Journal of Microbiology, Immunology and Infection 56 (2023) 802—814

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.e-jmii.com

Original Article

Whole genome characterization of methicillin resistant *Staphylococcus aureus* in an Egyptian Tertiary Care Hospital

Reem M. Hassan [a], Mervat G. Elanany [a], Mostafa M. Mostafa [b], Reham Hamed A. Yousef [a], Sherifa T. Salem [a,*]

[a] *Department of Clinical and Chemical Pathology, Faculty of Medicine, Cairo University, Egypt*
[b] *Faculty of Pharmacy, Ain Shams University, Egypt*

Received 16 December 2021; received in revised form 19 February 2023; accepted 20 April 2023
Available online 11 May 2023

KEYWORDS
MRSA;
NGS;
Healthcare-associated infections;
ST239;
spa t037

Abstract *Background:* Methicillin-resistant *Staphylococcus aureus* (MRSA) is a significant healthcare-associated (HA-MRSA) pathogen due to its increased morbidity and mortality rates. There is a paucity of data regarding MRSA clones circulating in the Middle East in the literature, especially from Egypt. We aimed to identify the pattern of resistance and virulence in the propagating clones using NGS technologies for the whole genome sequence.
Methods: From an 18-month surveillance program for MRSA-positive patients, 18 MRSA isolates from surgical healthcare associated infections were selected. The Vitek2 system was used to assess antimicrobial susceptibility. The whole genome sequencing was performed using the NovaSeq6000. The reads were mapped to the reference genome (Staphylococcus_aureus_ATCC_BAA_1680), used for variant calling, screened for virulence/resistance genes, and typed using multi-locus sequence typing and *spa* typing. Correlation between demographic and clinical data and molecular findings were performed.
Results: All the MRSA isolates were highly resistant to tetracycline followed by gentamicin (61%) and highly susceptible to trimethoprim/sulfamethoxazole. Most of the isolates showed a high virulence profile. ST239 was the predominant sequence type (6/18), while t037 was the predominant *spa* type (7/18). Five isolates shared the same ST239 and *spa* t037. ST1535, an emerging MRSA strain, was the second most prevalent in our study. One isolate showed a unique pattern of a high abundance of resistance and virulence genes.
Conclusion: WGS elucidated the resistance and virulence profiles of MRSA isolated from clinical samples of HAI patients with high-resolution tracking of clones predominant in our healthcare facility.

* Corresponding author.
E-mail addresses: Reem.mostafa@kasralainy.edu.eg (R.M. Hassan), miroislamic@yahoo.co.uk (M.G. Elanany), rehamhamed@kasralainy.edu.eg (R.H.A. Yousef), cherifa_tarek@kasralainy.edu.eg (S.T. Salem).

https://doi.org/10.1016/j.jmii.2023.04.005

Introduction

Staphylococcus aureus is a Gram-positive bacterium that causes a wide variety of clinical manifestations.[1] *S. aureus* is one of the most prevalent pathogens that spread both in the community as well as in hospital settings; treatment of *S. aureus* infections remains a challenge due to its ability to acquire resistance against multiple antibiotic classes.[2,3] In 2017, it was estimated that 119,000 Americans experienced *S. aureus* systemic infections and that approximately 20,000 died.[4] Among all *S. aureus* strains, MRSA strains are the most challenging and clinically debilitating.[5] MRSA strains pose a significant threat; according to a recent study published in the Journal of the American Board of Family Medicine, assessing the 11-year mortality risk of MRSA colonization in a community cohort study from the national database for adults aged 40—85, the mortality rate among MRSA carriers was 35.9%.[6] MRSA poses not only a clinical threat, but also a significant financial burden on hospitals and patients. According to a study conducted in Egypt, MRSA strains account for more than 70% of healthcare-associated *S. aureus* cases and 11.5% of community-acquired *S. aureus* infections.[7]

Healthcare facilities are the major source of MRSA infections.[8] MRSA has a high transmissibility and infectivity making cross-contamination facile and necessitate legitimate decontamination protocols.[8] Various techniques are currently available for investigating outbreaks and comparing evolving strains. These techniques include multilocus sequence typing (MLST), antibiograms, and pulse field gel electrophoresis (PFGE). However, these techniques cannot isolate closely related bacterial species with the desired resolution.[9]

At a rate of 10^{10} mutations per nucleotide base, *S. aureus* rapidly mutates using single-nucleotide variants and point mutations, which means 300 new mutations approximately every 10 h.[10] Traditional genotyping techniques are not adequate for tracking these rapid alterations, and the use of the novel whole-genome sequencing (WGS) is rudimentary.[9] WGS offers high resolution and identifies both small and large variants.[11] Furthermore, it can detect single mutations, insertions, and even deletions within the bacterium genome.[12] Therefore, using WGS techniques can increase our understanding of the pathogenicity of *S. aureus* by elucidating its determinants of virulence and resistance.

In this study, we used whole genome sequencing and bioinformatics techniques to identify novel resistance and virulence patterns in MRSA variants isolated from hospitalized patients in Egypt. Moreover, we assessed the relatedness and the diversity of the bacterial genome of the 18 MRSA isolates using multi-locus sequence typing and single nucleotide polymorphism (SNP) analysis.

Methods

Isolates

From January 2018 to September 2019, we collected all MRSA isolated from clinical samples sent to the microbiology laboratory of Kasr Al Ainy hospitals, faculty of medicine, Cairo university, Egypt (n = 210). MRSA isolates collected from patients suffering from healthcare-associated infections were selected (n = 110).

Twenty MRSA isolates were selected from surgically-admitted patients (ICU and wards) with MRSA infections. For sequencing, isolates were cultured from frozen stocks onto Blood agar (Oxoid, Basingstoke, UK), incubated aerobically at 35 °C overnight, and single colonies were picked for DNA extraction and further processing. Sequencing failed in two samples.

Characterization of the isolates

Phenotypic identification of isolates was made using conventional culture plates, Gram staining criteria, biochemical testing of isolates as catalase, coagulase, culture on mannitol salt agar, and DNase agar. According to the Clinical and Laboratory Standards Institute (CLSI), the antimicrobial resistance patterns of *S. aureus* strains were performed using the modified Kirby—Bauer disc diffusion test method. The antibiotic discs were purchased from Oxoid™ (ThermoFisher Scientific, USA), and *S. aureus* strains ATCC 25923 were used as quality control. MIC was determined using Vitek 2 system according to the manufacturer's instructions.[13] Isolates were confirmed for MRSA genotypically using multiplex PCR targeting femB and *mecA* genes.[14]

DNA extraction, library preparation, and sequencing

DNA was extracted using a DNA mini extraction kit (Qiagen) following the protocol of isolation of genomic DNA from gram-positive bacteria.[15] To investigate the sample's quality, we used the Fragment Analyzer to determine the concentration of the sample. The NEBNext® Ultra II DNA Library Prep kit for Illumina (cat# NEB #E7645 S/L) was used in the samples' processing according to the manufacturer's instructions. The DNA was then sequenced and clustered in concordance with the NovaSeq6000 manufacturer's protocols. The Illumina data analysis pipeline Bcl2fastq v2.20. and RTA3.4.4 were used for image analysis, base calling, and quality check.

Sequence data analysis

Contamination screening and raw data quality control were performed to assess the sequencing library. The reads were preprocessed with adapter trimming and quality filtering, followed by alignment of short reads and variant calling. To further genotype the strains, assembly into scaffolds, screening for resistance and virulence genes, and (pro) phage screening were performed (Supplemental Fig. 1).

