La Granja: Revista de Ciencias de la Vida
pISSN:1390-3799; eISSN:1390-8596
https://doi.org/10.17163/lgr.n44.2026.03

Special Issue / Edición Especial
Chirality across scales

CC BY NC SA

Tendency to Functional Asymmetry and Microsatellite-Induced Helical Chirality in *Staphylococcus aureus*

Tendencia a la asimetría funcional y quiralidad helicoidal inducida por microsatélites en *Staphylococcus aureus*

Ernesto Ríos-Willars[1], María Magdalena Delabra-Salinas*[2] and Mónica Itzel García Martínez[1]

[1]*Universidad Autónoma de Coahuila. Facultad de Sistemas Unidad Saltillo. Coahuila-México. [https://ror.org/00dpnh189]*
[2]*Universidad Autónoma de Coahuila. Facultad de Enfermería Unidad Saltillo. Coahuila-México. [https://ror.org/00dpnh189]*

*Corresponding author: magdalena-delabra-salinas@uadec.edu.mx

Article received on March 2nd, 2026. Accepted, after review, on July 30th, 2026. Published on September 1st, 2026.

Abstract

Staphylococcus aureus exhibits a microsatellite-rich genomic architecture that can induce functional asymmetries and helical chirality biases. We analyzed four genomes reported in the NCBI database and quantified ATAT, ACAC, and AGAG repetitive motifs in all genes using the MFind tool and surface plots. Reproducible differences in peak and valley topography between genomes were revealed, with a consistent pattern of similarity in Gn2–Gn4 and a marked deviation in Gn1. While ATAT predominated in Gn2–Gn4, Gn1 showed an atypical enrichment of ACAC. This change in the dominant motif, together with the displacement of maxima between genomes, suggests variation in the local arrangement of repeats compatible with chiral effects (asymmetries between strands, inverted repeats, and possible transitions between mirror patterns). Such configurations are plausible modulators of regulatory directionality, secondary structure formation, and recombination propensity. Methodologically, we highlight that the length of the regions and genes analyzed can influence the metric used. In the analysis of three motifs in four genomes, the observed patterns offer an initial comparative map of chiral dimension in the repetitive architecture of *S. aureus* and motivate mechanistic analyses that integrate structural and functional evidence to quantify the contribution of chirality to genomic variation between lineages. Genomes, microsatellite quantification tables, and calculations of the proposed chirality change trend index are available at https://www.usegalaxy.org/u/rioswillars/h/s-aureus-microsatellite.

Keywords: *Staphylococcus aureus*, chirality, microsatellite, genome.

©2026, Universidad Politécnica Salesiana, Ecuador.

Resumen

Staphylococcus aureus presenta una distribución genómica rica en microsatélites, lo que puede inducir asimetrías funcionales y sesgos de quiralidad helicoidal. Se analizaron cuatro genomas reportados en la base de datos NCBI y se cuantificaron los motivos repetitivos ATAT, ACAC y AGAG en todos los genes mediante la herramienta MFind y los gráficos de superficie. Se observaron diferencias reproducibles en la topografía de picos y valles entre genomas, con un patrón consistente de similitud en Gn2–Gn4 y un marcado desvío en Gn1. Mientras que ATAT predominó en Gn2–Gn4, Gn1 mostró un enriquecimiento atípico de ACAC. Este cambio en el motivo dominante, junto con el desplazamiento de máximos entre genomas, sugiere una variación en la disposición local de las repeticiones compatible con efectos quirales (asimetrías entre hebras, repeticiones invertidas y posibles transiciones entre patrones en espejo). Tales configuraciones son plausibles moduladoras de la direccionalidad regulatoria, de la formación de estructuras secundarias y de la propensión a la recombinación. Metodológicamente, se destaca que la métrica empleada puede verse influida por la longitud de las regiones y de los genes analizados. En el análisis de tres motivos en cuatro genomas, los patrones observados ofrecen un mapa comparativo inicial de la dimensión quiral en la arquitectura repetitiva de *S. aureus* y motivan análisis mecanísticos que integren evidencia estructural y funcional para cuantificar la contribución de la quiralidad a la variación genómica entre linajes. Se ponen a disposición los genomas, las tablas de cuantificación de microsatélites y los cálculos del índice propuesto de tendencia al cambio de quiralidad: https://www.usegalaxy.org/u/rioswillars/h/s-aureus-microsatellite.

Palabras clave: *Staphylococcus aureus*, quiralidad, microsatélite, genoma.

Suggested citation: Ríos-Willars, E., Delabra-Salinas, M. and García-Martínez, M. (2026). Tendency to Functional Asymmetry and Microsatellite-Induced Helical Chirality in *Staphylococcus aureus*. La Granja: Revista de Ciencias de la Vida. Vol. 44(2):42-58. https://doi.org/10.17163/lgr.n44.2026.03.

Orcid IDs:

Ernesto Ríos-Willars: https://orcid.org/0000-0002-7157-8884
María Magdalena Delabra-Salinas: https://orcid.org/0000-0002-4467-191X
Mónica Itzel García Martínez: https://orcid.org/0009-0000-2869-9747

©2026, Universidad Politécnica Salesiana, Ecuador.

1 Introduction

Staphylococcus aureus is a spherical Gram-positive bacterium that forms grape-like clusters. It is well known for its ability to cause a wide range of infections in humans, ranging from minor skin infections to severe diseases such as sepsis and endocarditis (Larkin et al., 2009; Gherardi, 2023; Tong et al., 2015). It has been isolated from wastewater, water, and contaminated soils (Heaton et al., 2020; Ramos et al., 2022). The genomic study of this bacterium is important because it supports bioinformatic approaches aimed at preventing infections in both animals and humans. In this context, microsatellites are repeated nucleotide sequences present in DNA; in bacteria, they occur in lower abundance than in other organisms because they are generally subject to mutational pressure that favors their elimination. Nevertheless, they are commonly found in regions of the bacterial genome where they contribute adaptive capabilities through evolution and natural selection. Among the main characteristics of *S. aureus* are its pathogenicity and virulence, as it employs a variety of factors, including toxin production, to cause infections. These toxins can trigger severe immune responses, such as fever and shock (Oliveira et al., 2018; Ahmad-Mansour et al., 2021). Additionally, *S. aureus* exhibits antibiotic resistance, which constitutes one of its most concerning characteristics. Examples include methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *S. aureus* (VRSA). This resistance complicates the treatment of infections and represents a growing problem in hospitals (Archer, 1998; Idrees et al., 2021). Associated infections: *S. aureus* is responsible for infections affecting multiple body systems, including skin infections, osteomyelitis, endocarditis, pneumonia, and others (Kwiecinski and Horswill, 2020). Furthermore, it is widely present in dust, feces, surfaces, and bedding on pig and poultry farms; the levels of staphylococci in air and dust can be very high near livestock buildings (Pedersen et al., 2024; Gladding et al., 2020; George et al., 2020) or animal housing facilities. In humans, it is a pathogen commonly associated with diabetic foot infections, contributing significantly to patient morbidity and mortality. This microorganism is the pathogen most frequently isolated from diabetic foot infections (Shettigar and Murali, 2020). Among its various toxins and virulence factors are TSST-1, leukocidins, enterotoxins, and exfoliatins, which play a crucial role in the colonization and infection of diabetic foot wounds (Dunyach-Remy et al., 2016). *S. aureus* strains isolated from diabetic foot ulcers can induce a significant inflammatory response, which is influenced by glucose levels and their metabolites (Wu et al., 2023). Regarding the relationship between *S. aureus* infection and wounds, it is known that this bacterium can induce an altered inflammatory response in the host, thereby inhibiting wound healing. For example, the induction of *miR-15b-5p* by *S. aureus* can suppress the inflammatory response and negatively affect healing (Ramirez et al., 2018). Farm dust may contain MRSA (methicillin-resistant *S. aureus*) and disperse over hundreds of meters, increasing the exposure of farmers, nearby residents, and their homes to resistant bacteria (Crespo-Piazuelo and Lawlor, 2021; Wang et al., 2023). Open, moist wounds containing necrotic tissue, created or enlarged by the screwworm, favor colonization by *S. aureus* and other pathogenic bacteria (Sayhood et al., 2022).

