sarU**	0.000	0.000	**set13**	0.000	0.002	**splB**	0.001	0.004
seg	0.000	0.000	**sek**	0.000	0.000	**hla**	0.000	0.002	**ccra**	0.001	0.004
sei	0.000	0.000	**seq**	0.000	0.000	**lpl1**	0.000	0.002	**cna**	0.001	0.004
sem	0.000	0.000	**set11**	0.000	0.000	**agrC**	0.001	0.003	**epiC**	0.001	0.004
sen	0.000	0.000	**ear**	0.000	0.000	**lukE**	0.001	0.003	**lpl4**	0.001	0.004
seo	0.000	0.000	**ssl11nm**	0.000	0.000	**hsdM**	0.001	0.004	**set7**	0.001	0.004
set14	0.000	0.000	**sarH2**	0.000	0.000	**pls**	0.001	0.004	**ssl6nm**	0.001	0.004
set15	0.000	0.000	**lukF-PV**	0.000	0.001	**set12**	0.001	0.004	**cadA**	0.002	0.007
spa	0.000	0.000	**bbp**	0.000	0.002	**speG**	0.001	0.004	**lpl11**	0.003	0.007
lukS-PV	0.000	0.000	**fhuD**	0.000	0.002	**splA**	0.001	0.004	**epiB**	0.003	0.009

https://doi.org/10.1016/j.heliyon.2018.e00674

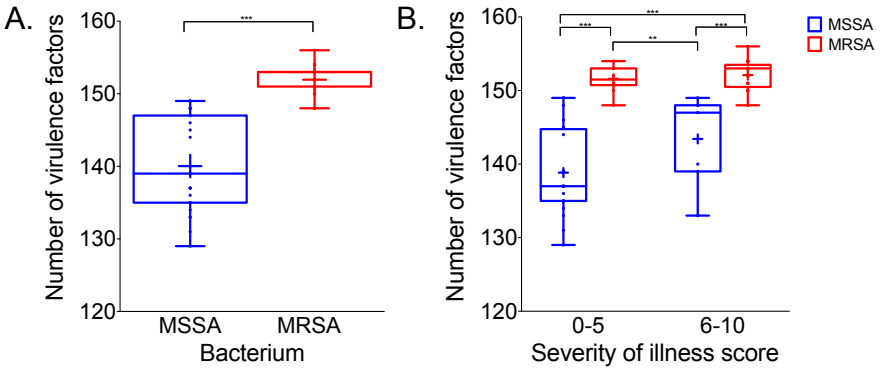

Fig. 2. Number of virulence genes identified among methicillin-sensitive (MSSA) or methicillin-resistant (MRSA) isolates (A). MRSA isolates had a significantly greater number of virulence genes. Number of virulence genes of MRSA or MSSA among isolates from children with severity of illness scores ranging from 0−5 compared with that of children with scores ranging from 6−10 (B). There was no significant difference of the number of virulence genes between severity categories.

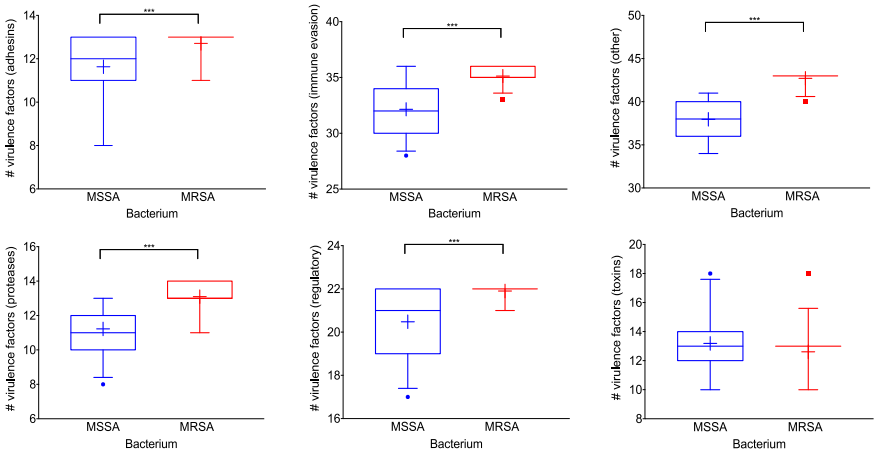

Fig. 3. MRSA isolates were found to have greater numbers of adhesion, immune evasion, protease, regulatory, and other virulence genes than were found in MSSA isolates.

