



OPEN Multiomics analysis of *Staphylococcus aureus* ST239 strains resistant to virulent *Herelleviridae* phages

M. Kornienko^{1✉}, D. Bespiatykh¹, N. Abdaimova¹, R. Gorodnichev¹, V. Gostev^{2,3}, D. Boldyreva¹, O. Selezneva¹, V. Veselovsky¹, O. Pobeguts⁴, I. Smirnov⁴, G. Arapidi^{4,5}, K. Klimina¹ & E. Shitikov¹

In the context of the antimicrobial therapy crisis, the significance of studying and implementing alternative treatment methods, particularly phage therapy, is increasingly evident. This study aimed to investigate the resistance of clinical *Staphylococcus aureus* ST239 strains to *Herelleviridae* phages through comparative genomics, transcriptomics, and proteomics. Analysis of resistant and sensitive *S. aureus* strains showed that resistant strains form a separate cluster on the phylogenetic tree, suggesting unique genetic traits underlying their phage resistance. Further in-depth analysis of the resistant SA191 strain infected with *Herelleviridae* phage, compared to an uninfected control, unveiled significant changes in the transcription of 462 genes (271↑ 191↓) at 5 min and 504 genes (276↑ 228↓) at 30 min post-infection. Proteomic analysis identified 184 differentially abundant proteins (41↑ 143↓) at 30 min. Functional analysis highlighted changes in the glycolysis, the tricarboxylic acid cycle, and transport systems; notable, changes were also observed in the transcription of prophage genes. Despite the observed metabolic shifts, classical resistance mechanisms related to teichoic acid synthesis, restriction-modification, and toxin-antitoxin systems were not identified, suggesting the existence of other mechanism. Our study contributes to the elucidation of *S. aureus* resistance mechanisms against *Herelleviridae* phages, highlighting the intricate nature of bacterial defense mechanisms.

In the realm of modern medical practice, there is a burgeoning interest in phage therapy—an approach for treating infectious diseases using virulent bacteriophages. This method gains particular relevance in light of the growing problem of multiple drug resistance (MDR) in microorganisms to conventional antibiotics¹. By reproducing within bacterial cells, virulent phages are capable of causing the death of bacterial cultures regardless of their antibiotic sensitivity, thereby facilitating the effective eradication of the pathogen.

The most promising results of phage therapy have been demonstrated in the treatment of infections caused by *Staphylococcus aureus*^{2–7}.

The currently described virulent phages targeting *S. aureus* belong to the *Herelleviridae* and *Rountreeviridae* families. Phages from the *Herelleviridae* family are most commonly used for therapeutic purposes due to their broader host range (up to 90% of strains, depending on the collection)^{8–10}.

Despite this, certain strains of *S. aureus*, particularly those belonging to sequence type (ST) 239, have demonstrated resistance to *Herelleviridae* phages, as previously shown by our research⁸, as well as by other studies^{9,11,12}. ST239 strains, associated with clonal complex (CC) 8, stand out as key epidemic strains of methicillin-resistant *S. aureus* (MRSA) identified worldwide^{13,14}. They reached peak prevalence in Asia during the 2000s, causing up to 90% of nosocomial MRSA infections in a region inhabited by over 60% of the world's population^{15–17}. Despite a decrease in their current frequency of identification¹⁷, ST239 strains continue to hold

¹Department of Biomedicine and Genomics, Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia. ²Pediatric Research and Clinical Center for Infectious Diseases, Saint Petersburg, Russia. ³North-Western State Medical University Named After I. I. Mechnikov, Saint Petersburg, Russia. ⁴Department of Post-Genomic Technologies, Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia. ⁵Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia. ✉email: kornienkomariya@gmail.com

clinical significance^{18–20}. They could produce a variety of toxins including enterotoxins, toxic shock syndrome toxin, exfoliative toxins A and B, and Panton-Valentine leukocidin (PVL)^{18,21,22}, and are capable of forming biofilms²¹, exacerbating infection progression and complicating treatment. Furthermore, ST239 strains exhibit high levels of resistance to multiple antimicrobial agents, often characterized by MDR²². The proportion of MRSA among ST239 strains can reach 100%²². Most commonly, MRSA ST239 strains have type-III staphylococcal cassette chromosome *mec*^{23,24}, although other cassette types have been reported^{22,25}. Additionally, ST239 strains with reduced sensitivity to vancomycin—the primary drug for treating MRSA infections^{26,27}—have been identified^{28,29}. Thus, ST239 strains, exhibiting multiple resistances to antibiotics and phages, could pose a significant challenge to modern medicine.

In the context of phage therapy against MDR *S. aureus* strains, comprehending the mechanisms behind their resistance to phages is pivotal for effective application. Some mechanisms of staphylococcal resistance to phages have been previously reported³⁰. Resistance typically arises at the adsorption level, primarily involving the interaction between the receptor-binding proteins of phages and specific cellular receptors. The primary receptors for staphylophages are teichoic acids, and for phages of the *Herelleviridae* family, it is usually the teichoic acid backbone³¹. Resistance can also emerge during intracellular stages of the phage life cycle. This can occur through various mechanisms, such as cellular restriction-modification systems or toxin-antitoxin systems. Additionally, resistance may manifest through the suppression of superinfection, a phenomenon facilitated by prophages³⁰.

In this study, we conducted a complex investigation of ST239 *S. aureus* strains resistant to the *Herelleviridae* family phages. Based on the data on phage adsorption to bacterial cells, it was determined that the bacteria implement an intracellular mechanism of resistance. However, genomic analysis of genes associated with phage resistance did not reveal significant differences between resistant and sensitive strains. Instead, a comprehensive analysis at the transcriptomic and proteomic levels showed significant differences in the response to phage infection between the resistant and sensitive strains, particularly impacting cellular energy acquisition pathways.

Results

Characterization of *S. aureus* ST239 strains and their interaction with phages of the *Herelleviridae* family

Five clinical ST239 strains included in the study were collected from three medical facilities and isolated from various loci (Table 1). All strains were MRSA and demonstrated MDR phenotype, however, remained susceptible to vancomycin and linezolid. Spa-typing revealed that four strains belonged to type t030, while one strain (SA0402r) was identified as type t233.

To assess the susceptibility of strains to *Herelleviridae* phages, four distinct phages were employed. The taxonomic classification of the phages was established based on whole-genome sequencing followed by subsequent phylogenetic analysis. (Figure S1). Phages caused nonspecific lysis of four ST239 strains, while SA0402r (t233 spa-type) exhibited sensitivity (Figure S2). The efficiency of plating (EOP) on SA0402r, relative to the host *S. aureus* strain SA515 (ST8, t008 spa-type), varied from 62.5 to 89.2% (Table 1).

The ability of *Herelleviridae* phages to adsorb to bacterial cells was performed for ST239 and the host strains. Ten minutes after infection, the percentage of phages that adsorbed onto the susceptible strains (SA515 and SA0402r) was 76.8–97.9%, while for the resistant strains, it varied from 82.3 to 99.6% (Fig. 1).

Comparative genomics of *Herelleviridae*-resistant strains of *S. aureus* ST239

To provide a detailed description of the studied strains and investigate potential reasons for resistance to phages, whole-genome sequencing (WGS) was conducted. The genome of SA191 was additionally sequenced using long-read technology. Also, the study included 32 fully assembled genomes sourced from NCBI, related to ST239 and isolated from patients across different continents.

To determine the relationships between the strains, a maximum likelihood phylogeny was constructed. The resulting phylogeny demonstrated a predominant clustering pattern according to different geographical regions, which aligns with previous publications^{14,32}. The resistant strains were grouped into one cluster, designated as cluster R in our analysis, along with samples from China. Meanwhile, the phage-sensitive strain showed closer proximity to samples from Brazil (Fig. 2).

For the SA191 strain, an analysis of key genes and systems potentially affecting resistance to phage infection was conducted (Fig. 3, Table S1)^{30,33}. Regarding teichoic acid synthesis genes, strain SA191, similar to other cluster R strains, harbored missense mutations in *tarK* (putative ribitol phosphate polymerase gene [RDJ18_RS01135]), *tagX* (putative glycosyltransferase gene [RDJ18_RS03615]), and *dltA* (D-alanine:D-alanyl carrier protein ligase gene [RDJ18_RS10555]) compared to the other investigated ST239 isolates.

The analysis of the restriction-modification (R-M) system genes of the SA191 strain revealed the presence of both type I and IV R-M systems. The type I R-M system had a standard profile, consisting of two *hsdMS* operons (RDJ18_RS02435–RDJ18_RS02440 and RDJ18_RS05955–RDJ18_RS05960), located on the genomic islands *vSaa* and *vSaβ*, respectively, and one *hsdR* gene (RDJ18_RS00865)³⁴. Additionally, the gene encoding the type I restriction endonuclease subunit R (RDJ18_RS14065) was identified on a *Tn6072* element. These genes and the corresponding mobile elements were identified in nearly all investigated ST239 genomes, including the sensitive strain SA0402r; however, strains of the cluster R harbored an additional missense mutation in RDJ18_RS14065 (Table S1, Figure S3). The type IV restriction-modification system was also found in almost all ST239 strains, including SA0402r, and was represented by the *sauUSI* (RDJ18_RS13240)^{35,36} and the *mcrBC* (RDJ18_RS00295–RDJ18_RS00300) system³⁷.

