

Research Article

Integrated Genomic, Virulence, and Immunoinformatic Characterization of Clinical *Staphylococcus aureus* Isolates for Identification and Prioritization of Conserved Vaccine Candidate Proteins

Ameer Hayder Razzaq Kaiwan ^{1*}

¹ College of Health and Medical Techniques, University of Alkafeel, Najaf, Iraq.

*Corresponding author: chiwan.ameer@alkafeel.edu.iq

Article Info

Keywords: MRSA, Antimicrobial resistance, *Staphylococcus aureus* Pan-genome, Reverse vaccinology, Vaccine candidates, Virulence factors, Whole-genome sequencing, Conserved antigens.

Received: 20.09.2026;

Accepted: 12.10.2026;

Published: 18.10.2026

 © 2026 by the author's. The terms and conditions of the Creative Commons Attribution (CC BY) license apply to this open access article.

Abstract

Background: *Staphylococcus aureus* is a major opportunistic pathogen associated with a wide range of clinical infections, while the increasing prevalence of methicillin-resistant *S. aureus* (MRSA), multidrug resistance, and diverse virulence determinants represents a significant therapeutic and public-health challenge. The genetic diversity of *S. aureus* further complicates the development of broadly applicable vaccine strategies. This study aimed to integrate phenotypic, molecular, genomic, and immunoinformatic approaches to characterize clinical *S. aureus* isolates and identify conserved proteins with potential as vaccine candidate antigens. **Methods:** A total of 150 clinical specimens were collected from different clinical sources in Babil Governorate, Iraq. Confirmed *S. aureus* isolates were subjected to antimicrobial susceptibility testing and PCR detection of antimicrobial resistance and virulence-associated genes. Representative isolates were confirmed by 16S rRNA gene sequencing, and 24 genetically and clinically representative isolates were selected for whole-genome sequencing (WGS). Genome assemblies were subjected to quality assessment, annotation, antimicrobial-resistance and virulence profiling, multilocus sequence typing, phylogenomic analysis, and pan-genome characterization. Conserved surface-associated and secreted proteins were subsequently screened using a reverse-vaccinology pipeline based on conservation, biological relevance, human-homology exclusion, predicted antigenicity, allergenicity, and toxicity. The prioritized proteins were further evaluated using immunoinformatic analyses of B-cell and T-cell epitopes, MHC class I and II interactions, and predicted population coverage. **Results:** Of the 150 clinical specimens, 112 (74.7%) were confirmed as *S. aureus*. Penicillin resistance was the most prevalent phenotype (81.3%), followed by erythromycin (46.4%) and ceftiofur (42.9%), with 48 isolates (42.9%) classified as MRSA and 43 (38.4%) exhibiting multidrug-resistant profiles. Among the investigated resistance determinants, *blaZ* was the most prevalent (83.0%), followed by *mecA* (41.1%). Virulence-associated genes were widely distributed, particularly *spa* (97.3%), *clfA* (94.6%), *clfB* (92.0%), *hla* (90.2%), and *icaD* (83.9%). The 24 WGS isolates represented multiple sequence types and clonal complexes, demonstrating substantial genomic diversity. A pan-genome analysis identified approximately 3,800 gene clusters: 2,250 core, 900 accessory, and 650 unique. Among the approximately 2,700 predicted proteins, 285 were classified as surface-associated or secreted; a tiered screening and prioritization process for these proteins resulted in the selection of 94 conserved candidates. **Conclusion:** An integrated genomic and immunoinformatic analysis identified conserved *Staphylococcus aureus* proteins with computational characteristics that support further vaccine research. Because clinical isolates show high antimicrobial resistance, virulence, and genetic diversity, continued molecular surveillance is necessary. Of the four candidates, Protein C5 had the most favorable computational profile. Experimental work is still needed to verify its expression, immunogenicity, safety, and protective efficacy before it can be considered a validated vaccine candidate.

1. Introduction

Staphylococcus aureus is a catalase- and coagulase-producing gram-positive organism commonly found in the nasal passages, mucous membranes, or skin of human carriers [1]. On culture, it is traditionally recognized by its characteristic golden-yellow colonies arranged in grape-like clusters (staphule-Greek for grapes; aureus Latin for gold) Figure 1. Among pathogens involved in surgical infections, it is arguably the most important, which accounts for the considerable attention given to it below. *S. aureus* causes most skin and soft tissue infections seen in surgical patients. It accounts for up to 20% of SSI isolates and 13% of all nosocomial infections [2, 3].

