JEPIBIO
Journal of Epidemiology and Infection Biology

Research Article

Insights from *S. aureus* Pangenome along with Agr-typing Revealed Agr operon Reduction in Chronic Elderly Patients

Syed Basit Ali Shah [1,2], Muhammad Makeen Alam[3]*

[1] Ministry of Education, Key Laboratory of Saline Alkali Vegetation Ecology Restoration, Northeast Forestry University, Harbin, China
[2] College of Life Science, Northeast Forestry University, Harbin, China
[3] School of Engineering, Ulster University, Belfast, Northern Ireland, United Kingdom

*Corresponding author:

Muhammad Makeen Alam
(e-mail: Alam-MM5@ulster.ac.uk)

Date of Receiving: 07/09/2024
Date of Acceptance 10/10/2024
Date of Publishing 10/12/2024

ABSTRACT *Staphylococcus aureus* (*S. aureus*) is a pathogen responsible for various infections, including skin and soft tissue infections (SSTI) and bloodstream infections. Methicillin-resistant *S. aureus* (MRSA) presents treatment challenges due to antibiotic resistance. To investigate the genetic diversity, pangenome structure, and the role of the agr operon in *S. aureus* using whole-genome analyses. A total of 460 complete *S. aureus* genomes were retrieved from the NCBI database. Whole-genome average nucleotide identity (ANI) and genome-to-genome distance were computed using FASTANI and GGDC, respectively. Metadata, including isolation host, country, and antibiotic resistance, were extracted. Genomes were annotated using Prokka. Pangenome analysis was conducted with Roary, categorizing genes into core, soft core, shell, and cloud. Recombination detection and phylogenetic analysis were performed using Gubbins and FastTree. Agr operon genes were analyzed using in-house scripts for local BLAST searches. The pangenome comprised 6,159 genes, with 2,093 core genes and 3,275 cloud genes. Pangenome analysis revealed an open pangenome. Core genome phylogeny identified significant genetic diversity and distinct population structures. Agr-typing showed dominance of agrI (59.47%) followed by agrII (25.27%) and agrIII (12.85%). One genome lacked the agr operon. The distribution of clonal complexes (CCs) varied, with CC8 being prominent in MRSA genomes. MRSA strains were dominant, especially within CC8, reflecting high methicillin resistance. The study highlights the genetic diversity and adaptability of *S. aureus*. The agr operon plays a critical role in virulence regulation. The absence of the agr operon in one strain suggests possible niche adaptation. Continuous surveillance and molecular typing are essential to monitor resistant strain's evolution and spread.

KEYWORDS *Staphylococcus*, Antimicrobial resistance, Genome Reduction, Evolution, Agr typing

Introduction

Staphylococcus aureus is an opportunistic pathogen that causes primarily skin and soft tissue infections in humans (Lai *et al*, 2024). Methicillin-resistant *S. aureus* (MRSA) is a multi-drug-resistant strain of *S. aureus*. The healthcare industry faces operational challenges from Methicillin-resistant *S. aureus* (MRSA) which represents multi-drug-resistant variants of S. *aureus*. Healthcare institutions and common environments both serve as locations where people acquire MRSA infections. Community-acquired strains have emerged as a specific threat due to their ability to infect with unexpected risk variables (Sano *et al*, 2022). New medical treatments for MRSA strains are being discovered (Motallebi *et al*, 2021; Rimi *et al*, 2024). Previous research shows that different antibiotic treatments effectively stop MRSA colonies from growing. MRSA infection rates found across various patient groups. Hospital-based age-specific testing on *S. aureus* shows elderly people aged above 50 who experience the most infections (Cavalcante *et al*, 2020). Majority of these elderly patients are reported to be infected with MRSA. Healthcare providers should use risk identification and therapy development with infection control practices to reduce MRSA burden thereby improving patient outcomes.

Previous studies demonstrated that the accessory gene regulators (agr) control *S. aureus* bacterial functions such as biofilm development, antibiotic defensiveness, and tissue damage potential (Javdan *et al*, 2019; Raghuram *et al*, 2022; Rimi *et al*, 2024). *Staphylococcus aureus* depends on the agr operon to manage its virulence factors along with controlling its quorum sensing processes (Raghuram *et al*, 2022). Extensive examination of this operon occurred across various clinical environments specifically human chronic infections (Tan *et al*, 2018; Javdan *et al*, 2019; Mossop *et al*, 2023). Four genes such as agrA, agrB, agrC and agrD constitute the agr operon which determines *S. aureus* strain adaptations across different environments (Giulieri *et al*, 2022). Furthermore, interspecies regulation between S. *caprae* and *S. aureus* suggests a complex network of interactions that may impact the pathogenicity of these bacteria (Ogura *et al*, 2022). Studies have shown that different strains of *S. aureus* can be characterized using comparison of agr operon sequence types. Existence of Agr-typing shows evolution from common ancestor and variation in virulence. Agr-types include agrI, agrII, agrIII, and agrIV that are further linked with clinically important phenotypes (Raghuram *et al*, 2022). These agr-types are often linked with particular niche environments such as chronic wounds etc. Agr typing play a critical role in understanding the pathogenesis and evolution of *S. aureus* in chronic human infections.

Genome reduction has also been reported in bacterial infections such as *Mycobacterium spp.*, *E. coli spp*. In a longitudinal study conducted on cystic fibrosis patients, genetic diversification of Mycobacterium abscesses within patients was observed (Lewin *et al*, 2021). It is established that phenotypic and genetic heterogeneity occurs in gram negative bacteria such as *Acinetobacter baumannii* strains during chronic infections, with variations in genome sizes and nucleotide sequences (Valcek *et al*, 2023). Variation in niche such as severe infections can act as selective pressure for *S. aureus* which undergoes genome degradation leading to adaptation (Giulieri *et al*, 2022). Additionally, *S. aureus* small-colony variants have been linked to chronic infections in human patients (Zhou *et al*, 2022). The current study was performed to understand the effects of genome reduction on agr operon using pangenome approaches and agr typing.

Materials and Methods

Data Collection

All the *S. aureus* genomes were retrieved from NCBI database (Awan *et al*, 2021) comprising of 460 complete genomes and chromosomes (Supplementary Table 1). All these genomes were included in whole-genome Average Nucleotide Identity (ANI) computation; performed by FASTANI (website and package). It utilizes Mashmap to calculate the ortholog sequence mappings and alignment identity. Furthermore, to perform the similar computations based on genome-to-genome genetic distance, Genome-to-Genome Distance Calculator (GGDC) was utilized (Meier-Kolthoff *et al*, 2013; Meier-Kolthoff *et al*, 2022).

Metadata isolation

Traits such as isolation host, country of origin, and antibiotic resistance (MRSA/MSSA) related to *S. aureus* were downloaded from NCBI and PATRIC (Gillespie *et al*, 2011). As the data in public databases is not complete, in-silico MLST was performed to find out the MLST groups and clonal complex groups of this dataset (https://github.com/tseemann/mlst-1). This metadata was used to associate patterns of gene presence and absence with phenotypes exhibited by groups of taxa.