Based on the genome annotation of the SA191 strain, we identified genes corresponding to both type I (RDJ18_RS05985 on *vSaβ*; RDJ18_RS12960) and type II (RDJ18_RS02350 on *vSaa*; RDJ18_RS04725 on prophage region 2; RDJ18_RS11020; RDJ18_RS11025; RDJ18_RS12825; RDJ18_RS12830; RDJ18_RS13100; RDJ18_RS13105)

Strain	City	Origin	Oxacillin	Ciprofloxacin	Clindamycin	Erythromycin	Gentamicin	Linezolid	Tetracycline	Vancomycin	MLST	Spa type	Phage titer (PFU/ml)			
													vB_SauM-515A1	vB_SauH-Mpol	vB_SauH-GE	vB_SauM-fruSau02
SA191	Moscow	Spinal fluid	R	I	R	R	R	S	R	S	ST 239	t030	Lysis from without			
SA177	Moscow	Endotracheal tube	R	I	R	R	R	S	R	S	ST 239	t030	Lysis from without			
SA4	St Petersburg	Wound infection	R	R	S	R	R	S	R	S	ST 239	t030	Lysis from without			
SA5	St Petersburg	Wound infection	R	R	S	I	R	S	R	S	ST 239	t030	Lysis from without			
SA0402r	Lipetsk	Bones and joints infection	R	R	R	R	R	S	R	S	ST 239	t233	5*10^8 (89.2%)	4*10^8 (62.5%)	3.2*10^8 (80%)	4.6*10^7 (76.7%)
SA515	Novosibirsk	Wound infection	S	S	S	S	S	S	S	S	ST 8	t008	5.6*10^8 (100%)	6.4*10^8 (100%)	4*10^8 (100%)	6*10^7 (100%)

Table 1. Characteristics of *S. aureus* strains and their sensitivity to bacteriophages.

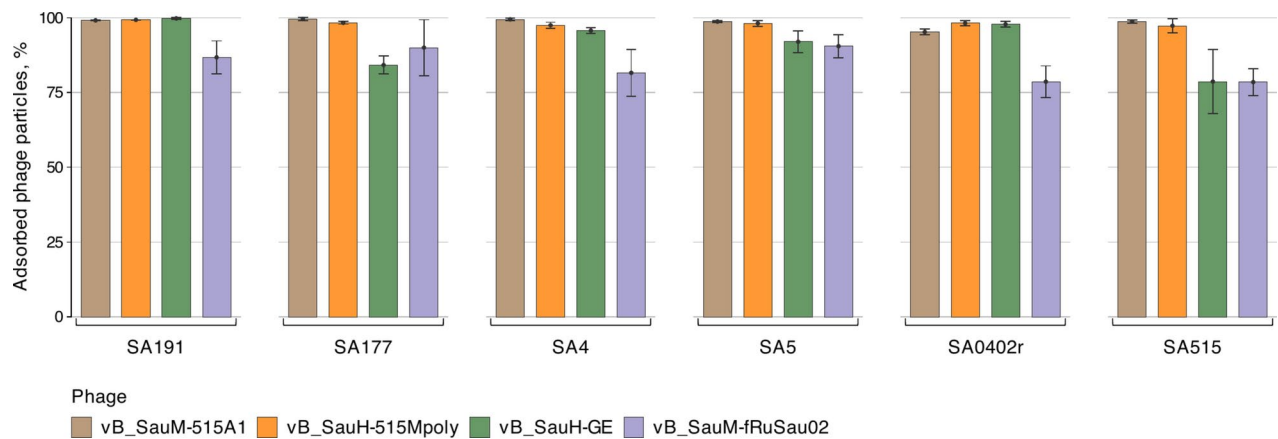


Fig. 1. Phage adsorption assay of *Herelleviridae* phages on *S. aureus* clinical isolates. The results are presented as the average percentage of adsorbed phage particles $\pm$ standard deviation (SD) of $n = 3$ biological replicates.

toxin-antitoxin systems. In *Herelleviridae*-resistant strains, no specific single nucleotide polymorphisms (SNPs) were detected in comparison to other members of ST239.

In the genome of strain SA191, two prophages, each approximately 40 kbp in length and belonging to the genus *Biseptimavirus*, were identified (Fig. 3, Figure S3). The prophage from region 1 showed the highest similarity to the recently published phage P630 (74% in length with 99.66% identity)³⁸. This prophage harbored a type 6 integrase and was found to be integrated into the lipase gene, which is a distinctive trait of Sa6int phages³⁹. The second prophage harbored a type 3 integrase and was integrated into the beta-hemolysin gene, which is also a typical feature of Sa3int phages³⁹. It harbored genes from the immune evasion cluster (IEC), including enterotoxin A (*sea*), staphylokinase (*sak*), and staphylococcal complement inhibitor (*scn*). Both prophages exhibited significant similarity to phage phiN315, with over 50% and 80% sequence identity in length for the first and second prophages, respectively. However, the prophages lacked the *tarP* gene associated with resistance to phages of the *Rountreeviridae* family⁴⁰.

Transcriptional and proteomic response of resistant *S. aureus* ST239 strain to phage infection

To describe the systemic response to phage infection, we selected strain SA191 and phage vB_SauM-515A1. Cells for transcriptomic analysis were harvested at 5- and 30-min post-infection, and for proteomic analysis at 30 min. Uninfected bacterial cells were used as a control.

RNA sequencing generated approximately 5.54 and 6.69 million high-quality reads per sample for uninfected and infected samples, respectively. Approximately 5.4 million reads per sample from uninfected culture and 910,000 reads per sample from infected culture were mapped to the SA191 strain genome. In the infected culture, approximately 5.57 million reads per sample were mapped to the phage genome. Over the course of infection, the proportion of reads mapped to the bacterial genome decreased from 18.9% at 5 min to 8% at 30 min post-infection per sample, while the percentage of reads mapped to the phage genome increased from 78.5 to 88.6%.

Comparison of the transcriptomic profile of infected strain SA191 to the control revealed 462 (271 $\uparrow$ 191 $\downarrow$) and 504 (276 $\uparrow$ 228 $\downarrow$) differentially expressed genes (DEGs) at 5 and 30 min, respectively (Table S2). Among these, 192 genes were upregulated, 123 genes were downregulated throughout the infection process, and 4 genes showed a shift in behavior, from upregulation at 5 min to downregulation at 30 min.

During proteomic studies of the phage-infected and uninfected bacterial culture 79,440 and 65,938 LC-MS/MS spectra were obtained, respectively. Qualitative analysis of the proteomic data revealed 1970 proteins in the case of phage infection and 1801 for the uninfected strain (Table S3). Quantitative analysis identified 184 proteins of SA191 strain (41 $\uparrow$ 143 $\downarrow$) with altered abundance levels in the infected strain compared to the control (Table S2). Changes in abundance for 6 ($\uparrow$) and 27 ($\downarrow$) proteins were consistent with the study at the transcriptomic level.

The Gene Ontology (GO) analysis of DEGs revealed their involvement in 32 GO categories: 17 Biological Processes (BP) terms, 14 Molecular Function (MF) terms, and 1 Cellular Components (CC) term. Specifically, for the early stage of infection (5 min), we observed significant enrichment ($p < 0.05$) in 18 categories (comprising 8 BP terms, 9 MF terms, and 1 CC term). At the later stages of infection, significant changes were identified in 20 categories (including 11 BP terms and 7 MF terms) (Fig. 4). Among these categories, two BP terms (GO: 0008645 and GO: 0009401) and four MF terms (GO: 0008982, GO: 0015149, GO: 0015078, GO: 0015294) exhibited significant enrichment both during the early and late stages of infection.

Functional analysis of differentially abundant proteins at the 30 min time point revealed changes compared to the uninfected control in only two GO terms related to molecular function: nucleotide binding (GO:0000166) and amino acid binding (GO:0016597).

In addition to GO analysis, prophage regions were analyzed, as prior studies have shown that lytic phages can activate prophage genes in bacteria^{41,42}. During phage infection, a notable upregulation of substantial number of genes from both prophages was observed: 33 genes at 5 min, and 43 genes at 30 min in the first prophage region, which comprises a total of 67 genes; similarly, 43 genes at 5 min, and 42 genes at 30 min in the second prophage