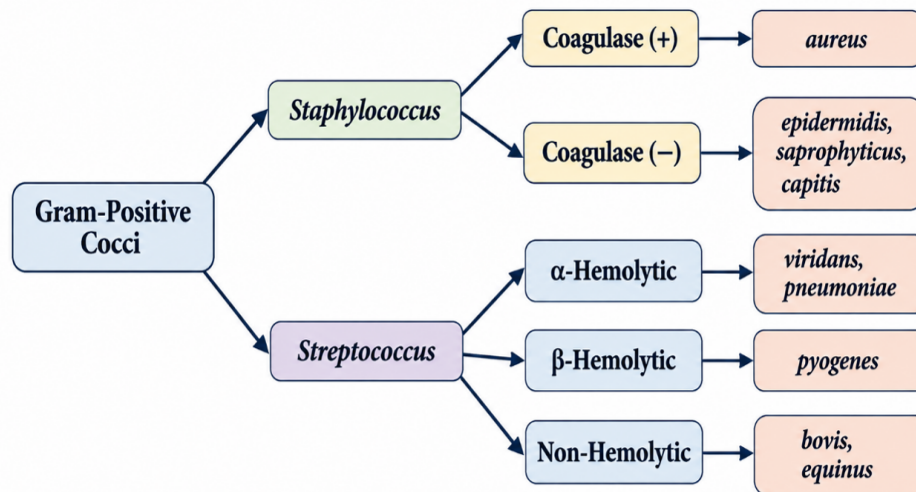


Figure 1: Gram-positive cocci

It also carries a broad set of virulence determinants, including adhesion proteins, biofilm-associated factors, iron-acquisition systems, toxins, and immune-evasion mechanisms. Together, these support colonization, persistence, tissue invasion, and survival within the host [4, 5]. Recent whole-genome studies of MDR clinical isolates have also shown that individual *S. aureus* genomes can carry multiple determinants associated with adhesion, biofilm formation, virulence, antimicrobial resistance, and mobile genetic elements [5]. Its considerable genomic plasticity adds to the organism's adaptability: resistance and virulence determinants can occur in both conserved and accessory genomic regions and may differ across bacterial lineages [2, 3]. Consequently, conventional characterization based solely on phenotypic antimicrobial susceptibility testing or a limited number of resistance and virulence markers may not fully capture the genomic complexity of clinical *S. aureus* populations. Whole-genome sequencing (WGS) provides a high-resolution approach for simultaneously investigating antimicrobial-resistance determinants, virulence-associated genes, sequence types, phylogenetic relationships, and genomic diversity, while also enabling comparison between phenotypic characteristics and their underlying genetic determinants [2, 6]. Recent genomic studies have demonstrated that WGS can reveal substantial diversity among clinical *S. aureus* isolates and can identify lineage-associated differences in antimicrobial resistance, virulence, biofilm-associated genes, and mobile genetic elements [3, 6]. In addition, pan-genome analysis enables bacterial genes to be classified into core, accessory, and strain-specific components, providing an effective framework for identifying conserved genetic features across genetically diverse clinical isolates [7]. Recent WGS-based studies of clinical MRSA isolates have reported thousands of gene clusters within the *S. aureus* pan-genome, together with considerable variation in antimicrobial-resistance and virulence-associated genes, emphasizing the value of pan-genomic analysis for understanding bacterial diversity and identifying conserved targets [7]. Such genomic information is particularly relevant to vaccine-oriented research because the development of a broadly effective *S. aureus* vaccine remains challenging. Despite extensive preclinical and clinical investigation, no licensed human vaccine has yet shown adequate protective efficacy. Recent reviews still highlight the challenges created by antigenic diversity, immune-evasion mechanisms, and the complex host immune response to *S. aureus* [8, 9]. Previous studies have demonstrated the feasibility of combining reverse vaccinology with immunoinformatics to identify potential *S. aureus* vaccine candidates, including conserved proteins and B- and T-cell epitopes [10–12]. More recent proteome-wide analyses have further shown that systematic computational screening can reduce the large *S. aureus* proteome to a limited number of promising candidate proteins for subsequent experimental evaluation [13]. In parallel, recent genome-informed vaccine-design strategies have emphasized the value of integrating pan-genome analysis with reverse vaccinology to account for genomic diversity when selecting conserved antigens [14]. Nevertheless, an important opportunity remains to connect the characteristics of locally circulating clinical *S. aureus* isolates including antimicrobial susceptibility, resistance determinants, virulence profiles, phylogenetic diversity, and pan-genomic variation with the systematic identification of conserved proteins that may have broad relevance as vaccine candidate antigens. Therefore, the present study was designed to establish an integrated and practical workflow beginning with the isolation and phenotypic characterization of clinical *S. aureus*, followed by antimicrobial susceptibility testing, molecular detection of selected resistance and virulence determinants, 16S rRNA sequencing, and WGS of representative isolates. The resulting genomic data will be used for comparative genomics, phylogenetic analysis, antimicrobial-resistance and virulence profiling, and pan-genome characterization. Conserved surface-associated and secreted proteins identified from the genomic dataset will subsequently be subjected to reverse-vaccinology and immunoinformatic screening based on conservation, subcellular localization, biological relevance, predicted antigenicity, absence of significant similarity to human proteins, allergenicity, toxicity, and the presence of favorable B-cell and T-cell epitopes. The final objective is to identify and prioritize a small number of conserved and immunologically promising *S. aureus* proteins for future experimental validation, while simultaneously providing an integrated molecular and genomic characterization of antimicrobial resistance, virulence, and genetic diversity among clinical isolates.

2. Methods

2.1. Study Design, Clinical Specimens, and Isolation of *Staphylococcus aureus*

A descriptive laboratory study was carried out to characterize clinical *Staphylococcus aureus* isolates at the phenotypic, antimicrobial-resistance, virulence, genomic, and immunoinformatic levels. Study setting and sample collection: Between January and June 2026, 150 clinical specimens were obtained from patients at selected healthcare facilities in Babylon Governorate, Iraq, including Marjan Teaching Hospital and Imam Al-Sadiq Teaching Hospital. The samples consisted mainly of wound and pus, blood, respiratory, and urine specimens, along with other relevant clinical materials. Standard microbiological methods were used to isolate and identify *Staphylococcus aureus*. When the same patient produced duplicate isolates, only one was retained. Standard microbiological procedures were used to process the specimens: culture, colony morphology, Gram staining, catalase, and coagulase testing, followed by additional biochemical tests when confirmation was needed. Confirmed *S. aureus* isolates were preserved appropriately for later antimicrobial-susceptibility and molecular analyses. Quality-control procedures were maintained throughout, and care was taken to prevent duplicate isolates from the same patient. Inclusion and exclusion criteria: Clinical samples from patients suspected of having bacterial infections were eligible when they could be processed for microbiological testing and produced isolates later confirmed as *Staphylococcus aureus* through standard microbiological and biochemical methods. The samples came from wounds or pus, blood, respiratory material, urine, and other relevant sources. Only one isolate per patient was counted; repeat isolates were excluded to prevent individual patients from being overrepresented. Samples unsuitable for processing, along with those that failed to produce confirmed *S. aureus* isolates, were excluded.

2.2. Antimicrobial Susceptibility Testing and Phenotypic Classification

Antimicrobial susceptibility was assessed using the Kirby–Bauer disk diffusion method on Mueller–Hinton agar, with bacterial suspensions standardized to 0.5 McFarland before inoculation and antibiotic disk application. The plates were incubated for 16–18 hours at approximately $35 \pm 2^\circ\text{C}$ under ambient atmospheric conditions. The agents tested were erythromycin, penicillin, cefoxitin, ciprofloxacin, tetracycline, clindamycin, trimethoprim–sulfamethoxazole, linezolid, gentamicin, and vancomycin. These agents were selected to provide a representative picture of the resistance patterns of the clinical isolates. Zones of inhibition were measured in millimeters, and results were interpreted based on current recommendations from the Clinical and Laboratory Standards Institute (CLSI). Details of the standard disk diffusion methodology are provided in document CLSI M02, while the corresponding interpretive criteria are available in document CLSI M100. Therefore, the latest available version of CLSI M100 should be consulted when performing the final laboratory analysis. For *S. aureus*, cefoxitin susceptibility served as the phenotypic indicator of methicillin resistance. Based on predefined criteria, isolates resistant to multiple classes of antimicrobial agents were classified as multidrug-resistant (MDR). When clinically relevant, it is preferable to determine vancomycin susceptibility using a standardized method based on the minimum inhibitory concentration (MIC) in accordance with CLSI M07/M100 recommendations—rather than inferring it solely from routine disk diffusion testing [15].

