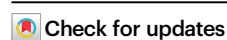


RNase-mediated defence and its association with *SCCmec* in *Staphylococcus aureus*

Received: 2 October 2025

Accepted: 2 July 2026

Published online: 20 August 2026



Kohei Kondo¹, Aa Haeruman Azam², Junzo Hisatsune¹, Shinjiro Ojima², Kotaro Chihara², Tomohiro Nakamura², Azumi Tamura², Wakana Yamashita², Satoshi Nakano¹, Shoko Kutsuno¹, Yo Sugawara¹, Jumpei Uchiyama³, Masayuki Amagai^{4,5}, Tetsuo Yamaguchi⁶, Yoshikazu Ishii⁶, Koichi Watashi^{1,2}, Yoshimasa Takahashi⁷, Kotaro Kiga² & Motoyuki Sugai¹✉

The rise of antibiotic resistance in *Staphylococcus aureus* is a major threat to global health. Resistance often emerges when bacteria acquire foreign DNA elements, such as the *SCCmec* cassette that makes strains resistant to nearly all β -lactam antibiotics. Yet, not all *S. aureus* lineages take up these elements equally, raising the question of what prevents some strains from becoming drug-resistant. Here we identify a previously unknown defence factor that explains this mystery in a common skin lineage (ST188), frequently isolated from patients with atopic dermatitis. Such defence factor is encoded by a gene, which we name *sadR*, that is present in methicillin-susceptible ST188 strains but absent from resistant counterparts. SadR exhibits RNase activity that provides anti-bacteriophage and anti-plasmid defence, thus acting as a barrier to the uptake of mobile genetic elements. Gene *sadR* is found in methicillin-susceptible strains in *S. aureus* lineages other than ST188, and is consistently located adjacent to *rlmH*, the canonical *SCCmec* integration site. Moreover, SadR homologues are widespread across staphylococci and other bacteria suggesting a previously unrecognised family of RNase-based defence systems.

Staphylococcus aureus is a versatile and opportunistic pathogen capable of causing a wide range of infectious diseases, ranging from skin and soft tissue infections to life-threatening systemic infections. The ability of *S. aureus* to cause such diverse diseases is attributed to its genomic diversity. The genome structure of *S. aureus* is characterised by a conserved core genome, into which various mobile genetic elements (MGEs) carrying exogenous genes, such as bacteriophages and pathogenicity (resistance) islands, are incorporated or acquired as plasmids^{1–3}. This results in the formation of various populations with unique virulence and antimicrobial resistance traits⁴. A notable example is methicillin-resistant *S. aureus* (MRSA), which emerges from

methicillin-susceptible *S. aureus* (MSSA) that has acquired the *SCCmec* element containing the *mecA* gene—whose product PBP2' exhibits decreased penicillin-binding ability—becomes resistant to virtually all β -lactams^{5,6}. The number of deaths caused by MRSA increased from 1990 to 2021^{7,8}, posing a public health concern. MRSA continues to be classified as a high-priority pathogen owing to its substantial estimated burden.

Although MGEs play crucial roles in the plasticity and pathogenicity of *S. aureus*, the mechanisms by which various *S. aureus* lineages accept or block MGEs are not fully understood. A thorough examination of *S. aureus* strains revealed that the resistance

¹Antimicrobial Resistance Research Center, National Institute of Infectious Diseases, Japan Institute for Health Security, Tokyo, Japan. ²Department of Drug Development, National Institute of Infectious Diseases, Japan Institute for Health Security, Tokyo, Japan. ³Department of Bacteriology, Graduate School of Medicine Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan. ⁴Department of Dermatology, Keio University School of Medicine, Tokyo, Japan. ⁵Laboratory for Skin Homeostasis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. ⁶Department of Microbiology and Infectious Diseases, Toho University School of Medicine, Tokyo, Japan. ⁷Research Center for Vaccine Development, National Institute of Infectious Diseases, Japan Institute for Health Security, Tokyo, Japan. ✉e-mail: sugai.m@jihs.go.jp

phenotype does not appear in all lineages at the same rate. Our comprehensive genomic surveillance of *S. aureus* isolates from lesions of patients with atopic dermatitis (AD) in Japan demonstrated that sequence type (ST)188 was one of the most frequently detected lineages, showing a strong correlation with AD⁹. The ST188 lineage exhibited a notable absence of the *mecA* gene, enterotoxin genes, and a paucity of other antimicrobial resistance genes (ARGs), which are thought to be acquired via MGEs. However, it remains unclear why ST188 strains maintain MSSA lineages. Recently, we discovered a new genome registry of MRSA-ST188 isolated in Vietnam in a public database¹⁰. MRSA-ST188 carries plasmids encoding multiple ARGs as well as prophages and staphylococcal pathogenicity islands (SaPIs). These strains offer a natural approach for dissecting the molecular barrier that shields ST188 from acquiring mobile genetic elements.

Here, we conducted a genomic comparison of MSSA-ST188 and MRSA-ST188, along with a search for the defence mechanisms of ST188 *S. aureus*. We identified an abortive infection C-terminal domain (Abi_C)-containing RNase (SadR, *S. aureus* defensive RNase) with an unknown function. Notably, this protein was present in all MSSA-ST188 isolates but not in MRSA-ST188 isolates. SadR blocked various phage infections and reduced plasmid transformation efficiency, suggesting its role in preventing the horizontal transfer of MGEs. This study highlights that SadR erects a lineage-specific barrier that limits MGEs and may help to preserve the methicillin-susceptible state of ST188 and related lineages.

Results

Comparative genomic analysis of MSSA- and MRSA-ST188

To identify candidate proteins responsible for defending against MGEs in *S. aureus* ST188, we conducted a comparative genomic analysis between MSSA-ST188 and MRSA-ST188.

First, we selected predicted and characterised protein from open reading frames exclusively present in MSSA-ST188 but not in MRSA-ST188 (Supplementary Data 1). We then characterised the domains of the candidate proteins (Fig. 1a). Among the candidate defence system proteins identified, an abortive infection C-terminal domain-containing protein was present in all MSSA-ST188 strains (Fig. 1a, and Supplementary Data 1) and was designated as SadR. In contrast, none of the MRSA-ST188 strains, including those from Vietnam and pig farms, harboured SadR. In this context, SadR is a candidate protein that restricts the acquisition of diverse MGEs, such as phages and plasmids, which can carry clinically important antimicrobial resistance and virulence genes.

We next constructed a maximum-likelihood phylogenetic tree based on core-genome alignments (core threshold $\geq 95\%$) of 193 ST188 isolates, including 106 from Japan and 87 from other countries (Fig. 1b; Supplementary Data 2). All 148 SadR-positive ST188 strains (black circles) were MSSA, with the exception of two strains (blue squares in Fig. 1b). All SadR-negative ST188 strains were MRSA. In addition, we quantified the number of insertion sequences (ISs) and ARGs between SadR-positive (SadR) and -negative ST188 strains (no-SadR). SadR-positive strains carried on average 0.8 ± 1.2 (CI: 0.6–1.0) ARGs per genome, whereas SadR-negative strains (no-SadR) carried 5.1 ± 1.1 (CI: 4.7–5.4) ARGs per genome (Fig. 1c). Similarly, SadR-positive strains harboured 30.5 ± 4.1 ISs per genome (CI: 29.8–31.2), compared with 43.4 ± 3.4 ISs per genome in SadR-negative strains (Fig. 1c; CI: 42.4–44.4). These results indicate that SadR-positive strains carry fewer ARGs and ISs per genome than SadR-negative strains. We further characterised the genetic context of the *sadR* gene in MSSA ST188 strains isolated from several countries (Fig. 1d). Notably, *sadR* was located downstream of *rlmH*, whose 3' terminus is the canonical integration site for the *SCCmec* element. In contrast, in MRSA-ST188 strains, the *SCCmec* element was located at the 3' end of *rlmH*, and the *sadR* gene together with its putative promoter region was absent,

whereas the downstream genes at this locus retained the same organisation as in MSSA-ST188 (Fig. 1d).