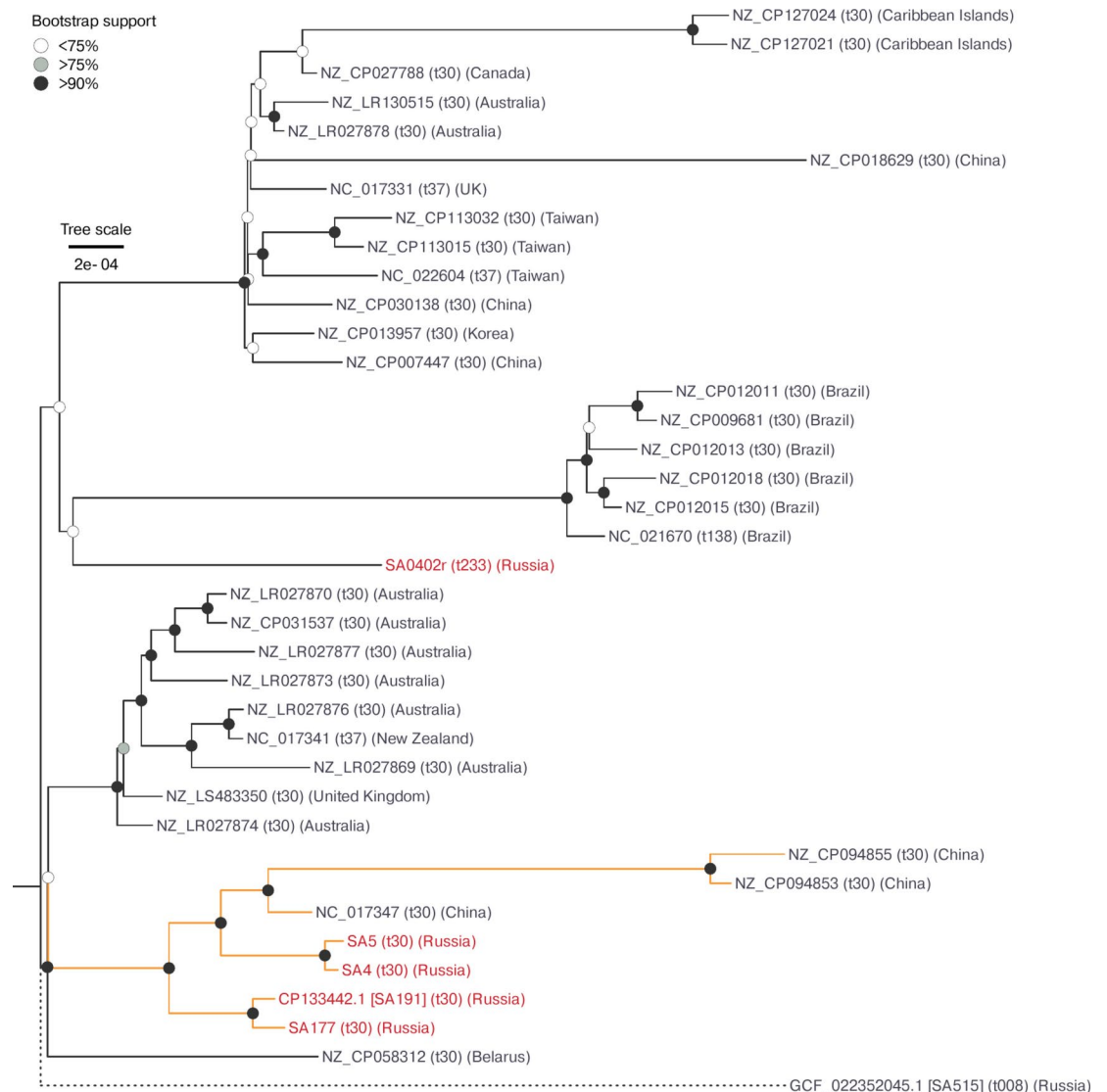


Fig. 2. Phylogeny of 38 *S. aureus* ST239 genomes. Maximum-likelihood phylogenetic tree constructed using 6,681 high-quality genome-wide SNPs from 38 *S. aureus* ST239 genomes and rooted on *S. aureus* SA515 (RefSeq assembly no. GCF_022352045.1 [branch length is omitted]). Bootstrap support values are shown as white <75%, grey ≥75% or black ≥90% dots on interior nodes. Isolates used in this study are colored red, resistant cluster R is colored orange. Spa type and origin of strains are indicated in parentheses.

region, totaling 63 genes. Also, an increase in the transcripts level for several virulence factor genes was observed (Table S2). No overrepresentation of prophage proteins was found at the proteomic level (Table S3).

In general, 27 out of 80 virulence factor genes were upregulated upon exposure of bacterial cells to phage (Table S2). Changes at the protein level were registered for MSCRAMM family adhesins (SdrD [RDJ18_RS03215], ClfA [RDJ18_RS04390], ClfB [RDJ18_RS14275]), type VII secretion proteins (EsbB [RDJ18_RS01310], EssC [RDJ18_RS01315]), exotoxins (staphylococcal protein A [RDJ18_RS00440]), as well as an immune modulator (RDJ18_RS12895) (Table S3).

To assess the transcriptional activity of known bacterial antiphage defense systems, the genes related to teichoic acid synthesis, restriction-modification systems, and toxin-antitoxin systems were examined. The analysis revealed significant decreases in the level of transcripts of key teichoic acid synthesis genes: *tarL* (downregulated at 5 and 30 min), *tarA*, *tarJ*, *tarK* (downregulated at 30 min), and glycosyltransferase *tarM* (downregulated at 30 min). Although the remaining genes in this cluster (*tarB*, *tarG*, *tarX*) were also downregulated, the decrease was less than two-fold.

Despite the observed changes in transcription, the early stages of phage infection in the resistant strain revealed an upregulation of genes associated with teichoic acid modification, specifically *dltA* and *dltB* (upregulated at 5 min). However, these changes were not corroborated by proteomic analysis.

The majority of genes involved in restriction-modification systems and toxin-antitoxin systems showed no significant changes on transcriptomic level during the infection. The proteomic analysis identified the

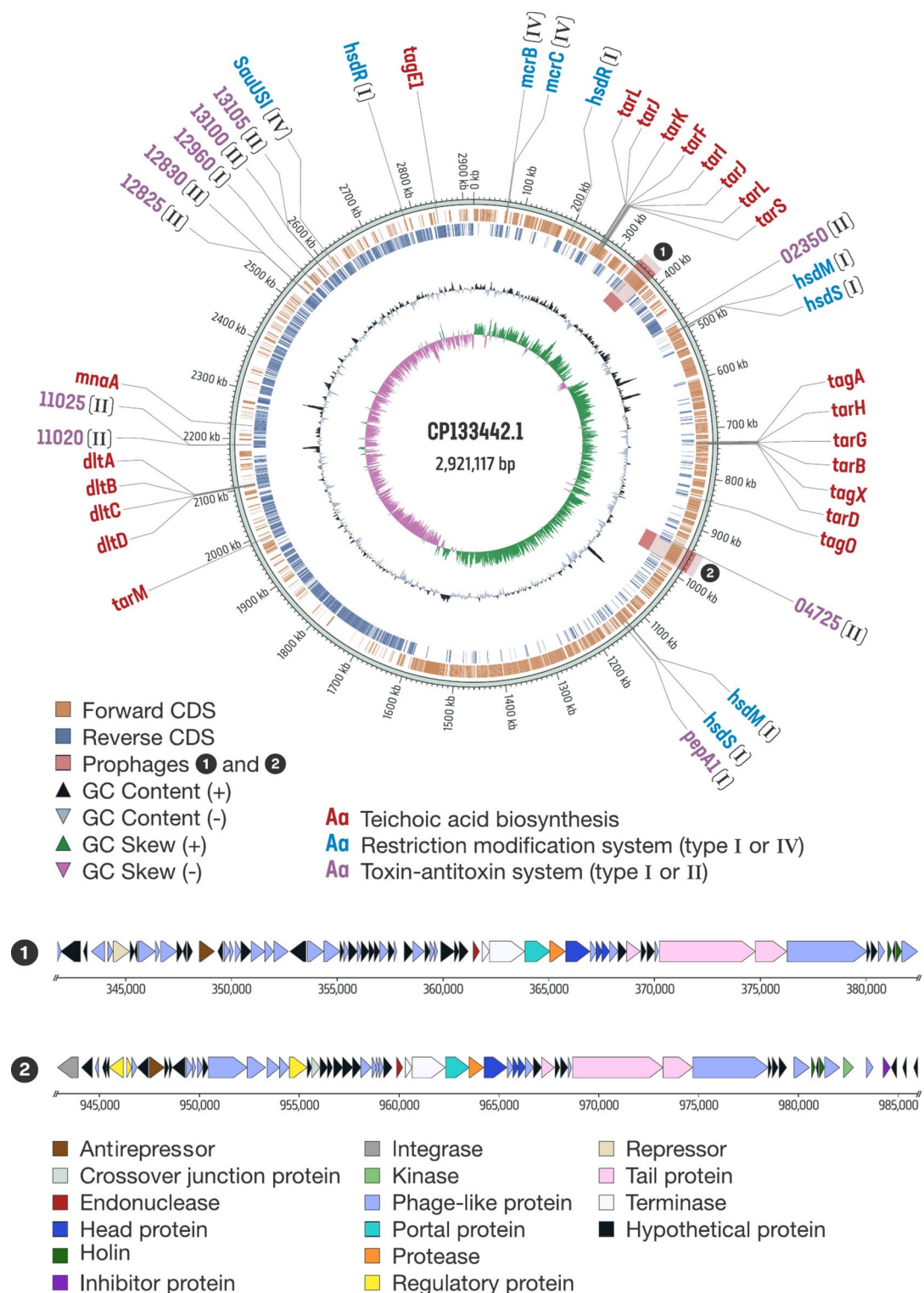


Fig. 3. Genomic organization of *S. aureus* strain SA191. Genes are referred to by their names when available; otherwise, numbered parts of locus tags are utilized. Prophage regions are designated with numerals (1 and 2). Genome maps highlight prophage regions structure and prophage genes are color-coded according to their predicted functions, as specified in the legend.