2.3. Genomic DNA Extraction and PCR Detection of Antimicrobial Resistance and Virulence Genes

Genomic DNA was extracted from confirmed *S. aureus* isolates with the QIAamp DNA Mini Kit (QIAGEN GmbH, Germany; Cat. No. 51304), following the manufacturer's protocol. DNA concentration and purity were checked by NanoDrop spectrophotometry, then adjusted to 35 ng/ μL . Its concentration and purity were verified via spectrophotometry, and only DNA preparations meeting purity criteria were used in the polymerase chain reaction (PCR) assays. Each PCR contained 2 μL of DNA template and targeted for selected antimicrobial-resistance genes *blaZ*, *mecA*, *mecC*, *ermA*, *ermB*, *ermC*, *tetK*, *tetM*, and *aac(6')*-*aph(2'')* associated with resistance to β -lactams, macrolides/lincosamides, tetracyclines, and aminoglycosides. Among these, *mecA* served as a key molecular marker for methicillin resistance, in accordance with established PCR methodology [16]. Each gene-specific reaction included appropriate primers and controls, and the resulting amplicons were verified by agarose gel electrophoresis [17]. PCR reactions were run in a total volume of 25 μL with a commercial 2 $\times$ PCR Master Mix containing Taq DNA polymerase, dNTPs, MgCl_2 , and reaction buffer. Each tube received 2 μL of DNA template, forward and reverse primers, and nuclease-free water. The program began with denaturation at 95° for 5 min, followed by 35 cycles of denaturation at 94° for 30 s, annealing at 57 – 58° for 30 s, and extension at 72° for 45–60 s. A final extension was then performed at 72° for 7 min. Every run included DNA from confirmed gene-positive *S. aureus* isolates as the positive control and a no-template reaction as the negative control.

2.4. 16S rRNA Gene Sequencing and Selection of Representative Isolates

Molecular identification involved 16S rRNA gene sequencing of 30 representative *S. aureus* isolates. For each isolate, a nearly complete ~ 1.5 -kb 16S rRNA region was amplified, purified, and sequenced using Sanger sequencing. Sequence quality was then assessed, and the resulting sequences were compared with reference sequences through BLASTn. Based on their phenotypic traits, antimicrobial susceptibility, resistance and virulence genes, and clinical sources, 24 representative isolates were chosen for whole-genome sequencing (WGS). The selection captured the observed phenotypic and genetic diversity for subsequent genome and pan-genome analyses.

2.5. Whole-Genome Sequencing, Genome Assembly, and Comparative Genomic Analysis

Genomic DNA from the 24 selected isolates was subjected to whole-genome sequencing using a suitable short-read sequencing platform. DNA libraries were prepared according to the manufacturer's recommendations, and sequencing was performed to obtain paired-end reads of adequate quality and coverage. Raw sequencing reads were initially assessed for sequence quality using FastQC. Low-quality sequences and adapter contamination were removed using a quality-control and read-trimming procedure. The filtered reads were subsequently assembled de novo using SPAdes, a widely used assembler for bacterial genome sequencing data [18]. Assembly quality was assessed using QUAST [19], including genome size, number of contigs, N50, and other relevant assembly statistics. Genome completeness and contamination were

evaluated using an appropriate genome-quality assessment approach. Assemblies that met the predefined quality criteria were retained for downstream analyses. The assembled genomes were expected to represent the chromosomal and, where adequately assembled, accessory genetic components of the respective isolates. Genome annotation was performed using Prokka or an equivalent standardized prokaryotic genome annotation pipeline. For the final version, the exact annotation software and version should be reported consistently, without listing alternative software options. Prokka automatically annotates bacterial genome assemblies, assigning features such as coding sequences, ribosomal RNA genes, transfer RNA genes, and other relevant elements [20]. Antimicrobial resistance determinants were identified with the Comprehensive Antibiotic Resistance Database (CARD) and the Resistance Gene Identifier (RGI), using appropriate sequence-identity and coverage thresholds [21]. Virulence-associated genes were screened against the Virulence Factor Database (VFDB) through its corresponding validated search procedure [22]. The resistance and virulence determinants detected were then compared with the phenotypic AST and PCR results to assess agreement between genotype and phenotype. Multilocus sequence typing (MLST) was performed using the established *S. aureus* MLST scheme. Sequence types (STs) were assigned based on the allelic profiles of the housekeeping genes, and the corresponding clonal complexes (CCs) were determined using the available MLST database. Comparative phylogenomic analysis was subsequently performed using conserved genomic regions or a core-genome alignment to assess genetic relationships among the sequenced isolates.

2.6. Pan-Genome Analysis and Identification of Conserved Surface-Associated Proteins

A pan-genome analysis was performed using the annotated genomes of the 24 WGS isolates. The analysis was designed to classify genes into core, accessory, and isolate-specific components. Panaroo was selected as the preferred pan-genome analysis pipeline because it is designed to reduce annotation-related errors and improve consistency in gene clustering across bacterial genomes [23]. The same annotation method was applied across the genomes before pan-genome analysis. Genes detected in all or nearly all of the analyzed isolates were classified as conserved/core genes according to the predefined similarity and presence thresholds. Genes present in a subset of isolates were classified as accessory genes, whereas genes restricted to individual or few isolates were considered strain-specific or unique components. The resulting core and accessory gene sets were further investigated to identify proteins potentially relevant to host interaction and vaccine development. Particular attention was given to proteins predicted to be localized on the bacterial surface or secreted into the extracellular environment, because such proteins may be more accessible to host immune recognition. Subcellular localization was predicted using an appropriate protein-localization prediction tool. Proteins predicted to be cytoplasmic or associated exclusively with intracellular metabolic functions were deprioritized. Surface-associated, membrane-associated, and secreted proteins were retained for further screening. Protein sequences were then assessed for conservation among the analyzed isolates and, where possible, against publicly available *S. aureus* genome sequences.

2.7. Reverse Vaccinology and Immunoinformatic Screening

A reverse vaccinology strategy was applied to prioritize conserved proteins with potential relevance as vaccine candidate antigens. The initial protein set obtained from genome annotation and pan-genome analysis was sequentially filtered according to conservation, predicted subcellular localization, biological relevance, antigenicity, host-homology exclusion, allergenicity, and toxicity.