On the other hand, microsatellites in bacteria are short tandemly repeated DNA sequences, also known as simple sequence repeats (SSRs), consisting of nucleotide motifs ranging from one to six base pairs. These sequences are highly mutable and are distributed throughout different locations in the bacterial genome. They are important genetic markers because their high mutation rate contributes to genetic variability and the evolution of the microbial genome (Saeed et al., 2016). In pathogenic bacteria, microsatellite instability can regulate gene expression through phase variation, a mechanism that enables rapid and reversible changes in phenotype expression, as observed in other microorganisms (Martin et al., 2005). However, bacteria exhibit a mutational bias toward microsatellite repeat contraction, which may lead to their extinction in the absence of natural selection, unlike eukaryotes, in which microsatellite repeats tend to expand (Metzgar et al., 2002). In the research conducted in this study, no studies related to microsatellites in *S. aureus* were found; however, in *Escherichia coli*, between 1.75 % and 2.85 % of all microsatellites have been identified as compound microsatellites, consisting of two or more adjacent repeat motifs. These compound microsatellites are significantly overrepresented in bacterial genomes. Although less common, bacterial microsatellites may play an important role in gene regulation and contribute adaptive

©2026, Universidad Politécnica Salesiana, Ecuador.

functional variability. This suggests that, although mutational pressure tends to eliminate them, those that persist may be maintained by selection because of their adaptive functions (Srivastava et al., 2019).

The bacterial genome consists of double-stranded right-handed B-DNA; this is the fundamental chemical chirality shared with other living organisms (Winogradoff et al., 2021; Yu et al., 2025). The right-handed double helix determines the formation of DNA crossovers between DNA segments and the stabilization of positive and negative supercoils (Timsit and Várnai, 2010). The chirality of right-handed versus left-handed crossovers modulates the physical stability of DNA and may have influenced the evolutionary selection that maintains predominantly negatively supercoiled genomes in mesophilic bacteria. At the enzymatic level, topoisomerases, such as DNA gyrase, recognize and capture DNA crossovers of a specific chirality in order to introduce or relax supercoils (Vayssières et al., 2024).

Finally, microsatellites can alter the three-dimensional architecture of DNA because, although they do not change its chemical chirality, they induce torsional stress, which may favor structural features such as A-DNA (compact, more tightly wound), Z-DNA (left-handed helix, opposite chirality), hairpin loops, cruciform structures, and strand slippage (Gadgil et al., 2020; Bagshaw, 2017). This does not alter the chemical chirality, but it does modify the helical conformation, causing certain DNA regions to behave differently because of their geometry, flexibility, or torsional properties (Bacolla et al., 2008). For example, the ATAT microsatellite is extremely flexible and therefore adopts slight curvatures, whereas ACAC and CGCG are more rigid, favoring an alternating geometry (Song et al., 2021).

In pathogens such as *S. aureus*, the spatial and structural instability of microsatellites is not merely a genomic artifact but rather acts as contingency loci that drive phase variation (Brooks and Jefferson, 2014). Conformational changes derived from chirality and torsional stress, such as the formation of loop structures or minidumbbells in tetranucleotide repeats such as ACAC, increase the rate of strand slippage during DNA replication (Harhay et al., 2019). This architectural reconfiguration facilitates reversible mutations (insertions/deletions) that rapidly switch gene expression on or off. Biologically, these structural differences enable human strains such as MRSA to adapt their virulence by controlling key processes such as biofilm formation (through the *ica* operon) to evade antibiotics or by modulating the regulatory system to survive as a "Trojan horse" within host macrophages (Gor et al., 2021). Therefore, the term "atypical case" in the context of microsatellite analysis refers to genomic arrangements that, by deviating from their standard chirality and geometry, confer greater phenotypic plasticity in terms of resistance and infectivity to certain lineages.

In this context, this research investigates the microsatellites identified in *Staphylococcus aureus* genomes reported in the NCBI database (Sayers et al., 2025) over different years. Subsequently, this information is used to estimate changes in genome chirality.

2 Materials and Methods

Four *S. aureus* genomes were downloaded, herein referred to as Gn1, Gn2, Gn3, and Gn4. General information on these genomes is provided below.

The Gn1 genome (strain MRSA_TCH60) is registered in the NCBI database by Baylor College of Medicine from a human skin sample (*Homo sapiens*). Clinically, TCH60 is a reference genome representative of the CC30-ST30-IV lineage. This clonal complex (CC30/ST30) has epidemiological importance because it descends from the pandemic phage 80/81 strain and has evolved into the "Southwest Pacific clone," a highly virulent and globally distributed community-associated MRSA variant responsible for severe invasive infections (Essien et al., 2022; Guillén et al., 2024). Bioinformatically, the assembly and validation of this strain were performed using Clusterflock (Narechania et al., 2016). A size of 2.8 megabases and two chromosomes are reported. It has a GC content of 33 %, as well as 2790 genes, of which 2700 are identified as protein-coding. It was sequenced using Illumina 454 GS FLX technology. The annotation details were registered on August 28, 2013. This record is accessible through the following link: https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000159535.2/

©2026, Universidad Politécnica Salesiana, Ecuador.

The genome designated Gn2 (strain MRSA252) is described below. It is a 2.9-megabase genome with a GC content of 33 %. Clinically, this strain is representative of the hospital epidemic clone EMRSA-16 (sequence type ST36) and was originally isolated from a case of fatal septicemia. Whole-genome analysis demonstrated that it is a highly divergent lineage characterized by rapid evolution toward virulence and multidrug resistance driven by the acquisition of mobile genetic elements, such as the type II SCC*mec* chromosomal cassette and multiple prophages (Holden et al., 2004). In the NCBI database, based on metadata from an animal-associated habitat, it is reported as corresponding to an original collection from 1998. In this case, 2746 genes were identified, of which 2656 were considered protein-coding. This record is publicly available at: https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000011505.1/

The genome designated Gn3 (strain CC239-MRSA-III var) corresponds to an isolate from a human host (*Homo sapiens*) in Trinidad and Tobago. Clinically, this lineage belongs to Clonal Complex 239 (CC239), one of the most widely distributed hospital epidemic clones worldwide and highly prevalent in clinical samples from the Caribbean islands. Its genome was sequenced and registered by the Leibniz Institute of Photonic Technology (Monecke et al., 2023). Oxford Nanopore MinION sequencing technology was used on June 5, 2023. A total of 3152 genes are reported, of which 2940 were identified as protein-coding. The GC content is 32.5 %, and the genome size is 3.1 megabases. This record is accessible through the following link: https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_030252755.1/

Finally, the genome designated Gn4 (strain HP20814-043) was analyzed. It has a size of 2.9 megabases and a GC content of 33 %. Clinically, it represents a linezolid-resistant isolate obtained from a patient with cystic fibrosis (*Homo sapiens*) at the University of Iowa Hospitals and Clinics, United States. Evolutionary analysis of this strain reveals profound adaptation under antibiotic stress, and the deletion of the DNA repair genes (*mutS* and *mutL*), which confers a hypermutable phenotype suitable for studying tandem-repeat variation. Its bioinformatic validation was processed according to RefSeq database representation strategies (Tatusova et al., 2015). A total of 2860 genes were reported in this genome, of which 2631 were identified as protein-coding. This genome is available at the following link: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_030409015.1

The exclusive selection of the tetranucleotide motifs ATAT, AGAG, and ACAC is based on thermodynamic considerations associated with the phase-variation mechanism. Microsatellites involved in strand-slippage events have been documented to be rich in adenine and thymine (A/T), which facilitates the local "melting" of the DNA double helix and produces hypermutability (Zhou et al., 2014). Unlike sequences composed exclusively of guanine and cytosine, such as CGCG, which exhibit high thermal stability that severely restricts their local opening, the ATAT, AGAG, and ACAC motifs combine the requirements necessary to permit this melting with a diversity of topological behaviors. Repetitions of these microsatellites act as a key evolutionary driver of the rapid adaptation and survival of pathogenic bacteria (Saeed et al., 2016); therefore, this specific selection of A/T-rich motifs is ecologically representative and consistent with the low GC content, approximately 33 %, characteristic of the *S. aureus* genome.