3.4. Phylogenetic analysis of UTSW55

Phylogenetic Analysis by maximum likelihood (PAML) depicts the relatedness of UTSW55 to the other clinical isolates from children with AHO caused by either MRSA or MSSA within our community. Construction of an unrooted radial dendrogram (Fig. 4) suggests a visible trend of higher severity of illness among predominantly MRSA isolates at one end of the dendrogram in comparison to lower severity of illness scores among predominantly MSSA isolates at the opposite end, with scattered exceptions of severity magnitude and antibiotic resistance throughout the dendrogram (Fig. 4). Using sequences publicly available in GenBank (supplementary table S1), PAML was used to depict the genetic distance of the study isolates to GenBank sequences.

https://doi.org/10.1016/j.heliyon.2018.e00674

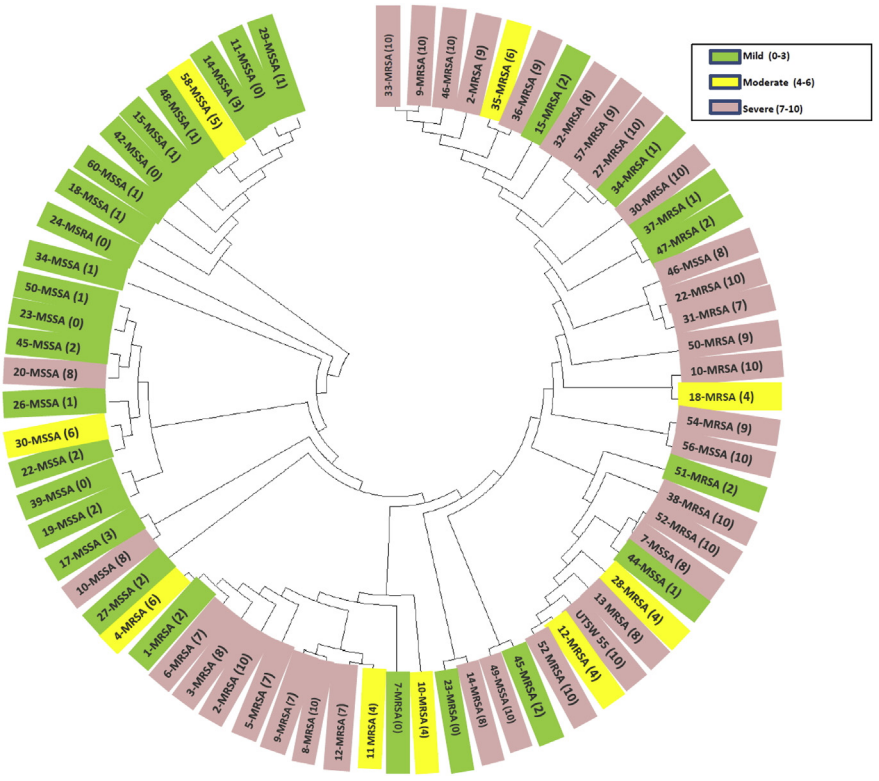

Fig. 4. Radial dendrogram depicting the genetic distance between isolates of children with acute hematogenous osteomyelitis caused by *Staphylococcus aureus*. Severity of illness scores of the affected children, ranging from 0 = mild to 10 = severe, are provided in parentheses for each isolate. Color coding of the isolates by severity of illness score is provided for visual reference (green − score 0 to 3 or mild; yellow − score 4 to 6 or moderate; and red − score 7−10 or severe).