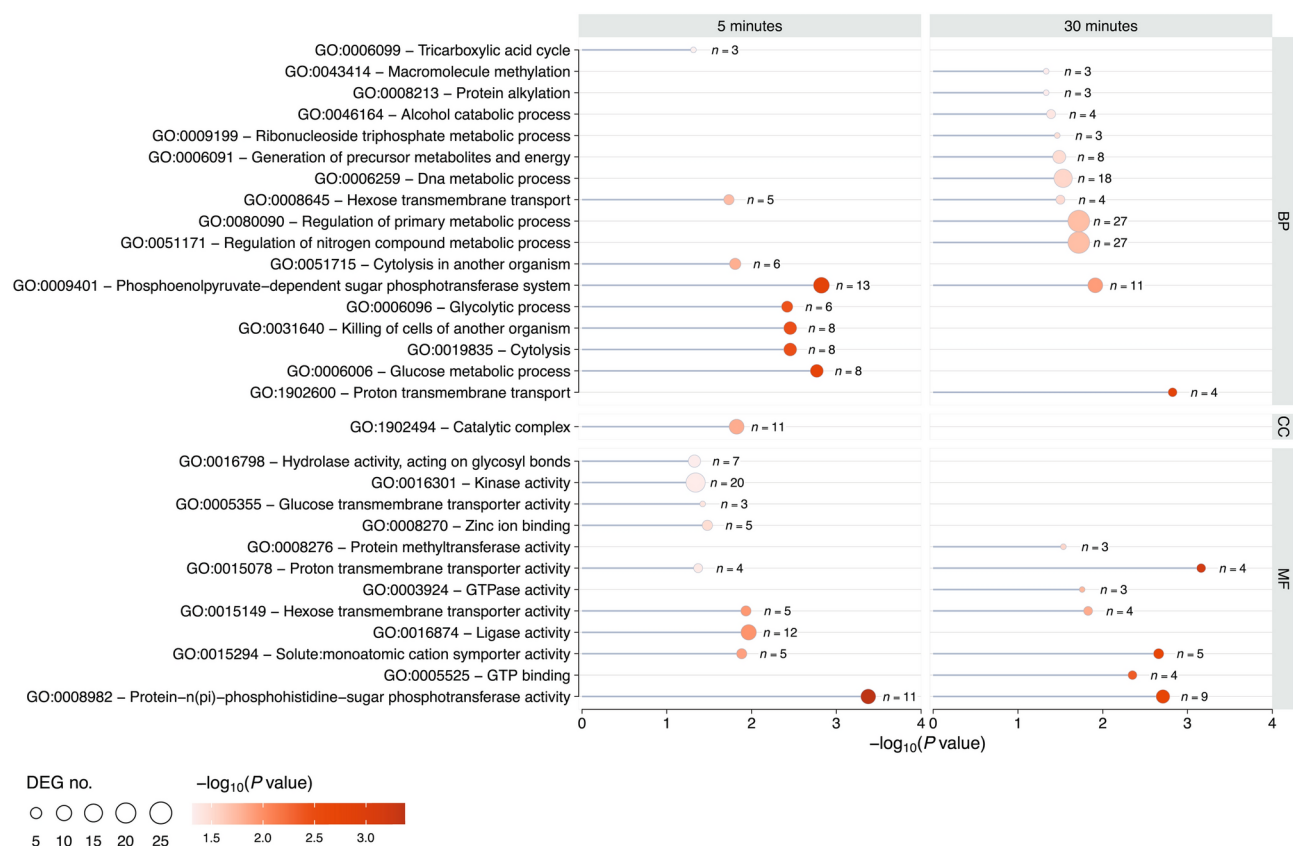


Fig. 4. Gene Ontology enrichment analysis during the infection process of *S. aureus* SA191 with *Herelleviridae* phage. The size of each circle represents the total number of Differentially Expressed Genes (DEGs) associated with the respective GO term, while the color intensity indicates the P-value ($-\log_{10}$).

underrepresentation of certain proteins from these systems (restriction endonuclease subunit S [RDJ18_RS0596] and the type II toxin-antitoxin system PemK/MazF family toxin [RDJ18_RS11020]). Noteworthy are the findings regarding the upregulation of genes *lrgA* (RDJ18_RS01180) and *lrgB* (RDJ18_RS01185) at 5 and 30 min, which encode the antiholin-like murein hydrolase modulator and antiholin-like protein, respectively, which was not confirmed at the proteomic level.

Discussion

In this study, our objective was to delve into the comprehension of potential mechanisms underlying the resistance of *S. aureus* ST239 strains to phages belonging to the *Herelleviridae* family. To achieve this, we selectively chose strains with both resistant and sensitive genotypes (Table 1).

Through comparative genomics, we have shown that the resistant strains under investigation, despite being isolated from hospitals in different cities (Table 1), are closely related and form a distinct phylogenetic cluster with strains from China (Fig. 2). Analysis of the teichoic acid genes of the cluster strains revealed differences in the *tarK*, *tagX*, and *dltA* genes, which are involved in the synthesis of poly(glycerol phosphate) wall teichoic acid, glycosylation, and D-alanylation of wall teichoic acid, respectively^{43,44}. In turn, no differences were found in the genes responsible for the synthesis of the poly (glycerol phosphate) wall teichoic acid backbone, which is the main receptor for *Herelleviridae* family phages. This explains the lack of differences in the phage adsorption study results for resistant and sensitive strains. This also aligns with previously published studies, which indicate that resistance to *Herelleviridae* phages occurs at the intracellular level⁴⁵. An examination of genes that may influence intracellular resistance to phages did not uncover significant differences between sensitive and resistant strains. The restriction-modification, toxin-antitoxin systems genes, and the overall content of mobile genetic elements were similar between the ST239 strains (Figure S3).

Strain SA191 was used for further transcriptomic and proteomic analyses. For transcriptomic studies, the 5 and 30 min time points, corresponding to the activity of the early and late promoters of the vB_SauM-515A1 phage on the host strain SA515, were selected⁴⁶. This strain belongs to ST8 and exhibits significant phylogenetic distance from ST239 strains (Fig. 2). Nevertheless, according to publications, ST8 strains are the closest ancestors to ST239 strains¹⁴. For proteomic analysis, a time point of 30 min was chosen to assess the overall effect of phage infection on the resistant strain. However, certain limitations of the proteomic approach should be noted, specifically the use of lysostaphin to facilitate more efficient lysis of *S. aureus* cells. Both the enzyme and the

duration of treatment may have an impact on the cells and the resulting protein profile, which is not specifically addressed in this study.

During analysis, we observed significant differences in the response of the resistant strain SA191 and the sensitive strain SA515 to infection with vB_SauM-515A1 phage. The sensitive strain displayed its most pronounced response to phage infection during the early stages, whereas the resistant strain exhibited comparable numbers of DEGs throughout both early and late stages of the infection. We did not identify any common enriched GO terms characterizing the response to phage infection in either the resistant or sensitive strains. Enriched GO terms were only identified in the sensitive strain during the initial stages of infection, primarily associated with nucleotide metabolism, specifically nucleoside monophosphate metabolism, as well as amino acids^{41,47}. The primary changes observed in cells of the resistant strain, caused by phage infection, were related to cellular energy metabolism. Significant alterations were detected in the tricarboxylic acid cycle, glycolytic process, and glucose metabolic process (Fig. 4). Differences from the uninfected control were also associated with the transcription level of genes encoding transport systems, such as the phosphoenolpyruvate-dependent sugar phosphotransferase system and hexose transmembrane transport.

A thorough analysis of changes associated with energy metabolism in the resistant strain revealed upregulation of genes involved in the transport of glucose and fructose into the cell at both early and late stages (RDJ18_RS00835, RDJ18_RS01075, RDJ18_RS01200, RDJ18_RS03920, RDJ18_RS13495, RDJ18_RS14335) (Table S4).

The key genes encoding enzymes involved in glycolysis exhibited upregulation: pyruvate kinase (RDJ18_RS06570), L-lactate dehydrogenase (RDJ18_RS01065), phosphoglycerate kinase (RDJ18_RS04310), triose-phosphate isomerase (RDJ18_RS04315), and 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (RDJ18_RS04320). However, glyceraldehyde-3-phosphate dehydrogenase (RDJ18_RS06625) and 2,3-diphosphoglycerate-dependent phosphoglycerate mutase (RDJ18_RS12880) were downregulated. The former catalyzes the conversion of glyceraldehyde-3-phosphate to 1,3-diphosphoglycerate, while the latter catalyzes the conversion of 3-phosphoglycerate to 2-phosphoglycerate. The reduced transcription level of these genes despite the upregulation of other glycolysis genes could be attributed to feedback inhibition.

The possible accumulation of pyruvate in the cell was facilitated by the downregulation of genes encoding enzymes involved in gluconeogenesis: phosphoenolpyruvate carboxykinase (ATP) (RDJ18_RS06075) and fructose-1,6-bisphosphatase (RDJ18_RS13375). The accumulated pyruvate in the infected strain was converted to lactate, as indicated by the upregulation of lactate dehydrogenase gene. Downregulation of genes encoding tricarboxylic acid cycle enzymes, including class II fumarate hydratase (RDJ18_RS05700), UDP-N-acetylmuramate-L-alanine ligase (RDJ18_RS06585), and malate dehydrogenase (quinone) (RDJ18_RS12615), argues for support of active conversion of pyruvate to lactate.