The processing method used for microsatellite analysis consisted of applying the MFind tool (Rios-Willars and Chirinos-Arias, 2024). Unlike generic methods based on multiple sequence alignment (MSA), MFind employs a parallel-processing "sliding window" algorithm that performs an exact count of motifs independently of their position. Additionally, its selection was based on the fact that it is a freely available tool with a simple graphical interface, whose computational efficiency allows complete genomes to be explored on conventional computer equipment without requiring supercomputing infrastructure. The MFind software is available in the author's GitHub repository: https://github.com/riosew/Mfind

Given the problem, a sequence PS in length ρ containing the subsequences (SP) to be searched is provided. $SP = \{SP_1 \dots SP_n\}$ and the set of sequences is $S = \{1 \dots N\}$ where N is the number of sequences in the set. Given an SP_x of size k, the problem size TP is defined as:

©2026, Universidad Politécnica Salesiana, Ecuador.

$$TP = (\rho - k)N \quad (1)$$

The number of operations N_e indicates the number of windows in each sequence of S with size k and is defined as:

$$N_e = \sum_{1}^{TP} \sum_{1}^{OC} (Ct + 1) \quad (2)$$

Where (Ct) is the hit counter and (OC) is the number of subsequences matching within the sliding window. The search process is executed through parallel iterations programmed in Java to maximize the use of hardware resources and reduce execution time. Six simultaneous evaluations are performed per iteration. This process produces a database consisting of the microsatellite count for each gene, which is subsequently used to generate graphs and conduct statistical analyses with Minitab 16 software. Subsequently, to quantify the tendency of a DNA region to adopt alternative helical conformations, the Chirality Change Propensity Index (IQZ) was defined. This index integrates three independent sequence-derived components: GC content, the presence of alternating purine–pyrimidine patterns, and the density of repetitive microsatellites. Each of these factors is associated with physicochemical and structural properties that favor the local transition from the B-DNA conformation to alternative states, such as Z-DNA.

The index is expressed as:

$$IQZ = w_1 \cdot GC_score + w_2 \cdot Alt_score + w_3 \cdot MS_score \quad (3)$$

Where:

$$GC_score = \frac{\%GC}{100} \quad (4)$$

This represents the proportion of G and C bases in the analyzed region. Regions with higher GC content exhibit greater torsional rigidity and a greater propensity to stabilize the Z-DNA conformation.

Alt_score is the number of bases forming alternating purine–pyrimidine patterns divided by the total sequence length, in this case, the gene being analyzed. This quantifies the proportion of bases organized into repeated alternating patterns, for example, ATAT and ACAC, which reduces the energy required for helicity inversion.

Finally:

$$MS_score = \frac{\text{Total length of microsatellites}}{Lmax} \quad (5)$$

This estimates the torsional effect resulting from the presence of microsatellites. The variable $Lmax$ corresponds to the maximum expected size of these repetitions for normalization purposes; the maximum number of *hits* counted by the MFind tool among the analyzed genes was used.

The coefficients w_1, w_2 and w_3 allow the relative contribution of each parameter to be adjusted. In this study, values of $w_1 = 0.4$, $w_2 = 0.4$ and $w_3 = 0.2$, were used, assigning greater weight to GC content and purine–pyrimidine alternation, which are the most relevant predictors of the transition to Z-DNA. This index provides a normalized measurement between 0 and 1, thereby facilitating comparisons among different genomic regions and allowing the identification of segments with a high propensity to modify their helical architecture.

3 Results and Discussion

The location of microsatellites in the bacterial genome is generally specific to functional regions, suggesting that they have evolved under selection to perform specific functions. In *S. aureus*, these functions may be related to its ability to cause skin infections and, consequently, diabetic foot wounds. The results obtained are described below. Figure 1 illustrates the surface generated from the analysis of the ATAT, AGAG, and ACAC microsatellites in genome 1 (Gn1). A peak is observed in the upper-right corner, corresponding to the highest microsatellite count values (*hits*) per gene in the genome.

Likewise, a decline is observed in the lower-left corner, suggesting that the microsatellites tend to disappear jointly in this region. Finally, there is a prominent flat region in the center of the surface, where it can be observed that, while the AGAG and ACAC microsatellites increase, the ATAT microsatellite remains absent. Regarding the distribution, Figure 2 shows the box plots of the three variables, in which outliers are observed for ATAT and ACAC. These correspond to genes that exhibited an extraordinary number of microsatellites and, at the same time, constitute the origin of the peak in the surface plot.

©2026, Universidad Politécnica Salesiana, Ecuador.

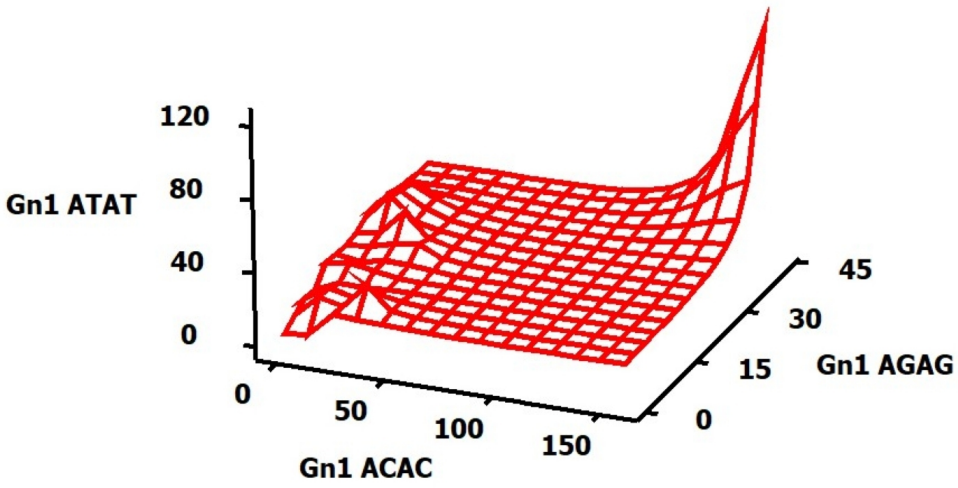


Figure 1. Surface plot of microsatellites in the Gn1 genome of *S. aureus*. The vertical axis represents the frequency of the ATAT motif as a function of the co-occurrence of the ACAC and AGAG motifs. The peaks indicate genes with an anomalous joint accumulation of these tetranucleotides.

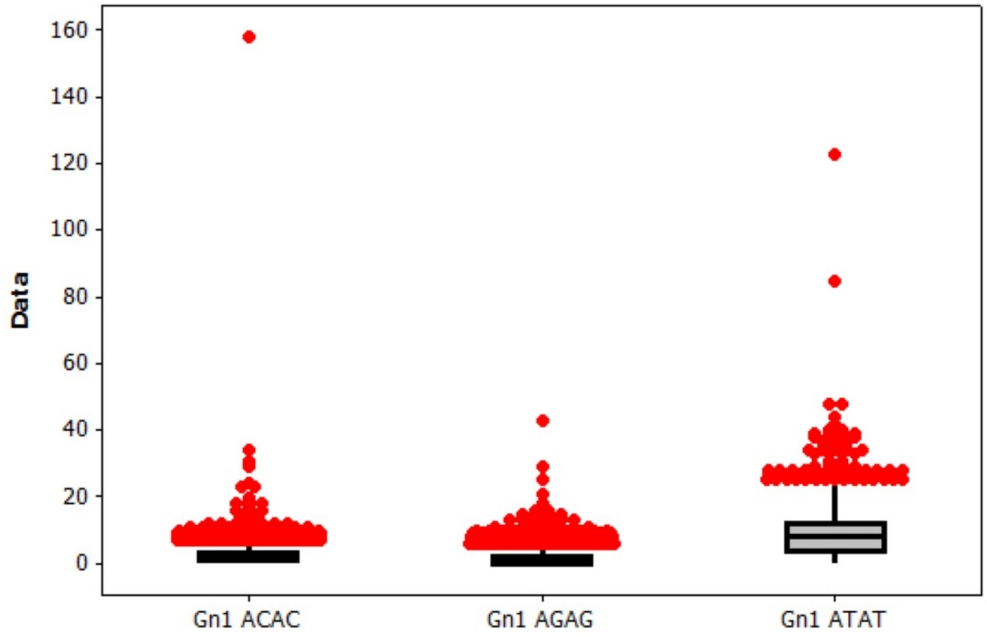


Figure 2. Box plot of the frequencies of the ACAC, AGAG, and ATAT motifs in the genes of the Gn1 genome. Individual points outside the whiskers represent genes that possess an extraordinary number of microsatellites compared with the central tendency of the genome.