Contamination screen

In order to scan the samples for contamination, the reads were classified using Centrifuge v1.0.4 based on a reference database which has been compressed using the Burrow-Wheeler transform and optimized by creating indices using the Ferragina-Manzini index. The reference database for Centrifuge consists of all viral, bacterial, and human sequences.

Raw data quality

The pipeline started with a quality control stage. In this stage, checks for possible sample and barcode contamination were performed, and a set of standard quality metrics for the raw data set was determined using third-party (FastQC v0.34) and in-house (FastQA v1.0) QC tools (Supplemental Table 1).

Adapter trimming and quality filtering

Before aligning the sequences, we trimmed the reads for adaptor sequences and set up filters for sequence quality. When the bases matched an adapter sequence with two or fewer mismatches, or when an alignment score of at least 12 was obtained using Trimmomatic v0.30, the presumed adapter sequences were eliminated. To reduce background noise caused by sequencing errors, we filtered and clipped the reads based on the quality scores. Within a sliding window, 5′ and 3′ bases with an average Phred score less than Q30 were removed. If the resulting reads were shorter than 20 bp, they have been removed altogether (both pairs in paired-end reads). The number and length of the reads and bases before and after each trimming step are shown in Supplemental Table 2.

Mapping to the reference sequence

Mapping the filtered reads to the reference sequences (HFH-29568, ATCC BAA-1680) was performed with BWA v0.7.4, a short-read aligner based on Burrows-Wheeler Transform, with default settings. The resulting mapping (BAM files sorted and indexed by the samtools v1.3 package) was used for downstream analysis such as SNP/indel calling, mutation analysis, or other structural variation analyses. The number and percentage of reads mapped to the reference sequence are reported in Supplemental Table 3.

Variant calling

Variant calling was based on the best practices of GATK in terms of their approach to variant calling using GATK's Haplotype Caller v3.7. Based on the mapped reads, single nucleotide variants and small indels in the sample data were detected when compared to the reference and reported.

Strain typing

Short Read Sequence Typing (SRST2 v0.2.0), which is an analysis pipeline designed to incorporate Illumina sequence data, MLST database, and/or gene sequencing database (e.g., virulence genes, and resistance genes) and the output sequence type. This tool was used in combination with the ARGannot database to search for resistance genes and the VFDB to search for virulence genes. Multi-locus sequence typing was performed using *S. aureus* dataset from Pubmlst.[16]

Assembly

Spades v3.14.1 was used for assembly. Spades is a de Bruijn graph-based assembler that can cope with a variety of k-mer sizes. The k-mer sizes was set to 21,33,55,77,89,99,111,127. With these k-mer sizes, the tool can better handle the large variations in coverage across the genome that are a characteristic of single-cell sequencing, as well as a novel method for handling paired-end information. The minimum coverage cutoff was automatically calculated by the tool. For mapping the edges of the assembly graph, the distances between the k-mers were estimated. Afterward, a paired assembly graph was constructed, and SPAdes outputted a set of scaffolds.

Gene annotation

The scaffolds produced from the de novo assembly were annotated using Prokka v1.14.5. The tool had several integrated feature prediction tools. The features predicted were coding sequences, ribosomal RNA genes, transfer genes, signal leader peptides, and non-coding RNA.

Phage screen

Virsorter_v1.0.6 was used to identify phage/prophage sequences based on the scaffolds resulting from the assembly step. Virsorter could detect a viral signal within a bacterial sequence data either with or without a reference, which further increased the chance of detecting novel viruses.

Resistance genes

A screen for common resistance genes was performed with SRST_v0.2. A heatmap was generated to enhance the visualization of the resistance pattern.

SCCmec typing, *Spa* typing

SCC*mec* types were determined using SCC*mec*Finder 1.2 to identify specific gene alleles of *ccr*, class of *mec* and differences in J regions for SCC*mec* type I-XI and subtyping for SCC*mec* type IV and V.[17] *Spa* type was determined using *spa*Typer 1.0 from assembled genome/contigs.[18]

Circular genome visualization

Circular genome visualization was performed on isolate 104064-001-001 using CGview. A blast comparison was performed against reference sequence ATCC BAA-1680, and the open reading frame was provided.

Results

Nature and characteristics of the isolates

In this study, we investigated 18 HAI-MRSA isolates recovered from patients admitted to surgery departments (ICU and wards), and all confirmed to express the methicillin resistance gene SCC*mec*A. The isolates were obtained from a variety of clinical sources; 55% of the samples were isolated from wounds, while 33% were extracted from blood (Table 1). Most samples were isolated from non-ICU hospitalized patients (61%); the majority showed signs of fever, while the remaining 39% were ICU patients (Table 1). Out of the 18 patients, only one died, while the rest were discharged from the hospital after recovery (Table 1). Given the limited number of isolates investigated in this study, a significant correlation between the sequence type of the sample and any of its clinical characteristics was not feasible.

Resistance profile analysis

Due to the importance of accurately determining the antibiotic resistance genes, in cases of MRSA, in preventing the spread of infection, we decided to investigate the correlation between the phenotypic antibiotic susceptibility patterns of our isolates and their resistance genes.

With the exception of isolate 15, which was phenotypically resistant to benzylpenicillin and did not carried the β-lactam *mecA* and *blaZ* resistance genes, all of the samples demonstrated phenotypic resistance to benzylpenicillin and carried the β-lactam *mecA* and *blaZ* resistance genes (Table 2 and Fig. 2 and Supplemental Table 4).

Isolate 1 showed a potent resistance pattern. Despite being susceptible to erythromycin, levofloxacin and trimethoprim/sulfamethoxazole (SXT) phenotypically, it carried erythromycin resistance gene (LnuA), levofloxacin resistance genes (norA and catA) and SXT resistance gene dfrC (Table 2 and Fig. 2 and Supplemental Table 4). Carrying the resistance genes for these antibiotics may indicate a higher probability for this isolate to develop phenotypical resistance against them later on. For improved visualization, a circular genome visualization of isolate 1 is provided (Fig. 1).

A total of ten isolates showed phenotypic resistance to gentamicin and carried at least four of the gentamicin resistance genes (aadD, aac(6′)-aph(2″), ant 6-Ia, Aph3″Ia, aph(3′)-III, Aadc, APH Stph, Sat4A, Aac3-IV). Isolates 6, 9, 14, 15, 17, and 19 demonstrated phenotypic susceptibility to gentamicin but carried at least one gentamicin resistance gene. The Aac3-IV and APH Stph were the most common gentamicin resistance genes expressed in all the isolates, followed by aac(6′)-aph(2″), which was expressed in 66% of the isolates (Table 2 and Fig. 2 and Supplemental Table 4).

Out of the 18 isolates analyzed, six isolates showed phenotypic resistance to levofloxacin and expressed at least three levofloxacin resistance genes (norA, Fex A, Dha1, and catA); norA and catA were expressed in all isolates. Moreover, 15 samples showed phenotypical resistance to tetracycline and carried at least two tetracycline resistance genes (tet K, tet M, tet-38). Tet-38 was expressed in all isolates (Table 2, Fig. 2, and Supplemental Table 4).

In comparison to other antibiotics, trimethoprim/sulfamethoxazole had the lowest profile of bacterial resistance, with only four resistant isolates (Table 2, Fig. 2, and Supplemental Table 4).