Genome annotation

To get the stable design in the dataset, Prokka was utilized to predict the genes and annotate all the genomes. Prokka utilizes the Prodigal in gene ORF prediction that further annotated by databases from UNIPROT and ISfinder (Seemann, 2014). MinCED package utilized by Prokka predicted the CRISPR sequences in the genomes (https://github.com/ctSkennerton/minced).

Pan-genome inference

The pan-genome of *S. aureus* was inferred with Roary (Page *et al*, 2015). The Prokka annotated genomes in gff format were used as input to Roary. These genomes were grouped by CD-Hit and gene presence-absence binary matrix was produced. Additionally, Roary also produced multi-fasta core gene alignment and first 4000 accessory genes-based tree using PRANK and FastTree 2 (version 2.1.9) respectively (Price *et al*, 2010; Löytynoja, 2014). Roary categorized the pan-genome into core, soft core, shell, and cloud. For further simple categorization, Epi Gene package in R was utilized that categorizes the binary matrix into core, accessory, and unique genes along with the identification of respective groups/clusters and genomes (github.com/furqan915/Epi-gene).

Core genome analysis

Recombination detection is an important step in prediction of accurate phylogeny. The core genome produced in previous step was used as input in gubbins package that detected the recombination sequences in the core genome alignment. The snp-sites package was utilized to filter the SNPs. This package identified both the substitution and insertions/deletions. Furthermore, FastTree was used to build a new tree while all visualizations of phylogenetic trees were produced with FigTree version 1.4.2 and iTOL version 3.5 (Letunic and Bork, 2021).

Bayesian Analysis of Population Structure (BAPS) analysis and antimicrobial resistance

Core gnome phylogeny was further utilized to estimate population structure using BAPS algorithm (Tonkin-Hill *et al*, 2019). The frequency of all genes in BAPS cluster relative to the rest of the clades were calculated. AMR genes and plasmid replicons were detected using ARIBA and the comprehensive antibiotic resistance database (CARD) and the Plasmid Finder databases. A similar approach was utilized to detect virulence factors, using the Virulence Factor Database (VFDB).

Agr-operon typing

Agr-operon genes (Agr A, B, C, D) were initially retrieved from NCBI gene database. Local databases consisting of these genes and genomes were setup. An in-house script was developed to run the local BLAST using USEARCH and the local databases (Edgar, 2010). To determine agr typing, all the possible agrD sequences were retrieved from NCBI and UNIPROT. Frequency of the Agr genes and agr-types in the included genomes was determined along with the existence of disrupted genes detected using SNIPPY and possible kmers with MUMers. Furthermore, disrupted agr operon was also identified.

Results

A total of 460 genomes were included in this study which were retrieved from NCBI. The distribution of genomes analyzed in the study showed that the USA contributed the highest number of genomes, over 100, followed by France (n=65) (Fig. 1). Spain (n=23), UK (n=20), and Canada (n=25) also provided notable contributions. Some less developed countries, such as Mali, Venezuela, and South Korea, contributed very few genomes, often only one or two. In brief, a broad geographic distribution of genomes, including contributions from a wide range of countries across different continents.

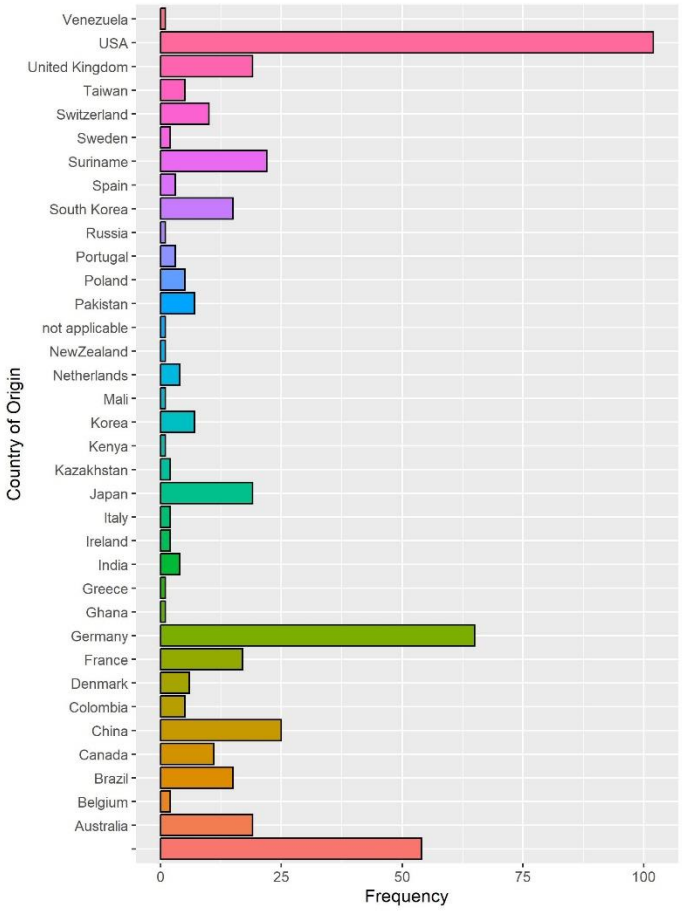

Fig. 1: Distribution of genome included from different countries around the world.

Distribution of Clonal complexes and methicillin resistance status

In-silico MLST of genomes revealed wide range of strain types (n=68). These strain types were further estimated into different 10 clonal complexes (CCs) as per mentioned guidelines. The distribution of clonal complexes was related to methicillin resistance status. Antibiotic resistance analysis revealed total of 180 methicillin-susceptible *Staphylococcus aureus* (MSSA) genomes and 279 methicillin-resistant *Staphylococcus aureus* (MRSA) genomes. The total number of isolates is significantly higher in the MRSA group compared to MSSA. Within both groups, certain clonal complexes such as CC8 and CC30 are prominently represented. Unclassified isolates also contribute a notable portion in both MSSA and MRSA categories. Additionally, other clonal complexes like CC5 and CC22 are present but in lower numbers in MSSA. It was determined that the CC8 complex is particularly dominant in the MRSA group, suggesting its prevalence in methicillin resistance among the analysed genomes (Fig. 2). The distribution of untypeable isolates shows that some strains could not be classified into specific clonal complexes, highlighting the genetic diversity within both MSSA and MRSA populations.