©2026, Universidad Politécnica Salesiana, Ecuador.

In the analysis of Gn2, a surface plot of the ATAT, ACAC, and AGAG variables was obtained. Figure 3 shows the surface, in which a peak can be observed near the right edge. Likewise, another peak is observed near the central region. This suggests a redistribution of the genome and the corresponding relocation of microsatellites. A valley or flat region of smaller extent than that observed in the previous result can also be seen. This also indicates microsatellite dispersion.

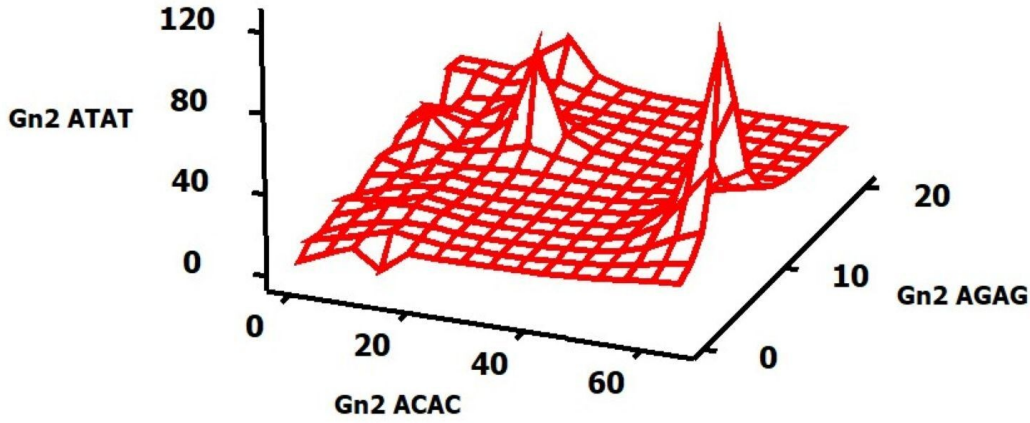


Figure 3. Surface plot of microsatellites in the Gn2 genome of *S. aureus*. The vertical axis represents the frequency of the ATAT motif as a function of the co-occurrence of the ACAC and AGAG motifs. The peaks indicate genes with an anomalous joint accumulation of these tetranucleotides.

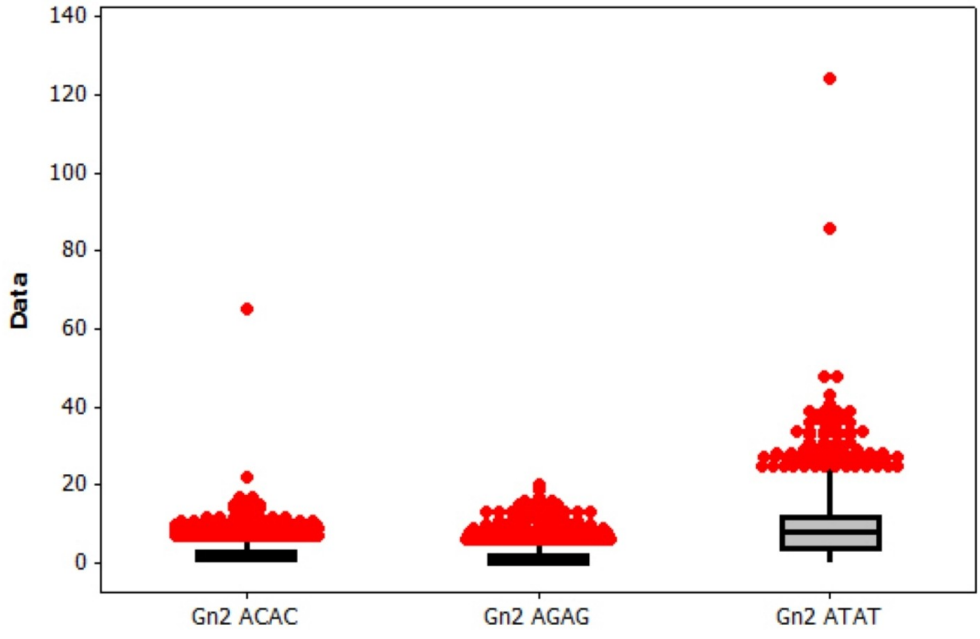


Figure 4. Box plot of the frequencies of the ACAC, AGAG, and ATAT motifs in the genes of the Gn2 genome. Individual points outside the whiskers represent genes that possess an extraordinary number of microsatellites compared with the central tendency of the genome.

©2026, Universidad Politécnica Salesiana, Ecuador.

On the other hand, the box plot in Figure 4, corresponding to the microsatellites in Gn2, illustrates the findings. Outliers can be observed for ATAT, which in turn generate the peaks observed in the surface plot. It should be noted that, for AGAG and ACAC, most points are located near zero, indicating that most Gn2 genes contain few or no microsatellites of this type.

Regarding Gn3, a surface plot of the ATAT, AGAG, and ACAC variables was also obtained. Figure 5 illustrates this representation.

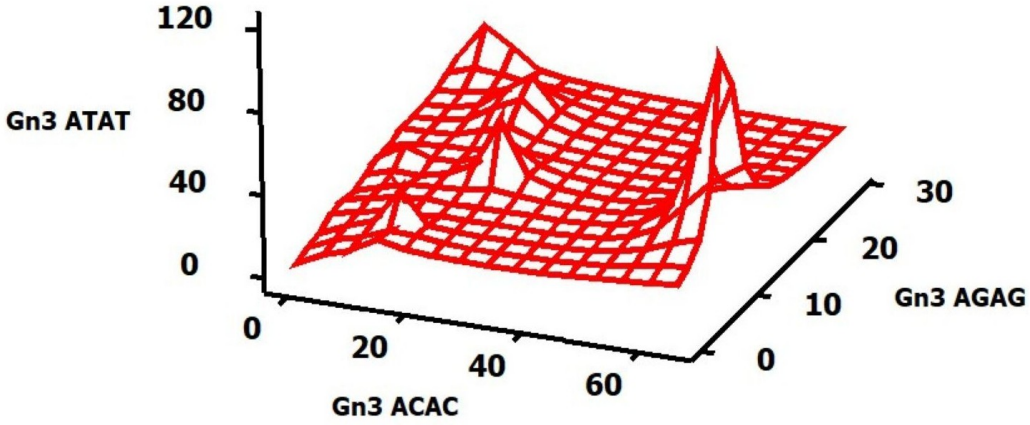


Figure 5. Surface plot of microsatellites in the Gn3 genome of *S. aureus*. The vertical axis represents the frequency of the ATAT motif as a function of the co-occurrence of the ACAC and AGAG motifs. The peaks indicate genes with an anomalous joint accumulation of these tetranucleotides.

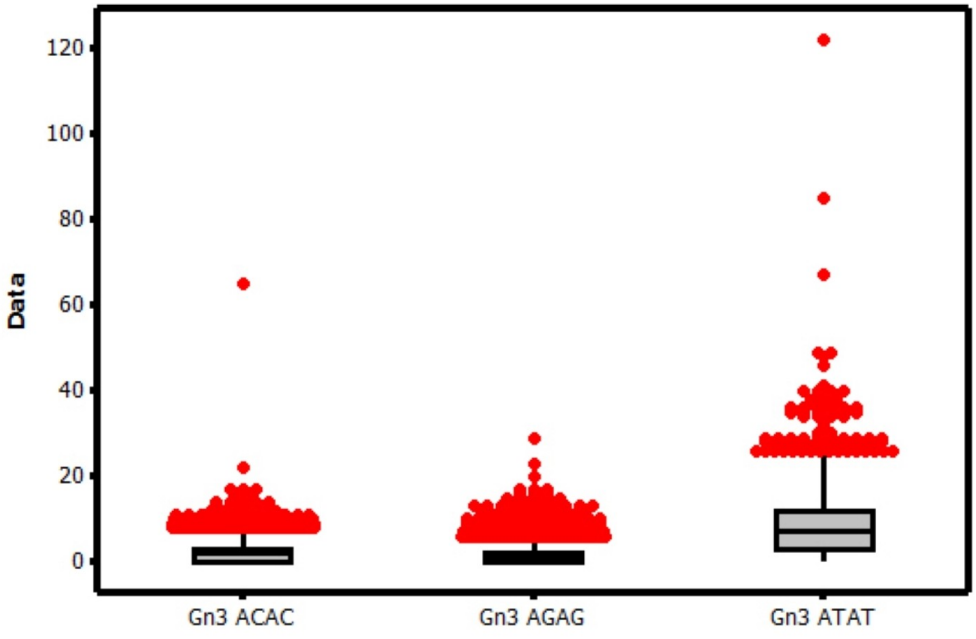


Figure 6. Box plot of the frequencies of the ACAC, AGAG, and ATAT motifs in the genes of the Gn3 genome. Individual points outside the whiskers represent genes that possess an extraordinary number of microsatellites compared with the central tendency of the genome.