First, proteins predicted to be surface-associated or secreted were retained. Proteins showing high conservation among the analyzed *S. aureus* isolates were prioritized, whereas highly variable or strain-restricted proteins were excluded. Candidate proteins were further evaluated for their potential association with bacterial survival, adhesion, colonization, immune interaction, or virulence. Potential similarity to human proteins was evaluated using a suitable sequence-comparison approach. Proteins showing substantial similarity to human proteins were excluded or deprioritized to reduce the possibility of undesirable cross-reactivity. The remaining proteins were evaluated for predicted antigenicity using VaxiJen or an equivalent antigenicity-prediction method. VaxiJen provides an alignment-independent approach for predicting potential protective antigens based on physicochemical properties of protein sequences [24]. Allergenicity and toxicity were subsequently evaluated using established computational prediction resources.

For MHC class I analysis, candidate peptides were screened according to predicted binding affinity to representative HLA class I alleles. MHC class II analysis was performed similarly using representative HLA class II alleles [25]. Peptides with favorable predicted binding characteristics were retained for further evaluation. Population coverage was estimated for the selected T-cell epitopes based on their predicted HLA restrictions. The IEDB population coverage methodology was used to estimate the theoretical proportion of a target population expected to present the selected epitopes [24]. Population coverage was considered an additional prioritization criterion rather than direct evidence of vaccine efficacy. The final candidate antigens were prioritized by integrating the results of conservation analysis, subcellular localization, biological relevance, antigenicity, host-homology screening, allergenicity, toxicity, B-cell epitope prediction, MHC-I and MHC-II binding, and population coverage. The highest-ranked proteins were considered candidate vaccine antigens or potential vaccine candidates for future experimental validation. Importantly, computational prediction of antigenicity or favorable epitope characteristics was not interpreted as proof of protective immunity. Experimental validation using recombinant proteins, immunological assays, animal models, or other appropriate approaches would be required to establish whether the prioritized proteins can induce protective responses.

2.8. Statistical Analysis and Integration of Phenotypic and Genomic Data

Descriptive statistical analysis was performed using frequencies and percentages for categorical variables. Antimicrobial resistance rates were calculated as the proportion of resistant isolates among the confirmed *S. aureus* isolates. The distribution of resistance and virulence genes was summarized according to the total number of isolates examined. Where appropriate, phenotypic antimicrobial resistance was compared with the presence or absence of corresponding resistance determinants. Similarly, associations between virulence-gene profiles, clinical source, antimicrobial resistance, and genomic lineage were explored. For genomic data, descriptive statistics were used to summarize genome size, GC content, predicted coding sequences, assembly metrics, sequence types, clonal complexes, and pan-genome composition. The relationships between isolates were evaluated using phylogenomic clustering and comparative genomic analysis. The final immunoinformatic ranking was based on the combined computational criteria described above rather than on a single score. Candidate proteins satisfying a greater number of predefined favorable criteria were assigned higher priority for subsequent experimental investigation.

3. Results

3.1. Isolation and Identification of Clinical *Staphylococcus aureus* Isolates

A total of 150 clinical specimens were collected from patients with suspected bacterial infections. Following culture, Gram staining, catalase and coagulase testing, and conventional biochemical identification, 112 isolates (74.7%) were confirmed as *Staphylococcus aureus*. The highest number of isolates was recovered from wound and pus specimens, followed by blood, respiratory, and urine specimens Table 1.

Table 1: Distribution of Clinical Specimens and Confirmed *Staphylococcus aureus* Isolates

Clinical specimen	No. examined	<i>S. aureus</i> isolates	Isolation rate (%)
Wound/pus	60	45	75.0
Blood	30	20	66.7
Respiratory	25	18	72.0
Urine	20	14	70.0
Other specimens	15	15	100
Total	150	112	74.7

The recovery of *S. aureus* from different clinical sources provided a representative collection for subsequent antimicrobial-resistance, virulence, and genomic characterization.

3.2. Antimicrobial Susceptibility and Resistance Determinants

Antimicrobial susceptibility testing demonstrated considerable variation among the 112 isolates. The highest resistance rate was observed against penicillin (81.3%), followed by erythromycin (46.4%) and ceftiofur (42.9%). Based on ceftiofur resistance, 48 isolates (42.9%) were classified as MRSA, while 43 isolates (38.4%) fulfilled the predefined MDR criterion Table 2.

Table 2: Antimicrobial Resistance Profile of Clinical *Staphylococcus aureus* Isolates

Antimicrobial agent	Resistant isolates, n (%)
Penicillin	91 (81.3)
Erythromycin	52 (46.4)
Ceftiofur	48 (42.9)
Clindamycin	38 (33.9)
Ciprofloxacin	35 (31.3)
Tetracycline	28 (25.0)
Gentamicin	21 (18.8)
Trimethoprim	17 (15.2)
–sulfamethoxazole	
Linezolid	0
Vancomycin	0

PCR analysis showed that blaZ was the most prevalent resistance determinant, detected in 93 isolates (83.0%), followed by mecA in 46 isolates (41.1%). The ermC, tetK, and ermA genes were detected in 24 (21.4%), 19 (17.0%), and 15 (13.4%) isolates, respectively Table 3.

Table 3: Distribution of Selected Antimicrobial-Resistance Genes among Clinical *S. aureus* Isolates

Resistance gene	Positive isolates, n (%)
blaZ	93 (83.0)
mecA	46 (41.1)
ermC	24 (21.4)
tetK	19 (17.0)
ermA	15 (13.4)
aac(6′)-aph(2′′)	14 (12.5)
tetM	11 (9.8)
ermB	7 (6.3)
mecC	2 (1.8)

The predominance of blaZ and mecA was consistent with the observed beta-lactam resistance profile. The combination of phenotypic susceptibility testing and molecular detection provided complementary information regarding antimicrobial resistance, as recommended in genomic investigations of clinical *S. aureus* [2, 3].

3.3. Virulence-Associated Gene Profile

The 112 clinical isolates demonstrated a broad distribution of virulence-associated genes. The most prevalent determinants were *spa* (97.3%), *clfA* (94.6%), *clfB* (92.0%), and *hla* (90.2%). The biofilm-associated genes *icaD* and *icaA* were detected in 83.9% and 78.6% of isolates, respectively Table 4.