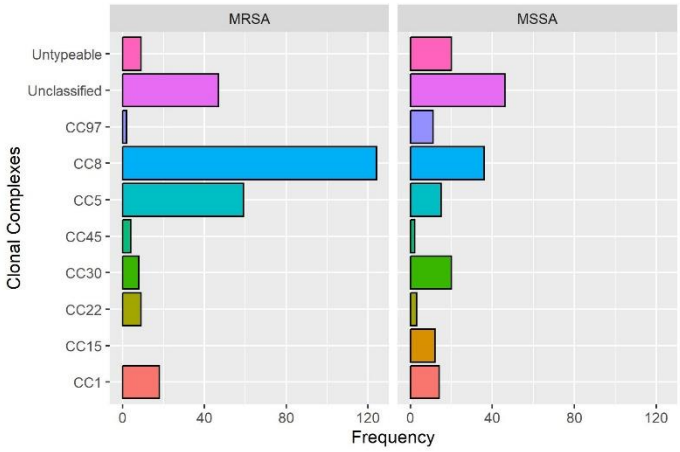

Fig. 2: Distribution of clonal complexes along with the methicillin resistance status.

Pan-genome Analysis

In initial ANI computation and GGDC analyses, one genome was removed from the initially included genomes. Further downstream pangenome analyses, remaining genomes (n=459) were utilized. The pangenome consists of 6,159 total genes, categorized into core, soft core, shell, and cloud genes based on their presence across strains. Core genes (99% to 100% genomes) and soft-core genes (95% to 99% genomes) contributed fairly in pangenome (n= 2,093, n =30, respectively). While shell genes (15% to 95% of genomes) and cloud genes (less than 15% genomes) make up higher contribution (n =761 and n =3,275, respectively). This distribution highlights the varying levels of gene conservation and variability among the strains. A binary pangenome clustering and dendrogram shows a variation in presence and absence of genes among all the genomes (Fig. 3a). The varying patterns off-diagonal indicated accessory genes, reflecting genetic diversity and potential adaptations among the strains. On the other hand, relationship between number of genomes and number of genes revealed the

current pangenome as open pangenome (Fig. 3b). The divergence between these two highlights the presence of both conserved and variable genes across the genome dataset. It underscores the dynamic nature of the pangenome, with a relatively small core genome and a large, diverse accessory genome.

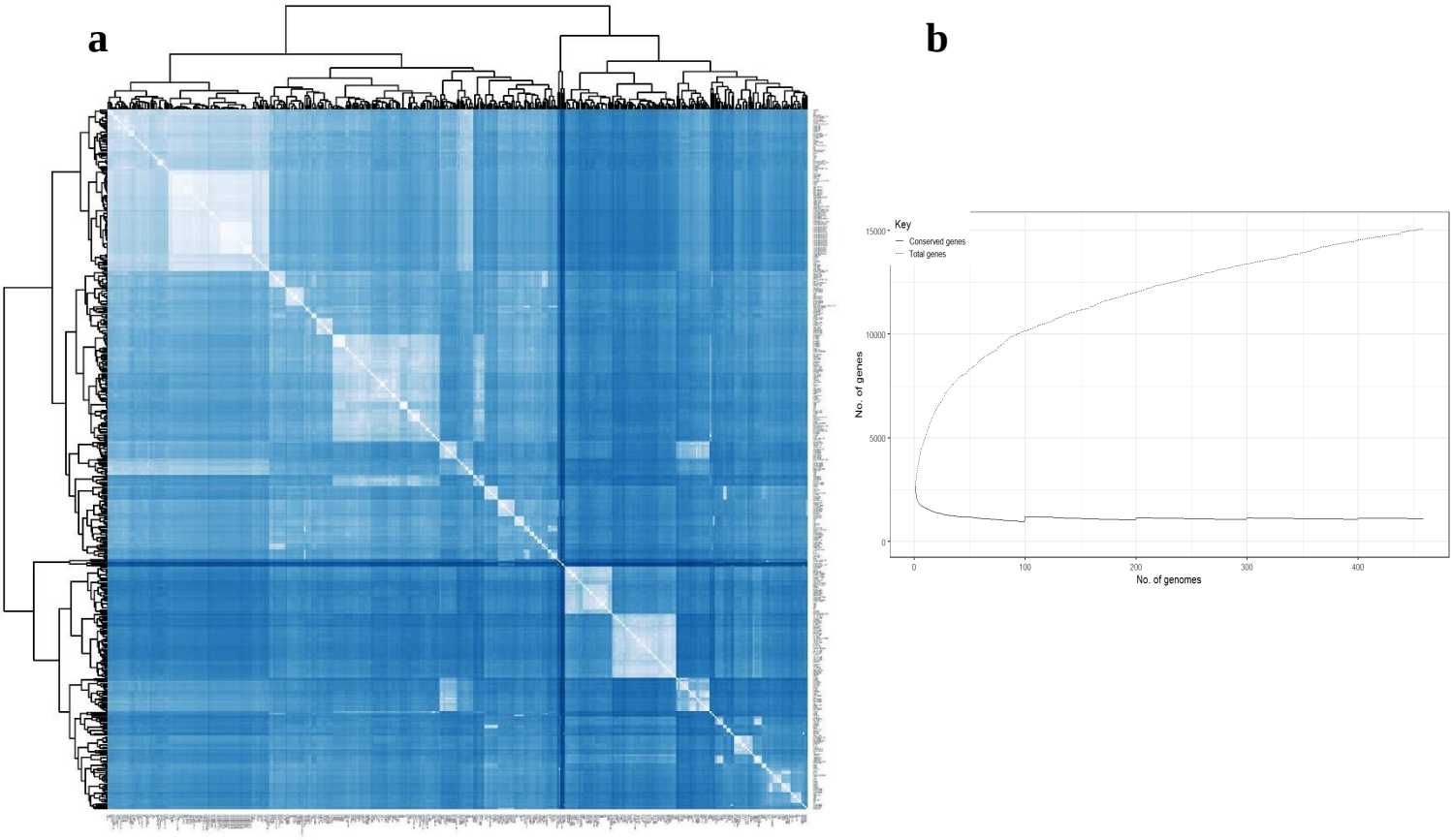

Fig. 3: Pangenome analysis revealed open-pangenome with varying levels of presence or absence of genes. (a) Heatmap illustrates the binary presence or absence of genes. The heatmap uses blue to indicate the presence of genes and white to indicate their absence across the genomes. **(b)** This figure depicts the relationship between the number of genomes and the number of genes. The solid line represents conserved genes, which remain constant as more genomes are added, indicating the core genome. The dotted line represents the total number of genes, which increases with the inclusion of additional genomes, reflecting the expanding pangenome.

Pan-genome and sequence similarity

Another downstream analysis was performed using protein sequence similarity ranging from 45 % -95% among the genomes in pangenome (Fig. 4). This analysis highlights the diversity in gene sequence similarity, indicating both conserved and divergent genetic regions among the genomes. This analysis can identify sequence variation among core, accessory, and unique gene clusters across the bacterial pangenome.