Virulence profile analysis

Using whole genome sequencing, we were able to identify the virulence genes classes carried by each of the 18 MRSA isolates (Table 3, Fig. 3, and Supplemental Table 5). Genes encoding Adenosine synthase A, Autolysin, Aureolysin, Hemolysins, Leukocidins (luk D and E) and *Staphylococcus* Complement Inhibitor were detected in all our MRSA isolates (Table 3, Fig. 3, and Supplemental Table 5). In 94% of the isolates, we detected genes belonging to the Haemolysins, Leucocidins, Staphylokinase, and Serine Proteases. Interestingly, virulence genes associated with epidermal cell differentiation inhibition were only found in isolate 14; an ST-80 MRSA isolate; which was extracted from a surgical ward patient. Moreover, toxic shock syndrome toxin (TSST) and staphylococcal super-antigen-like 9 (Ssl9) were also each expressed in one isolate. Isolate 1 (ST-121) carried the TSST gene, while Ssl9 was detected in isolate 14 (ST-80), both of which were isolated from the blood of an ICU patient. The Panton-Valentine leukocidin (PVL)-encoding genes lukS and lukF were detected in ST-1535 Isolate 13 (wound, ward) and ST-80 isolate 14 (blood, ICU). While ST-121 isolate 1 (blood, ICU) only carried lukF virulence gene (Table 3, Fig. 3, and Supplemental Table 5).

Multi-locus sequence typing

To further understand the bacterial genome diversity of the 18 MRSA isolates, multi-locus sequence typing was carried out using seven loci specific to *S. aureus,* and a total of 8 Sequence types (ST) were identified (Table 1). ST 239 was the most prevalent sequence type, corresponding to 33% of all isolates, followed by ST 1535, which accounted for 16% of all isolates. The eight sequence types were clustered into three clonal complexes (CC) using eBURST, while 3 STs were singletons. The most common clonal complex was CC8 (50%), followed by CC5 (16%) and CC15 (16%). A phylogeny dendrogram was generated to facilitate the visualization of the related sequence types (Fig. 4).

Spa and SCC*mec* typing from WGS

SCC*mec* types were analyzed to identify specific gene alleles of *ccr* and *mec*. Within the investigated 18 MRSA isolates, three SCC*mec* types were identified (Table 1). Type III (38%) was the most common, followed by type IV (33%) and type V (27%). In addition, the 18 isolates were clustered into 10 *Spa* types (Supplement Table 4). *Spa* type t037 was the most prevalent, corresponding to 38% of all the isolates. Interestingly, all the t037 isolates belonged to the same MLST clonal complex, CC8, which further strengthened the relatedness of these isolates.

Table 1 Nature and characteristics of the isolates.

Isolate No.	Sex	Age	Speci-men	ICU/non-ICU	Fever	Duration of hospitalization[a]	Antibiotics	Device[b]	Duration of device[a]	Outcome	SCC*mec* type gene	ST	*spa* Type
1	F	24	Bl	Y	Y	8	Y	N	N	Dis	SCC*mec*_type_IV(2B&5)	121	t314
2	F	40	W	N	Y	10	Y	N	N	Dis	SCC*mec*_type_V(5C2)	1535	t084
3	F	60	W	Y	Y	9	N	N	N	Dis	SCC*mec*_type_III(3A)	239	t037
4	M	55	Bl	Y	N	10	Y	Y	9	Dis	SCC*mec*_type_III(3A)	239	t037
5	F	63	Bl	Y	Y	21	Y	Y	14	Died	SCC*mec*_type_III(3A)	239	t037
6	F	40	BAL	N	Y	4	Y	N	N	Dis	SCC*mec*_type_V(5C2)	8	t008
7	F	44	W	N	Y	5	Y	N	N	Dis	SCC*mec*_type_V(5C2)	1535	t084
8	F	38	W	N	N	20	Y	N	N	Dis	SCC*mec*_type_III(3A)	241	t037
9	M	51	W	N	Y	15	Y	N	N	Dis	SCC*mec*_type_III(3A)	239	t037
10	M	36	W	Y	Y	6	Y	N	N	Dis	SCC*mec*_type_III(3A)	239	t037
12	M	61	Bl	N	Y	30	Y	N	N	Dis	SCC*mec*_type_III(3A)	239	t138
13	F	42	W	N	Y	7	Y	Y	30	Dis	SCC*mec*_type_V(5C2)	1535	t2036
14	M	22	Bl	Y	Y	21	Y	N	N	Dis	SCC*mec*_type_IVc(2B)	80	t044
15	M	31	W	N	Y	5	Y	N	N	Dis	SCC*mec*_type_IVa(2B)	6	t304
16	F	28	W	N	Y	7	Y	N	N	Dis	SCC*mec*_type_IVa(2B)	241	t037
17	F	67	Bl	N	Y	7	N	N	N	Dis	SCC*mec*_type_VI(4B)	5	t954
19	F	48	Sp	Y	Y	7	Y	Y	5	Dis	SCC*mec*_type_VI(4B)	5	t688
20	M	30	W	N	N	5	Y	N	N	Dis	SCC*mec*_type_V(5C2)	121	t314

[a] Duration in days at time of sample collection.
[b] Device = Presence of external device e.g. central venous line.
M = Male; F=Female; Bl = Blood; W=Wound; BAL= Bronchoalveolar lavage; Sp = Sputum; Y=Yes, N=No; Dis = Discharged.

Table 2 Genotypic and phenotypic resistance identified in MRSA isolates.