4. Discussion

The pathogenesis of AHO in children is related to the virulence capability of the causative organism, *Staphylococcus aureus* [3, 4, 7, 9, 10, 11, 12]. In order for otherwise healthy children to become infected, the organism must gain access to the circulation, survive in human blood, evade the host defenses, and adhere to deep tissues in a cascade of events supported by hemolysins, leukocidins, toxins, and adhesins [13, 14, 15, 16, 17, 18, 19]. Variation in the clinical spectrum of this disease is not surprising given the array of virulence genes within the organism's armamentarium which have been discovered to date [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. Following the initial publication of the gene sequence of N315 several investigations were undertaken to explore single gene determinants and genes in combination through polymerase chain reaction (PCR) technology [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. However, discovery has been slow due to the limited discriminatory power of these methods as well as the lack of a relevant reference strain for genomic studies [32, 33]. As new technologies have emerged, it is necessary to reassess the virulence

https://doi.org/10.1016/j.heliyon.2018.e00674

capability of *Staphylococcus aureus* using next generation tools with improved discriminatory power while utilizing sequence data of a relevant reference strain. Within this context, the addition of complete sequence data of UTSW55 to GenBank (Genbank: NZ_CP013231.1) (Fig. 5) provides necessary foundation for this investigation. Given the evolutionary nature of the *Staphylococcus aureus* genome, the study findings are enhanced by the geographic and chronologic proximity of the reference strain to the community isolates. This study represents the most exhaustive effort to date to accurately annotate the virulence genes of *Staphylococcus aureus*. More importantly, this study has identified 40 virulence genes which are significantly associated with severity of illness among children with AHO [5, 6].

As expected, the majority of MRSA isolates within our community are more closely related to USA 300 than to other available reference strains. An unexpected finding was the phylogenetic diversity among the MSSA strains. Other investigators have similarly found CA-MRSA to be relatively homogenous while MSSA isolates from their study population demonstrated greater heterogeneity [32]. The MRSA isolates of this study were found to possess more virulence genes than MSSA isolates. Recent clinical experience in several communities suggest that MRSA causes

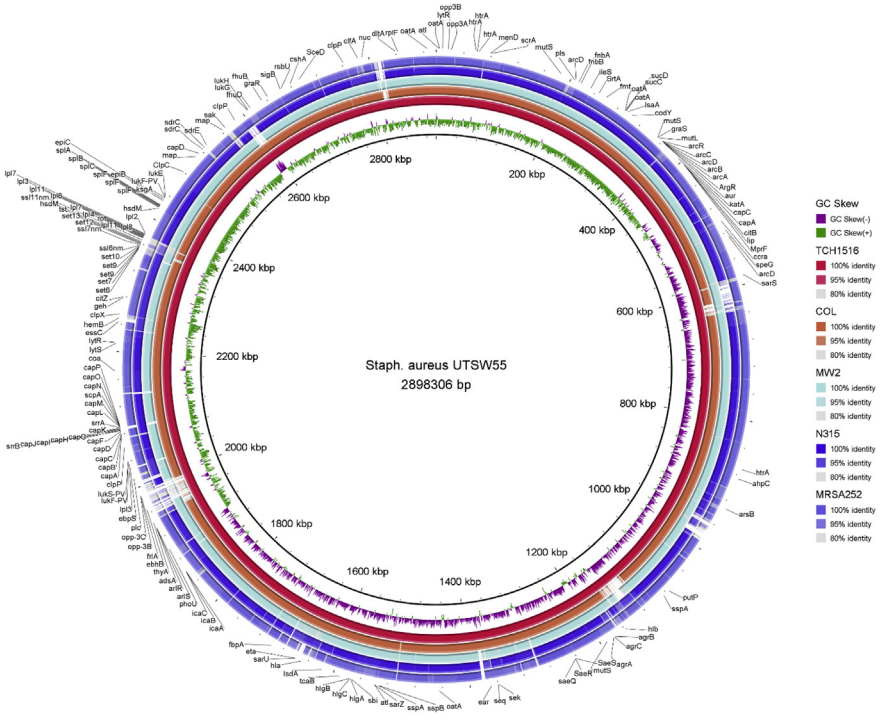

Fig. 5. Alignment results of commonly used *Staphylococcus aureus* genomes (TCH1516, COL, MW2, N315, and MRSA252) against the newly assembled UTSW55 reference genome. Annotation locations for 201 genes (Supplemental Table 1) are depicted in the outer circle along with GC content in the inner circle. Alignment identity cutoff of 0.8 was used to determine missing regions in the query genomes compared to the UTSW55 reference.