BAPS-core SNP analysis

Population Structure calculated using BAPS algorithm based on core SNP analysis revealed significant nucleotide variation and population structure among the strains. The analysis identified several distinct groups or clusters (n=9), indicated by horizontal divisions, each displaying unique patterns of nucleotide variation (Fig. 5). BAPS cluster number 5 had highest number of strains (n=150) while cluster number 9 had the least strains (n=3). These patterns reflected underlying genetic differences among the groups, with certain regions showing high SNP diversity, indicated by a mix of colors, while other regions appear more conserved within specific clusters. The focus on core SNPs highlights the shared genetic backbone of these genomes while showcasing the variations that define different populations.

Agr-Types and Clonal complexes

Agr-typing revealed 5 categories (agrI, agrII, agrIII, agrIV, and absent). Expectedly, agrI category was found highest (n=273, 59.47 %) among genomes followed by agrII (n=116, 25.27%) and agrIII (n=59, 12.85%) showed the backbone and ancestral sequence of all the agr types. AgrIV type was found in few genomes (n=10, 2.18%). Only one genome was found having no agr type which later revealed the absence of agr operon (Fig. 6a). Furthermore, distribution of these agr types across different clonal complexes (CCs) revealed distinct patterns (Fig. 6b). Among the agrI type, unclassified isolates and CC8 showed notable representation. For agrII, CC5 dominated with a high count, while agrIII type is mainly represented by CC30. The agrIV type showed minimal representation, primarily through unclassified strains.

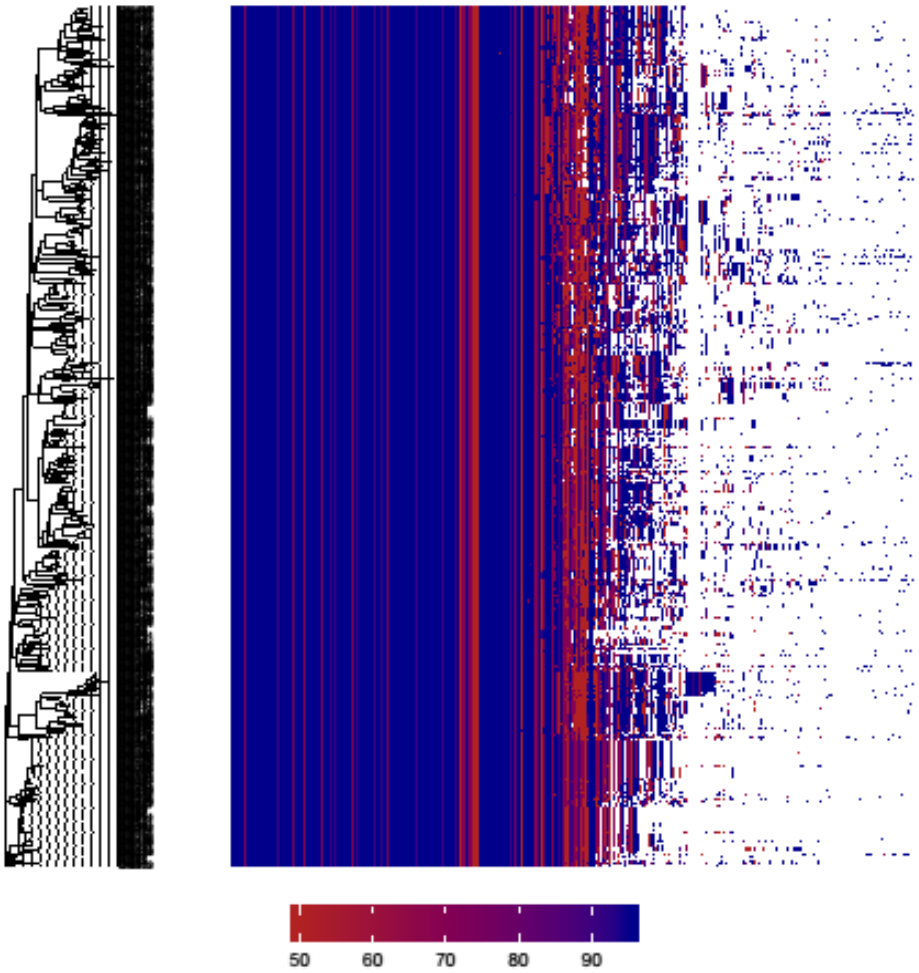

Fig. 4: Diversity of genome sequence in the pangenome. The heatmap uses a color gradient, with red representing lower similarity (50%) and blue representing higher similarity (90%). The dendrogram on the left clusters the genomes based on their gene sequence similarities.

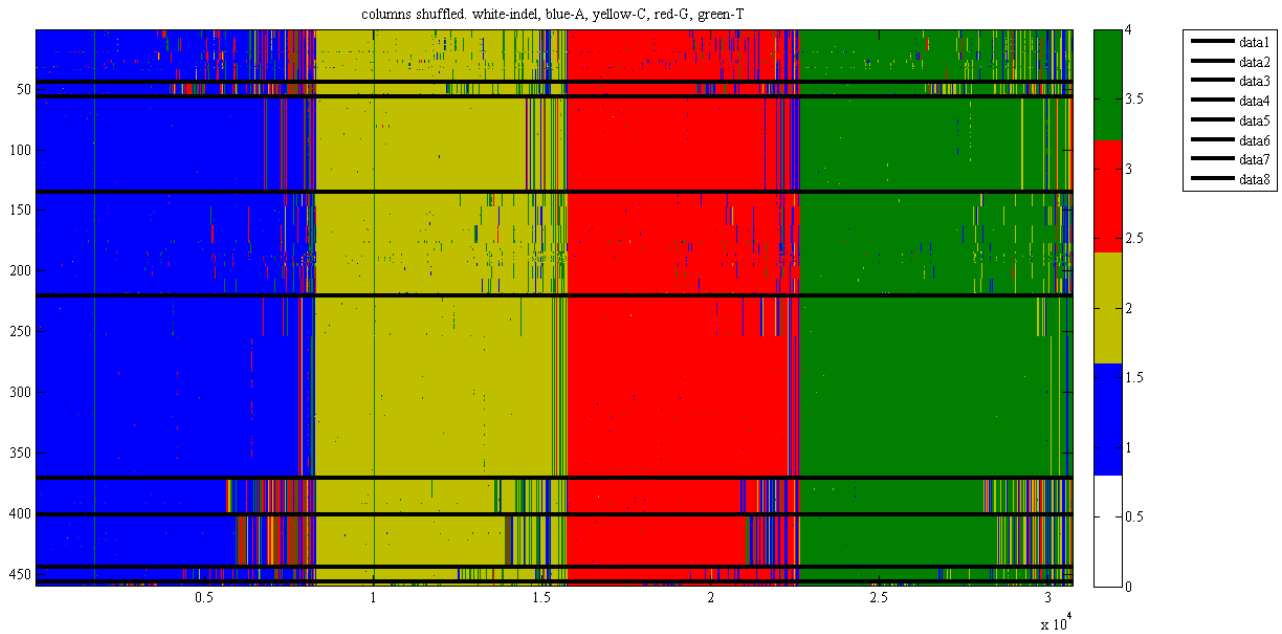

Fig. 5: BAPS Population structure based on core genome phylogeny. The heatmap visualizes this variation, with the x-axis representing SNP positions and the y-axis showing individual genomes. Distinct nucleotide bases are color-coded: blue for adenine (A), yellow for cytosine (C), red for guanine (G), green for thymine (T), and white for indels (insertions or deletions).