The proteomic analysis data confirmed alterations in the tricarboxylic acid cycle, evidenced by reduced abundance of dihydrolipoyllysine-residue succinyltransferase (RDJ18_RS08040) and succinate-CoA ligase subunit alpha (RDJ18_RS08935), along with changes in gluconeogenesis attributed to underrepresentation of the key enzyme, phosphoenolpyruvate carboxykinase (RDJ18_RS06075). Compensating for decreased energy production through the tricarboxylic acid cycle, there was an increased abundance of pyruvate dehydrogenase (acetyl-transferring) E1 component subunit alpha (RDJ18_RS09720) at the protein level, which participates in the synthesis of acetyl-CoA from pyruvate. Considering the presented changes at the transcriptomic and proteomic levels, it is possible to make an assumption about the probable shift of infected cells of the resistant strain from aerobic to anaerobic respiration, which is significantly different from the response to phage infection in the case of the susceptible strain SA515 (Table S4).

In both instances involving strains resistant to *Herelleviridae* family phages and sensitive strains during phage infection, DNA metabolism is impacted. However, in the case of a resistant strain, the ribonucleoside-triphosphate metabolic process (GO: 0009199) category was enriched, which is predominantly associated with cellular energy processes. In contrast, for the sensitive strain, enriched categories are linked to nucleoside monophosphate formation.

In the transcriptional response to phage infection of the resistant strain, an enrichment of GO terms associated with cytolysis was observed. This category encompassed genes encoding hemolysins with various localizations (prophage/chromosome), and their transcript levels were elevated irrespective of localization. For the sensitive strain, only the gamma-hemolysin genes located outside the prophage region were upregulated.

Regarding other virulence factors, akin to the sensitive strain, genes associated with the Effector delivery system (specifically, the WXG100 family type VII secretion effector EsxA) and Immune modulation (including immunoglobulin G-binding protein SBI and staphylococcal complement inhibitor SCIN) were upregulated in the resistant strain. The staphylokinase gene was downregulated in the sensitive strain compared to the resistant strain.

In discussions concerning the classical mechanisms of *S. aureus* resistance to phages associated with teichoic acid synthesis, restriction-modification, and toxin-antitoxin systems, no alterations were observed at the transcriptional and proteomic levels that could explain the strain's resistance to *Herelleviridae* phages. One exception is the consistent upregulation of the *lrgA* (RDJ18_RS01180) and *lrgB* (RDJ18_RS01185) genes throughout the infection process. These genes encode the modulator of anti-cholinergic murine hydrolase LrgA and the anti-cholinergic protein LrgB^{48,49}, both involved in programmed cell death, which could potentially be associated with the observed resistance to *Herelleviridae* phages in this strain. Notably, previous research has demonstrated that upregulation of these genes is not observed in sensitive strain⁴¹.

Another potential explanation for the development of resistance to *Herelleviridae* phages in the examined strain could involve mechanisms that prevent superinfection. In our study, we observed the upregulation of genes linked to prophage regions, which may play a role in preventing superinfection. These genes include the phage antirepressor KilAC domain-containing protein (RDJ18_RS01540 [first prophage], RDJ18_RS01550 [first prophage], RDJ18_RS04765 [second prophage]), and the siphovirus Gp157 family protein (RDJ18_RS01595

[first prophage]^{41,50}. The phenomenon of gene upregulation within prophages during phage infection has been documented in various studies, encompassing infections of *S. aureus* strains with *Herelleviridae* phages^{41,42}, as well as in other bacteria and phage families^{50,51}. Nevertheless, investigating the contribution of the listed genes to resistance seems to be quite challenging due to the high level of mosaicism of the prophages. For example, RDJ18_RS01595 was not found in the genome of strain SA515, and the other genes showed less than 90% identity. In turn, for the genome of the sensitive strain SA402r, the sequence identity of the four genes ranged from 90 to 100%. This rather reduces the role of the described genes in the formation of resistance but at the same time requires further investigation.

Conclusion

The conducted study demonstrated that ST239 *S. aureus* strains can exhibit resistance to virulent phages of the *Herelleviridae* family, implemented at the intracellular level. Comparative genomic analysis revealed a close relationship among resistant strains, forming a separate cluster on the phylogenetic tree. This suggests their unique genetic features and highlights their potential significance in elucidating the dynamics of interactions between phages and bacteria. Through a systems analysis of one of the resistant strains, it was revealed that the response to phage infection significantly differs from that of a sensitive strain. This contrast is evident in gene transcription level and cell metabolism, particularly in energy metabolism. Noteworthy alterations were observed in glycolysis, the tricarboxylic acid cycle, and transport systems. Despite these data, our analysis revealed no significant alterations in the representation of systems linked to classical resistance mechanisms, implying the existence of undiscovered resistance pathways. The absence of anticipated mechanisms underscores the intricate nature of phage-bacterial interactions and suggests the possible presence of alternative strategies employed by *S. aureus* to combat phage infection.

Materials and methods

S. aureus strains and phages

The strains of *S. aureus* used in this study were obtained from the collection of the Laboratory of Molecular Genetics of Microorganisms at the Lopukhin Federal Research and Clinical Center of Physical–Chemical Medicine of the Federal Medical Biological Agency (Table 1). Among these strains, five (SA191, SA177, SA4, SA5, SA0402r) belonged to ST239, while strain SA515 belonged to ST8.

STs and spa-types of the strains were determined using standard MLST scheme (<https://pubmlst.org/organisms/staphylococcus-aureus>) and spa typing (<https://spaserver.ridom.de>), respectively. The antibiotic resistance profiles of the strains (oxacillin, ciprofloxacin, clindamycin, erythromycin, gentamicin, linezolid, tetracycline, and vancomycin) were determined according to the Clinical and Laboratory Standards Institute (CLSI) guidelines. All strains were cultured in Luria–Bertani (LB) broth or on LB agar plates at 37 °C.

To characterize the interaction of the selected strains with phages of the *Herelleviridae* family, four phages were used. Phage vB_SauM-515A1 (GenBank accession no. MN047438.1) had been previously characterized⁸. Phages vB_SauH-Mpol (GenBank accession no. OR513906.1) and vB_SauH-GE (GenBank accession no. OR513904.1) were isolated using strain SA515 as the host from commercial preparations, employing the enrichment culture method according to the standard protocol⁵¹. Phage vB_SauM-fRuSau02 (GenBank accession no. MF398190.1) was generously provided by S. Kiljunen and K. Leskinen from the Department of Bacteriology and Immunology, Medicum, Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Finland.

Host range determination and adsorption assay

The sensitivity of *S. aureus* strains to phages was evaluated using the efficiency of plating (EOP) method, as previously described⁸. EOP was calculated as the ratio of the average effective phage titer on the tested strain to the titer on the host.

The ability of phages to adsorb on bacterial cells was assessed, following a previously described protocol⁵², with slight modifications. Specifically, LB broth was used instead of brain heart infusion medium for growing test strains and as the negative control.

Whole-genome sequencing

The DNA of phages and *S. aureus* was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA) following the manufacturer's protocol. The DNA concentration and quality were assessed using a Qubit 4 Fluorometer and Nanodrop ND-1000 (Thermo Fisher Scientific Inc.).

For WGS sequencing on the Illumina platform, 100 ng of the extracted DNA underwent library preparation with the KAPA HyperPlus Kit (Roche, Switzerland) as per the manufacturer's instructions. Sequencing was performed on the HiSeq 2500 platform (Illumina, USA) following the manufacturer's recommendations. Reagent kits utilized included: HiSeq Rapid PE Cluster Kit v2, HiSeq Rapid SBS Kit v2 (200 cycles), and HiSeq Rapid PE FlowCell v2. A 2% PhiX spike-in control was also included. The results of whole-genome sequencing have been deposited in the NCBI Sequence Read Archive (SRA) database under accession number PRJNA1027597.

For WGS sequencing on the Oxford Nanopore platform, the library was prepared according to the manufacturer's protocol using NEB reagents (New England Biolabs, Inc.). Long reads were generated with PromethION sequencing (Oxford Nanopore Technologies, UK) using the ligation sequencing kit SQK-LSK109 and native barcoding expansion kit EXP-NBD196, and run in a FLO-PRO002 flow cell. Basecalling was performed using Guppy v6.5.7 with default parameters (high accuracy model (HAC), minimum quality value ≥ 7).

RNA sequencing

For RNA sequencing, the SA191 bacterial culture (early log phase; OD₆₀₀=0.12) in LB was infected with the vB_SauM-515A1 phage at a multiplicity of infection (MOI) of 10 and incubated at 37 °C with shaking (220 rpm). Samples were collected at 5, 15, and 30 min intervals following the addition of the phage. The cells were centrifuged, harvested, and immediately stored at −70 °C. These procedures were replicated in three biological repeats.

The RNA extraction procedure and subsequent sequencing have been described previously⁵³. A total of 350 ng of total RNA was used for library preparation. Ribosomal RNA was removed using the NEBNext rRNA Depletion Kit (New England Biolabs, Inc.), and libraries were prepared with the KAPA RNA Hyper Kit (Roche, Switzerland) following the manufacturer's protocol. RNA cleanup was then performed using the RNA Clean XP kit (Beckman Coulter, Brea, USA). The library underwent a final cleanup with KAPA HyperPure Beads (Roche, Switzerland), and its size distribution and quality were assessed using a high-sensitivity DNA chip (Agilent Technologies). Libraries were quantified using the Quant-iT DNA Assay Kit, High Sensitivity (Thermo Fisher Scientific, Inc.). Finally, equimolar quantities of all libraries (10 pM) were sequenced in a high-throughput run on the Illumina HiSeq 2500 using 50 bp reads and a 2% Phix spike-in control. The RNA-Seq dataset was deposited in the NCBI SRA database under accession number PRJNA1066200.