SadR confers resistance to phages and reduces plasmid acquisition

Next, we assessed the defence activity of SadR against MGEs such as phages and plasmids. In spot assays, expression of SadR in RN4220 (RN4220 pLI50/*sadR*) reduced plaque formation of multiple phages, including 80 α , ϕ 55 and S24-1, by approximately 10–10,000-fold (Fig. 2a). We also assessed a strain carrying chromosomally integrated *sadR* (RN4220/*sadR*(chr)) to evaluate defence activity from a single-copy locus. RN4220/*sadR*(chr) showed resistance to 80 α and ϕ 55, but not to S24-1 (Fig. 2a). Overall, SadR conferred defence against multiple phages at both plasmid-borne and single-copy chromosomal expression levels (Fig. 2a).

We further assessed SadR's defence activity by inducing prophages in 197 *S. aureus* strains isolated from various clinical strains of diverse MLST (Supplementary Data 3). We found that SadR blocked several clinically derived temperate phages. Both RN4220 pLI50/*sadR* and RN4220/*sadR*(chr) reduced the efficiency of plating (EOP) against temperate phages isolated from several STs (ST88, ST5, and ST45) by 10–1000-fold. In contrast, SadR did not reduce the EOP of temperate phages isolated from ST20, ST380, and ST1516 (Fig. 2b). Our data indicated that SadR can partially block several temperate phages from various MLST strains that were isolated from clinical settings.

To identify phage determinants associated with escape from SadR-mediated defence, we isolated an 80 α escape mutant (e80 α) that was less sensitive to SadR-mediated defence. Whereas the parental 80 α phage exhibited an approximately 10,000-fold reduction in EOP on RN4220/*sadR*(chr), e80 α showed only an approximately 10-fold reduction (Supplementary Fig. 1a). Whole-genome sequencing identified a nonsynonymous substitution (D143Y) in a gene encoding a single-strand annealing protein (Supplementary Fig. 1b).

We compared the plasmid transformation efficiency between wild-type RN4220 and RN4220/*sadR*(chr). Using plasmids with different origins of replication (pLI50, pLC1, and pKOR), the normalised transformation efficiency of WT RN4220 was set to 100%, and RN4220/*sadR*(chr) (SadR+) showed reduced transformation efficiencies for all three plasmids ($3.4 \pm 3.8\%$ for pLI50, $4.9 \pm 7.0\%$ for pLC1, and $3.2 \pm 7.1\%$ for pKOR). Overall, the RN4220/*sadR*(chr) strain retained ~3–5% of the transformation efficiency of RN4220 (Fig. 2c). We attempted to demonstrate *SCCmec* exclusion by introducing the *SCCmec* element via natural transformation; however, we were unable to reproduce this natural transformation in *S. aureus*.

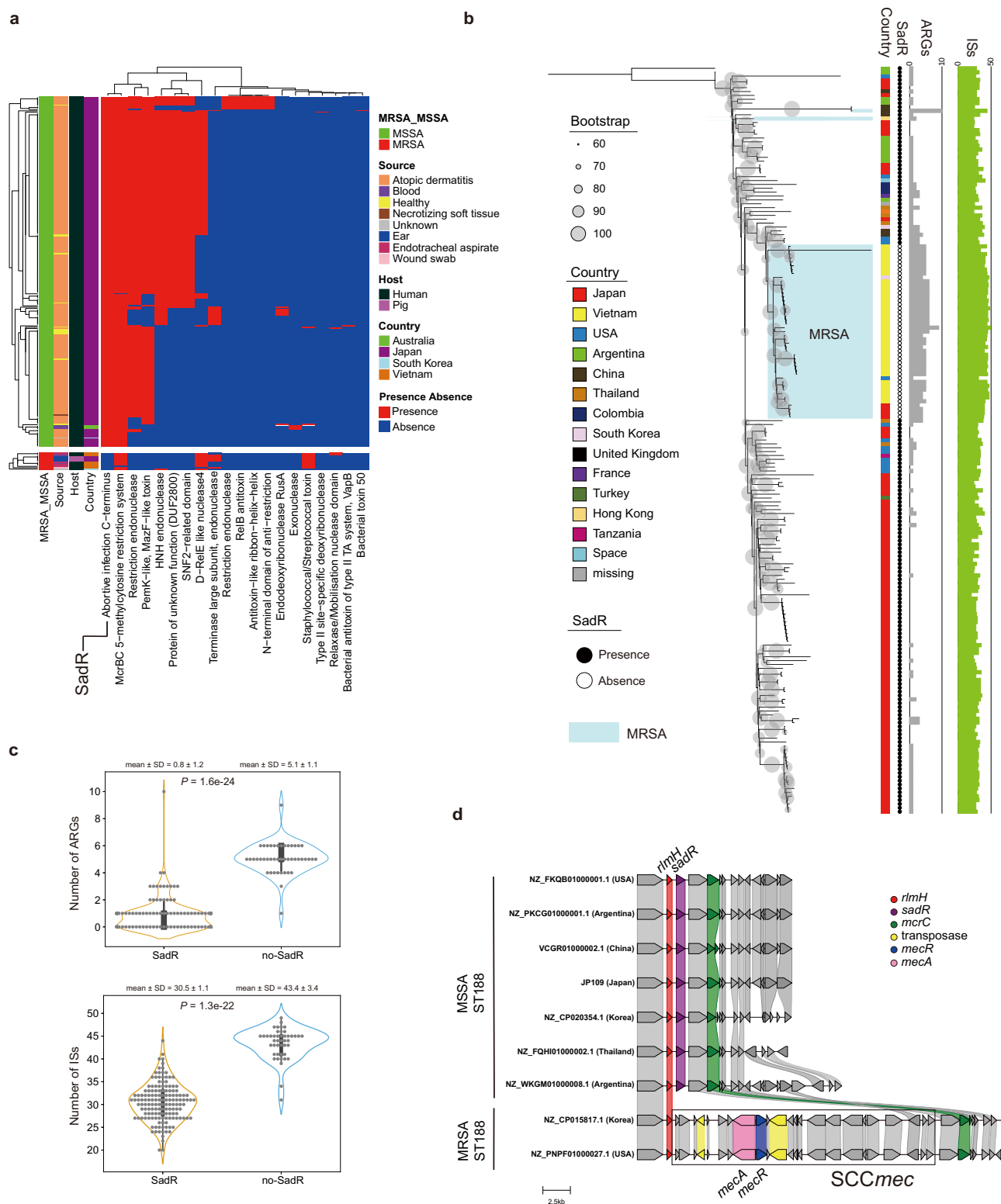
Overall, our results demonstrate that SadR provides defence against various phages and restricts plasmid transformation.

SadR exhibits RNase activity

Structural prediction suggested that SadR adopts a helix-turn-helix fold (Supplementary Fig. 2). Structural comparison identified that SadR shares structural similarities with DNA helicase and/or CRISPR-Cas systems, which have been observed to exhibit DNase and/or RNase activity and predicted a Higher Eukaryotes and Prokaryotes Nucleotide-binding (HEPN) domain in the C-terminus of SadR (Supplementary Data 4), implying its potential as a nucleotide-binding protein.

Our in vitro study demonstrated that SadR degraded both ssRNA and total RNA in the presence of Mg²⁺ ions (Fig. 3a) but did not degrade dsDNA, ssDNA, or dsRNA (Fig. 3b). The SadR protein encoded by MSSA-ST188 is 30.4 kDa (266 aa). Our Native-PAGE analysis revealed that the purified SadR band appeared at approximately 66 kDa (Fig. 3c).

We searched the HEPN-type RNase motif (RΦX[3–5]H), where “Φ” represents polar amino acids, such as N, H, T, or Q, and “X” represents



any 3–5 amino acid residues, with histidine being highly conserved. We found the presence of the HEPN-RNase motif “RNEYGDAH” within the *Abi_C*-terminal domain region of SadR. The three mutant SadR proteins (R225A, N226A, and H232A) demonstrated a loss of the capacity to degrade single-stranded RNA and total RNA (Fig. 3d). Overall, SadR possesses an RNase motif similar to that of other HEPN-type RNases. Furthermore, we found that the three-SadR mutants lost their defence activity (Fig. 3e). Overall, these findings confirmed that the SadR HEPN-

type RNase motif is essential for both the RNase and phage defence activities.