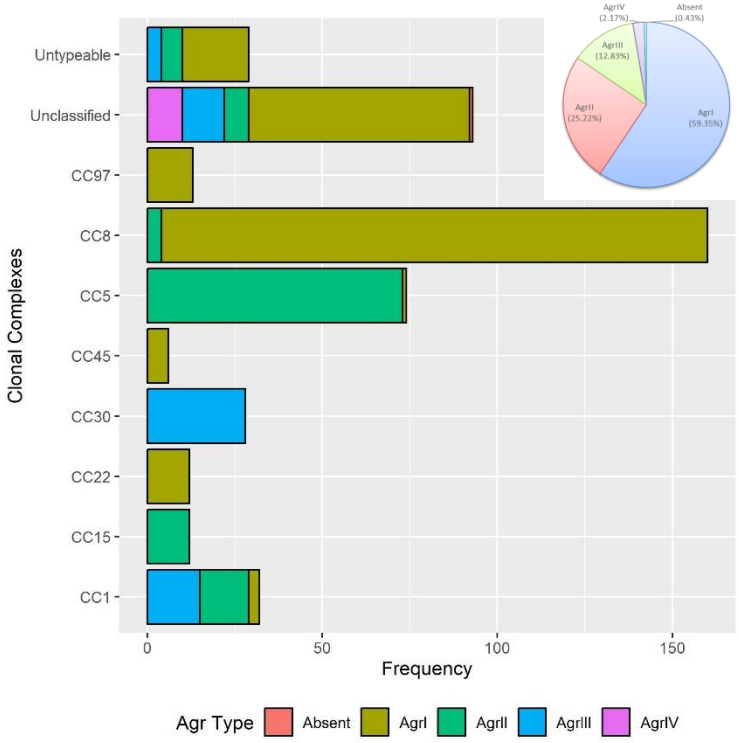

Fig. 6: Agr typing frequency and distribution of clonal complexes in the pangenome. (a) Pie chart showcasing the contribution of agr types. **(b)** Distribution of CCs and agr types.

Core genome with agr types

Core genome phylogeny revealed clustering of common origins and agr types. Phylogenetic tree illustrated the distribution of agr types across different branches suggests independent evolutionary events leading to the variation among these types. It also explains the genetic diversity and adaptability of *S. aureus* in response to environmental pressures and host interactions (Fig. 7). Majority of the strains belong to human origin followed by animal origin strains. Meta-data of few strains lack the origin related information. Most of the human origin strains were clustered together in the phylogeny while animal origin strains were clustered together. Similarly, agrI type was found clustered as mostly human origin strains were clustered. But in majority, agr typing showed no or little correlation with origin.

A comparative analysis reveals one genome lacking this agr operon. This absence of agrA and agrB genes in this strain may have significant implications for the regulatory capabilities of the agr system, potentially affecting quorum sensing and virulence. This strain belonged to human origin but clustered among agrI and agr II types.

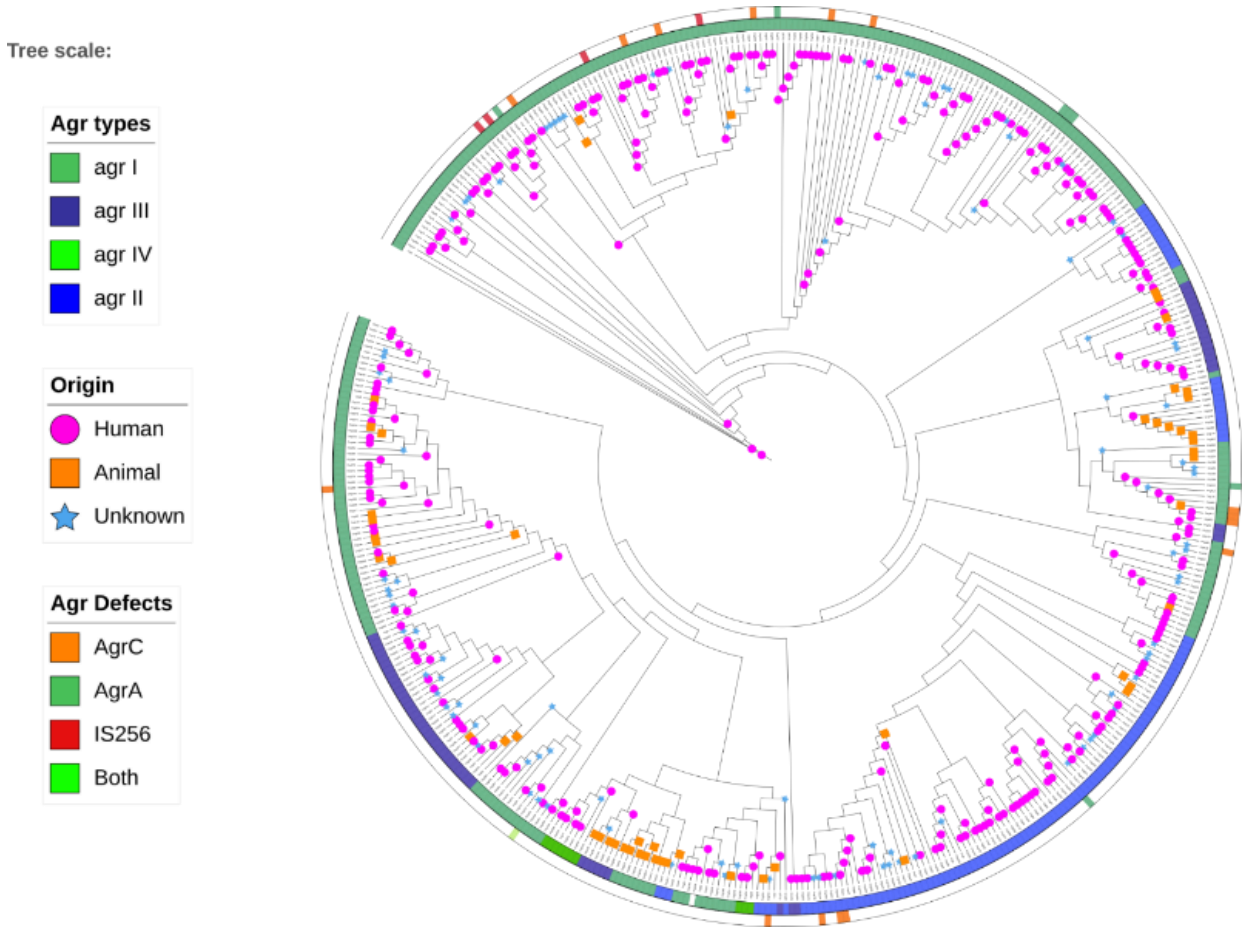

Fig .7: The circular phylogenetic tree illustrates the genetic relationships among 459 *S. aureus* genomes, highlighting variations in the agr operon.