Proteomic analysis

Aliquots for proteomic analysis were sampled at a point 30 min post-infection in parallel with the RNAseq experiment and then centrifuged at 5000 rpm for 5 min. Initially, the cells were treated with lysostaphin (Sigma-Aldrich, Inc.) (200 µg/mL lysostaphin in 20 mM Tris-HCl, pH 7.5, 10 mM EDTA) for 30 min at 37 °C. After adding 10% sodium deoxycholate (reaching a final concentration of 1%) and a mixture of nucleases, the samples were incubated for 30 min at 4 °C. Tris(2-carboxyethyl)phosphine and chloroacetamide were added to final concentrations of 5 mM and 30 mM, respectively. Cysteine reduction and alkylation were achieved by incubating the sample at 80 °C for 10 min. Proteins were precipitated using chloroform-methanol. The protein pellet was then resuspended in 100 µL of 50 mM Tris-HCl buffer (pH 8.5), and the protein concentration was determined using a BCA Assay Kit (Thermo Fisher Scientific Inc.). Trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega) was added to the protein samples at a ratio of 1/50 w/w and incubated for 16 h at 37 °C. Proteolysis was stopped by adding trifluoroacetic acid to a final concentration of 1%. Peptides were dried in a SpeedVac (Thermo Fisher Scientific Inc.) and resuspended in 20 µL of 3% ACN, 0.1% TFA in mQ. The peptide concentration was determined using the BCA Assay Kit (Thermo Fisher Scientific Inc.). The samples were obtained in three biological replicates.

Liquid chromatography of proteome samples was performed using a Dionex Ultimate 3000 nano-LC system (Thermo Fisher Scientific, Waltham, MA) equipped with laboratory made C-18 column (length 40 cm with 150 µm ID) packed with Kinetex C18, 2.4 µm (Phenomenex, Torrance, CA). The flow rate was set 400 µL/min at 60°C. The Buffer A composition was 0.1% formic acid in LC/MS-grade water and buffer B was 80% acetonitrile, 0.1% formic acid in LC/MS-grade water. Gradient was run from 3 to 56% of Buffer B for 128 min. A total of 2 µg of tryptic digest was injected onto the enrichment column for concentration/purification. MS analysis was performed using Orbitrap Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) equipped with nano-spray source (+ 2.2 kV, capillary temperature 300°C). The range for MS1 scans was set 350–1500 m/z with 45 ms maximum injection time, AGC target 3e6 and resolution of 60000. HCD was set to 30 eV. MS2 scans were acquired with range 200–2000 m/z and isolation window 2.0 Da, resolution 30000, AGC 2e5 and 50 ms of maximum injection time. Full MS/dd-MS2 method (Top12) was used in this analysis.

Bioinformatics analysis

Assembly and annotation of *S. aureus* genomes

The quality assessment of short-reads was performed using FastQC v0.12.1 (<https://github.com/s-andrews/FastQC>) and MultiQC v1.17⁵⁴. Short-read assemblies of the SA4, SA5, SA0402r, and SA177 strains were conducted using Unicycler v0.5.0⁵⁵ and SPAdes v3.15.5⁵⁶. PGAP 2023–05-17.build6771⁵⁷ was utilized for annotating the assemblies. BWA MEM v0.7.17⁵⁸ was employed for mapping reads to the assemblies, with SAMtools v1.17⁵⁹ and mosdepth v0.3.4⁶⁰ used to compile mapping statistics.

Long-read quality was evaluated with Nanoq v0.10.0⁶¹. Prior to assembly long reads were filtered with Filtlong v0.2.1 (<https://github.com/rrwick/Filtlong>). Hybrid assemblies of the SA191 strain were created using Tricycler v0.5.4⁶², Flye v2.9.2-b1786⁶³, Canu v2.2⁶⁴, Raven v1.8.1⁶⁵, Medaka v1.7.3 (<https://github.com/nanoporetech/medaka>), Polypolish v0.5.0⁶⁶, and Polca v4.1.0 from MaSuRCA toolkit⁶⁷. PGAP 2022–12-13.build6494 was applied for annotating hybrid assemblies. Minimap2 v2.26-r1175⁶⁸ was used to map long reads to the assemblies, and SAMtools v1.17 along with mosdepth v0.3.4 were utilized to collect mapping statistics. SeqKit v2.4.0⁶⁹ was employed for sorting and filtering sequences. *S. aureus* strain SA191 was submitted to the NCBI GenBank database under accession number CP133442.1.

Prophage sequences within SA191 were identified using the PHASTEST web server⁷⁰. To search for genomic islands, IslandViewer 4 was used⁷¹. SCCmecFinder was used to determine the type of SCCmec cassette⁷². The visualization of the SA191 chromosome and mobile elements was accomplished using Circos v0.69–8 (<https://circos.ca>), in conjunction with the following R (R Core Team. R: A Language and Environment for Statistical Computing. 2023) packages: scales v1.2.1 (<https://CRAN.R-project.org/package=scales>), tidyverse v2.0.0⁷³, cowplot v1.1.1 (<https://CRAN.R-project.org/package=cowplot>), gggenes v0.5.1 (<https://CRAN.R-project.org/package=gggenes>), and ggplot2 v3.4.4⁷⁴.

Assembly and annotation of phage genomes

The quality assessment of short-reads was conducted using FastQC v 0.11.9 and MultiQC v v1.13. Adapter removal and quality trimming were performed using fastp v0.23.2. Phage genomes were assembled using Unicycler v0.5.0 and SPAdes v3.15.5. Prokka v1.14.6⁷⁵, supplemented with the pVOG⁷⁶ and PHROG v3⁷⁷ databases, was employed to annotate the resulting phage assemblies.

Retrieval of assembled genomes of *S. aureus* ST239

Assembled genomes of *S. aureus* ST239 were obtained using NCBI Datasets command-line tool (<https://github.com/ncbi/datasets>) with the following parameters: “taxon 1280 –assembly-level chromosome,complete”. MLST types were determined using MLST v2.23.0 (<https://github.com/tseemann/mlst>). Variants within these complete genomes were identified using NUCmer and show-snps from the MUMmer v4.0.0rc1⁷⁸ package. The effects of these variants were annotated using SnpEff v5.1d⁷⁹.

Phylogeny construction

Complete ST239 genomes retrieved from NCBI were used to generate simulated FASTQ read files using the wgsim v1.17 from SAMtools⁵⁹ package with the following parameters: “-1 150 -2 150 -N 500000 -S 1984 -r 0 -R 0 -X 0 -e 0”. Subsequently, all FASTQ reads were aligned to the reference *S. aureus* NCTC 8325 genome (RefSeq accession no. NC_007795.1) employing the BWA MEM v0.7.17 algorithm. The aligned reads were sorted by coordinates, converted to BAM format, and indexed using SAMtools v1.17. Duplicate reads were then eliminated using Samblaster v0.1.26⁸⁰. Mapping quality was evaluated with SAMtools v1.17 and mosdepth v0.3.4. Variant calling was conducted using BCFtools v1.17⁵⁹. The core SNP alignment was generated by consolidating variants from all samples, converting them to tab-delimited format with GATK VariantsToTable v4.5.0.0⁸¹, and extracting the SNP alignment using a Python script. Recombinant regions were filtered out using Gubbins v3.3.1⁸², and the resulting alignment was refined with SNP-sites v2.5.1⁸³. The maximum likelihood (ML) phylogeny was inferred from 45 sequences with 6,837 core SNP sites using IQ-TREE 2 v2.3.4⁸⁴. Support values were determined from 10,000 ultrafast bootstrap replicates (UFBoot)⁸⁵ with the “-bnni” parameter and from 10,000 replicates for Shimodaira-Hasegawa (SH) approximate likelihood ratio test with the “-altr” parameter. The best-fit model was identified by ModelFinder⁸⁶ with the “-m MFP” parameter, and the model selected based on the Bayesian information criterion (BIC) was TVMe + ASC + R2. *S. aureus* SA515 (RefSeq assembly no. GCF_022352045.1) was used as an outgroup. Phage genomes were acquired using ncbi-acc-download v0.2.8 (<https://github.com/kblin/ncbi-acc-download>), and the phylogeny of the phages was inferred using ViPTreeGen v1.1.2⁸⁷. The phylogenies were visualized using the following R packages: tidytree v0.4.5⁸⁸, phylotools v0.2.2 (<https://CRAN.R-project.org/package=phylotools>), ape v5.7-1⁸⁹, ggnewscale v0.4.9 (<https://CRAN.R-project.org/package=ggnewscale>), ggplot2 v3.4.4, ggtree v3.8.2⁹⁰, ggtext v0.1.2 (<https://CRAN.R-project.org/package=ggtext>), and glue v1.6.2 (<https://CRAN.R-project.org/package=glue>).

Differential gene expression analysis

Sequenced reads were aligned to the reference *S. aureus* SA191 genome (GenBank accession no. CP133442.1) using STAR v2.7.11a⁹¹. SAMtools v1.17 software was utilized for sorting and converting SAM files to BAM format, followed by indexing and subsequent statistical analysis. Mapping quality and coverage along genes were evaluated with QualiMap v2.2.2⁹², and individual reports were merged using MultiQC v1.15. Reads were assigned to genes using featureCounts v2.0.6⁹³. Differential gene expression analysis utilized the edgeR v3.42.4⁹⁴ package for R. Genes with a false discovery rate (FDR) cutoff of 0.05 and a fold change (log2FC) threshold of |1| (i.e., ≥ 2 -fold change) were deemed differentially expressed.