Isolate number	ST	Benzylpenicillin	Gentamicin	Erythromycin	Levofloxacin	Tetracycline	Trimethoprim/ Sulfamethoxazole
1	121	BlaZ, *mecA*/R	aadD, aac(6′)-aph(2″), Aac3-lV, APH-Stph/R	-/S	norA/S	tetK, tet-38/R	dfrC/S
2	1535	BlaZ, *mecA*/R	aadD, aac(6′)-aph(2″), Aac3-lV/R	ermC/S	norA/S	tetK, tet-38/R	-/S
3	239	BlaZ, *mecA*/R	aac(6′)-aph(2″), ant 6-Ia, Aph3″Ia, Aac3-lV, Aadc, APH-Stph/R	ermC/S	norA/S	tetK, tetM, tet-38/R	-/S
4	239	BlaZ, *mecA*/*R*	aac(6′)-aph(2″), ant 6-Ia, Aph(3′)-III, Aac3-lV, APH-Stph/R	-/S	norA/R	tetK, tetM, tet-38/R	-/S
5	239	BlaZ, *mecA*/R	aac(6′)-aph(2″), ant 6-Ia, Aac3-lV, aph(3′)-III, APH-Stph/R	ermC/R	norA/R	tetK, tetM, tet-38/R	-/S
6	8	BlaZ, *mecA*/R	Aac3-lV, APH-Stph/S	-/S	norA/S	tet-38/S	-/S
7	1535	BlaZ, *mecA*/*R*	aadD, aac(6′)-aph(2″), Aac3-lV, APH-Stph/R	-/S	norA/S	tetK, tet-38/R	-/S
8	241	BlaZ, *mecA*/R	aac(6′)-aph(2″), APH-Stph, aph(3′)-III, Aac3-lV/R	ermA/R	norA/R	tetK, tetM, tet-38/R	dfrC/R
9	239	BlaZ, *mecA*/R	aac(6′)-aph(2″), ant 6-Ia, Aac3-lV, Aadc, APH-Stph/S	ermC/S	norA/S	tetK, tetM, tet-38/R	-/S
10	239	BlaZ, *mecA*/R	aac(6′)-aph(2″), Aac3-lV, APH-Stph/R	-/S	norA/R	tetM, tet-38/R	-/S
12	239	BlaZ, *mecA*/R	aac(6′)-aph(2″), ant 6-Ia, APH-Stph, aph(3′)-III, Aac3-lV, Aadc/R	ermC/R	norA/R	tetK, tetM, tet-38/R	-/S
13	1535	BlaZ, *mecA*/R	aac(6′)-aph(2″), Aac3-lV, APH-Stph, Aadc, aph(3′)-III, ant 6-Ia/R	-/S	norA/S	tetK, tetM, tet-38/R	-/S
14	80	BlaZ, *mecA*/R	aadD, Aac3-lV, APH-Stph/S	-/S	norA/S	tetK, tet-38/R	-/S
15	6	*mecA*/R	Aac3-lV, ant 6-Ia, APH-Stph, aph(3′)-III/S	-/S	norA/S	tet-38/S	-/S
16	241	BlaZ, *mecA*/R	APH-Stph, aph(3′)-III, Aac3-lV, aac(6′)-aph(2″),/R	ermA, ermC/R	norA/R	tetK, tetM, tet-38/R	dfrC/R
17	5	BlaZ, *mecA*/R	Aac3-lV, APH-Stph/S	ermC/S	norA/S	tetK, tetM, tet-38/R	-/S
19	5	BlaZ, *mecA*/R	Aac3-lV, APH-Stph/S	-/S	norA/S	tetM, tet-38/R	dfrC/R
20	121	BlaZ, *mecA*/R	Aac3-lV, APH-Stph/S	ermC/S	norA/S	tetK, tet-38/S	-/S

S=Sensitive; R = Resistant.
(Phenotyic resistance was determined using modified Kirby–Bauer disc diffusion test method).

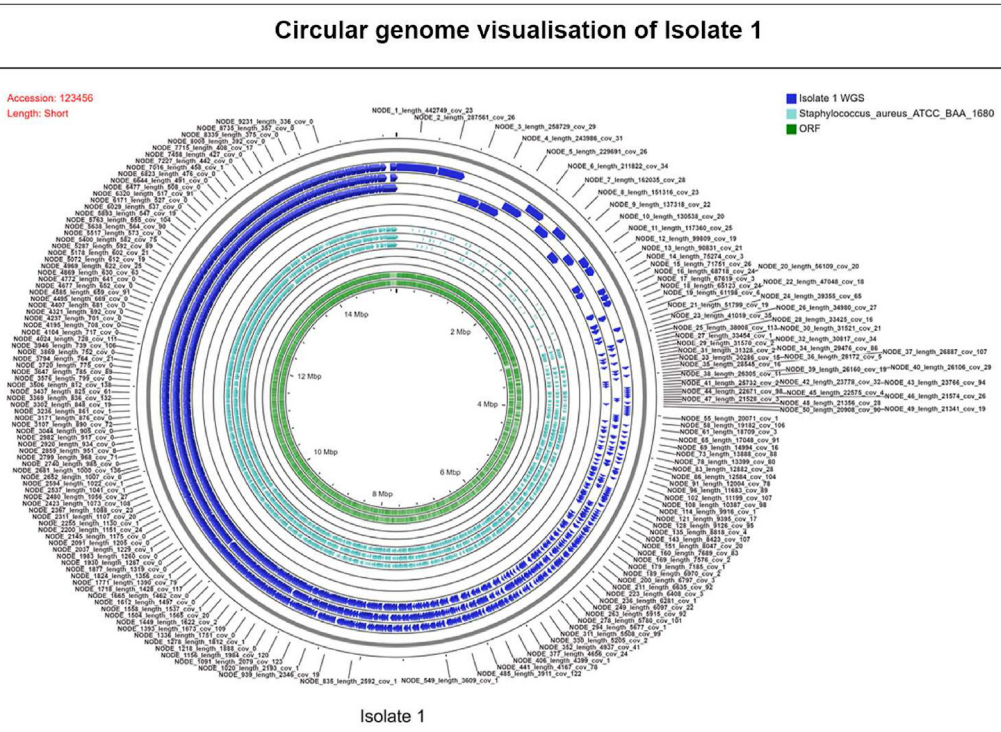

Figure 1. A CGView map of the isolate 1 genome, with nodes marked as arrows. Both the forward and reverse genes are marked with blue. The genome was blasted against the reference *Staphylococcus aureus* labeled with cyan. The open reading frame is marked with green.

Single nucleotide polymorphisms analysis

SNPs derived from whole genome sequencing could contribute to assess the relatedness of strains in a suspected outbreak when combined with other epidemiological data due to their high resolution compared to other methods. Due to its significance in determining strain relatedness, we started to investigate the robustness of the SNPs across all our MRSA isolates. First, the Isolate genomic sequences were aligned to the reference genome (Staphylococcus_aureus_ATCC_BAA_1680), and 285337 positions were found in all analyzed genomes. Additionally, 9.6% of the reference genome was covered by all the isolates. Then, we assessed the genomic distribution of SNPs; the positions of the identified SNPs were plotted on the *S. aureus* ATCC_BAA_1680 genome using a sliding window of 1000 bp. SNPs were distributed throughout regions included in the analysis and did not indicate any locations for mutational hotspots (Fig. 5A). Subsequently, we identified the number of genomic SNPs shared between any two isolates, and the similarity SNP count was provided (Fig. 5B). SNP counts ranged between 6 and 1459 shared SNP (Fig. 5B). To visualize the closely related s. *S. aureus* strains, a phylogeny tree was constructed (Fig. 5C). Based on the relatedness of the strains, the strains clustered into two large clusters and one outlier. Isolates[1 and 10] and Isolates[19 and 17] were the most closely related based on the SNP analysis, while isolate 6 showed the least correlation with any other strain (Fig. 5C).

Discussion

The diversity of the MRSA molecular epidemiology across diverse geographic regions has resulted in a distinct susceptibility pattern within each health care setting.[19] While the significant expansion of next-generation sequencing usage has made thousands of MRSA genomic data publicly available, these public genomic sequences lack the required global diversity and thus are not representative of the infective variants worldwide. For instance, using the bacterial bioinformatics resource center PATRIC database,[20] we found the genomic data for 17993 *S. aureus* isolates. However, while 6090 isolates were isolated from the United States and 1805 isolates were isolated from the UK, only 68 isolates were from Egypt. This finding highlights the disparate representation of this data in relation to the global dynamics of infection. Consequently, this study aimed to use whole genome sequencing to characterize the genomic resistance and virulence of 18 MRSA isolates isolated from hospitalized Egyptian patients and compare it to their phenotypic resistance. Moreover, we investigated the relationship between these samples using MLST and SNP analysis.