https://doi.org/10.1016/j.heliyon.2018.e00674

more severe clinical disease in children [4, 7]. Although this severity is attributed to the antibiotic resistance profile of the bacteria, a more plausible explanation is the increased number of virulence genes among MRSA isolates. The findings of this study support the idea that the antibiotic resistance of the isolate is not the primary driver of severity, but rather a bystander effect.

No single virulence gene has been conclusively identified as responsible for the invasiveness of *Staphylococcus aureus*. It is suspected that combinations of virulence genes may be involved [11]. Previous investigators have shown that the genes fnbA, cna, sdrE, sej, eta, hlg, and ica are more common among invasive isolates [11]. After an extensive search of the literature and thorough gene annotation, this study reports 201 virulence genes of *Staphylococcus aureus* (see Supplementary table S1). Among these are 40 which are significantly associated with severity of illness of children with AHO (Table 5). These specific genes will guide future studies of gene regulation (*agr*C), hemolysins (*hla*), immune evasion (*cap*E, *luk*S-PV, *luk*F-PV, *luk*E), and the Microbial Surface Components Recognizing Adhesive Matrix Molecultes (MSCRAMMs) including adhesive surface protein genes of collagen adhesion (*cna*) and bone sialo-protein binding (*bbp*). A substantial percentage of children with AHO are bacteremic at initial presentation which suggests that *Staphylococcus aureus* has acquired mechanisms by which to survive in human blood by liberating free iron through hemolysis (*hla*) and evading the host neutrophil defense (*cap*E, *luk*S-PV, *luk*F-PV, *luk*E). The organism then must attach to deep structures of collagen and bone (*can*, *bbp*) in order to establish the deep infection of bone.

Important limitations of this study include the lack of gene expression data or a mechanistic explanation of relationship of the *Staphylococcus aureus* genome to the clinical manifestations of infection. It is not possible to equate any of the virulence genes identified as specifically responsible for bone infections or surmise what cascade of events might be responsible for these infections by specific isolates. While the bacterial transcriptome was beyond the scope of this study, promising research using a murine osteomyelitis model demonstrates a differential gene expression of *Staphylococcus aureus* between *in vitro* and *in vivo* conditions [39, 40]. These investigators found 180 genes which were significantly expressed during acute infection with actions involved gluconeogenesis, proteolysis, iron acquisition, evasion of host immune defense and stress response [40].

This study represents the most current and extensive analysis of virulence genes of *Staphylococcus aureus* to date based on extensive literature review, detailed gene annotation, and phylogenetic analysis. By including all existing GenBank strains for which genomic sequence data is accessible, we have identified a comprehensive list of virulence genes and their penetration across 71 local and 49 remote *Staphylococcus aureus* strains. Utilizing a framework to objectively categorize severity of illness enhances the clinical relevance of this study specifically to children with

https://doi.org/10.1016/j.heliyon.2018.e00674

AHO. The addition of the whole genome of UTSW55 to GenBank (GenBank: NZ_CP013231.1) will facilitate future investigations of the genome, transcriptome and proteotome of *Staphylococcus aureus* isolates from children with AHO. An important question to be addressed in future investigations is whether there is any association between specific virulence factors and long-term clinical outcomes of disease, specifically recurrence of infection, osteonecrosis, growth arrest, or skeletal deformity.

Declarations

Author contribution statement

Angela Collins, Naureen G. Tareen: Performed the experiments.

Edward K. Wakeland: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Prithvi Raj: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Min S. Kim, Jiwoong Kim: Analyzed and interpreted the data.

Lawson A. B. Copley: Conceived and designed the experiments; Wrote the paper.