For gene ontology (GO) enrichment analysis, GO categories were annotated using the PANNZER2 tool⁹⁵ with a Positive Predictive Value (PPV) cutoff of 0.5. GO enrichment analysis was conducted in R using the topGO v2.52.0⁹⁶ package and visualized with ggsci v3.0.0 (<https://CRAN.R-project.org/package=ggsci>), ggh4x v0.2.6 (<https://CRAN.R-project.org/package=ggh4x>), cowplot v1.1.1, and ggplot2 v3.4.4 R packages.

Proteome data processing

Raw LC-MS/MS data obtained from the Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific Inc.) were processed into MGF peaklists using MSConvert (ProteoWizard Software Foundation). The conversion process utilized the following parameters: “-mgf –filter peakPicking true 1–2”. To ensure comprehensive protein identification, the resulting peak lists were then queried against a concatenated protein database using the MASCOT v2.5.1 and X! Tandem (ALANINE, 2017.02.01) search engines.

The concatenated protein database included sequences from *S. aureus* (annotated from bacterial genome data), phages (annotated from phage genome data), as well as potential contaminants such as trypsin digestion by-products and proteins from bovine and yeast cell culture media (sourced from the UniProt knowledgebase, taxa *Bos taurus* and *Saccharomyces cerevisiae*, respectively), along with a reverse decoy dataset. Search parameters were set with a precursor mass tolerance of 20 ppm and a fragment mass tolerance of 0.04 Da. Other parameters encompassed tryptic digestion with allowance for one missed cleavage, static modification for carbamidomethyl (C), and dynamic/flexible modifications for oxidation (M).

Subsequently, result files underwent validation and meta-analysis using Scaffold 5 v5.1.0 software. The local false discovery rate scoring algorithm was applied with standard experiment-wide protein grouping, and a false discovery rate of 5% was utilized for both peptide and protein hits. False positive identifications were determined through reverse database analysis.

For protein quantification, spectral counting, a straightforward label-free quantification strategy, was employed^{97,98}. Differential protein abundance analysis was performed using the limma package⁹⁹. No data were excluded from the analyses.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD047971.

Data availability

The results of whole-genome sequencing of *Staphylococcus aureus* strains have been deposited in the NCBI Sequence Read Archive database under accession number PRJNA1027597. Genomes of strain SA191 and bacteriophages vB_SauH-Mpol and vB_SauH-GE were submitted to the NCBI GenBank database under accession numbers CP133442.1, OR513906.1, OR513904.1 respectively. The RNA-Seq dataset was deposited in the NCBI SRA database under accession number PRJNA1066200. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD047971. (Project DOI: 10.6019/PXD047971). Reviewer account details: Username: reviewer_pxd047971@ebi.ac.uk; Password: sV0Nqoqo.

Received: 11 March 2024; Accepted: 22 November 2024

Published online: 26 November 2024

References

1. Uytendaele, S. et al. Safety and efficacy of phage therapy in difficult-to-treat infections: a systematic review. *Lancet Infect. Dis.* **22**, e208–e220 (2022).
2. Plumet, L. et al. Bacteriophage therapy for *Staphylococcus aureus* infections: A review of animal models, treatments, and clinical trials. *Front. Cell. Infect. Microbiol.* **12**, 907314 (2022).
3. Mendes, J. J. et al. In vitro design of a novel lytic bacteriophage cocktail with therapeutic potential against organisms causing diabetic foot infections. *J. Med. Microbiol.* **63**, 1055–1065 (2014).
4. Ooi, M. L. et al. Safety and tolerability of bacteriophage therapy for chronic rhinosinusitis Due to *Staphylococcus aureus*. *JAMA Otolaryngol. Neck Surg.* **145**, 723 (2019).
5. Petrovic Fabijan, A. et al. Safety of bacteriophage therapy in severe *Staphylococcus aureus* infection. *Nat. Microbiol.* **5**, 465–472 (2020).
6. Ferry, T. et al. Phage therapy as adjuvant to conservative surgery and antibiotics to salvage patients with relapsing *S. aureus* prosthetic knee infection. *Front. Med.* **7**, 570572 (2020).
7. Mulzer, J., Trampuz, A. & Potapov, E. V. Treatment of chronic left ventricular assist device infection with local application of bacteriophages. *Eur. J. Cardiothorac. Surg.* **57**, 1003–1004 (2020).
8. Kornienko, M. et al. Contribution of Podoviridae and Myoviridae bacteriophages to the effectiveness of anti-staphylococcal therapeutic cocktails. *Sci. Rep.* **10**, 18612 (2020).
9. Cui, Z. et al. Characterization and complete genome of the virulent Myoviridae phage JD007 active against a variety of *Staphylococcus aureus* isolates from different hospitals in Shanghai, China. *Virol. J.* **14**, 26 (2017).
10. Leskinen, K. et al. Characterization of vB_SauM-fRuSau02, a twort-like bacteriophage isolated from a therapeutic phage cocktail. *Viruses* **9**, 258 (2017).
11. Botka, T. et al. Lytic and genomic properties of spontaneous host-range Kayvirus mutants prove their suitability for upgrading phage therapeutics against staphylococci. *Sci. Rep.* **9**, 5475 (2019).
12. Ajuebor, J. et al. Comparison of *Staphylococcus* phage K with close phage relatives commonly employed in phage therapeutics. *Antibiotics* **7**, 37 (2018).
13. Smyth, D. S. et al. Population structure of a hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. *PLoS ONE* **5**, e8582 (2010).
14. Gill, J. L., Hedge, J., Wilson, D. J. & MacLean, R. C. Evolutionary processes driving the rise and fall of *Staphylococcus aureus* ST239, a dominant hybrid pathogen. *mBio* **12**, e02168–21 (2021).
15. Feil, E. J. et al. Rapid detection of the pandemic methicillin-resistant *Staphylococcus aureus* Clone ST 239, a dominant strain in Asian hospitals. *J. Clin. Microbiol.* **46**, 1520–1522 (2008).
16. De Backer, S. et al. Remarkable geographical variations between India and Europe in carriage of the staphylococcal surface protein-encoding sasX/sesI and in the population structure of methicillin-resistant *Staphylococcus aureus* belonging to clonal complex 8. *Clin. Microbiol. Infect.* **25**(628), e1–628.e7 (2019).
17. Lee, Y.-H., Chen, C.-J., Lien, R.-I. & Huang, Y.-C. A longitudinal molecular surveillance of clinical methicillin-resistant *Staphylococcus aureus* isolates in neonatal units in a teaching hospital, 2003–2018. *J. Microbiol. Immunol. Infect.* **55**, 880–887 (2022).
18. Patil, S. et al. Molecular epidemiology and characterization of multidrug-resistant MRSA ST398 and ST239 in Himachal Pradesh, India. *Infect. Drug Resist.* **16**, 2339–2348 (2023).
19. Liu, J. et al. Antimicrobial resistance, SCCmec, virulence and genotypes of MRSA in Southern China for 7 years: Filling the gap of molecular epidemiology. *Antibiotics* **12**, 368 (2023).
20. Wang, B. et al. Methicillin-resistant *Staphylococcus aureus* in China: a multicentre longitudinal study and whole-genome sequencing. *Emerg. Microbes Infect.* **11**, 532–542 (2022).
21. Goudarzi, H., Goudarzi, M., Sabzehali, F., Fazeli, M. & Salimi Chirani, A. Genetic analysis of methicillin-susceptible *Staphylococcus aureus* clinical isolates: High prevalence of multidrug-resistant ST239 with strong biofilm-production ability. *J. Clin. Lab. Anal.* **34**, e23494 (2020).
22. Abd El-Hamid, M. I. et al. Clonal diversity and epidemiological characteristics of ST239-MRSA strains. *Front. Cell Infect. Microbiol.* **12**, 782045 (2022).
23. Shang, W. et al. Comparative fitness and determinants for the characteristic drug resistance of ST239-MRSA-III-t030 and ST239-MRSA-III-t037 strains isolated in China. *Microb. Drug Resist.* **22**, 185–192 (2016).
24. Moncke, S. et al. Molecular typing of ST239-MRSA-III from diverse geographic locations and the evolution of the SCCmec III element during its intercontinental spread. *Front. Microbiol.* **9**, 1436 (2018).
25. Cha, H. Y. et al. Prevalence of the ST239 clone of methicillin-resistant *Staphylococcus aureus* and differences in antimicrobial susceptibilities of ST239 and ST5 clones identified in a Korean hospital. *J. Clin. Microbiol.* **43**, 3610–3614 (2005).
26. Van Hal, S. J. et al. Performance of various testing methodologies for detection of heteroresistant vancomycin-intermediate *Staphylococcus aureus* in bloodstream isolates. *J. Clin. Microbiol.* **49**, 1489–1494 (2011).
27. Zhang, X. et al. First report of a sequence type 239 vancomycin-intermediate *Staphylococcus aureus* isolate in Mainland China. *Diagn. Microbiol. Infect. Dis.* **77**, 64–68 (2013).
28. Choo, E. J. & Chambers, H. F. Treatment of methicillin-resistant *Staphylococcus aureus* bacteremia. *Infect. Chemother.* **48**, 267 (2016).
29. Lambert, M. IDSA guidelines on the treatment of MRSA infections in adults and children. *Am. Fam. Phys.* **84**, 455–463 (2011).