La Granja: *Revista de Ciencias de la Vida* 44(2) 2026:42-58.
©2026, Universidad Politécnica Salesiana, Ecuador.

Similar features are observed compared with the previous figure; however, differences indicating changes in the microsatellites are also present. For example, the location of the central peaks in the image. Figure 6 below illustrates this distribution using a box plot.

Scattered points corresponding to ATAT and AGAG are observed in the figure. This indicates that other genes, compared with the previous genome, contain a greater number of microsatellites. Finally, for Gn4, Figure 7 shows the surface plot corresponding to the three microsatellites: ATAT, AGAG, and ACAC.

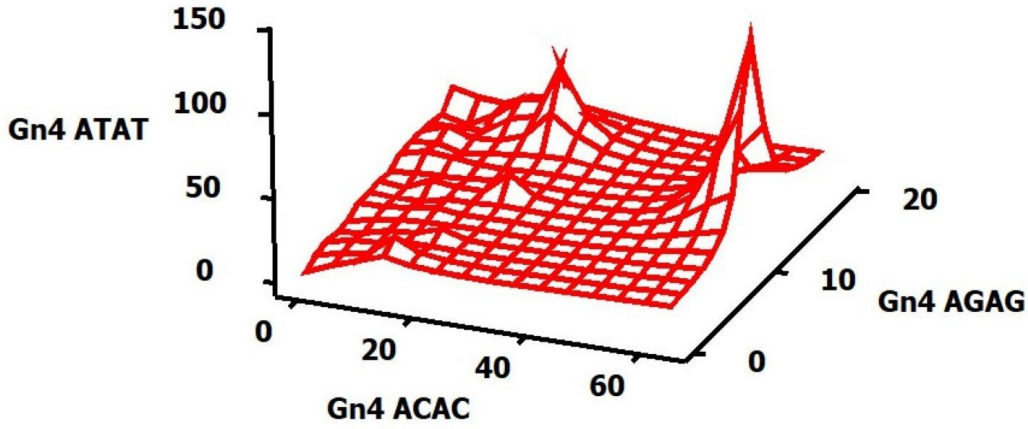


Figure 7. Surface plot of microsatellites in the Gn4 genome of *S. aureus*. The vertical axis represents the frequency of the ATAT motif as a function of the co-occurrence of the ACAC and AGAG motifs. The peaks indicate genes with an anomalous joint accumulation of these tetranucleotides.

It is noteworthy that a peak is present on the right side, although at a location closer to the upper corner. This suggests new microsatellite arrangements. Likewise, the peak near the upper-left corner exhibits a greater elevation than that of the previous genome. On the other hand, Figure 8 shows the box plot of this genome for microsatellite counts.

The microsatellite count for AGAG remained practically below those of the other motifs, with no outliers. The distribution also changed for ACAC. In addition, ATAT appears to be increasing in abundance compared with the other genomes. Table 1 below presents the descriptive statistics for the mean, minimum, and maximum values of the three variables in the Gn1, Gn2, Gn3, and Gn4 genomes. It is noteworthy that, in all cases, the mean values of the ATAT variable are higher. This indicates a greater presence of ATAT microsatellites in all cases. However, the highest value was observed in the Gn1 genome for the ACAC variable, with a value of 158.

Regarding the results obtained, a folder containing the Gn1, Gn2, Gn3, and Gn4 genomes in .fna format is available for download on the Galaxy platform. Likewise, the .csv files generated from the processing performed with the MFind tool for each microsatellite in the genomes are made available. Thus, for example, the files Gn1 ACAC.csv, Gn1 ATAT.csv, and Gn1 AGAG.csv are available for download for genome 1. Three .csv files corresponding to each of the remaining genomes are presented below. The tables corresponding to the calculation of the IQZ for each genome are also included. Available at: https://www.usegalaxy.org/u/rioswillars/h/s-aureus-microsatellite.

©2026, Universidad Politécnica Salesiana, Ecuador.

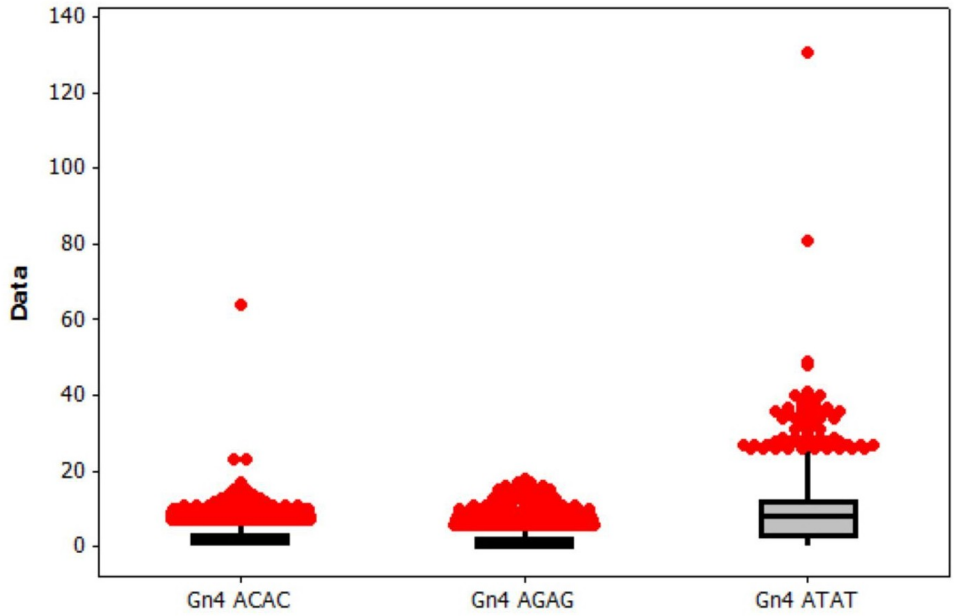


Figure 8. Box plot of the frequencies of the ACAC, AGAG, and ATAT motifs in the genes of the Gn4 genome. Individual points outside the whiskers represent genes that possess an extraordinary number of microsatellites compared with the central tendency of the genome.

Table 1. Descriptive statistics of the ATAT, AGAG, and ACAC variables in the Gn1, Gn2, Gn3, and Gn4 genomes.

Genome	Variable	Mean	SD	Min	Max
Gn1	ACAC	2.47	4.03	0	158
Gn1	AGAG	1.5688	2.4365	0	43
Gn1	ATAT	8.722	7.286	0	123
Gn2	ACAC	2.3752	2.7245	0	65
Gn2	AGAG	1.5338	2.2454	0	20
Gn2	ATAT	8.705	7.26	0	124
Gn3	ACAC	2.3016	2.6768	0	65
Gn3	AGAG	1.5826	2.348	0	29
Gn3	ATAT	8.618	7.36	0	122
Gn4	ACAC	2.3844	2.7234	0	64
Gn4	AGAG	1.5654	2.2614	0	18
Gn4	ATAT	8.484	7.162	0	131

In this study, we calculated the IQZ composite index for four *Staphylococcus aureus* genomes (Gn1–Gn4) using three normalized components: GC_norm, Alt_Score, and MS_score, weighted at 0.4, 0.4, and 0.2, respectively. This combination integrates nucleotide composition (GC_norm) with quantitative metrics derived from hit patterns (Alt_Score and MS_score), thereby providing a unified metric for sequence prioritization. The IQZ calculation was applied to all genes in each genome, and quantification tables were generated. In all cases, we observed that the highest IQZ values were

©2026, Universidad Politécnica Salesiana, Ecuador.

associated with segments that simultaneously exhibited moderate to high GC profiles, together with a high microsatellite count in terms of both density and relative prominence. This indicates that IQZ prioritizes regions in which pattern-based metrics and GC composition converge.