In this study, 72%—83% of MRSA isolates exhibited similar phenotypic susceptibility patterns as shown through disc diffusion method or broth micro-dilution method and carriage of resistance genes for most tested antibiotics. This finding aligned with other studies that reported a similar pattern in 76—87% of isolates.[21,22] Interestingly, a higher

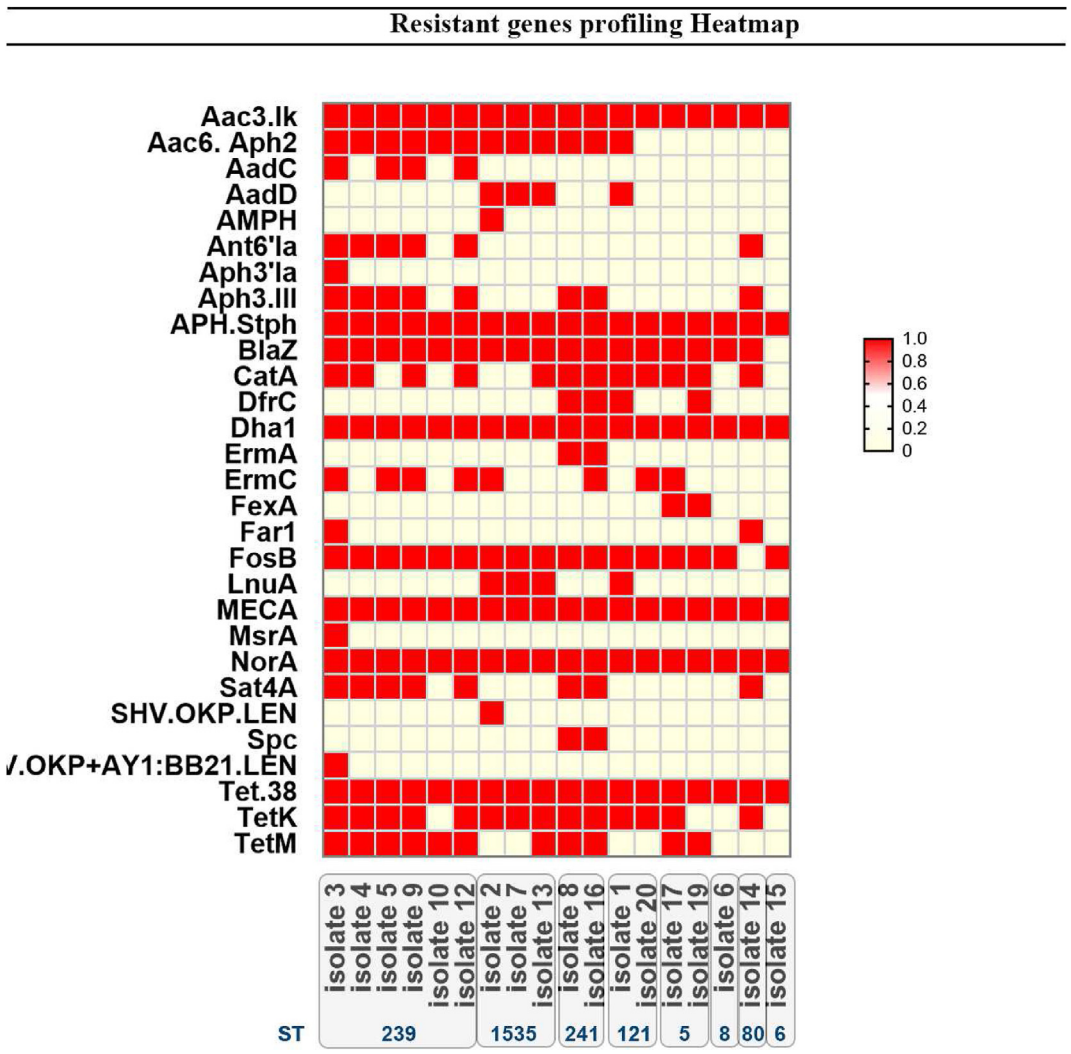

Figure 2. Resistant genes profiling heatmap. Red indicates the expression of the gene, while yellow indicates the absence of the genes. The isolates were clustered based on their sequence type.

dissimilarity was found between gentamicin phenotypic susceptibility and genomic resistance profile. While 39% of the isolates were phenotypically susceptible to the drug, it contained at least two gentamicin resistance genes. This dissimilarity could be due to the presence of unexpressed resistance genes. Given that these *S. aureus* isolates were selected for their methicillin resistance, each isolate contained the *mecA* gene, which encodes the PBP2a protein essential for β-lactamase activity. Furthermore, 33% of the isolates demonstrated phenotypic resistance to levofloxacin and expressed at least three levofloxacin resistance genes. Additionally, 83% of the samples exhibited phenotypic resistance to tetracycline and carried at least two tetracycline resistance genes. Finally, 22% of isolates showed phenotypic resistance to trimethoprim/sulfamethoxazole. These findings were consistent with the previously published reports.[22,28]

The investigation into the virulence profile of the isolates revealed the existence of various virulence classes. Adenosine synthase A, Autolysin, Aureolysin, Staphylococcus Complement Inhibitor, Haemolysins and Leucocidins (lukD and lukE) were the most commonly detected in all our MRSA isolates, followed by genes belonging to the groups of Staphylokinase, and Serine Proteases (in 94% of the isolates). These virulence profiles were concordant with other countries' patterns, as Hemolysin A was detected in 98.7% and 99% within more than 50% of isolates harboring more than 6 virulence genes.[23,24]

Moreover, the Panton-Valentine leukocidin (PVL)-encoding, lukS, and lukF, genes were detected in 16% of the isolates, which is in concordance with previously published studies in the middle east.[25,26] PVL is a common community-acquired MRSA strain exotoxin.[27] Interestingly, two of the three isolates that carried the PVL encoding genes were isolated from ICU patients, which supports the previous notion about the emergence of new PVL-positive hospital-acquired MRSA.[23] Moreover, despite the association between the non-severe MRSA infection and PVL-positive MRSA strains,[23,24] our data confirmed the existence of PVL-positive MRSA in two severely infected ICU patients out of the three patients who were infected with PVL-positive MRSA. Our results revealed the presence of

Table 3 Virulence profile identified in MRSA isolates. Y indicates the expression of the gene within the isolate while N indicates the absence of the gene.

	ST	Adenosine Synthase A	Autolysin	Aureolysin	Epidermal cell	Haemolysins	Toxic shock syndrome	Staphylococcal superantigen	Leucocidins				Staphylokinase	Staphylococcus complement inhibitor	Serine proteases		
		adsA	atl	aur	edin-B	hlb	TSST	ssl9	lukF-PV	lukS-PV	lukD	lukE	sak	scn	splA	splB	splE
Isolate 1	121	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 2	1535	Y	Y	Y	N	Y	N	N	N	N	Y	Y	N	Y	Y	Y	Y
Isolate 3	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 4	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 5	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 6	8	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 7	1535	Y	Y	Y	N	Y	N	N	N	N	Y	Y	N	Y	Y	Y	Y
Isolate 8	241	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 9	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 10	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 12	239	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 13	1535	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
Isolate 14	80	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	N	N
Isolate 15	6	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 16	241	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Isolate 17	5	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	N
Isolate 19	5	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	Y	N
Isolate 20	121	Y	Y	Y	N	Y	N	N	N	N	Y	Y	Y	Y	N	Y	Y

ST= Sequence type; Y=Yes; N=No.