SadR is distributed in the genus *Staphylococcus*

To explore the evolutionary distribution and genomic context of SadR, we analysed its prevalence across the *Staphylococcaceae* family and its association with *SCCmec* elements and lineage types. We first analysed the distribution of SadR homologues in *Staphylococcaceae*. We then

Fig. 1 | Discovery of the SadR protein. **a** Identification of candidate defence systems in Methicillin-susceptible *Staphylococcus aureus* (MSSA)-ST188 ($n = 275$) isolates and comparator Methicillin-resistant *Staphylococcus aureus* (MRSA)-ST188 isolates ($n = 12$) from a Vietnamese ICU and ($n = 8$) from pigs in Japan. Defence-associated protein domains were detected using HMMER searches against the Pfam-A v35 database. A clustered heatmap shows the presence (red) or absence (blue) of candidate defence determinants across isolates, with associated metadata (isolation source, host, and country) indicated alongside. **b** Phylogenetic analysis of 193 ST188 core genomes. Circle size on internal branches indicates bootstrap support. Blue squares denote MRSA isolates. In the SadR track, black-filled circles indicate strains harbouring SadR, whereas open circles indicate strains lacking SadR. Bars show the number of antimicrobial resistance genes (ARGs) and insertion sequences (ISs) for each strain. **c** The numbers of ARGs and ISs were compared between SadR-positive ($n = 148$) and SadR-negative ($n = 45$) *S. aureus* strains shown in Fig. 1b. Here, n denotes the number of strains analysed in each group. The y axis indicates the number of ARGs or ISs in each *S. aureus* strain. SadR-positive and

SadR-negative strains correspond to the black-filled and open circles in Fig. 1b, respectively. Statistical analysis was performed using the two-sided Mann–Whitney U test to compare ARG and IS counts between SadR-positive and SadR-negative strains. Violin plots show the distribution of the data, with individual points representing each data point. Box plots embedded within the violin plots summarise the data: boxes indicate the interquartile range, with the lower and upper bounds representing the 25th and 75th percentiles, respectively. Whiskers extend to the minimum and maximum values. **d** Comparison of the genomic organisation of the *sadR* locus among nine globally sourced ST188 isolates. Seven MSSA-ST188 isolates from different countries and two MRSA-ST188 isolates from Korea (NZ_CP015817.1) and the USA (NZ_PNPF01000027.1) were included. Key genes (*sadR*, *rlmH*, *mecA*, and *mecR*) are shown as purple, red, magenta, and blue arrows, respectively. Syntenic regions flanking *sadR* are highlighted to illustrate conserved gene order and structural variation between MSSA and MRSA isolates. Source data are provided as a Source Data file.

examined 27,655 genomes from the *Staphylococcaceae* family and other related genera, including *Abyssiococcus*, *Aliicoccus*, *Cytobacillus*, *Gemella*, *Jeotgalicoccus*, *Macroccoccus*, *Nosocomiicoccus*, *Salinicoccus*, *Pseudobacillus*, and *Bacillus*. We detected SadR homologues in approximately 3.5% (961/27,655) of genomes. Although *S. aureus* was the species most frequently found to harbour SadR, homologues were also identified in other species, including more than 15 *Staphylococcus* species (eg., *S. epidermidis* and *S. argenteus*) (Supplementary Data 5 and Supplementary Fig. 3). Based on this analysis, we focused on SadR-positive *Staphylococcaceae* genomes to determine whether the presence of SadR is associated with reduced acquisition of mobile genetic elements, including ARGs and SCCmec—and whether the genomic synteny surrounding *sadR* is conserved or variable across clades.

Phylogenetic analysis of SadR amino acid sequences revealed that the homologues clustered into seven distinct clades (A–G) (Fig. 4a). Clade A was strongly associated with the ST188 lineage but was also detected in several other STs. Within each ST, SadR prevalence was calculated as the number of SadR-carrying strains divided by the total number of strains in that ST. Clade A SadR was detected in ST188 (135/139, 97.1%), ST25 (64/65, 98.5%), and ST3926 (28/29, 96.6%), as well as in several additional STs (Supplementary Data 6). Among clade A-positive genomes, 261/274 (95.3%) were MSSA (Fig. 4b). SadR homologues in clades E and F were also enriched in MSSA backgrounds (among clade-positive genomes: clade E, 31/38 (81.6%) MSSA; clade F, 27/27 (100%) MSSA) (Fig. 4b). Within-ST prevalence of clade E included ST692 (4/14, 28.6%), *S. auricularis* (5/15, 33.3%), and *S. cohnii* (3/80, 3.8%), with additional low-frequency occurrences in other STs (Supplementary Data 6). Clade F was detected in ST479 (8/9, 88.9%), ST1651 (5/5, 100%), with additional low-frequency occurrences in other STs (Supplementary Data 6).

In contrast, clades B–D were predominantly found in MRSA backgrounds (among clade-positive genomes: clade B, 87/87 (100%) MRSA; clade C, 223/225 (99.1%) MRSA; clade D, 292/306 (95.4%) MRSA) (Fig. 4b, c). The associated ST backgrounds were diverse. For example, within-ST prevalence of clade B included ST30 (35/333, 10.5%) and *S. argenteus* ST2250 (17/131, 13.0%), with additional low-frequency occurrences in other STs (Fig. 4c). Clade C was detected in ST22 (43/1,486, 2.9%) and ST125 (42/44, 95.5%), with additional low-frequency occurrences in other STs (Fig. 4c). Clade D was detected mainly in ST8 (80/2434, 3.3%), ST5 (76/2,927, 2.6%), and ST80 (39/52, 75.0%) (Fig. 4c, Supplementary Data 6). Clade B SadR almost always co-occurred with SCCmec IV (Supplementary Data 5). SCCmec IV was also frequent in other SadR-positive MRSA lineages, occurring in 71.1% of clade C and 91.8% of clade D genomes.

As confirmed by functional assays, clade B SadR demonstrated the ability to defend against phage 80 α (Supplementary Fig. 4a). Quantitative PCR revealed that transcription levels of clade A and B *sadR* were similar in the presence or absence of phage 80 α , indicating that their

expression is independent of phage induction under the tested conditions (Supplementary Fig. 4b).

Complete genome sequence data of the ST188 strain JP109 indicated that the SadR-coding gene was located just downstream of *rlmH* (formerly annotated as *orfX*), through which SCCmec was inserted to convert MSSA to MRSA (Fig. 1d). We therefore analysed the 30-kb upstream and downstream sequences of *sadR* genes in representative strains and complete *S. aureus* genomes available in the NCBI database. Our analysis provided a clear view of the genomic locations of *sadR* gene homologues, as shown in Fig. 4d. *sadR* genes belonging to clades A, E, and F were located just downstream of the *rlmH* gene, while *sadR* genes in clades B, C, and D were downstream of the *rlmH* gene and at the 3' end of the SCCmec element (Fig. 4d). To investigate whether *sadR* is located within or outside the SCCmec element, we searched for SCCmec-specific direct repeats (DR; AGAAGCATATCATAA) and inverted repeats (IR; CATATCATAAATGATGCCGTTTTT). The *sadR* gene was located within the SCCmec boundaries defined by the DR and IR sequences (Fig. 4d), indicating that *sadR* was co-transferred with SCCmec upon chromosomal integration in MRSA lineages. Overall, SadR clades A/E/F were predominantly found in MSSA strains and were located downstream of *rlmH*, whereas SadR clades B/C/D were enriched in MRSA strains and were located within SCCmec.

Given the distinct genomic contexts of SadR, we quantified ARG and IS counts as proxies for overall MGE content across strains carrying different SadR clades. We compared the number of MGEs, including ARGs and ISs, across the strains encoding different SadR clades. Average ARG counts differed among strains harbouring each SadR clades: A, 0.8 ± 1.0 ; B, 2.7 ± 1.0 ; C, 4.4 ± 1.9 ; D, 3.7 ± 1.9 ; E, 2.3 ± 3.6 ; F, 0.5 ± 1.3 ; and G, 4.5 ± 3.3 (Supplementary Fig. 5a), indicating that strains carrying SadR clades A/E/F tend to harbour fewer ARGs than strains carrying other clades. IS counts showed a similar pattern (Supplementary Fig. 5b): A, 28.0 ± 6.0 ; B, 31.0 ± 5.5 ; C, 36.5 ± 12.3 ; D, 36.6 ± 11.5 ; E, 32.9 ± 11.9 ; F, 31.3 ± 16.0 ; and G, 40.3 ± 6.9 . Furthermore, when clades A, E, and F were considered together, genomes carrying these clades harboured 0.9 ± 1.5 (CI: 0.76–1.09) ARGs, whereas genomes lacking clades A/E/F harboured 4.1 ± 2.7 (CI: 4.0–4.1) ARGs, corresponding to an approximately 4.6-fold difference (Supplementary Fig. 5c). Taken together, our results indicate that strains carrying SadR clades A/E/F tend to have a lower MGE burden than strains carrying other SadR clades or lacking SadR.