In Gn1, the sequences with the highest IQZ correspond to extended regions with high Alt_Score (a metric of pattern density per unit length) and high MS_score (a metric of normalized intragenomic prominence). In Gn2 and Gn3, the maximum IQZ values exhibit similar behavior: the top candidates combine GC_norm values above the average, a favorable hits-to-length ratio, and remarkable relative prominence within the genome. In Gn4, the profile of the maximum IQZ values is consistent with that of Gn2 and Gn3, reinforcing the reproducibility of the composite criterion across different genomic contexts of *S. aureus*. Comparatively among genomes, although MS_score is intrinsically intragenomic (normalized to the maximum value within each file), the structure of IQZ allows the qualitative assessment of the robustness of priority regions in each genome. Furthermore, the stability of the IQZ ranking under moderate variations in GC_norm suggests that the pattern-based metrics (Alt_Score and MS_score) adequately capture both the density and the prominence of the motifs, thereby avoiding excessive bias due to nucleotide composition.

For traceability and practical use, Table 2 presents the lists of the five sequences (genes) with the highest IQZ values for each genome (Gn1–Gn4). These lists support the conclusions presented above and provide a basis for subsequent decisions, such as experimental validation, probe/primer design, or annotation prioritization. Additionally, Figure 9 presents a visual representation of the IQZ index for the sequences of each genome. Overall, IQZ behaves as a synthetic and stable metric for highlighting candidate genomic segments by integrating GC composition and the quantitative metrics of the pattern in a balanced manner.

Table 2. List of the five genes with the highest IQZ values in each genome.

Dataset	seq num	Gene length	Total hits	ACAC hits	AGAG hits	ATAT hits	IQZ
Gn1	1861	31303	324	158	43	123	0.35334
Gn1	1337	74	1	0	1	0	0.270888
Gn1	1325	76	1	0	1	0	0.269038
Gn1	1336	76	1	0	1	0	0.269038
Gn1	1391	72	2	2	0	0	0.262346
Gn2	1447	32241	195	65	6	124	0.34976
Gn2	172	7176	121	22	13	86	0.277447
Gn2	1967	84	2	2	0	0	0.27348
Gn2	1929	89	3	2	1	0	0.27274
Gn2	1983	74	1	1	0	0	0.271296
Gn3	1519	31266	196	65	9	122	0.349261
Gn3	194	7176	120	22	13	85	0.275793
Gn3	2043	84	2	2	0	0	0.273469
Gn3	2060	74	1	1	0	0	0.271291
Gn3	624	76	1	0	1	0	0.269441
Gn4	1491	31875	205	64	10	131	0.348869
Gn4	1962	84	2	2	0	0	0.27338
Gn4	1914	89	3	2	1	0	0.27259
Gn4	1978	74	1	1	0	0	0.271246
Gn4	174	7176	119	23	15	81	0.270445

©2026, Universidad Politécnica Salesiana, Ecuador.

It is noteworthy that the length of the region is strongly associated with the IQZ (correlation of 0.98, $p = 0.00$). Longer regions tend to exhibit higher IQZ values. This is consistent with the fact that IQZ includes components that depend on both the number of hits and the sequence length, as well as the associated normalization procedures.

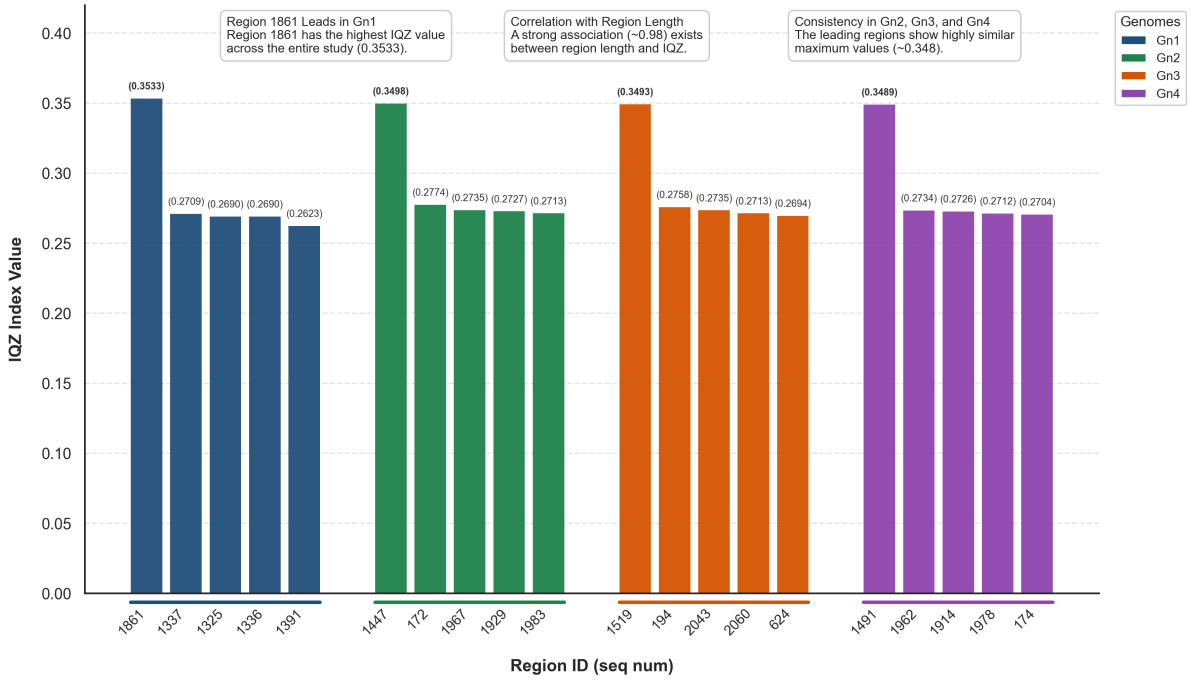


Figure 9. Comparison of the Chirality Change Propensity Index (IQZ) for the five genes with the highest values in each genome. The graph illustrates the magnitude of the index grouped by genome, allowing the differences in the distribution of the maximum values of the Gn1 genome to be observed relative to the remaining lineages.

Finally, the electronic folder on the Galaxy platform mentioned above contains tables showing the calculations for the *MS_score*, *Alt_score*, *GC_norm* and *IQZ* metrics for each of the genes in the Gn1, Gn2, Gn3 and Gn4 genomes.

4 Conclusions

Microsatellite analysis is a strategy for understanding the evolution and adaptation of organisms of interest. In this study, four *S. aureus* genomes were described in detail by analyzing each of their genes and identifying and counting the ATAT, AGAG, and ACAC microsatellites using the MFind tool, a freely available software package that can be downloaded and executed in Java and is available on GitHub. Among the findings, Gn1 exhibited clear differences compared with the other three genomes. Whereas Gn2, Gn3, and Gn4 displayed similarities that allowed comparable regions to be identified in the surface plots, the peaks and valleys of the plots were located at different positions, suggesting a reorganization of the genetic information. On the other hand, microsatellite counts were higher for the ATAT motif in the Gn2, Gn3, and Gn4 genomes, whereas Gn1 represented an exception, in which the maximum value was observed for the ACAC variable, with a value of 158. This indicates an atypical case due to the presence of this number of microsatellites within a gene of this genome.