Table 4: Distribution of Virulence-Associated Genes among Clinical *S. aureus* Isolates

Virulence gene	Positive isolates, n (%)
Spa	109 (97.3)
clfA	106 (94.6)
clfB	103 (92.0)
Hla	101 (90.2)
Hld	96 (85.7)
icaD	94 (83.9)
fnbA	91 (81.3)
icaA	88 (78.6)
isdA	87 (77.7)
Hlb	82 (73.2)
Sbi	81 (72.3)
Sak	76 (67.9)
Scn	69 (61.6)
lukS-PV/lukF-PV	24 (21.4)

The results demonstrated that adhesion, biofilm formation, hemolysis, iron acquisition, and immune-evasion determinants were widely distributed among the isolates, whereas the PVL-associated genes were detected in a smaller subset. Such diversity is consistent with recent WGS-based investigations showing extensive variation in the virulence repertoire of clinical *S. aureus* populations [3, 5].

3.4. 16S rRNA Confirmation and Selection of WGS Isolates

To confirm species identification at the molecular level, 16S rRNA sequencing was performed for 30 representative isolates. The approximately 1.5-kb sequences showed 99.2–100% identity with reference *S. aureus* sequences. Based on clinical source, antimicrobial susceptibility, resistance-gene profiles, virulence characteristics, and molecular diversity, 24 representative isolates were selected for whole-genome sequencing (WGS). The assembled genomes ranged from 2.65 to 2.90 Mb, with GC contents of 32.5–33.0%. MLST analysis demonstrated considerable genomic diversity among the sequenced isolates Table 5 & Table 6.

The assembled genomes ranged from 2.65 to 2.90 Mb, with GC contents of 32.5–33.0%. MLST analysis demonstrated considerable genomic diversity among the sequenced isolates.

Table 5: General Genomic Characteristics of the Whole-Genome-Sequenced *S. aureus* Isolates

Parameter	Observed range
WGS isolates	24
Genome size	2.65–2.90 Mb
GC content	32.5–33.0%
Predicted CDS	2,400–2,850
N50	40–150 kb
Genome completeness	>98%
Contamination	<2%

Table 6: Multilocus Sequence Typing and Clonal Complex Distribution of the Whole-Genome-Sequenced Isolates

Sequence type/clonal complex	No. isolates
ST239/CC8	5
ST22/CC22	4
ST5/CC5	3
ST30/CC30	3
ST8/CC8	3
ST15/CC15	2
Other STs	4
Total	24

The presence of multiple sequence types supported the genetic heterogeneity of the clinical population. Recent WGS studies similarly demonstrate the utility of genome sequencing for simultaneously characterizing antimicrobial resistance, virulence, MLST, and phylogenetic diversity in *S. aureus* [6].

3.5. Pan-Genome Analysis and Identification of Conserved Proteins

Pan-genome analysis of the 24 sequenced isolates identified approximately 3,800 gene clusters, including 2,250 core genes, 900 accessory genes, and 650 unique genes Table 7.

Table 7: Pan-Genome Composition of the Whole-Genome-Sequenced *S. aureus* Isolates

Genomic component	Gene clusters
Core genome	2,250
Accessory genome	900
Unique genome	650
Total	3,800

The predicted proteomes contained approximately 2,700 proteins. Screening for surface-associated or secreted proteins identified 285 candidates, of which 94 proteins were conserved among the sequenced isolates. Further functional screening identified 31 proteins associated with bacterial survival, adhesion, colonization, or virulence. This progressive reduction provided a practical genomic basis for selecting conserved proteins for subsequent vaccine-oriented analysis. Recent studies have similarly used genome- and proteome-based approaches to identify conserved *S. aureus* proteins as potential vaccine targets [11, 13].

3.6. Reverse Vaccinology and Immunoinformatic Prioritization

The 31 conserved and biologically relevant proteins were subjected to sequential reverse-vaccinology screening based on conservation, predicted localization, biological relevance, antigenicity, allergenicity, toxicity, and similarity to human proteins Table 8. The screening resulted in 18 biologically important proteins, followed by 15 proteins without significant human homology, 11 predicted antigenic proteins, 8 non-allergenic proteins, and finally 4 prioritized candidate antigens, designated C1, C2, C5, and C6.

Table 8: Sequential Filtering and Prioritization of Conserved Proteins for Vaccine Candidate Selection

Screening stage	Remaining proteins
Predicted proteins	2,700
Surface-associated/secreted proteins	285
Conserved proteins	94
Virulence/survival-associated proteins	31
Biologically important proteins	18
Without significant human homology	15
Predicted antigenic proteins	11
Non-allergenic proteins	8
Favorable toxicity profile	6
Final prioritized candidates	4

B-cell and T-cell epitope prediction was subsequently performed for the four final candidates. The candidates contained multiple predicted B-cell epitopes and MHC class I and class II binding peptides Table 9.

Table 9: Immunoinformatic Characteristics of the Prioritized *S. aureus* Vaccine Candidate Antigens

Candidate	B-cell epitopes	MHC-I epitopes	MHC-II epitopes	Population coverage
C1	4	5	7	88%
C2	3	4	6	85%
C5	5	8	10	94%
C6	3	3	5	83%

Among the four candidates, C5 demonstrated the most favorable overall immunoinformatic profile, with the highest number of retained B-cell and T-cell epitopes and the highest estimated population coverage. The use of antigenicity, allergenicity, human-homology, B-cell, T-cell, MHC-binding, and population-coverage criteria is consistent with established reverse-vaccinology strategies for *S. aureus* vaccine candidate discovery.

3.7. Integrated Summary of the Main Findings

The integrated analysis identified 112 clinical *S. aureus* isolates, including 48 MRSA isolates, and demonstrated substantial antimicrobial-resistance and virulence diversity. WGS of 24 representative isolates revealed multiple genetic lineages and a pan-genome comprising conserved, accessory, and strain-specific components. The subsequent comparative genomic and reverse-vaccinology analyses reduced the initial protein dataset to four conserved candidate antigens (C1, C2, C5, and C6). Among these, C5 was the highest-priority computational candidate based on its conservation, predicted antigenicity, epitope profile, and population coverage. These candidates should be regarded as potential vaccine antigens requiring subsequent experimental validation, rather than confirmed protective antigens.