Abi_C domain-containing proteins are widespread in bacterial classes and possess a conserved HEPN motif in the C-terminal domain

To investigate the conservation and functional characteristics of Abi_C domain-containing proteins, we analysed their distribution and assessed the presence of additional domains (Supplementary Data 7). A phylogenetic analysis of 2,815 Abi_C domain-containing proteins,

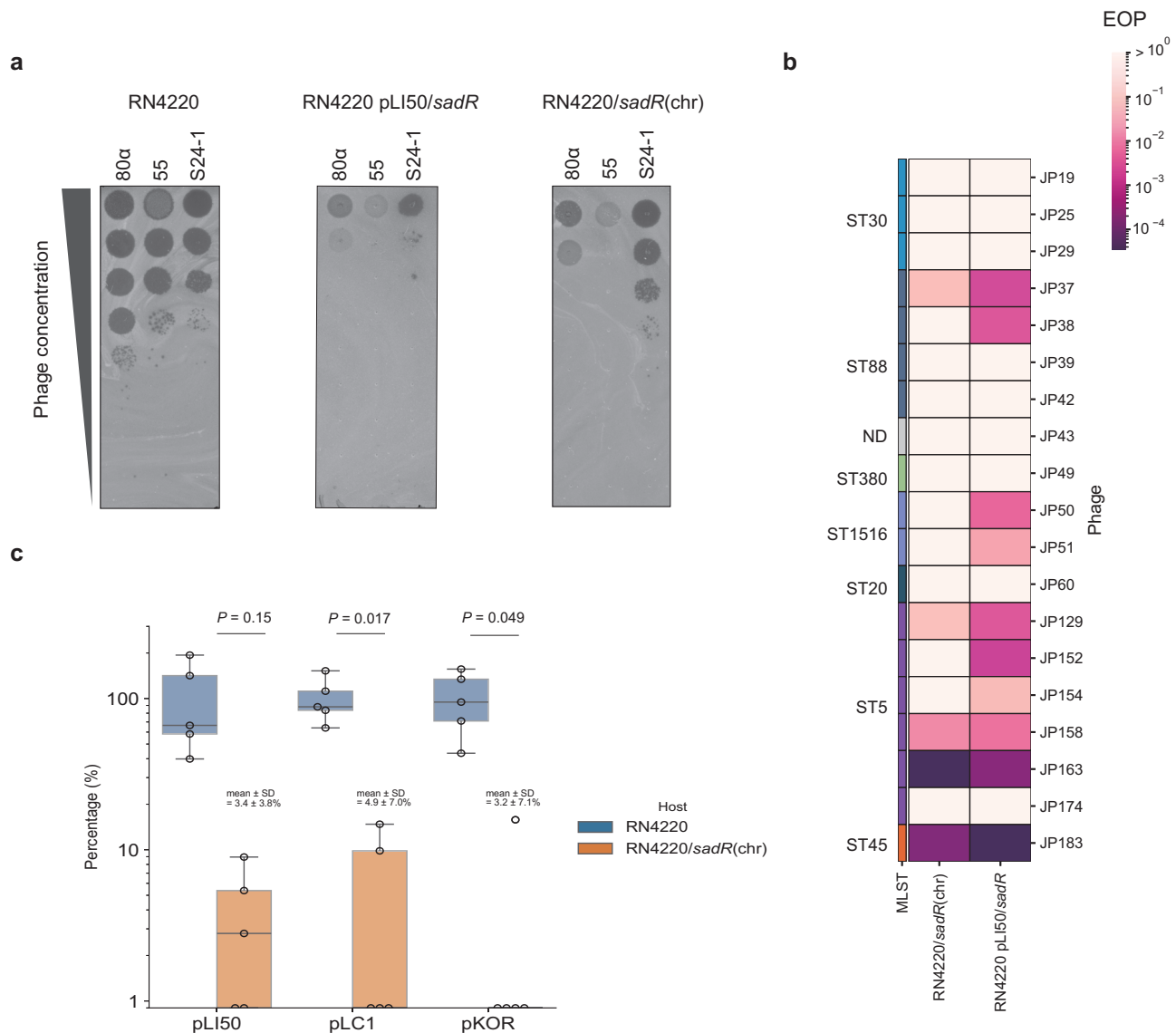


Fig. 2 | Phage- and plasmid-blocking ability of SadR. **a** Spotting assay for several types of phages (80α, φ55, and S24-1) against RN4220, RN4220 pLI50/sadR, and RN4220/sadR(chr). A 10-fold dilution series of phages was used in the spotting assay. **b** Defence activity against temperate phage preparations induced from 197 clinical *Staphylococcus aureus* isolates in Japan. Temperate phage lysates were induced with mitomycin C from each isolate. JP identifiers (e.g., JP19) denote the prophage lysate induced from the corresponding isolate (Japan clone 19). Colours indicate the MLST (sequence type) of the *S. aureus* isolate from which each temperate phage was obtained. Efficiency of plating (EOP) on RN4220/sadR(chr) was calculated as the number of plaques formed on RN4220/sadR(chr) divided by the number of plaques formed on the parental RN4220 strain. EOP on RN4220 pLI50/sadR was calculated as the number of plaques formed on RN4220 pLI50/sadR divided by the number of plaques formed on the control strain (RN4220 pLI50). Experiments shown in a and b were

performed once. **c** Plasmid transformation assay to assess SadR-mediated defence using *E. coli*-*S. aureus* shuttle vectors (pLI50, pKOR, and pLC1). Transformation efficiency of RN4220/sadR(chr) was normalised to the mean transformation efficiency of RN4220 measured within the same experimental run (RN4220 = 100%), and RN4220/sadR(chr) was compared with this normalised control. The experiment was performed independently five times. Violin plots show the distribution of the data, with individual points representing each data point. Box plots embedded within the violin plots summarise the data: boxes indicate the interquartile range, with the lower and upper bounds representing the 25th and 75th percentiles, respectively. Whiskers extend to the minimum and maximum values. *P* values were calculated using a two-way ANOVA (host: $P = 9.6 \times 10^{-8}$; plasmid: $P = 0.92$; host $\times$ plasmid interaction: $P = 0.99$), followed by the Games-Howell post hoc test. Source data are provided as a Source Data file.

performed using an approximately maximum-likelihood approach, showed that these proteins are broadly distributed across diverse bacterial classes, including Bacilli, Gammaproteobacteria, Alphaproteobacteria, Betaproteobacteria, and Actinomycetes, as well as several additional classes (Fig. 5a). Abi_C domain-containing proteins were classified into five distinct types based on their N-terminal domains (NTDs) (Fig. 5b). Proteins with no NTD or AbiJ-NTD3 domain-containing proteins were broadly distributed across multiple bacterial classes, and Sdh5 domain-containing proteins were exclusively found in Bacilli.

These results suggest that the presence of specific NTDs reflects bacterial-specific adaptation or functions.

To classify SadR proteins, we used two widely used tools for phage defence system detection: Defence-finder and PADLOC. Defence-finder classified 1,593 of the 2,815 Abi_C proteins (56.6%) as members of the AbiJ family, leaving the remaining unassigned, whereas PADLOC did not match any Abi_C sequence to a known defence system. Thus, although more than half of the Abi_C repertoire was sufficiently similar to be labelled as AbiJ, a substantial fraction still fell outside current

classifiers. This suggests that the Abi_C domain-containing protein and AbiJ represent either a single but highly divergent defence family or several closely related systems (Fig. 5a). The function of the Sdh5 domain (succinate dehydrogenase assembly factor 2) in phage defence systems is not yet fully understood.