Agr operon-lacking strain

A single genome (strain name = MS4, Genbank ID=GCA_001456215.1) found to be having no agr A and agr B genes when compared with reference strain genome (strain name = NCTC_8325, Genbank ID=GCA_000013425.1) (Fig. 8). Metadata of this strain showed that this strain was isolated from 76 years old patient of broken bone residing in Shenzhen, China (Fig. 8). This strain was found methicillin resistant and belonged to SCC-mec type Vb (5C2&5). It could be the special case of chronic disease leading to genome reduction. This comparison emphasizes the genetic divergence between the strains, particularly in the agr operon, which plays a crucial role in quorum sensing and virulence regulation.

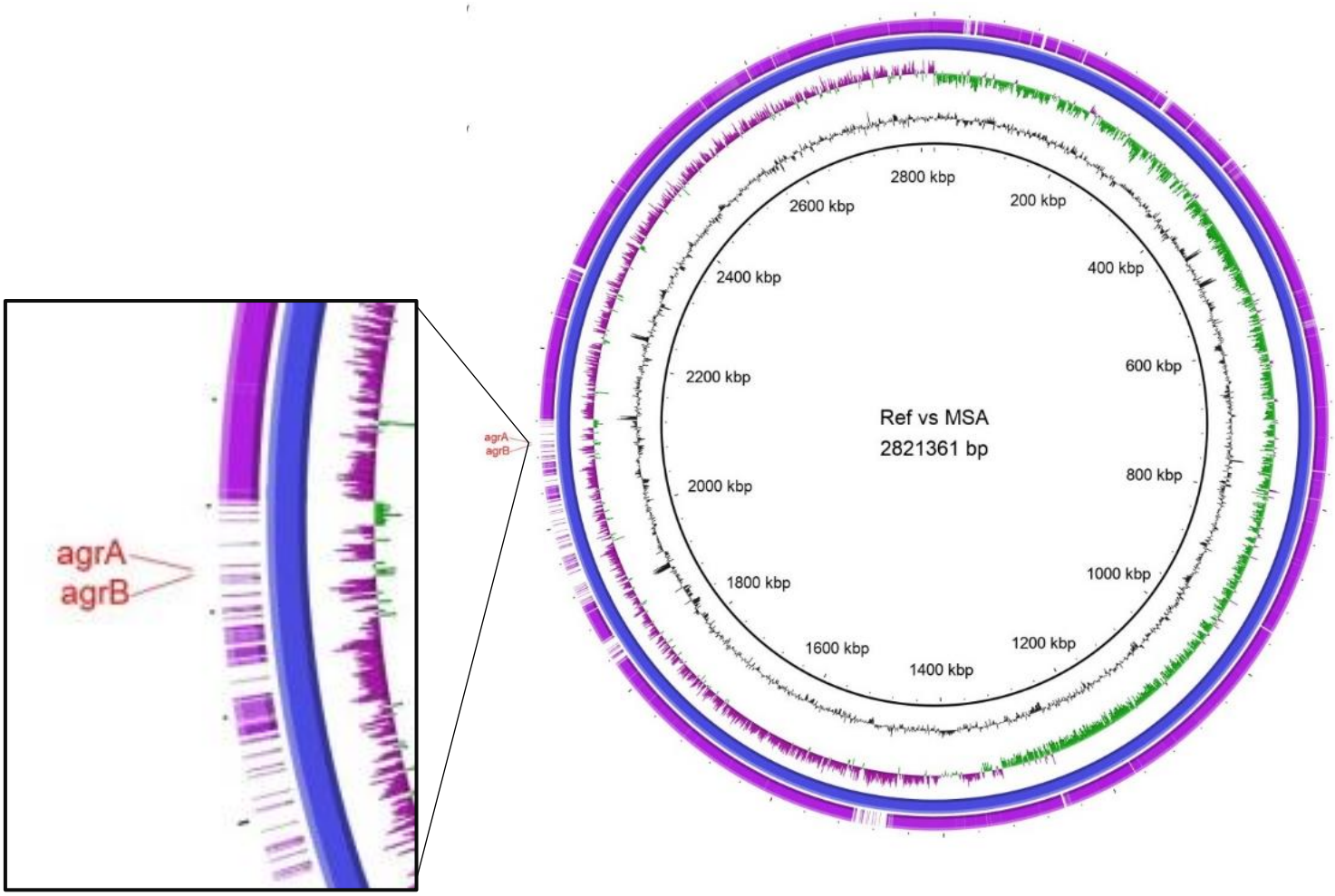

Fig. 8: Ring diagram showed the comparison of reference genome and MS4 strain which revealed the variation in agr operon and other genes. The circular genome comparison highlights the genetic differences between two Staphylococcus aureus strains.

Discussion

Current study has revealed essential findings relating to both *S. aureus* evolutionary adaptations and genetic variation. Pangenome analysis determined the estimation of an open pangenome. Previous studies examining *S. aureus* pangenomes observed these genomic collections to be open pangenomes (John *et al*, 2019). The ability of *S. aureus* to demonstrate genetic flexibility and adaptability becomes evident through examinations of core genes and shell or cloud genes analysis (Park *et al*, 2022). The pathogen requires this genetic variability to effectively adjust itself both to numerous hosts and ecological spaces especially during prolonged infections (Park *et al*, 2022; Sivakumar *et al*, 2023). The extensive accessory genome contains genetic sequences which belong specifically to particular growth environments. Accessory genes confer beneficial traits to pathogen survival while providing advantageous traits (Giulieri *et al*, 2022). The pathogens genetic components specifically enable persistent infections by affecting their ability to infect elderly patients with ongoing medical conditions.

The present study demonstrated significant differences regarding both clonal complex group distributions and methicillin resistance detection patterns. The major role of MRSA strains indicates that clinical isolates resist methicillin and its related antibiotics very commonly. The data from this study supports previous studies showing CC8 MRSA represents the primary lineage linked to hospital-acquired infections (Smith *et al*, 2021; Khatoon *et al*, 2024). The research uncovered untypeable microorganisms

that existed inside both MSSA and MRSA classifications while showing genetic variability indicating possible generation of novel ancestral groups. Monitoring resistant strain evolution requires ongoing surveillance with molecular testing because these findings demonstrate resistance evaluation needs must continue (Syed *et al*, 2021).