4. Discussion

4.1. Antimicrobial Resistance and Virulence Profiles of *Clinical Staphylococcus aureus* Isolates

Staphylococcus aureus was isolated from 112 of the 150 clinical specimens, representing 74.7% of the samples examined. Its greater frequency in wound and pus specimens aligns with its established role in colonizing and infecting skin and soft tissue. The antimicrobial susceptibility results demonstrated particularly high resistance to penicillin (81.3%), followed by erythromycin (46.4%), whereas resistance to cefoxitin was detected in 42.9% of isolates, corresponding to the MRSA group. In addition, 38.4% of isolates were classified as multidrug-resistant. The high prevalence of penicillin resistance was accompanied by a high frequency of the *blaZ* gene (83.0%), supporting the contribution of β -lactamase-mediated resistance to the observed phenotype. Similarly, *mecA* was identified in 41.1% of isolates, which was broadly comparable with the proportion of cefoxitin-resistant isolates. Such phenotype–genotype associations have also been reported in recent genomic investigations of clinical *S. aureus*, although complete concordance is not always expected because resistance phenotypes may be influenced by gene expression, regulatory mechanisms, mutations, or resistance determinants not included in targeted PCR assays [2, 26, 27]. The lower frequencies of *erm*, *tet*, and aminoglycoside-resistance determinants suggest that these mechanisms were distributed heterogeneously among the investigated isolates rather than being uniformly present.

The virulence-gene analysis demonstrated a high prevalence of several genes associated with adhesion, colonization, biofilm formation, and tissue damage. The detection of *spa* (97.3%), *clfA* (94.6%), *clfB* (92.0%), *hla* (90.2%), *hld* (85.7%), and *icaD* (83.9%) indicates that these determinants were widely represented within the studied clinical population. The relatively high prevalence of *fibA* (81.3%) and *icaA* (78.6%) further supports the frequent occurrence of genes involved in host-cell adhesion and biofilm-associated processes. In contrast, the *lukS-PV/lukF-PV* genes were detected in 21.4% of isolates, indicating a more restricted distribution of this particular leukocidin determinant. Recent whole-genome studies have similarly demonstrated considerable variation in the distribution of virulence-associated genes among clinical *S. aureus* isolates, with combinations of adhesion, biofilm, toxin, and immune-evasion determinants occurring across different genomic lineages [5, 28]. However, detecting a virulence-associated gene does not, by itself, show that the gene is expressed or that clinical disease is more severe. Instead, these findings suggest that the isolates have the genetic capacity for particular pathogenic mechanisms. Because antimicrobial-resistance and virulence determinants can occur within the same clinical population, both features need to be characterized when evaluating the genomic diversity of *S. aureus*.

4.2. Genomic Diversity, Whole-Genome Sequencing, and Pan-Genome Characteristics

The 16S rRNA analysis verified the identities of the selected isolates, which showed 99.2–100% sequence identity relative to reference *S. aureus* sequences. As a confirmation step, 16S rRNA sequencing was suitable for supporting species identification, but its limited ability to distinguish closely related *S. aureus* lineages meant that it could not provide detailed epidemiological or genomic characterization. Whole-genome sequencing was therefore carried out on 24 representative isolates. The resulting genomes were approximately 2.65–2.90 Mb in size, with GC contents of 32.5–33.0%, high completeness (>98%), and low contamination (<2%). These metrics indicated that the selected genomes were suitable for downstream comparative genomic analyses. MLST analysis identified several sequence types and clonal complexes, including ST239/CC8, ST22/CC22, ST5/CC5, ST30/CC30, ST8/CC8, and ST15/CC15. The presence of multiple sequence types supports the genetic heterogeneity observed among the clinical isolates. Similar diversity has been reported in recent WGS-based investigations of clinical *S. aureus*, in which several sequence types and clonal complexes were identified within individual clinical or geographical populations [29]. The presence of both a relatively conserved core genome and a substantial accessory component reflects the genomic plasticity characteristic of *S. aureus*. Accessory genomic regions may help explain why isolates differ in antimicrobial resistance, virulence, host adaptation, and other phenotypic traits. Conserved genes, by contrast, offer a steadier genomic basis for finding targets that could be shared across several isolates. That difference matters especially in vaccine-focused research: an antigen found in only a small fraction of strains may provide limited coverage. Earlier genomic studies showed that WGS and pan-genome analyses can uncover substantial diversity within *S. aureus* populations while also identifying conserved genomic components that may support subsequent target prioritization [30]. Here, approximately 2,700 predicted proteins were screened. This process identified 285 surface-associated or secreted proteins, including 94 conserved across the genomes analyzed. Those conserved proteins were taken forward for additional evaluation, rather than selecting targets only from proteins linked to individual strains. Conservation among the 24 sequenced isolates does not, by itself, demonstrate that the feature is conserved throughout the global *S. aureus* population. Broader genomic datasets would be needed to assess how broadly these findings apply.

4.3. Reverse Vaccinology and Immunoinformatic Prioritization of Candidate Antigens

The subsequent reverse-vaccinology analysis progressively reduced the number of potential targets from 31 biologically relevant proteins to four prioritized candidates, designated C1, C2, C5, and C6. The filtering strategy considered conservation, biological relevance, human homology, predicted antigenicity, allergenicity, and toxicity. This sequential approach is consistent with previous reverse-vaccinology studies in *S. aureus*, in which genomic or proteomic datasets were progressively screened to identify proteins with characteristics potentially suitable for vaccine development [11, 12]. The identification of 11 predicted antigenic proteins after the initial filtering and the subsequent prioritization of eight non-allergenic and six favorable candidates demonstrate the utility of combining several computational criteria rather than relying on a single antigenicity score. Such multi-parameter prioritization can reduce the number of candidates requiring subsequent experimental investigation.

The immunoinformatic analysis identified B-cell and MHC-I/MHC-II-associated epitopes within all four prioritized proteins. Predicted population coverage ranged from 83% for C6 to 94% for C5, with C5 also showing the highest number of predicted B-cell, MHC-I, and MHC-II epitopes among the four candidates. These characteristics led to the prioritization of C5 as the leading computational candidate in the present study. Similar immunoinformatic approaches have been used to identify potential *S. aureus* vaccine candidates by combining reverse vaccinology with B- and T-cell epitope prediction and population-coverage analysis [13]. However, these predictions should be interpreted cautiously. Antigenicity scores, epitope predictions, and theoretical population coverage provide evidence for computational prioritization

but do not demonstrate that the corresponding proteins are expressed during infection, accessible to the immune system, immunogenic in vivo, or capable of providing protection. Recent reviews of *S. aureus* vaccine development have emphasized the continuing difficulty of translating computationally attractive targets into effective vaccines and the importance of experimental validation [8, 9]. Accordingly, C1, C2, C5, and C6 may be considered candidate vaccine antigens, but not confirmed protective antigens. Within this dataset, C5 is the highest-priority computational candidate. Determining whether any of these prioritized proteins can provide genuine protection against infection requires experimental work, including the expression of recombinant proteins, the assessment of antigen-specific immune responses, and ultimately the use of appropriate infection models.