To further explore the functional characteristics of SadR-containing proteins, we examined whether they possessed the HEPN-type RNase motif. Our analysis of 120 amino acid residues within the Abi_C-terminal domain region led to the identification of a conserved HEPN RNase motif. A highly conserved RΦXH RNase motif, in which X mainly consists of three to five variable residues, was detected across all bacterial classes (Fig. 5c). We analysed the length of the “X” region within the RNase motif and found that Abi_C domain-containing proteins exhibited a distribution spanning 2–6 residues, with five residues being the most frequent across all bacterial classes examined (Fig. 5d).

Discussion

Based on our in vitro evidence that SadR restricts phage infection and plasmid acquisition, together with comparative genomic patterns across lineages, we suggest that SadR in clades A, E, and F contributes to reduced permissiveness to MGEs. Although these in vitro assays capture only a subset of the selective pressures operating in natural settings, the functional data, together with the observation that several MSSA lineages retain conserved SadR homologues, are consistent with a role for SadR in constraining the acquisition of MGEs that carry ARGs and virulence genes, and may also influence SCCmec acquisition. However, the presence of SadR within SCCmec in some strains indicates that it does not universally prevent cassette acquisition, highlighting that its effect may be context-dependent. We found that *Macrococcus hajekii* carries a SadR homologue located immediately downstream of the *rlmH* gene but does not carry SCCmec-related elements (Supplementary Fig. 6). *Macrococcus* belongs to the family *Staphylococcaceae*, and *Staphylococcus* and *Macrococcus* diverged from a common ancestor¹¹. This suggests that strains harbouring SadR protein in clades A, E, and F may have originated from a common ancestor carrying the *sadR* gene just downstream of the *rlmH* gene. Conversely, MRSA lineages, in which *sadR* is typically located within most type IVc SCCmec elements, may have acquired *sadR* more recently via a similar SCCmec element and gained the defence activity of SadR following an earlier phase of increased MGE permissiveness in the absence of SadR. It has been reported that the SCCmec region accumulates various defence systems, and notably, AbiJ, which is a homologue of SadR, is the most prevalent defence system after Type-I RM¹². However, the evolutionary mechanism of such packaging of antimicrobial resistance, that is, *mecA* and defence mechanisms in the SCCmec element, remains unclear. The presence of *sadR* immediately downstream of *rlmH* in MSSA, in contrast to its integration within the SCCmec element in MRSA, suggests distinct evolutionary histories for these loci and warrants further investigation. Given our discovery that the *sadR* gene confers defence against both plasmid transformation and phage infection, it likely plays a critical role in shaping horizontal gene transfer dynamics and bacterial fitness. Its incorporation into SCCmec may indicate selective pressure favouring the colocalisation of defence and antimicrobial resistance mechanisms, potentially enhancing the adaptability of MRSA in clinical settings. It remains unclear how SadR maintained non-toxicity under normal growth conditions while restricting incoming MGEs. The identification of an 80α escape mutant (Supplementary Fig. 1) carrying an amino-acid substitution in a single-strand annealing protein raises the possibility that SadR is activated during phage infection through interaction with one or more phage-encoded proteins or through recognition of infection-associated DNA replication or

recombination intermediates. A similar principle may operate during SCCmec entry or integration, where replication intermediates, ssDNA-containing structures, or replication-associated processes could provide molecular cues for SadR activation. Further studies will be needed to elucidate the molecular mechanism by which SadR is activated and contributes to the restriction of diverse incoming MGEs.

Recently, more than 150 novel defence systems against phages and plasmids have been discovered in *E. coli*^{13,14}, *Pseudomonas aeruginosa*¹⁵, and *S. aureus*¹² as well as in comprehensive bacterial metagenomic databases¹⁶, with their mechanisms being increasingly elucidated. Defence systems are likely associated with the acquisition and exclusion of MGEs harbouring ARGs, virulence genes, and other accessory genes. Therefore, the distribution of defence systems is a key factor in understanding increased pathogenicity, niche adaptation, and colonisation in clinical settings. However, the relationship between diverse defence systems and acquisition of virulence traits in clinical isolates remains unexplored¹⁷. Phage defence systems have often been discovered on genomic islands such as prophages and integrons^{18,19}. The distribution of defence systems varies depending on the strain and/or lineage. For ST188, our data suggest that SadR is a key factor in the low acceptance rate of MGEs. It has been reported that other STs, such as ST97 and ST15, exhibit very low MRSA rates and prevalence of pathogenicity genes^{9,20}, but they do not possess the SadR defence system. These observations imply that lineage-specific defence systems play a central role in shaping the antimicrobial and pathogenic potential of *S. aureus*.

The HEPN RNase motif of Abi_C domain-containing proteins predominantly exhibited five unique residues in this region, suggesting that Abi_C domain-containing proteins constitute a novel superfamily within the HEPN-type RNase family. These findings highlight the evolutionary and functional diversity of Abi_C-containing proteins across bacterial phyla. The Abi_C-terminal domain was first identified in the Abi-859 protein encoded on the plasmid pND859, carried by *Lactococcus lactis*, and Abi-859 was named as “AbiJ” in 2005²¹. Given the nonspecific RNA degradation activity of most HEPN-type RNases, Abi_C appears to be a nonspecific RNase (Fig. 3).

A key strength of this study is the use of matched comparisons within the same genomic background across an extensive set of clinical isolates. By comparing closely related MSSA and MRSA strains within the same ST, we minimised confounding by broader lineage structure and identified defence determinants linked to reduced MGE “acceptance”. This within-ST framework is particularly important because many defence systems are themselves MGE-encoded, and analyses that do not control for genomic background can be biased by co-mobilisation.

This study also has limitations. We attempted to directly test whether SadR inhibits SCCmec acquisition using a natural competence-based approach; however, we were unable to reproducibly establish natural competence under our experimental conditions. Moreover, although our comparative genomics and in vitro assays support a role for SadR in restricting MGEs, we cannot fully disentangle its contribution from these contextual drivers, and our in vitro assays do not capture the full complexity of MGE transmission and selection in hospital environments. Finally, whether SadR can confer phage-defence activity across diverse bacterial species remains an important open question from both evolutionary and applied perspectives.

In summary, this study marks the first step toward uncovering the intricate relationships between various defence systems and the acquisition of MGEs that confer pathogenicity and antimicrobial resistance to microbes, and these findings not only deepen our understanding of lineage-specific defence mechanisms, but also provide a foundation for exploring new therapeutic strategies targeting bacterial virulence and antibiotic resistance.