Funding statement

This work was supported by a grant from the Texas Scottish Rite Hospital for Children (TSRHC), grant # 02-14-1130. The bioinformatics analyses were funded by CPRIT. The funders had no role in study design, data collection, data analysis, interpretation and writing of this paper or in the decision to submit the paper for publication.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2018.e00674.

References

[1] S.R. Arnold, D. Elias, S.C. Buckingham, E.D. Thomas, E. Novais, A. Arkader, C. Howard, Changing patterns of acute hematogenous osteomyelitis and septic

arthritis: emergence of community associated methicillin-resistant *Staphylococcus aureus*, J. Pediatr. Orthop. 26 (6) (2006) 703−708.

[2] O.A. Gafur, L.A. Copley, S.T. Hollmig, R.H. Browne, L.A. Thornton, S.E. Crawford, The impact of the current epidemiology of pediatric musculoskeletal infection on evaluation and treatment guidelines, J. Pediatr. Orthop. 28 (7) (2008) 777−785.

[3] R.M. Klevens, M.A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L.H. Harrison, R. Lynfield, G. Dumyati, J.M. Townes, A.S. Craig, E.R. Zell, G.E. Fosheim, L.K. McDougal, R.B. Carey, S.K. Fridkin, Active Bacterial Core Surveillance (ABCs) MRSA Investigators, Invasive methicillin-resistant Staphylococcus aureus infections in the United States, J. Am. Med. Assoc. 298 (15) (2007) 1763−1771.

[4] J.J. Hawkshead 3rd, N.B. Patel, R.W. Steele, S.D. Heinrich, Comparative severity of pediatric osteomyelitis attributable to methicillin-resistant versus methicillin-sensitive Staphylococcus aureus, J. Pediatr. Orthop. 29 (1) (2009) 85−90.

[5] L.A. Copley, T. Barton, C. Garcia, D. Sun, C. Gaviria-Agudelo, W.T. Gheen, R.H. Browne, A proposed scoring system for assessment of severity of illness in pediatric acute hematogenous osteomyelitis using objective clinical and laboratory findings, Pediatr. Infect. Dis. J. 33 (1) (2014) 35−41.

[6] A.G. Athey, M.E. Mignemi, W.T. Gheen, E.A. Lindsay, C.H. Jo, L.A. Copley, Validation and modification of a severity of illness score for children with acute hematogenous osteomyelitis, J. Pediatr. Orthop. (2016 Oct 12) [Epub ahead of print].

[7] K.L. Ju, D. Zurakowski, M.S. Kocher, Differentiating between methicillin-resistant and methicillin-sensitive Staphylococus aureus osteomyelitis in children: an evidence-based clinical prediction algorithm, J. Bone Jt. Surg. Am. 93 (18) (2011) 1693−1701.

[8] D.A. Tuason, T. Gheen, D. Sun, R. Huang, L. Copley, Clinical and laboratory parameters associated with multiple surgeries in children with acute hematogenous osteomyelitis, J. Pediatr. Orthop. 34 (5) (2014) 565−570.

[9] C. Gaviria-Agudelo, C. Aroh, N. Tareen, E.K. Wakeland, M. Kim, L.A. Copley, Genomic heterogeneity of methicillin resistant *Staphylococcus aureus* associated with variation in severity of illness among children with acute hematogenous osteomyelitis, PLoS One 10 (6) (2015) e0130415.

[10] M.C. Wehrhahn, J.O. Robinson, E.M. Pascoe, G.W. Coombs, J.C. Pearson, F.G. O'Brien, H.L. Tan, D. New, P. Salvaris, R. Salvaris, R.J. Murray, Illness

https://doi.org/10.1016/j.heliyon.2018.e00674

severity in community-onset invasive Staphylococcus aureus infection and the presence of virulence genes, J. Infect. Dis. 205 (2012) 1840–1848.

[11] S.J. Peacock, C.E. Moore, A. Justice, M. Kantzanou, L. Story, K. Mackie, G. O'Neill, N.P.J. Day, Virulent combinations of adhesion and toxin genes in natural populations of *Staphylococcus aureus*, Infect. Immun. 70 (9) (2002) 4987–4996.