Recently developed technique known as Agr-typing allows scientists to track different evolutionary *S. aureus* strains. The agrI group dominated the studied DNA sequences followed by agrII and agrIII types whereas agrIV group represented the rarest genomes based on Agr typing results. The observed distribution indicates the relevance of agr types to virulence control (Tan *et al*, 2018). Studied evidence showed that a single bacterium did not contain an agr operon which suggests both possible niche-developing traits and genetic simplification (Giulieri *et al*, 2022). A strain lacking the agr operon might reveal significant consequences for both quorum sensing and virulence. Research needs to explore how S. aureus strains perform clinically without the agr operon during disease maintenance.

The core genome phylogenetic analysis together with BAPS produced understanding of genetic diversity and population division between *S. aureus* strains. The analysis revealed clusters with distinct patterns of single nucleotide polymorphisms that demonstrated hidden genetic variations. Previous studies have utilized distinctive patterns to conduct epidemiological investigations (Piper *et al*, 2024). The pathogenic characteristics along with antibiotic response of multiple strains depend on genetic variations within their lineages (Gostev *et al*, 2021). Population structures of *S. aureus* need to be comprehensively understood to create effective targeted treatments in healthcare environments.

Numerous questions emerge from the identification of this agr operon-deficient strain discovered in a long-term illness elderly patient. The organisms' genome has probably become smaller because they evolved to succeed in permanent infectious conditions. When virulence decreases, the bacteria typically succeed in surviving inside the body to create persistent infections which become chronic. Previous study by (Giulieri *et al*, 2022) demonstrated how S. aureus utilizes evolutionary adaptations as strategies for remaining in host environments. Further research is needed to evaluate how lack of agr operon functions affects infection outcomes during clinical conditions.

In conclusion, the results demonstrate why pangenome analyses combined with agr typing play essential roles in estimating the evolutionary patterns of *S. aureus.* The emphasis in future research should be directed towards understanding how genetic variations function biologically. The obtained results may help understand infection control complications and their relevant therapeutic approaches.

Declaration of Competing Interest

The authors declare that they have no competing or conflict of interests.

Author Contributions

SABS: Conceptualization, Methodology, formal analysis, Writing—original draft preparation. **MMA:** Formal analysis, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

References

1. Awan, F., M. M. Ali, Y. Dong, Y. Yu, Z. Zeng, and Y. Liu. (2021). *In Silico Analysis of Potential Outer Membrane Beta-Barrel Proteins in Aeromonas hydrophila Pangenome*. International journal of peptide research and therapeutics 27(4):2381-2389. doi: 10.1007/s10989-021-10259-z
2. Cavalcante, F. S., C. Alvarenga, S. Saintive, E. Dios Abad, D. Carvalho Ferreira, and K. R. Netto Dos Santos. (2020). *Staphylococcus aureus nasal isolates may have the same genetic profile in atopic dermatitis paediatric patients and their close contacts*. J Med Microbiol 69(6):850-853. doi: 10.1099/jmm.0.001197
3. Edgar, R. C. (2010). *Search and clustering orders of magnitude faster than BLAST*. Bioinformatics 26(19):2460-2461. doi: 10.1093/bioinformatics/btq461 %J Bioinformatics
4. Gillespie, J. J., A. R. Wattam, S. A. Cammer, J. L. Gabbard, M. P. Shukla, O. Dalay, T. Driscoll, D. Hix, S. P. Mane, C. Mao, E. K. Nordberg, M. Scott, J. R. Schulman, E. E. Snyder, D. E. Sullivan, C. Wang, A. Warren, K. P. Williams, T. Xue, H. S. Yoo, C. Zhang, Y. Zhang, R. Will, R. W. Kenyon, and B. W. Sobral. (2011). *PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species*. Infect Immun 79(11):4286-4298. doi: 10.1128/iai.00207-11
5. Giulieri, S. G., R. Guérillot, S. Duchene, A. Hachani, D. Daniel, T. Seemann, J. S. Davis, S. Y. C. Tong, B. C. Young, D. J. Wilson, T. P. Stinear, and B. P. Howden. (2022). *Niche-specific genome degradation and convergent evolution shaping Staphylococcus aureus adaptation during severe infections*. Elife 11:e77195. doi: 10.7554/eLife.77195
6. Gostev, V., S. Leyn, A. Kruglov, D. Likholetova, O. Kalinogorskaya, M. Baykina, N. Dmitrieva, Z. Grigorievskaya, T. Priputnevich, L. Lyubasovskaya, A. Gordeev, and S. Sidorenko. (2021). *Global Expansion of Linezolid-Resistant Coagulase-Negative Staphylococci*. Frontiers in Microbiology 12doi: 10.3389/fmicb.2021.661798
7. Javdan, S., T. Narimani, M. Shahini Shams Abadi, and A. Gholipour. (2019). *Agr typing of Staphylococcus aureus species isolated from clinical samples in training hospitals of Isfahan and Shahrekord*. BMC Res Notes 12(1):363. doi: 10.1186/s13104-019-4396-8
8. John, J., S. George, S. R. C. Nori, and S. Nelson-Sathi. (2019). *Phylogenomic Analysis Reveals the Evolutionary Route of Resistant Genes in Staphylococcus aureus*. Genome Biology and Evolution 11(10):2917-2926. doi: 10.1093/gbe/evz213 %J Genome Biology and Evolution
9. Khatoon, A., S. F. Hussain, S. M. Shahid, S. K. Sidhwani, S. A. Khan, O. A. Shaikh, and A. J. Nashwan. (2024). *Emerging novel sequence types of Staphylococcus aureus in Pakistan*. Journal of Infection and Public Health 17(1):51-59. doi: https://doi.org/10.1016/j.jiph.2023.10.036
10. Lai, C.-H., M. Y. Wong, T.-Y. Huang, C.-C. Kao, Y.-H. Lin, C.-H. Lu, and Y.-K. Huang. (2024). *Exploration of Agr Types, Virulence-associated Genes, and Biofilm Formation Ability in Staphylococcus Aureus Isolates from*