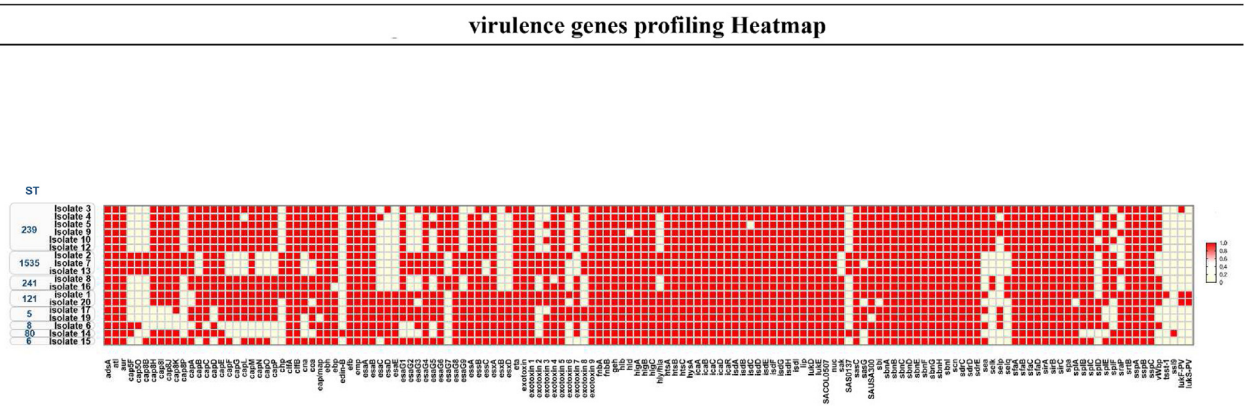

Figure 3. Virulence genes profiling heatmap. Red indicates the expression of the gene while yellow indicate the absence of the genes. The isolates were clustered based on their sequence type.

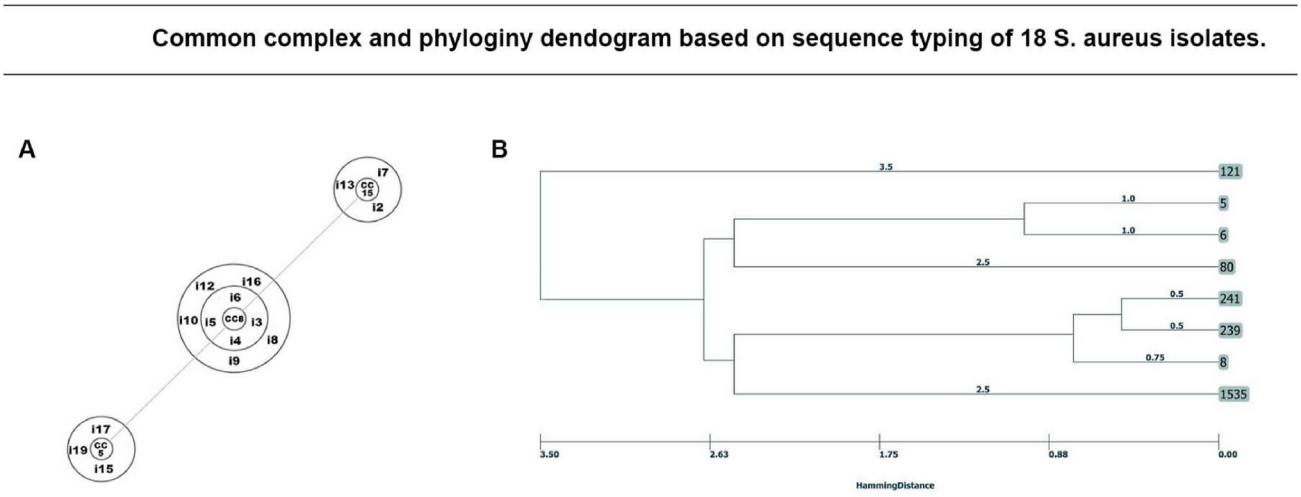

Figure 4. A) EBURST analysis of *S. aureus* STs. The assigned common complex is within the central ring. The isolated complexed are within the outer rings. B) A phylogeny dendrogram of the related sequence types based on sequence typing of 18 *S. aureus* isolates. The unweighted pair-group method with arithmetic means tree of the seven gene fragments examined during multi-locus sequence typing of *Staphylococcus aureus* isolates.

virulence genes TSST (ST-121 isolate 13) and Ssl9 (ST-80 isolate 19), each in only one isolate, which is in concordance with previous reports.[28]

In the present study, ST-239 was the most prevalent sequence type, corresponding to 33% of all isolates, followed by ST-1535 which accounted for 16% of all isolates. Since the 1970s, ST-239 has been considered the oldest pandemic MRSA strain causing severe infections in many countries.[29] This strain is wide-spreading, and the fact that it is still capable of causing hospital-borne pandemics further highlights the need for stricter infection control and screening protocols.

The previous report has also highlighted the isolation of the ST-239 MRSA strain in Egypt.[30] Despite being ubiquitously characterized in the literature as methicillin-susceptible *S. aureus* strains, the clonal complex 15 ST-1535 strain has recently emerged as a rare MRSA strain31. In 2016, Saudi Arabia published the first report identifying this nasal colonizer MRSA strain31. Multiple reports from various demographic regions had since reported the emergence of this rare strain.[30] It was found that ST-1535 was the second most prevalent sequence type in our data. The three ST-1535 isolates were isolated from wounds of non-ICU patients. ST-121 isolate 1 (singleton), harbored a potent antibiotic resistance genes pattern, compared to all other isolates.

In the current study, type III (38%) was the most common SCC*mec* type within our MRSA isolates, followed by SCC*mec*-IV (33%) and SCC*mec*-V (27%). This finding is in line with the previous reports from Egypt and the middle east demonstrating SCC*mec*-IV, SCC*mec*-V, and SCC*mec*-III to be the most abundant among their MRSA isolates.[24,32] However, according to these studies, SCC*mec*-IV is the most prevalent of the three types in the Middle East, whereas our research indicates that SCC*mec*-III is the most prevalent strain. This finding may be due to the fact that all of the MRSA isolates in this study were obtained from hospitalized patients, and SCC*mec*-III is known to be the most prevalent among HA-MRSA strains.[33,34]

The 5 isolates of interest that harboured SCC*mec* V (5C2), showed diverse STs; 3 of them (Isolates 2, 7, and 13)