Finally, this work provides an overview of microsatellites in four reported *S. aureus* genomes, this bacterium being the species most commonly identified in studies of wounds in humans and farm animals. In addition, this study characterizes the organization of repetitive motifs in the four *Staphy-*

©2026, Universidad Politécnica Salesiana, Ecuador.

lococcus aureus genomes analyzed, focusing on the chirality of the patterns and their variation among genomes. The comparative profiles show that Gn1 differs from the behavior observed in Gn2–Gn4, not only because of the predominant motif but also because of the displacement of peaks and valleys in the density surfaces, suggesting chirally relevant differences in the local arrangement of the repeats. In Gn2–Gn4, the ATAT motif predominates, whereas Gn1 exhibits an atypical enrichment of the ACAC motif. This asymmetry in the predominant motif, together with the displacement of maxima among genomes, is consistent with a reconfiguration of the repetitive architecture with a chiral component (differences in relative orientations and possible transitions between mirror motifs along the genome). Such patterns may affect genomic directionality and plasticity, insofar as inverted or strand-biased repeats modulate properties such as secondary structure formation or the propensity for recombination events.

The interpretation of these findings should consider certain limitations: only three motifs (ATAT, AGAG, and ACAC) were evaluated, the metrics depend on the size of the analyzed region, and the MFind detection tool was employed. Future studies should incorporate an index based on three-dimensional simulations because, unlike the current statistical predictive index (IQZ), which is based on a linear (2D) sequence, a 3D simulation would allow the torsional stress energy to be measured accurately and confirm whether alternating chirality induces the folding of secondary structures *in vivo*. Such biophysical validation is crucial because these three-dimensional conformational changes promote polymerase slippage and activate the phase-variation switch that regulates virulence and adaptation in *S. aureus*. Overall, this study provides a comparative map highlighting the chiral dimension of *S. aureus* microsatellites and motivates mechanistic analyses to quantify their contribution to genomic variation among lineages. In this regard, future studies should determine whether these regions occur in specific genomic areas, such as low-complexity regions, or whether they are distributed in a delimited manner within them.

Acknowledgments

The authors wish to thank the Faculty of Systems of the Autonomous University of Coahuila, Saltillo Campus, for the assistance provided during the development of this work.

Funding

The authors would like to thank the IPICyT National Supercomputing Center for supporting this research with computing time grant TKII-E-1024-I-011024-47.

Declaration on the use of artificial intelligence

The authors DECLARE that, during the preparation of the article entitled "Tendency to Functional Asymmetry and Microsatellite-Induced Helical Chirality in *Staphylococcus aureus*", no generative artificial intelligence tools or automated assistance systems were used for the writing, analysis, data interpretation, content generation, translation, or editing of the manuscript.

The authors assume full responsibility for the content, originality, integrity, and final version of the article.

Author's contribution

E.R.W.: Conceptualization, Investigation, Methodology, Software. **M.M.D.S.:** Writing – original draft, Investigation, Data curation. **M.I.G.M.:** Writing – review & editing.

References

Ahmad-Mansour, N., Loubet, P., Pouget, C., Dunyach-Remy, C., Sotto, A., Lavigne, J.-P., and Molle, V. (2021). Staphylococcus aureus toxins: An update on their pathogenic properties and potential treatments. *Toxins*, *13*(10), 677. Online: https://doi.org/10.3390/toxins13100677

Archer, G. L. (1998). Staphylococcus aureus: A well–armed pathogen. *Clinical Infectious Diseases*, *26*(5),

©2026, Universidad Politécnica Salesiana, Ecuador.

1179–1181. Online: https://doi.org/10.1086/520289

Bacolla, A., Larson, J. E., Collins, J. R., Li, J., Milosavljevic, A., Stenson, P. D., ... Wells, R. D. (2008). Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties. *Genome Research*, *18*(10), 1545–1553. Online: https://doi.org/10.1101/gr.078303.108

Bagshaw, A. T. M. (2017). Functional mechanisms of microsatellite adn in eukaryotic genomes. *Genome Biology and Evolution*, *9*(9), 2428–2443. Online: https://doi.org/10.1093/gbe/evx164

Brooks, J., and Jefferson, K. (2014). Phase variation of poly-n-acetylglucosamine expression in staphylococcus aureus. *PLoS Pathogens*, *10*(7), e1004292. Online: https://doi.org/10.1371/journal.ppat.1004292

Crespo-Piazuelo, D., and Lawlor, P. G. (2021). Livestock-associated methicillin-resistant staphylococcus aureus (la-mrsa) prevalence in humans in close contact with animals and measures to reduce on-farm colonisation. *Irish Veterinary Journal*, *74*(1), 21. Online: https://doi.org/10.1186/s13620-021-00200-7

Dunyach-Remy, C., Ngba Essebe, C., Sotto, A., and Lavigne, J.-P. (2016). Staphylococcus aureus toxins and diabetic foot ulcers: Role in pathogenesis and interest in diagnosis. *Toxins*, *8*(7), 209. Online: https://doi.org/10.3390/toxins8070209

Essien, U. C., Boswihi, S. S., Agbakoba, N. R., and Udo, E. E. (2022). Description of methicillin-susceptible staphylococcus aureus clonal complex 30 related to the pandemic phage type 80/81 isolated from patients in three tertiary hospitals in jos, north central nigeria. *Medical Principles and Practice*, *31*, 269–275. Online: https://doi.org/10.1159/000524755

Gadgil, R. Y., Romer, E. J., Goodman, C. C., Rider, S. D., Damewood, F. J., Barthelemy, J. R., ... Leffak, M. (2020). Replication stress at microsatellites causes dna double-strand breaks and break-induced replication. *Journal of Biological Chemistry*, *295*(45), 15378–15397. Online: https://doi.org/10.1074/jbc.RA120.013495

George, A. N., Stewart, J. R., Evans, J. C., and Gibson, J. M. (2020). Risk of antibiotic-resistant staphylococcus aureus dispersion from hog farms: A critical review. *Risk Analysis*, *40*(8), 1645–1665. Online: https://doi.org/10.1111/risa.13495

Gherardi, G. (2023). Staphylococcus aureus infection: Pathogenesis and antimicrobial resistance. *International Journal of Molecular Sciences*, *24*(9), 8182. Online: https://doi.org/10.3390/ijms24098182

Gladding, T. L., Rolph, C. A., Gwyther, C. L., Kinnersley, R., Walsh, K., and Tyrrel, S. (2020). Concentration and composition of bioaerosol emissions from intensive farms: Pig and poultry livestock. *Journal of Environmental Management*, *272*, 111052. Online: https://doi.org/10.1016/j.jenvman.2020.111052

Gor, V., Ohniwa, R., and Morikawa, K. (2021). No change, no life? what we know about phase variation in staphylococcus aureus. *Microorganisms*, *9*(2), 244. Online: https://doi.org/10.3390/microorganisms9020244

Guillén, R., Salinas, C., Mendoza-Álvarez, A., Rubio Rodríguez, L. A., Díaz-de Usera, A., Lorenzo-Salazar, J. M., ... Rodríguez, F. (2024). Genomic epidemiology of the primary methicillin-resistant staphylococcus aureus clones causing invasive infections in paraguayan children. *Microbiology Spectrum*, *12*(2), e03012-23. Online: https://doi.org/10.1128/spectrum.03012-23

Harhay, G., Harhay, D., Bono, J., Capik, S., Dedonder, K., Apley, M., ... Smith, T. (2019). A computational method to quantify the effects of slipped strand mispairing on bacterial tetranucleotide repeats. *Scientific Reports*, *9*, 17192. Online: https://doi.org/10.1038/s41598-019-53866-z

Heaton, C. J., Gerbig, G. R., Sensius, L. D., Patel, V., and Smith, T. C. (2020). Staphylococcus aureus epidemiology in wildlife: A systematic review. *Antibiotics*, *9*(2), 89. Online: https://doi.org/10.3390/antibiotics9020089

Holden, M. T. G., Feil, E. J., Lindsay, J. A., Peacock, S. J., Day, N. P. J., Enright, M. C., ... Parkhill, J. (2004). Complete genomes of two clinical staphylococcus aureus strains: Evidence for the rapid evolution of virulence and drug resistance. *Proceedings of the National Academy of Sciences*, *101*(26), 9786–9791. Online: https://doi.org/10.1073/pnas.0402521101

©2026, Universidad Politécnica Salesiana, Ecuador.