5. Conclusions

The present study found substantial phenotypic, molecular, and genomic diversity among clinical *S. aureus* isolates. Of the 150 clinical specimens examined, 112 isolates were confirmed as *S. aureus*; 48 were MRSA and 43 were MDR. *blaZ* and *mecA* occurred at high frequencies, while genes linked to adhesion, biofilm formation, toxin production, iron acquisition, and immune evasion were broadly distributed. WGS of 24 representative isolates identified multiple sequence types and clonal complexes, along with a complex pan-genome composed of conserved, accessory, and unique genetic components. Subsequently, the comparative analysis narrowed the focus to a set of conserved proteins whether surface-associated or secreted likely to play a role in bacterial survival, adhesion, colonization, and virulence. Sequential reverse vaccinology and immune informatics screening identified four candidate antigens C1, C2, C5, and C6. C5 had the most favorable computational profile, with predicted B-cell and MHC-I/MHC-II epitopes and the highest theoretical population coverage. Overall, the combined use of microbiology, antimicrobial-resistance and virulence profiling, WGS, pangenomics, reverse vaccinology, and immune informatics made it possible to prioritize conserved *S. aureus* vaccine candidates. Still, C1, C2, C5, and C6 are computationally predicted candidates, not confirmed protective antigens. They require experimental validation, including recombinant protein production, immunological assays, and suitable infection models. This integrated approach offers a practical basis for identifying conserved targets for future vaccine, diagnostic, and therapeutic development against multidrug-resistant *S. aureus*.

Article Information

Ethical Approval: The ethical approval for this study was granted by the College of Health and Medical Technologies, University of Alkafeel, Najaf, Iraq, in early September 2026. The study was conducted in accordance with the applicable institutional ethical guidelines and relevant principles governing research involving clinical bacterial isolates. The approval was issued on 1 September 2026.

Informed Consent: Informed consent was obtained from all participants, where applicable, in accordance with the approved study protocol and institutional ethical guidelines. Where informed consent was waived, the waiver was granted by the relevant ethics committee.

Consent for Publication: This study does not contain any information that could identify the patients from whom the clinical samples were obtained.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Clinical Trial Registration: Not applicable. This study does not involve a clinical trial.

Conflict of Interest: The author declares no competing interests.

Funding: The author received no external funding.

Acknowledgments: The authors would like to express their sincere gratitude to all individuals and institutions who contributed to and supported the completion of this study. Special thanks are extended to the laboratory staff and participants for their valuable assistance and cooperation.

AI Use Disclosure: Generative artificial intelligence tools, including ChatGPT, were used to assist with language editing and improving the clarity of the manuscript. The authors reviewed and verified the content and take full responsibility for the final version of the manuscript.

References

- [1] S. Y. C. Tong, J. S. Davis, E. Eichenberger, T. L. Holland, and V. G. Jr. Fowler. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev*, 28(3):603–661, 2015. URL <https://doi.org/10.1128/cmr.00134-14>.
- [2] J. Hisatsune, S. Kutsuno, Y. Iwao, K. Ishida-Kuroki, et al. *Staphylococcus aureus* ST764-SCCmecII high-risk clone in bloodstream infections revealed through national genomic surveillance integrating clinical data. *Nat Commun*, 16(1):2698, 2025. URL <https://doi.org/10.1038/s41467-025-57575-2>. PMID:40108131.
- [3] D. G. Han, D. H. Kim, J. S. Yoo, and S. M. Jeon. Comprehensive analysis of antimicrobial resistance and virulence determinants in clinical methicillin-resistant *Staphylococcus aureus* isolates collected in South Korea. *J Infect Public Health*, 19(5):103215, 2026. URL <https://doi.org/10.1016/j.jiph.2026.103215>. PMID: 41945996.
- [4] T. J. Foster, J. A. Geoghegan, V. K. Ganesh, and M. Höök. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat Rev Microbiol*, 12(1):49–62, 2014.