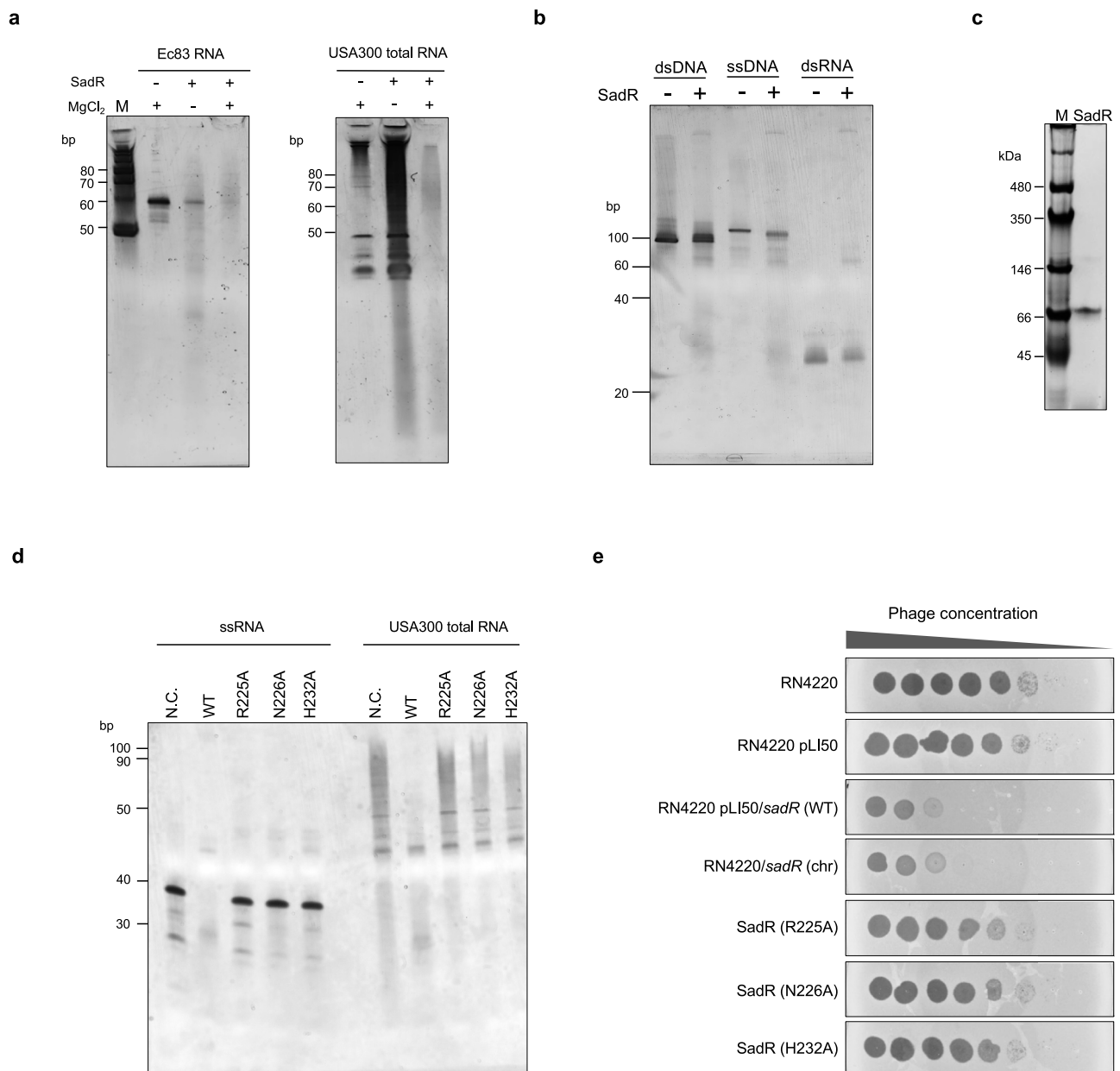


Fig. 3 | SadR RNase activity assay. **a** RNase activity of SadR with or without MgCl₂ in PBS buffer. Retron-Ec83 mRNA and total RNA derived from *S. aureus* USA300 strain were used to assess SadR RNase activity. The lanes were derived from the same gel; intervening lanes were removed, and splice boundaries are indicated by a vertical line. Unprocessed gel image was provided in the Source Data. **b** SadR nucleotide degradation activity for dsDNA, ssDNA, and dsRNA. Purified SadR protein was mixed with each type of nucleotide. **c** Native-PAGE (polyacrylamide gel electrophoresis) analysis of purified SadR; 12.5% acrylamide precast

gel was used in this analysis. The lane labelled “M” indicates the protein marker. **d** Degradation activity of three SadR mutants (R225A, N226A, and H232A) against ssRNA and total RNA. Experiments shown in (a–d) were performed once. **e** Phage plating assay of SadR mutants against phage 80α. RN4220 carrying pLI50 expressing SadR or the indicated mutants was challenged with phage 80α to assess defence activity. Three independent experiments were performed. Source data are provided as a Source Data file.

Methods

Media and buffers

E. coli was grown in Luria–Bertani (LB) broth (BD Difco, Franklin Lakes, NJ, USA) or LB agar (BD Difco). *S. aureus* was grown in tryptic soy broth (TSB) and Tryptic soy agar (TSA) (BD Difco), unless otherwise specified. Phosphate-buffered saline (PBS) (8.0 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄, and 0.24 g KH₂PO₄ per litre) was used to dilute bacterial solutions, and sodium–magnesium (SM) buffer (5.8 g NaCl, 2.0 g MgSO₄·7H₂O, 50 mL 1 M Tris-HCl (pH 7.5), and 5 mL of 2% (w/v) gelatin solution per litre) was used for phage storage and phage dilution. When required, antibiotics were used at the following final concentrations: ampicillin (100 µg/mL), carbenicillin (100 µg/mL),

chloramphenicol (10 µg/mL), and anhydrous tetracycline (1 µg/mL). Unless otherwise specified, antibiotics were purchased from FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan). All primers used in this study were synthesised by Eurofins (SE, Luxembourg, France) and are listed in Supplementary Data 10.

Bacteria and bacteriophages used in this study

Phages 80α and φ55 were kindly provided by José R. Penadés (Imperial College London). Phage S24-1 was previously isolated from sewage samples in Japan²².

Temperate phage lysates referred to here as JP19, JP25, JP29, JP37, JP38, JP39, JP42, JP43, JP49, JP50, JP51, JP60, JP129, JP152, JP154, JP158,