[12] S.R. Gill, L.M. McIntyre, C.L. Nelson, B. Remortel, T. Rude, L.B. Reller, V.G. Fowler Jr., Potential associations between severity of infection and the presence of virulence-associated genes in clinical strains of *Staphylococcus aureus*, PLoS One 6 (4) (2011) e18673.

[13] F. Alonzo III, M.A. Benson, J. Chen, R.P. Novick, B. Shopsin, V.J. Torres, *Staphylococcus aureus* leukocidin ED contributes to systemic infection by targeting neutrophils and promoting bacterial growth *in vivo*, Mol. Microbiol. 83 (2) (2012) 423–435.

[14] C.E. Bocchini, K.G. Hulten, E.O. Mason Jr., B.E. Gonzalez, W.A. Hammerman, S.L. Kaplan, Panton-valentine leukocidin genes are associated with enhanced inflammatory response and local disease in acute hematogenous *Staphylococcus aureus* osteomyelitis in children, An. Pediatr. 117 (2) (2006) 433–440.

[15] M.O. Elasri, J.R. Thomas, R.A. Skinner, J.S. Blevins, K.E. Beenken, C.L. Nelson, M.S. Smeltzer, *Staphylococcus aureus* collagen adhesion contributes to the pathogenesis of osteomyelitis, Bone 30 (1) (2002) 275–280.

[16] R.C. Massey, M.N. Kantzanou, T. Fowler, N.P.J. Day, K. Schofield, E.R. Wann, A.R. Berendt, M. Hook, S.J. Peacock, Fibronectin binding protein A of *Staphylococcus aureus* has multiple, substituting, binding regions that mediate adherence to fibronectin and invasion of endothelial cells, Cell Microbiol. 3 (12) (2001) 839–851.

[17] S.J. Peacock, N.P.J. Day, M.G. Thomas, A.R. Berendt, T.J. Foster, Clinical isolates of *Staphylococcus aureus* exhibit diversity in fnb genes and adhesion to human fibronectin, J. Infect. 41 (2000) 23–31.

[18] B. Loffler, M. Hussain, M. Grundmeier, M. Bruck, D. Holzinger, G. Varga, J. Roth, B.C. Kahl, R.A. Proctor, G. Peters, *Staphylococcus aureus* panton-valentine leukocidin is a very potent cytotoxic factor for human neutrophils, PLoS Pathog. 6 (1) (2010) e1000715.

[19] S. Tamber, D. Reyes, N.P. Donegan, J.D. Schwartzman, A.L. Cheung, G. Memmi, The *Staphylococcus*-specific gene *rsr* represses *agr* and virulence in *Staphylococcus aureus*, Infect. Immun. 78 (10) (2010) 4384–4391.

https://doi.org/10.1016/j.heliyon.2018.e00674

[20] R. Tewhey, C.R. Cannavino, J.A.D. Leake, V. Bansal, E.J. Topol, A. Torkamani, J.S. Bradley, N. Schork, Genetic structure of community acquired methicillin-resistant *Staphylococcus aureus* USA300, BMC Genom. 13 (2012) 508.

[21] U. Nubel, J. Dordel, K. Kurt, et al., A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant Staphylococcus aureus, PLoS Pathog. 6 (4) (2010) e1000855.

[22] M. Fraunholz, J. Bernhardt, J. Schuldes, R. Daniel, M. Hecker, B. Sinha, Complete genome sequence of Staphylococcus aureus 6850, a highly cytotoxic and clinically virulent methicillin-sensitive strain with distant relatedness to prototype strains, Genome Announc. 1 (5) (2013).

[23] Y. Chen, S.S. Chatterjee, S.F. Porcella, Y.S. Yu, M. Otto, Complete genome sequence of a Panton-Valentine leukocidin-negative community-associated methicillin-resistant Staphylococcus aureus strain of sequence type 72 from Korea, PLoS One 8 (8) (2013) e72803.