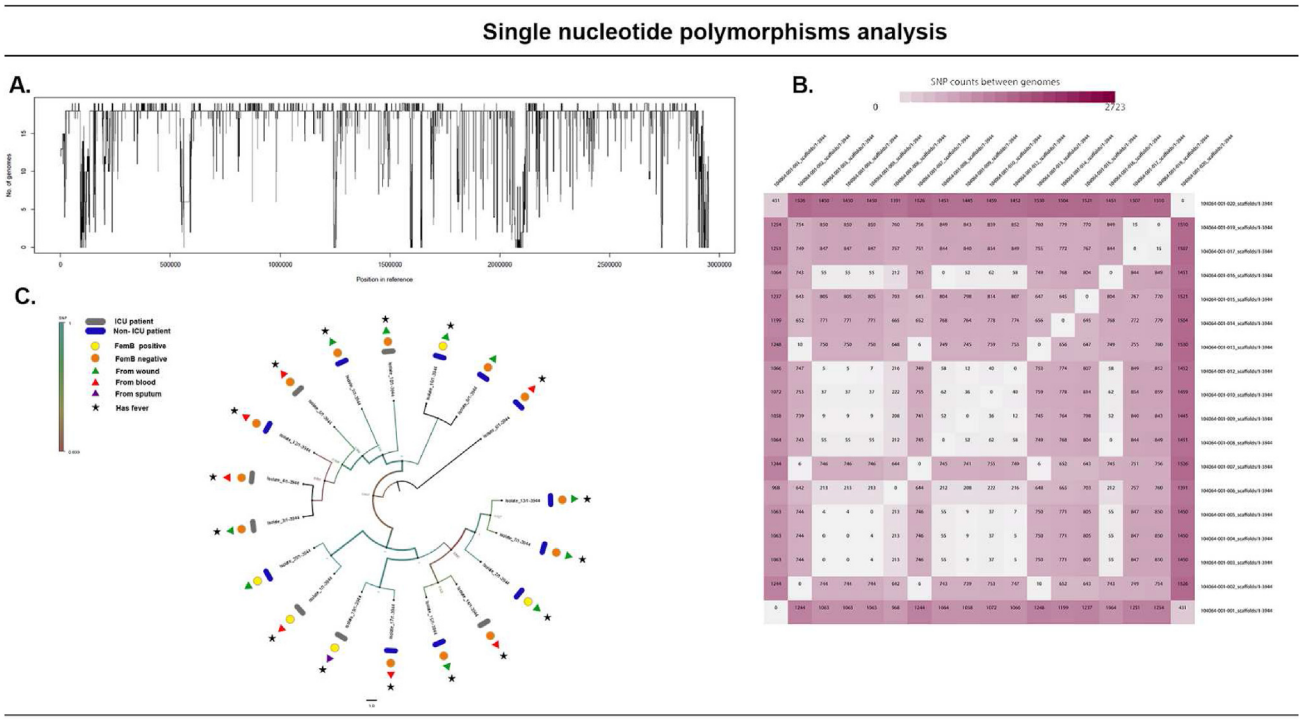

Figure 5. Single nucleotide polymorphisms analysis. A. Genomic distribution of SNPs. The position of the identified SNPs was plotted on the *Staphylococcus aureus* ATCC_BAA_1680 genome using a sliding window of 1000 bp. SNPs were distributed throughout regions included in the analysis and did not indicate any locations for mutational hotspots. B. Main SNP matrix showing the number of genomic SNPs shared between any two isolates. C. Maximum parsimony SNP tree of MRSA isolates.

were assigned as ST1535 and only one isolate was assigned as ST8 (Isolate 6, USA300) and one isolate as ST121 (Isolate 20). This may indicate the expansion of a previously known community-acquired SCC*mec*-V MRSA[35] into multiple ST lineages of HA-MRSA.

In another study in Egypt conducted on 18 MRSA isolates from a tertiary care hospital, the ST1535-SCC*mec* V MRSA clone was the most frequently isolated (16.6%), followed by ST5-SCC*mec* IV, ST1-SCC*mec* V and ST239-SCC*mec* III (11.1% each).[36] Also, other studies showed this rising prevalence of SCC*mec* V in variable STs of MRSA.[37,38]

Lastly, in this study, we performed a SNP analysis to accurately investigate if an epidemiological association exists within the isolated MRSA strains. We identified 285337 parsimony informative SNPs in all analyzed genomes. In addition, 9.6% of the reference genome was covered by all the isolates. We observed a distinct clustering of the isolates into two groups and one single outlier. Isolates[1 and 10] and isolates[19 and 17] were the most closely related, indicating that these isolates shared a common origin (Fig. 5C). Interestingly, these isolates had a similar antibiotic resistance pattern. Isolate 6 showed the least correlation with any other strain, indicating a distinct lineage from the other isolate.

Conclusion

This study characterized 18 confirmed HAI-MRSA isolates isolated from surgically hospitalized patients (ICU & wards) in a tertiary care hospital in Egypt. The analysis revealed high resistance to tetracycline followed by gentamicin, while trimethoprim/sulfamethoxazole demonstrated the lowest bacterial resistance profile. Most of the isolated MRSA showed a high virulent profile of Adenosine synthase A, Autolysin, Aureolysin, Hemolysin, Leukocidins D, E and Staphylococcus Complement Inhibitor genes expressed in almost all of them. ST239 was the predominant sequence type (6/18), while t037 was the prevalent *spa* type (7/18). Five isolates shared the same ST239 and *spa* t037. ST1535, an emerging MRSA strain, was our cohort's second most prevalent sequence type. Isolate 1 showed a unique pattern of a high abundance of resistant and virulent genes. SNP analysis showed that isolates comprise two distinct groups, and one singleton.

Acknowledgement

The authors sincerely acknowledge the Science, Technology & Innovation Funding Authority (STDF) for funding this research (as a part of young researcher grant no. 25657).

References

1. Chambers HF, Deleo FR. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* 2009;**7**(9): 629–41. https://doi.org/10.1038/nrmicro2200.
2. Schmidt A, Bénard S, Cyr S. Hospital cost of staphylococcal infection after cardiothoracic or orthopedic operations in France: a retrospective database analysis. *Surg Infect* 2015; **16**(4):428–35. https://doi.org/10.1089/sur.2014.045.