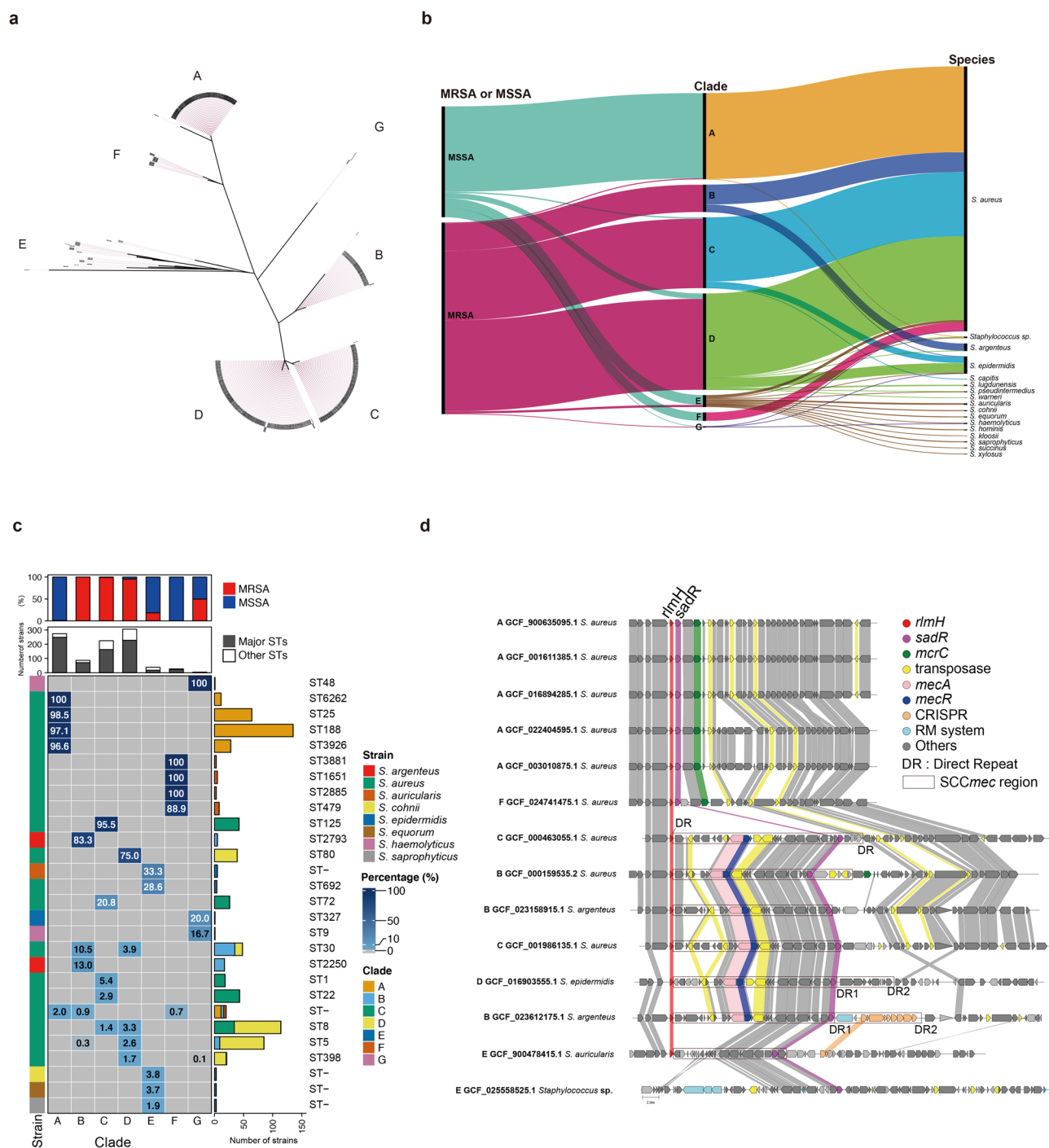
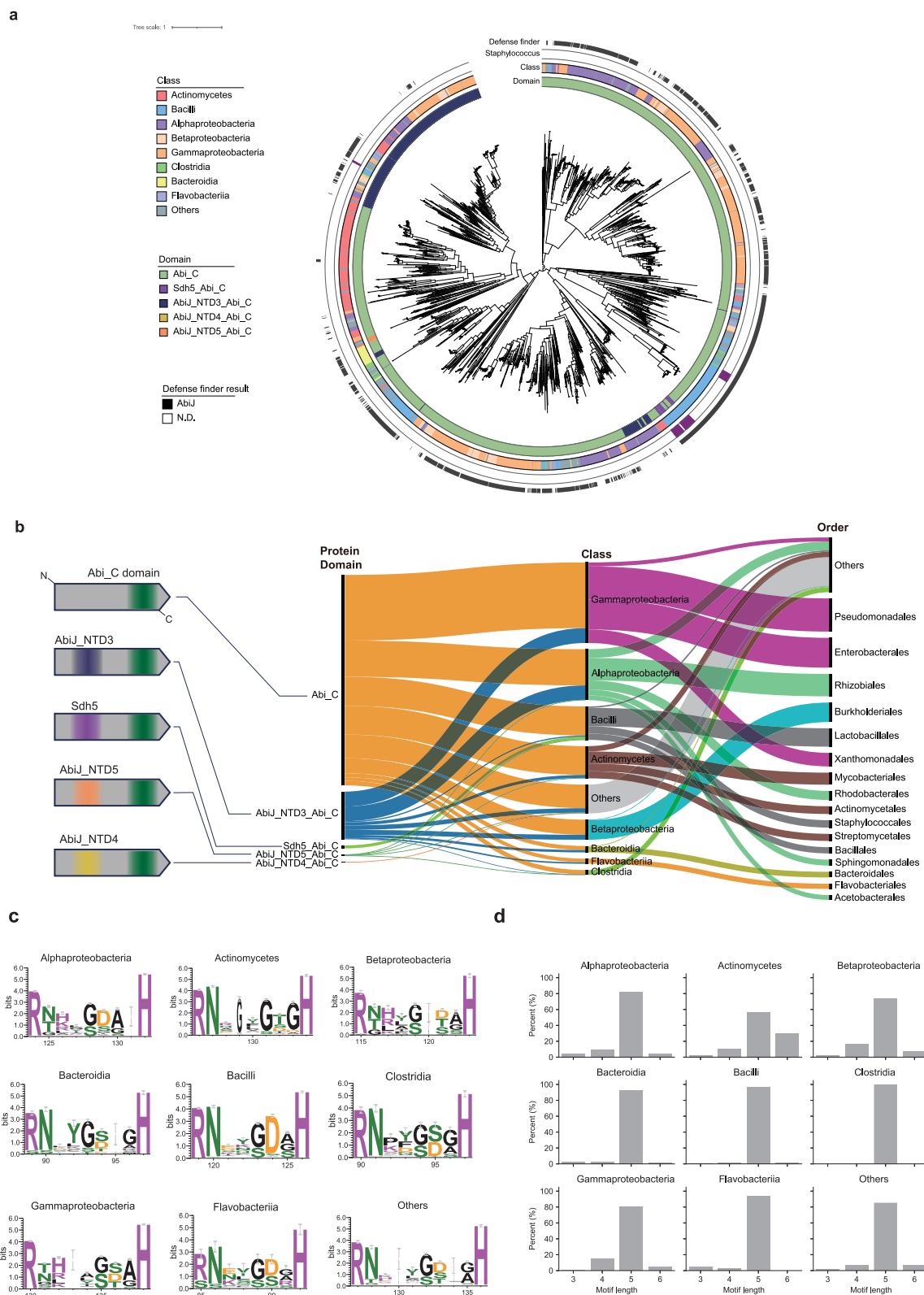


Fig. 4 | SadR homologous proteins in *Staphylococcaceae*. **a** Amino acid phylogenetic tree of 961 SadR homologues identified from 27,655 *Staphylococcaceae* and other related species genomes deposited in the NCBI database. SadR homologues were detected using a local BLASTP search with an identity threshold of >80% and coverage of >80%. SadR proteins were assigned to seven clades (A–G) based on phylogenetic clustering: clade A ($n = 274$), clade B ($n = 87$), clade C ($n = 225$), clade D ($n = 306$), clade E ($n = 38$), clade F ($n = 27$), and clade G ($n = 4$). **b** Sankey diagram depicting the relationship among MRSA/MSSA status, SadR classification, and species. **c** Detailed heatmap of SadR-harboring species and sequence types. The percentages shown in the heatmap represent the proportion of SadR-harboring strains within each sequence type (ST). The top bar chart shows the percentage of MRSA (red) and MSSA (blue) within each clade. The second bar chart shows the

total number of strains in each clade, partitioned into “major STs” (STs displayed in the heatmap) and “minor STs” (all other STs not shown in the heatmap). The right bar chart shows the number of strains in each ST, with colours indicating the contribution of individual clades. The left colour strip indicates the species assignment for each ST. “ST–” indicates strains or lineages for which a sequence type could not be assigned using the MLST scheme applied in this study. All information including other STs was described in Supplementary Data 6. **d** Genomic organisation of the *sadR* locus across representative *Staphylococcus* species. A 30-kb region upstream and downstream of *sadR* was extracted from assemblies in which the locus was contiguous, and representative genomes were visualised using Clinker (<https://github.com/gamcil/clinker>). DR indicates direct-repeat sequences flanking *SCCmec*. Source data are provided as a Source Data file.



JP163, JP174, JP183, and JP193 were isolated from clinical strains by induction with 1 μ M/ml mitomycin C (Nacalai tesque, Kyoto, Japan; Catalogue: 20898-21).

Plasmid cloning and chromosomal integration

The *sadR* gene together with its predicted promoter region was amplified from *S. aureus* strain JP109 (ST188) using KOD FX Neo (TOYOBO,

Osaka, Japan; Catalogue: KFX-201). PCR was performed under the following conditions: an initial denaturation at 95 $^{\circ}$ C for 3 min, followed by 35 cycles of 98 $^{\circ}$ C for 1 min, 55 $^{\circ}$ C (annealing temperature) for 30 s, and 68 $^{\circ}$ C for 3 min. The primers used for amplification were as follows: forward, 5'-ctggcttttataaggaggagattaataactaaaaataatattaaac-3'; reverse, 5'-cattatctcatattactttttgtatgaagataaaagaattcg-3'. Both primers contained 5' overhangs homologous to the pLI50 vector. The amplified *sadR*

Fig. 5 | Distribution of Abi_C-containing proteins. **a** Approximate maximum-likelihood phylogeny of 2,815 Abi_C-containing proteins retrieved by NCBI CD-search. Additional domains detected by CD-search and the bacterial class of each source organism are indicated. Proteins were classified into five types based on their N-terminal domains (NTDs): no NTD ($n = 2,257$), AbiJ-NTD3 ($n = 510$), AbiJ-NTD4 ($n = 3$), AbiJ-NTD5 ($n = 13$), and Sdh5 ($n = 32$). Known defence systems were identified using der; black shading indicates proteins annotated as part of the AbiJ system. **b** Co-localisation of additional domains with Abi_C-containing proteins. The Sankey plot shows the distribution of Abi_C-containing proteins carrying other domains across bacterial classes and orders. The schematic on the left illustrates the domain organisation of representative Abi_C-containing proteins, indicating

the relative positions of the Abi_C domain and co-occurring domains. **c** Conserved predicted HEPN-type RNase motif in the C-terminal region of Abi_C-containing proteins. The HEPN-type RNase motif was defined as RΦXH, where Φ denotes a polar residue and X denotes a variable spacer of ≥ 3 amino acids between Φ and the conserved histidine (H). Sequence logos highlight conserved residues within the predicted motif across major bacterial classes. **d** Length distribution of the variable spacer (X) within predicted HEPN-type RNase motifs detected in Abi_C-containing proteins. For each bacterial class, percentages were calculated as $100 \times$ the number of motif instances with each X-spacer length divided by the total number of predicted RNase motif instances detected in that class. Source data are provided as a Source Data file.