30. Moller, A. G., Lindsay, J. A. & Read, T. D. Determinants of phage host range in *Staphylococcus* species. *Appl. Environ. Microbiol.* **85**, e00209–e219 (2019).
31. Leprince, A. & Mahillon, J. Phage adsorption to gram-positive bacteria. *Viruses* **15**, 196 (2023).
32. Botelho, A. M. N. et al. Local diversification of methicillin-resistant *Staphylococcus aureus* ST239 in South America after its rapid worldwide dissemination. *Front. Microbiol.* **10**, 82 (2019).
33. Jurado, A., Fernández, L., Rodríguez, A. & García, P. Understanding the mechanisms that drive phage resistance in *Staphylococci* to prevent phage therapy failure. *Viruses* **14**, 1061 (2022).
34. Cooper, L. P. et al. DNA target recognition domains in the Type I restriction and modification systems of *Staphylococcus aureus*. *Nucleic Acids Res.* **45**, 3395–3406 (2017).
35. Corvaglia, A. R. et al. A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical *Staphylococcus aureus* strains. *Proc. Natl. Acad. Sci.* **107**, 11954–11958 (2010).
36. Xu, S., Corvaglia, A. R., Chan, S.-H., Zheng, Y. & Linder, P. A type IV modification-dependent restriction enzyme SauUSI from *Staphylococcus aureus* subsp. *aureus* USA300. *Nucleic Acids Res.* **39**, 5597–5610 (2011).
37. Sadykov, M. R. Restriction-modification systems as a barrier for genetic manipulation of *Staphylococcus aureus*. In *The Genetic Manipulation of Staphylococci* (ed. Bose, J. L.) (Springer, 2014).
38. Kraushaar, B. et al. Acquisition of virulence factors in livestock-associated MRSA: Lysogenic conversion of CC398 strains by virulence gene-containing phages. *Sci. Rep.* **7**, 2004 (2017).
39. Goerke, C. et al. Diversity of prophages in dominant *Staphylococcus aureus* clonal lineages. *J. Bacteriol.* **191**, 3462–3468 (2009).
40. Gerlach, D. et al. Methicillin-resistant *Staphylococcus aureus* alters cell wall glycosylation to evade immunity. *Nature* **563**, 705–709 (2018).
41. Kuptsov, N. et al. Global transcriptomic response of *Staphylococcus aureus* to virulent bacteriophage infection. *Viruses* **14**, 567 (2022).
42. Finstřlová, A. et al. Global transcriptomic analysis of bacteriophage-host interactions between a Kayvirus therapeutic phage and *Staphylococcus aureus*. *Microbiol. Spectr.* **10**, e00123–e222 (2022).
43. Brown, S., Santa Maria, J. P. & Walker, S. Wall teichoic acids of gram-positive bacteria. *Annu. Rev. Microbiol.* **67**, 313–336 (2013).
44. Winstel, V., Sanchez-Carballo, P., Holst, O., Xia, G. & Peschel, A. Biosynthesis of the unique wall teichoic acid of *Staphylococcus aureus* lineage ST395. *mBio* **5**, e00869–14 (2014).
45. Kuntová, L. et al. *Staphylococcus aureus* prophage-encoded protein causes abortive infection and provides population immunity against Kayviruses. *mBio* **14**, e02490–22 (2023).
46. Kornienko, M. et al. Transcriptional landscape of *Staphylococcus aureus* Kayvirus bacteriophage vB_SauM-515A1. *Viruses* **12**(11), 1320 (2020).
47. Kornienko, M., Bespiatykh, D., Gorodnichev, R., Abdraimova, N. & Shitikov, E. Transcriptional landscapes of Herelleviridae bacteriophages and *Staphylococcus aureus* during phage infection: An overview. *Viruses* **15**, 1427 (2023).
48. Groicher, K. H., Firek, B. A., Fujimoto, D. F. & Bayles, K. W. The *Staphylococcus aureus* *lrgAB* operon modulates murein hydrolase activity and penicillin tolerance. *J. Bacteriol.* **182**, 1794–1801 (2000).
49. Bayles, K. W. Are the molecular strategies that control apoptosis conserved in bacteria? *Trends Microbiol.* **11**, 306–311 (2003).
50. Ruiz-Cruz, S. et al. Lysogenization of a lactococcal host with three distinct temperate phages provides homologous and heterologous phage resistance. *Microorganisms* **8**, 1685 (2020).
51. Twest, R. & Kropinski, A. M. Bacteriophage enrichment from Water and Soil. In *Bacteriophages* (eds Clokie, M. R. J. & Kropinski, A. M.) (Humana Press, 2009).
52. Kelly, D., McAuliffe, O., O'Mahony, J. & Coffey, A. Development of a broad-host-range phage cocktail for biocontrol. *Bioeng. Bugs* **2**, 31–37 (2011).
53. Shitikov, E. et al. Genome-wide transcriptional response of mycobacterium smegmatis MC2155 to G-Quadruplex ligands BRACO-19 and TMPyP4. *Front. Microbiol.* **13**, 817024 (2022).
54. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
55. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).
56. Pribelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes De Novo assembler. *Curr. Protoc. Bioinform.* **70**, e102 (2020).
57. Tatusova, T. et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624 (2016).
58. Li, H. & Durbin, R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
59. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
60. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics* **34**, 867–868 (2018).
61. Steinig, E. & Coin, L. Nanoq: ultra-fast quality control for nanopore reads. *J. Open Source Softw.* **7**, 2991 (2022).
62. Wick, R. R. et al. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol.* **22**, 266 (2021).
63. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
64. Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
65. Vaser, R. & Šikić, M. Time- and memory-efficient genome assembly with Raven. *Nat. Comput. Sci.* **1**, 332–336 (2021).
66. Wick, R. R. & Holt, K. E. Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLoS Comput. Biol.* **18**, e1009802 (2022).
67. Zimin, A. V. et al. The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
68. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
69. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
70. Wishart, D. S. et al. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.* **51**, W443–W450 (2023).
71. Bertelli, C. et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **45**, W30–W35 (2017).
72. Kaya, H. et al. SCC *mec* finder, a web-based tool for typing of staphylococcal cassette chromosome *mec* in *Staphylococcus aureus* using whole-genome sequence data. *mSphere* **3**, e00612–17 (2018).
73. Wickham, H. et al. Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
74. Wickham, H. *Ggplot2* (Springer, 2016).
75. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
76. Grazzietti, A. L., Koonin, E. V. & Kristensen, D. M. Prokaryotic virus orthologous groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.* **45**, D491–D498 (2017).
77. Terzian, P. et al. PHROG: families of prokaryotic virus proteins clustered using remote homology. *NAR Genomics Bioinform.* **3**, lqab067 (2021).
78. Marçais, G. et al. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **14**, e1005944 (2018).

79. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w 1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
80. Faust, G. G. & Hall, I. M. SAMBLASTER : fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503–2505 (2014).
81. Van Der Auwera, G. A. et al. From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 1 (2013).
82. Croucher, N. J. et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15–e15 (2015).
83. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genomics* **2**, e000056 (2016).
84. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
85. Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
86. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
87. Nishimura, Y. et al. ViPTree: the viral proteomic tree server. *Bioinformatics* **33**, 2379–2380 (2017).
88. Yu, G. *Data Integration, Manipulation and Visualization of Phylogenetic Trees* (Chapman and Hall/CRC, 2022).
89. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
90. Xu, S. et al. Ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. *iMeta* **1**, e56 (2022).
91. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
92. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
93. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
94. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
95. Törönen, P., Medlar, A. & Holm, L. PANNZER2: a rapid functional annotation web server. *Nucleic Acids Res.* **46**, W84–W88 (2018).
96. Alexa, Adrian. *JR topGO* (Bioconductor, 2017).
97. Geis-Asteggiane, L., Ostrand-Rosenberg, S., Fenselau, C. & Edwards, N. J. Evaluation of spectral counting for relative quantitation of proteoforms in top-down proteomics. *Anal. Chem.* **88**(22), 10900–10907 (2016).
98. Dowle, A. A., Wilson, J. & Thomas, J. R. Comparing the diagnostic classification accuracy of iTRAQ, peak-area, spectral-counting, and Empai methods for relative quantification in expression proteomics. *J. Proteome Res.* **15**(10), 3550–3562 (2016).
99. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**(7), e47 (2015).

Acknowledgements

This work was conducted using the core facilities of the Lopukhin FRCC PCM “Genomics, Proteomics, Metabolomics” (<http://rcpcm.org/?p=2806>).

Author contributions

M.K., E.S. conceived and designed the experiments; M.K., N.A., R.G., D. Boldyreva., V.G., O.S., V.V., O.P., I.S., K.K. performed the experiments; M.K., D.Bespiatykh., E.S., A.G. analyzed the data; M.K., D.Bespiatykh, E.S. wrote the main manuscript text. All authors reviewed the manuscript.

Funding

This research was funded by the Russian Science Foundation (research project no. 22–15- 00443, <https://rscf.ru/en/project/22-15-00443/>).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-80909-x>.

Correspondence and requests for materials should be addressed to M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024