- [5] E. Noumi, T. Zmantar, N. Bouali, A. S. Bazaid, K. F. Alabbosh, K. Chaieb, H. N. Altayb, V. Feo, and M. Snoussi. Whole-genome sequencing and biofilm gene characterization in multidrug-resistant *Staphylococcus aureus* clinical strains. *J Genet Eng Biotechnol*, 23(3), 2025. URL <https://doi.org/10.1016/j.jgeb.2025.100521>. PMID:40854640.
- [6] T. Inagawa, J. Hisatsune, S. Kutsuno, Y. Iwao, Y. Koba, S. Kashiya, K. Ota, N. Shime, and M. Sugai. Genomic characterization of *Staphylococcus aureus* isolated from patients admitted to intensive care units of a tertiary care hospital: epidemiological risk of nasal carriage of virulent clone during admission. *Microbiol Spectr*, 12(6), 2024. URL <https://doi.org/10.1128/spectrum.02950-23>. PMID:38709078.
- [7] A. Kuddus, M. R. Islam, A. Hossain, K. A. Bongulto, K. Watanabe, and M. A. Muhit. Whole-genome sequencing to explore genomic profile, pathogenicity, virulence motifs, and pangenome of clinical MRSA isolates. *Sci Rep*, 16(1), 2025. URL <https://doi.org/10.1038/s41598-025-32897-9>. PMID:41402559.
- [8] R. Lapeyre, N. Rezk, S. McClean, and A. C. Moore. Systematic mapping review of preclinical and clinical studies of *Staphylococcus aureus* vaccines. *Vaccine*, 61:127333, 2025. URL <https://doi.org/10.1016/j.vaccine.2025.127333>.
- [9] K. Shinwari, B. Vieira, I. Ciaparin, A. P. Hakansson, M. Darrieux, and T. R. Converso. Overcoming immune evasion in *Staphylococcus aureus*: strategies for rational vaccine design. *ACS Infect Dis*, 11(10), 2025. URL <https://doi.org/10.1021/acsinfecdis.5c00569>.
- [10] S. J. Goodswen, P. J. Kennedy, and J. T. Ellis. A guide to current methodology and usage of reverse vaccinology towards in silico vaccine discovery. *FEMS Microbiol Rev*, 47(2), 2023. URL <https://doi.org/10.1093/femsre/fuad004>.
- [11] M. A. Soltan, D. Magdy, S. M. Solymann, and A. Hanora. Design of *Staphylococcus aureus* new vaccine candidates with B and T cell epitope mapping, reverse vaccinology, and immunoinformatics. *OMICS*, 24(4):195–204, 2020. URL <https://doi.org/10.1089/omi.2019.0183>.
- [12] N. N. Goodarzi, N. Bolourchi, S. Fereshteh, A. S. Shirazi, M. R. Pourmand, and F. Badmasti. Investigation of novel putative immunogenic targets against *Staphylococcus aureus* using a reverse vaccinology strategy. *Infect Genet Evol*, 96:105149, 2021. URL <https://doi.org/10.1016/j.meegid.2021.105149>.
- [13] A. Salemi, M. M. Pourseif, Y. Masoudi-Sobhanzadeh, R. Ansari, and Y. Omid. Proteome-wide reverse vaccinology to identify potential vaccine candidates against *Staphylococcus aureus*. *Mol Immunol*, 183:296–312, 2025. URL <https://doi.org/10.1016/j.molimm.2025.05.016>.
- [14] M. S. Ibrahim, J. A. Fahmy, M. A. Helal, F. T. Hakim, L. R. Sidarous, N. E. Attallah, S. Abouelhadid, M. Elhadidy, and E. Badr. AntiPan: a genome-informed in silico pipeline for advancing subunit vaccine discovery against *Staphylococcus aureus*. *Sci Rep*, 16(1), 2026. URL <https://doi.org/10.1038/s41598-025-33445-1>. PMID:41663438.
- [15] *Performance Standards for Antimicrobial Susceptibility Testing*. Clinical and Laboratory Standards Institute (CLSI), 36th edition. CLSI supplement M100. Wayne, PA: Clinical and Laboratory Standards Institute; 2026. The current M100 edition provides updated antimicrobial susceptibility breakpoints and quality-control criteria used with CLSI M02 and M07.
- [16] K. Murakami, W. Minamide, K. Wada, E. Nakamura, H. Teraoka, and S. Watanabe. Identification of methicillin-resistant strains of staphylococci by polymerase chain reaction. *J Clin Microbiol*, 29(10):2240–2244, 1991. URL <https://doi.org/10.1128/jcm.29.10.2240-2244.1991>.
- [17] J. A. McClure, J. M. Conly, V. Lau, S. Elsayed, T. Louie, W. Hutchins, and K. Zhang. Novel multiplex PCR assay for detection of the staphylococcal virulence marker Pantone-Valentine leukocidin genes and simultaneous discrimination of methicillin-susceptible from -resistant staphylococci. *J Clin Microbiol*, 44(3):1141–1144, 2006. URL <https://doi.org/10.1128/jcm.44.3.1141-1144.2006>.
- [18] A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*, 19(5):455–477, 2012. URL <https://doi.org/10.1089/cmb.2012.0021>.
- [19] A. Gurevich, V. Saveliev, N. Vyahhi, and G. Tesler. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8):1072–1075, 2013. URL <https://doi.org/10.1093/bioinformatics/btt086>.
- [20] T. Seemann. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14):2068–2069, 2014. URL <https://doi.org/10.1093/bioinformatics/btu153>.
- [21] B. P. Alcock, A. R. Raphenya, T. T. Y. Lau, K. K. Tsang, M. Bouchard, A. Edalatmand, et al. CARD 2020: antibiotic resistance surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.*, 48(D1):D517–D525, 2020. URL <https://doi.org/10.1093/nar/gkz935>.
- [22] B. Liu, D. Zheng, S. Zhou, L. Chen, and J. Yang. VFDB 2019: a comparative pathogenomic platform with an update to the database. *Nucleic Acids Res.*, 47(D1):D687–D692, 2019.
- [23] G. Tonkin-Hill, N. MacAlasdair, C. Ruis, A. Weimann, G. Horesh, J. A. Lees, R. A. Gladstone, S. Lo, C. Beaudoin, R. A. Floto, S. D. W. Frost, J. Corander, S. D. Bentley, and J. Parkhill. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol.*, 21(1):136, 2020. URL <https://doi.org/10.1186/s13059-020-02090-4>.

- [24] I. A. Doytchinova and D. R. Flower. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8:4, 2007. URL <https://doi.org/10.1186/1471-2105-8-4>.
- [25] W. Fleri, S. Paul, S. K. Dhanda, S. Mahajan, X. Xu, B. Peters, and A. Sette. The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design. *Front Immunol*, 8:278, 2017. URL <https://doi.org/10.3389/fimmu.2017.00278>.
- [26] L. Liu, Z. Yan, F. He, J. Chen, L. Kuang, X. Liu, Y. Cui, X. Wang, C. Miao, H. Li, and Y. Jiang. Antibiotic susceptibility and molecular characterization based on whole-genome sequencing of *Staphylococcus aureus* causing invasive infection in children and women living in Southwest China during 2018–2023. *BMC Microbiology*, 25:47, 2025. URL <https://doi.org/10.1186/s12866-025-03758-2>.
- [27] M. Snoussi, E. Noumi, N. Bouali, A. S. Bazaid, M. M. Alreshidi, H. N. Altayb, and K. Chaieb. Antibiotic Susceptibility Profiling of Human Pathogenic *Staphylococcus aureus* Strains Using Whole Genome Sequencing and Genome-Scale Annotation Approaches. *Microorganisms*, 11(5), 2023. URL <https://doi.org/10.3390/microorganisms11051124>. PMID: 37317098.
- [28] A. Shelenkov, A. Slavokhotova, M. Yunusova, V. Kulikov, Y. Mikhaylova, and V. Akimkin. Genomic typing, antimicrobial resistance gene, virulence factor and plasmid replicon database for the important pathogenic bacteria *Staphylococcus aureus*. *BMC Genomic Data*, 26:65, 2025. URL <https://doi.org/10.1186/s12863-025-01363-w>.
- [29] A. M. Che Hamzah, C. H. Chew, E. I. Al-Trad, S. M. Puah, K. H. Chua, N. I. A. Rahman, S. Ismail, T. Maeda, P. Palittapongarnpim, and C. C. Yeo. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus* clinical isolates from Terengganu, Malaysia, indicates the predominance of the EMRSA-15 (ST22-SCCmec IV) clone. *Sci Rep*, 14(1):3485, 2024. URL <https://doi.org/10.1038/s41598-024-54182-x>.
- [30] J. Yu, Y. Ding, X. Zhang, Y. Fang, S. Tai, E. Yuan, and Y. Duan. Whole-genome sequencing analysis of *Staphylococcus aureus* isolated from female patients with mastitis in Henan, China. *Journal of Global Antimicrobial Resistance*, 44:15–20, 2025. URL <https://doi.org/10.1016/j.jgar.2025.05.019>.