Hemodialysis Patients with Vascular Access Infections. FRONTIERS IN CELLULAR AND INFECTION MICROBIOLOGY
11. Letunic, I., and P. Bork. (2021). *Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation*. Nucleic acids research 49(W1):W293-W296. doi: 10.1093/nar/gkab301 %J Nucleic Acids Research
12. Lewin, A., E. Kamal, T. Semmler, K. Winter, S. Kaiser, H. Schäfer, L. Mao, P. Eschenhagen, C. Grehn, J. Bender, and C. Schwarz. (2021). *Genetic diversification of persistent Mycobacterium abscessus within cystic fibrosis patients*. Virulence 12(1):2415-2429. doi: 10.1080/21505594.2021.1959808
13. Löytynoja, A. (2014). *Phylogeny-aware alignment with PRANK*. In: Russell, D. J., Multiple Sequence Alignment Methods. Humana Press, Totowa, NJ. 155-170 p.
14. Meier-Kolthoff, J. P., A. F. Auch, H. P. Klenk, and M. Göker. (2013). *Genome sequence-based species delimitation with confidence intervals and improved distance functions*. BMC Bioinformatics 14:60. doi: 10.1186/1471-2105-14-60
15. Meier-Kolthoff, J. P., J. S. Carbasse, R. L. Peinado-Olarte, and M. Göker. (2022). *TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes*. Nucleic acids research 50(D1):D801-d807. doi: 10.1093/nar/gkab902
16. Mossop, M., L. Robinson, J. H. Jiang, A. Y. Peleg, L. V. Blakeway, N. Macesic, A. Perry, S. Bourke, F. R. Ulhuq, and T. Palmer. (2023). *Characterisation of key genotypic and phenotypic traits of clinical cystic fibrosis Staphylococcus aureus isolates*. J Med Microbiol 72(6)doi: 10.1099/jmm.0.001703
17. Motallebi, M., Z. Alibolandi, Z. F. Aghmiyuni, W. B. van Leeuwen, M. R. Sharif, and R. Moniri. (2021). *Molecular analysis and the toxin, MSCRAMM, and biofilm genes of methicillin-resistant Staphylococcus aureus strains isolated from pemphigus wounds: A study based on SCCmec and dru typing*. Infection, Genetics and Evolution 87:104644. doi: https://doi.org/10.1016/j.meegid.2020.104644
18. Ogura, K., H. Furuya, N. Takahashi, K. Shibata, M. Endo, S. Watanabe, L. Cui, T. Miyoshi-Akiyama, S. Okamoto, K. Ogai, and J. Sugama. (2022). *Interspecies Regulation Between Staphylococcus caprae and Staphylococcus aureus Colonized on Healed Skin After Injury*. 13doi: 10.3389/fmicb.2022.818398
19. Page, A. J., C. A. Cummins, M. Hunt, V. K. Wong, S. Reuter, M. T. G. Holden, M. Fookes, D. Falush, J. A. Keane, and J. Parkhill. (2015). *Roary: rapid large-scale prokaryote pan genome analysis*. Bioinformatics 31(22):3691-3693. doi: 10.1093/bioinformatics/btv421 %J Bioinformatics
20. Park, S., D. Jung, B. O'Brien, J. Ruffini, F. Dussault, A. Dube-Duquette, É. Demontier, J. F. Lucier, F. Malouin, S. Dufour, and J. Ronholm. (2022). *Comparative genomic analysis of Staphylococcus aureus isolates associated with either bovine intramammary infections or human infections demonstrates the importance of restriction-modification systems in host adaptation*. Microb Genom 8(2)doi: 10.1099/mgen.0.000779
21. Piper, K. R., O. O. Ikhimiukor, S. S. R. Souza, T. Garcia-Aroca, and C. P. Andam. (2024). *Evolutionary dynamics of the accessory genomes of Staphylococcus aureus*. mSphere 9(4):e00751-00723. doi: doi:10.1128/msphere.00751-23
22. Price, M. N., P. S. Dehal, and A. P. Arkin. (2010). *FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments*. PLOS ONE 5(3):e9490. doi: 10.1371/journal.pone.0009490
23. Raghuram, V., A. M. Alexander, H. Q. Loo, R. A. Petit, J. B. Goldberg, and T. D. Read. (2022). *Species-Wide Phylogenomics of the Staphylococcus aureus Agr Operon Revealed Convergent Evolution of Frameshift Mutations*. 10(1):e01334-01321. doi: doi:10.1128/spectrum.01334-21
24. Rimi, S. S., M. N. Ashraf, S. H. Sigma, M. T. Ahammed, M. P. Siddique, M. A. Zinnah, M. T. Rahman, and M. S. Islam. (2024). *Biofilm formation, agr typing and antibiotic resistance pattern in methicillin-resistant Staphylococcus aureus isolated from hospital environments*. PLOS ONE 19(8):e0308282. doi: 10.1371/journal.pone.0308282
25. Sano, M., Y. Shindo, K. Takahashi, J. Okumura, T. Sakakibara, Y. Murakami, M. Iguchi, T. Yagi, S. Matsui, and Y. Hasegawa. (2022). *Risk factors for antibiotic resistance in hospital-acquired and ventilator-associated pneumonia*. J Infect Chemother 28(6):745-752. doi: 10.1016/j.jiac.2022.02.012
26. Seemann, T. (2014). *Prokka: rapid prokaryotic genome annotation*. Bioinformatics 30(14):2068-2069. doi: 10.1093/bioinformatics/btu153 %J Bioinformatics
27. Sivakumar, R., P. S. Pranav, M. Annamanedi, S. Chandrapriya, S. Isloor, J. Rajendhran, and N. R. Hegde. (2023). *Genome sequencing and comparative genomic analysis of bovine mastitis-associated Staphylococcus aureus strains from India*. BMC Genomics 24(1):44. doi: 10.1186/s12864-022-09090-7
28. Smith, J. T., E. M. Eckhardt, N. B. Hansel, T. R. Eliato, I. W. Martin, and C. P. Andam. (2021). *Genomic epidemiology of methicillin-resistant and -susceptible Staphylococcus aureus from bloodstream infections*. BMC Infectious Diseases 21(1):589. doi: 10.1186/s12879-021-06293-3
29. Syed, M. A., B. Jamil, H. Ramadan, M. Rukan, S. Ali, S. A. Abbasi, T. A. Woodley, and C. R. Jackson. (2021). *Genetic Diversity of Staphylococcus aureus Strains from a Tertiary Care Hospital in Rawalpindi, Pakistan*. Microorganisms 9(11):2301.
30. Tan, L., S. R. Li, B. Jiang, X. M. Hu, and S. Li. (2018). *Therapeutic Targeting of the Staphylococcus aureus Accessory Gene Regulator (agr) System*. Front Microbiol 9:55. doi: 10.3389/fmicb.2018.00055
31. Tonkin-Hill, G., J. A. Lees, S. D. Bentley, S. D. W. Frost, and J. Corander. (2019). *Fast hierarchical Bayesian analysis of population structure*. Nucleic acids research 47(11):5539-5549. doi: 10.1093/nar/gkz361 %J Nucleic Acids Research
32. Valcek, A., C. Philippe, C. Whiteway, E. Robino, K. Nesporova, M. Bové, T. Coenye, T. De Pooter, W. De Coster, M. Strazisar, and C. Van der Henst. (2023). *Phenotypic Characterization and Heterogeneity among Modern Clinical Isolates of Acinetobacter baumannii*. Microbiol Spectr 11(1):e0306122. doi: 10.1128/spectrum.03061-22
33. Zhou, S., Y. Rao, J. Li, Q. Huang, and X. Rao. (2022). *Staphylococcus aureus small-colony variants: Formation, infection, and treatment*. Microbiol Res 260:127040. doi: 10.1016/j.micres.2022.127040