Idrees, M., Sawant, S., Karodia, N., and Rahman, A. (2021). Staphylococcus aureus biofilm: Morphology, genetics, pathogenesis and treatment strategies. *International Journal of Environmental Research and Public Health*, *18*(14), 7602. Online: https://doi.org/10.3390/ijerph18147602

Kwiecinski, J. M., and Horswill, A. R. (2020). Staphylococcus aureus bloodstream infections: pathogenesis and regulatory mechanisms. *Current Opinion in Microbiology*, *53*, 51–60. Online: https://doi.org/10.1016/j.mib.2020.02.005

Larkin, E., Carman, R., Krakauer, T., and Stiles, B. (2009). Staphylococcus aureus: The toxic presence of a pathogen extraordinaire. *Current Medicinal Chemistry*, *16*(30), 4003–4019. Online: https://doi.org/10.2174/092986709789352321

Martin, P., Makepeace, K., Hill, S. A., Hood, D. W., and Moxon, E. R. (2005). Microsatellite instability regulates transcription factor binding and gene expression. *Proceedings of the National Academy of Sciences*, *102*(10), 3800–3804. Online: https://doi.org/10.1073/pnas.0406805102

Metzgar, D., Liu, L., Hansen, C., Dybvig, K., and Wills, C. (2002). Domain-level differences in microsatellite distribution and content result from different relative rates of insertion and deletion mutations. *Genome Research*, *12*(3), 408–413. Online: https://doi.org/10.1101/gr.198602

Monecke, S., Akpaka, P. E., Smith, M. R., Unakal, C. G., Thoms Rodriguez, C.-A., Ashraph, K., … Ehricht, R. (2023). Clonal complexes distribution of staphylococcus aureus isolates from clinical samples from the caribbean islands. *Antibiotics*, *12*(6), 1050. Online: https://doi.org/10.3390/antibiotics12061050

Narechania, A., Baker, R., DeSalle, R., Mathema, B., Kolokotronis, S.-O., Kreiswirth, B., and Planet, P. J. (2016). Clusterflock: a flocking algorithm for isolating congruent phylogenomic datasets. *GigaScience*, *5*(1), 43. Online: https://doi.org/10.1186/s13742-016-0152-3

Oliveira, D., Borges, A., and Simões, M. (2018). Staphylococcus aureus toxins and their molecular activity in infectious diseases. *Toxins*, *10*(6), 252. Online: https://doi.org/10.3390/toxins10060252

Pedersen, K., Nielsen, M. W., Fertner, M. E., Espinosa-Gongora, C., and Bækbo, P. (2024). Development of livestock-associated methicillin-resistant staphylococcus aureus (la-mrsa) loads in pigs and pig stables during the fattening period. *Veterinary Sciences*, *11*(11), 558. Online: https://doi.org/10.3390/vetsci11110558

Ramirez, H. A., Pastar, I., Jozic, I., Stojadinovic, O., Stone, R. C., Ojeh, N., … Tomic-Canic, M. (2018). Staphylococcus aureus triggers induction of mir-15b-5p to diminish dna repair and deregulate inflammatory response in diabetic foot ulcers. *Journal of Investigative Dermatology*, *138*(5), 1187–1196. Online: https://doi.org/10.1016/j.jid.2017.11.038

Ramos, B., Rosalino, L. M., Palmeira, J. D., Torres, R. T., and Cunha, M. V. (2022). Antimicrobial resistance in commensal staphylococcus aureus from wild ungulates is driven by agricultural land cover and livestock farming. *Environmental Pollution*, *303*, 119116. Online: https://doi.org/10.1016/j.envpol.2022.119116

Rios-Willars, E., and Chirinos-Arias, M. C. (2024). Mfind: a tool for dna barcode analysis in angiosperms and its relationship with microsatellites using a sliding window algorithm. *Planta*, *259*(6), 134. Online: https://doi.org/10.1007/s00425-024-04420-3

Saeed, A. F., Wang, R., and Wang, S. (2016). Microsatellites in pursuit of microbial genome evolution. *Frontiers in Microbiology*, *6*, 1462. Online: https://doi.org/10.3389/fmicb.2015.01462

Sayers, E. W., Beck, J., Bolton, E. E., Brister, J. R., Chan, J., Connor, R., … Pruitt, K. D. (2025). Database resources of the national center for biotechnology information in 2025. *Nucleic Acids Research*, *53*(D1), D20–D29. Online: https://doi.org/10.1093/nar/gkae979

Sayhood, M. H., Mohammed, A. L., Abdulhameed, M. F., and Jori, M. M. (2022). Classical and molecular identification of staphylococcus aureus isolated from infestation cattle wounds with myiasis in basrah governorate, iraq. *Iraqi Journal of Veterinary Sciences*, *36*(3), 641–646. Online: https://doi.org/10.33899/ijvs.2021.131196.1930

Shettigar, K., and Murali, T. S. (2020). Virulence factors and clonal diversity of staphylococcus

©2026, Universidad Politécnica Salesiana, Ecuador.

aureus in colonization and wound infection with emphasis on diabetic foot infection. *European Journal of Clinical Microbiology & Infectious Diseases*, *39*(12), 2235–2246. Online: https://doi.org/10.1007/s10096-020-03984-8

Song, X., Yang, T., Zhang, X., Yuan, Y., Yan, X., Wei, Y., … Zhou, C. (2021). Comparison of the microsatellite distribution patterns in the genomes of euarchontoglires at the taxonomic level. *Frontiers in Genetics*, *12*, 622724. Online: https://doi.org/10.3389/fgene.2021.622724

Srivastava, D., Ahmad, M. M., Shamim, M., Maurya, R., Srivastava, N., Pandey, P., … Siddiqui, M. H. (2019). Modulation of gene expression by microsatellites in microbes. En *New and future developments in microbial biotechnology and bioengineering* (pp. 209–218). Elsevier. Online: https://doi.org/10.1016/B978-0-444-63503-7.00012-7

Tatusova, T., Ciufo, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2015). Refseq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Research*, *43*(7), 3872–3872. Online: https://doi.org/10.1093/nar/gkv278

Timsit, Y., and Várnai, P. (2010). Helical chirality: a link between local interactions and global topology in dna. *PLoS ONE*, *5*(2), e9326. Online: https://doi.org/10.1371/journal.pone.0009326

Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L., and Fowler, V. G. (2015). Staphylococcus aureus infections: Epidemiology, pathophysiology, clinical manifestations, and management. *Clinical Microbiology Reviews*, *28*(3), 603–661. Online: https://doi.org/10.1128/CMR.00134-14

Vayssières, M., Marechal, N., Yun, L., Lopez Duran, B., Murugasamy, N. K., Fogg, J. M., … Lamour, V. (2024). Structural basis of dna crossover capture by escherichia coli dna gyrase. *Science*, *384*(6692), 227–232. Online: https://doi.org/10.1126/science.adl5899

Wang, Y., Zhang, P., Wu, J., Chen, S., Jin, Y., Long, J., … Yang, H. (2023). Transmission of livestock-associated methicillin-resistant staphylococcus aureus between animals, environment, and humans in the farm. *Environmental Science and Pollution Research*, *30*(37), 86521–86539. Online: https://doi.org/10.1007/s11356-023-28532-7

Winogradoff, D., Li, P., Joshi, H., Quednau, L., Maffeo, C., and Aksimentiev, A. (2021). Chiral systems made from dna. *Advanced Science*, *8*(5), 2003113. Online: https://doi.org/10.1002/advs.202003113

Wu, Y., Chen, T., Wang, Y., Huang, M., Wang, Y., and Luo, Z. (2023). New insight into the virulence and inflammatory response of staphylococcus aureus strains isolated from diabetic foot ulcers. *Frontiers in Cellular and Infection Microbiology*, *13*, 1234994. Online: https://doi.org/10.3389/fcimb.2023.1234994

Yu, S., Xie, Y., Jiao, Y., Li, N., and Ding, B. (2025). Nucleic acid-based chiral nanostructures and their biomedical applications. *Nanoscale Horizons*, *10*(9), 1802–1814. Online: https://doi.org/10.1039/D5NH00140D

Zhou, K., Aertsen, A., and Michiels, C. (2014). The role of variable dna tandem repeats in bacterial adaptation. *FEMS Microbiology Reviews*, *38*(1), 119–141. Online: https://doi.org/10.1111/1574-6976.12036

©2026, Universidad Politécnica Salesiana, Ecuador.