3. Li B, Webster TJ. Bacteria antibiotic resistance: new challenges and opportunities for implant-associated orthopedic infections. *J Orthop Res* 2018;**36**(1):22—32. https://doi.org/10.1002/jor.23656.
4. Kourtis AP, Hatfield K, Baggs J, Mu Y, See I, Epson E, et al. Vital signs: epidemiology and recent trends in methicillin-resistant and in methicillin-susceptible *Staphylococcus aureus* bloodstream infections - United States. *MMWR Morb Mortal Wkly Rep* 2019;**68**(9):214—9. https://doi.org/10.15585/mmwr.mm6809e1.
5. Zhen X, Lundborg CS, Zhang M, Sun X, Li Y, Hu X, et al. Clinical and economic impact of methicillin-resistant *Staphylococcus aureus*: a multicentre study in China. *Sci Rep* 2020;**10**(1):3900. https://doi.org/10.1038/s41598-020-60825-6.
6. Mainous AG, Rooks BJ, Carek PJ. Methicillin-resistant *Staphylococcus aureus* colonization and mortality risk among community adults aged 40-85. *J Am Board Fam Med* 2021;**34**(2):439—41. https://doi.org/10.3122/jabfm.2021.02.200394.
7. Abdel-Maksoud M, El-Shokry M, Ismail G, Hafez S, El-Kholy A, Attia E, et al. Methicillin-resistant *Staphylococcus aureus* recovered from healthcare- and community-associated infections in Egypt. *Int J Bacteriol* 2016;**2016**:5751785. https://doi.org/10.1155/2016/5751785.
8. 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**(5):609—18. https://doi.org/10.1093/cid/cit807.
9. Köser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012;**366**(24):2267—75. https://doi.org/10.1056/NEJMoa1109910.
10. Pray L. Antibiotic resistance, mutation rates and MRSA. *Nature Education* 2008;**1**(1):30.
11. Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, et al. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc Natl Acad Sci USA* 2007;**104**(22):9451—6. https://doi.org/10.1073/pnas.0609839104.
12. Price J, Gordon NC, Crook D, Llewelyn M, Paul J. The usefulness of whole genome sequencing in the management of *Staphylococcus aureus* infections. *Clin Microbiol Infect* 2013;**19**(9):784—9. https://doi.org/10.1111/1469-0691.12109.
13. *Performance standards for antimicrobial susceptibility testing. Twenty-ninth Informational Supplement (M100-S29).* Wayne, PA: Clinical and Laboratory Standards Institute; 2019.
14. Jonas D, Speck M, Daschner FD, Grundmann H. Rapid PCR-based identification of methicillin-resistant *Staphylococcus aureus* from screening swabs. *J Clin Microbiol* 2002;**40**(5):1821—3. https://doi.org/10.1128/JCM.40.5.1821-1823.2002.
15. Qiagen. *QIAamp DNA mini and blood mini handbook*. 2016. Retrieved from, https://www.qiagen.com/us/resources/resourcedetail?id=62a200d6-faf4-469b-b50f-2b59cf738962&lang=en.
16. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 2018;**3**:124. https://doi.org/10.12688/wellcomeopenres.14826.1.
17. International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (IWG-SCC). Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrob Agents Chemother* 2009;**53**(12):4961—7. https://doi.org/10.1128/AAC.00579-09.
18. Bartels MD, Petersen A, Worning P, Nielsen JB, Larner-Svensson H, Johansen HK, et al. Comparing whole-genome sequencing with Sanger sequencing for spa typing of methicillin-resistant Staphylococcus aureus. *J Clin Microbiol* 2014;**52**(12):4305—8.
19. Alkharsah KR, Rehman S, Alkhamis F, Alnimr A, Diab A, Al-Ali AK. Comparative and molecular analysis of MRSA isolates from infection sites and carrier colonization sites. *Ann Clin Microbiol Antimicrob* 2018;**17**(1):7. https://doi.org/10.1186/s12941-018-0260-2.
20. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, et al. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* 2020;**48**(D1):D606—12. https://doi.org/10.1093/nar/gkz943.
21. Mišić M, Čukić J, Vidanović D, Šekler M, Matić S, Vukašinović M, et al. Prevalence of genotypes that determine resistance of staphylococci to macrolides and lincosamides in Serbia. *Front Public Health* 2017;**5**:200. https://doi.org/10.3389/fpubh.2017.00200.
22. Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, et al. Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing. *J Clin Microbiol* 2014;**52**(4):1182—91. https://doi.org/10.1128/JCM.03117-13.
23. Li X, Huang T, Xu K, Li C, Li Y. Molecular characteristics and virulence gene profiles of *Staphylococcus aureus* isolates in Hainan, China. *BMC Infect Dis* 2019;**19**(1):873. https://doi.org/10.1186/s12879-019-4547-5.
24. Li X, Fang F, Zhao J, Lou N, Li C, Huang T, et al. Molecular characteristics and virulence gene profiles of Staphylococcus aureus causing bloodstream infection. *Braz J Infect Dis* 2018;**22**(6):487—94. https://doi.org/10.1016/j.bjid.2018.12.001.
25. Hadyeh E, Azmi K, Seir RA, Abdellatief I, Abdeen Z. Molecular characterization of methicillin resistant *Staphylococcus aureus* in west bank-Palestine. *Front Public Health* 2019;**7**:130. https://doi.org/10.3389/fpubh.2019.00130.
26. Enany S, Yaoita E, Yoshida Y, Enany M, Yamamoto T. Molecular characterization of Panton-Valentine leukocidin-positive community-acquired methicillin-resistant *Staphylococcus aureus* isolates in Egypt. *Microbiol Res* 2010;**165**(2):152—62. https://doi.org/10.1016/j.micres.2009.03.005.
27. Bastidas CA, Villacrés-Granda I, Navarrete D, Monsalve M, Coral-Almeida M, Cifuentes SG. Antibiotic susceptibility profile and prevalence of mecA and lukS-PV/lukF-PV genes in *Staphylococcus aureus* isolated from nasal and pharyngeal sources of medical students in Ecuador. *Infect Drug Resist* 2019;**12**:2553—60. https://doi.org/10.2147/IDR.S219358.
28. Hassan R, Barwa R, EL-Sokkary MM, Ashraf D. Virulence characteristics of methicillin resistant *Staphylococcus aureus* isolated from different clinical sources. *N. Egypt J. Microbiol.* 2017;**48**:14—29.
29. Monecke S, Slickers P, Gawlik D, Müller E, Reissig A, Ruppelt-Lorz A, et al. Molecular typing of st239-MRSA-III from diverse geographic locations and the evolution of the SCCmec III element during its intercontinental spread. *Front Microbiol* 2018;**9**:1436. https://doi.org/10.3389/fmicb.2018.01436.
30. Abdulgader SM, Shittu AO, Nicol MP, Kaba M. Molecular epidemiology of Methicillin-resistant *Staphylococcus aureus* in Africa: a systematic review. *Front Microbiol* 2015;**6**:348. https://doi.org/10.3389/fmicb.2015.00348.
31. Senok AC, Somily AM, Slickers P, Raji MA, Garaween G, Shibl A, et al. Investigating a rare methicillin-resistant *Staphylococcus aureus* strain: first description of genome sequencing and molecular characterization of CC15-MRSA. *Infect Drug Resist* 2017;**10**:307—15. https://doi.org/10.2147/IDR.S145394.
32. Sarkar A, Raji A, Garaween G, Soge O, Rey-Ladino J, Al-Kattan W, et al. Antimicrobial resistance and virulence markers in methicillin sensitive *Staphylococcus aureus* isolates associated with nasal colonization. *Microb Pathog* 2016;**93**:8—12. https://doi.org/10.1016/j.micpath.2016.01.008.
33. Abou Shady HM, Bakr AE, Hashad ME, Alzohairy MA. *Staphylococcus aureus* nasal carriage among outpatients attending

primary health care centers: a comparative study of two cities in Saudi Arabia and Egypt. *Braz J Infect Dis* 2015;**19**(1):68—76. https://doi.org/10.1016/j.bjid.2014.09.005.

34. Lee CY, Fang YP, Chang YF, Wu TH, Yang YY, Huang YC. Comparison of molecular epidemiology of bloodstream methicillin-resistant *Staphylococcus aureus* isolates between a new and an old hospital in central Taiwan. *Int J Infect Dis* 2019;**79**:162—8. https://doi.org/10.1016/j.ijid.2018.12.002.
35. Zetola N, Francis JS, Nuermberger EL, Bishai WR. Community-acquired meticillin-resistant *Staphylococcus aureus*: an emerging threat. *Lancet Infect Dis* 2005;**5**:275—86.
36. Soliman MS, Soliman NS, El-Manakhly AR, ElBanna SA, Aziz RK, El-Kholy AA. Genomic characterization of methicillin-resistant *Staphylococcus aureus* (MRSA) by high-throughput sequencing in a tertiary care hospital. *Genes* 2020;**11**(10):1219. https://doi.org/10.3390/genes11101219.
37. Bakthavatchalam YD, Vasudevan K, Rao S, Varughese S, Rupali P, Gina M, et al. Genomic portrait of community-associated methicillin-resistant Staphylococcus aureus ST772-SCCmec V lineage from India. *Gene Reports* 2021:24.
38. Rakonjac B, Lepšanović Z, Suljagić V, Jovčić B, Kojić M, Larsen AR, et al. Predominance of t355/ST152/SCCmec V clonal type among PVL-positive MRSA isolates in a tertiary care hospital in Belgrade, Serbia. *PLoS One* 2022;**17**(9): e0273474. https://doi.org/10.1371/journal.pone.0273474. 2022.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmii.2023.04.005.