[24] M.T. Holden, L.Y. Hsu, K. Kurt, et al., A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic, Genome Res. 23 (4) (2013) 653−664.

[25] M.M. Mwangi, S.W. Wu, Y. Zhou, et al., Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing, Proc. Natl. Acad. Sci. U. S. A. 104 (22) (2007) 9451−9456.

[26] K. Chua, T. Seemann, P.F. Harrison, et al., Complete genome sequence of Staphylococcus aureus strain JKD6159, a unique Australian clone of ST93-IV community methicillin-resistant Staphylococcus aureus, J. Bacteriol. 192 (20) (2010) 5556−5557.

[27] T.W. Huang, F.J. Chen, W.C. Miu, et al., Complete genome sequence of Staphylococcus aureus M013, a pvl-positive, ST59-SCCmec type V strain isolated in Taiwan, J. Bacteriol. 194 (5) (2012) 1256−1257.

[28] M.T. Holden, E.J. Feil, J.A. Lindsay, et al., Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance, Proc. Natl. Acad. Sci. U. S. A. 101 (26) (2004) 9786−9791.

[29] T. Baba, F. Takeuchi, M. Kuroda, et al., Genome and virulence determinants of high virulence community-acquired MRSA, Lancet (London, England) 359 (9320) (2002) 1819−1827.

[30] M. Kuroda, T. Ohta, I. Uchiyama, et al., Whole genome sequencing of meticillin-resistant Staphylococcus aureus, Lancet (London, England) 357 (9264) (2001) 1225−1240.

 https://doi.org/10.1016/j.heliyon.2018.e00674

[31] T. Baba, T. Bae, O. Schneewind, F. Takeuchi, K. Hiramatsu, Genome sequence of Staphylococcus aureus strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands, J. Bacteriol. 190 (1) (2008) 300–310.

[32] S.K. Highlander, K.G. Hulten, X. Qin, et al., Subtle genetic changes enhance virulence of methicillin resistant and sensitive Staphylococcus aureus, BMC Microbiol. 7 (2007) 99.

[33] B.A. Diep, S.R. Gill, R.F. Chang, et al., Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus, Lancet (London, England) 367 (9512) (2006) 731–739.

[34] D. Coil, G. Jospin, A.E. Darling, A5-miseq: an updated pipeline to assemble microbial genomes from illumina miseq data, Bioinformatics 31 (4) (2014) 587–589.

[35] B. Buchfink, C. Xie, D.H. Huson, Fast and sensitive protein alignment using DIAMOND, Nat. Methods 12 (2015) 59–60.

[36] Wallis Kruskal, Use of ranks in one-criterion variance analysis, J. Am. Stat. Assoc. 47 (260) (1952) 583–621.

[37] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. Roy. Stat. Soc. Ser. B Method 57 (1) (1995) 289–300. http://www.jstor.org/stable/2346101.

[38] S. Kurtz, A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S.L. Salzberg, Versatile and open software for comparing large genomes, Genome Biol. 5 (2004) RI2.

[39] A.K. Szafranska, A.P.A. Oxley, D. Chaves-Moreno, S.A. Horst, S. Roblenbroich, G. Peters, O. Goldmann, M. Gohde, B. Sinha, D.H. Pieper, B. Loffler, R. Jauregui, M.L. Wos-Oxley, E. Medina, High-resolution transcriptomic analysis of the adaptive response of *Staphylococcus aureus* during acute and chronic phases of osteomyelitis, mBio 5 (6) (2014) 1–13.

[40] S.A. Horst, V. Hoerr, A. Beineke, C. Kreis, L. Tuchscherr, J. Kalinka, S. Lehne, I. Schleicher, G. Kohler, T. Fuchs, M.J. Raschke, M. Rohde, G. Peters, C. Faber, B. Loffler, E. Medina, A novel mouse model of *Staphylococcus aureus* chronic osteomyelitis that closely mimics the human infection, Am. J. Pathol. 181 (2012) 1206–1214.

https://doi.org/10.1016/j.heliyon.2018.e00674