gene was cloned into the *E. coli*-*S. aureus* shuttle vector pLI50 plasmid²³ using the NEBuilder HiFi DNA Assembly Kit (New England Biolabs, Ipswich, MA, USA, Catalogue: E2621S). The promoter region was predicted using Softberry software (<http://www.softberry.com/berly.phtml>). Constructed plasmids were first transformed into *E. coli* using the heat shock method and incubated on LB supplemented with ampicillin (100 µg/mL). Plasmid constructs in *E. coli* DH5α were extracted using the NucleoSpin Plasmid EasyPure kit (Takara Bio, Shiga, Japan; Catalogue: U0727) according to the manufacturer's instructions and subsequently electroporated into *S. aureus* RN4220. The electroporated RN4220 cells were incubated in TSB medium for 2 h at 37 °C, followed by plating onto TSA supplemented with 10 µg/mL chloramphenicol. A *sadR* chromosomal integration transformant was constructed using the temperature-sensitive vector pIMAY as previously described²⁴. Briefly, *sadR* gene and its promoter region from JP109 (ST188) were cloned into the multi-cloning site of pIMAY together with 500 to 600 bp homology arms targeting the region downstream of *rlmH* in RN4220. The homology arms were cloned into pIMAY using NEBuilder. The constructed pIMAY plasmid was transformed into *E. coli*. Thereafter, the plasmid was extracted and transformed into RN4220 via electroporation, followed by incubation on TSA at 37 °C in the presence of chloramphenicol to promote plasmid integration into chromosome. Subsequently, the overnight culture was grown at 28 °C in antibiotic-free media to allow for plasmid excision from the chromosome. Finally, the culture was plated on TSA containing 1 µg/mL anhydrous tetracycline to screen for bacteria that had lost the plasmid. Correct integration and plasmid curing were verified by PCR and confirming the loss of chloramphenicol resistance.

Computational screening of candidate defence proteins

Amino acid sequences from ST188 *S. aureus* genomes from Japan were queried against those from Vietnamese ST188 isolates using local BLASTP v2.12.0 (<https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>)²⁵ with an e-value threshold of $<1e-20$. Sequences with no detectable homologues ("0 hits found") were collected using SeqKit (v.2.10.0) (<https://bioinf.shenwei.me/seqkit/>) with the "faidx" sub-command from JP124 strain. Protein domains in these group exclusive sequences were assigned using HMMER v3.3.2²⁶ against the Pfam-A35 database with an e-value cutoff of 0.1. Proteins possessing domains involved in antiphage defence systems (e.g. abortive infection and nuclease) were manually defined as candidate defence systems in MSSA-ST188. The tertiary structure of SadR was predicted using AlphaFold3 (<https://alphafoldserver.com/welcome>)²⁷. For structural comparison, Foldseek (accessed in January 2023) (<https://search.foldseek.com/search>)²⁸ was used with the ESMFold database. Furthermore, to refine and complement the structural prediction and functional analysis of the protein, the HHPred platform (accessed in January 2024) (<https://toolkit.tuebingen.mpg.de/tools/hhpred>) was employed with several databases implemented in HHPred, including PDB_mmCIF70, SCOPe, Pfam-A, and PHROG as default settings. The defence systems were detected using defence-finder (<https://defencefinder.mdmlab.fr/>)²⁹ v1.2.2 and PADLOC³⁰ v1.1.0 with default settings.

S. aureus isolation and genome sequencing

The *S. aureus* strains used in this study were previously isolated and reported⁹. Pig isolates were obtained from ear swab samples. Briefly, *S. aureus* was lysed using 2 µL of lysostaphin solution (5 mg/mL; FUJIFILM Wako, Osaka, Japan; Cat. No. 126-06613), and a minimum activity of ≥ 30 U per reaction based on the manufacturer's specification, and incubated overnight at 37 °C. *S. aureus* DNA was purified using the Agencourt AMPure XP kit (Beckman Coulter, Brea, CA, USA) according to the manufacturer's protocol. Sequencing libraries were prepared using Enzymatic 5 × WGS reagents (BioStream, Tokyo, Japan) on a Biomek i7 Workstation (Beckman Coulter). Paired-end reads (2 × 150 bp) were generated using an Illumina HiSeq X Five platform (Macrogen, Tokyo, Japan). Raw reads were quality-filtered and assembled de novo with Shovill v1.0.9 (<https://github.com/tseemann/shovill>) using default settings.

Core genome extraction and phylogenetic tree construction

Following assembly, contigs were annotated using Prokka³¹ (v1.14.6) (<https://github.com/tseemann/prokka>) (parameter --kingdom Bacteria) and the core genome was determined using Panaroo³² (v1.3.4) (<https://github.com/gtonkinhill/panaroo>) with standard settings. A maximum-likelihood tree was generated from the core gene alignment file using RAXML-NG²⁹ v1.1 (<https://github.com/amkozlov/raxml-ng>) with a GTR + Γ substitution model. Bootstrap values were calculated using 1000 replicates. The tree was visualised using iTOL (<https://itol.embl.de/>)³³ with midpoint rooting. Metadata, including MLST, SCCmec element, and ARGs, were detected using mlst v2.23.0 (<https://github.com/tseemann/mlst>), SCCmec-staphopia v1.0.0³⁴, and Resfinder³⁵ implemented with Abricate v1.0.1 (<https://github.com/tseemann/abricate>). For ARG detection, thresholds were set at $\geq 80\%$ coverage and $\geq 80\%$ sequence identity (default settings). The SCCmec-specific direct repeat (DR; AGAAGCATATCATATA) and inverted repeat (IR; CATATCATAAATGATGCGGTTTTT) sequences were manually searched within the SCCmec region, as previously reported^{36–38}.

S. aureus competent cells and transformation

S. aureus electrocompetent cells were prepared as described in Sato et al.³⁹, with slight modifications. Each RN4220 and RN4220/*sadR*(chr) *S. aureus* colony was inoculated into B2low medium (per litre: 25 g yeast extract, 10 g tryptone, 10 g NaCl, 0.5% (w/v) glucose, and 1 g K₂HPO₄; pH 7.5) and incubated overnight at 37 °C with shaking at 200 rpm. The overnight culture was transferred into 100 mL of fresh B2low medium and grown at 30 °C with shaking at 150 rpm until the OD₆₀₀ reached 0.4–0.5. Cells were harvested by centrifugation at $4,000 \times g$ for 10 min at 4 °C and resuspended in 30 mL of ice-cold 1.5 M NaCl. Cells were pelleted again ($4,000 \times g$ for 10 min at 4 °C), resuspended in 10 mL of ice-cold 10% (v/v) glycerol, and centrifuged as described above. The wash with ice-cold 10% (v/v) glycerol was repeated at least five times. Competent cells were stored at -80 °C until use.

For transformation, competent cells were thawed at room temperature for 5 min, pelleted at $5,000 \times g$ for 1 min, and resuspended in electroporation buffer (0.5 M sucrose and 10% (v/v) glycerol). Plasmid

DNA was added, and cells were electroporated using an ELEPO21 electroporator (NEPA GENE, Chiba, Japan) under the following conditions: Poring Pulse (voltage: 2,000 V; pulse length: 2.5 ms; pulse interval: 50 ms; number of pulses: 1; polarity: +) and Transfer Pulse (voltage: 150 V; pulse length: 50 ms; pulse interval: 50 ms; number of pulses: 5; polarity: +/−). After electroporation, cells were recovered in B2 medium (per litre: 25 g yeast extract, 10 g tryptone, 25 g NaCl, 0.5% (w/v) glucose, and 1 g K₂HPO₄; pH 7.5) and plated on TSA containing the appropriate antibiotics.

Natural transformation

Natural transformation was performed as previously described, with modifications⁴⁰. Briefly, three biofilm-forming recipient strains (FK300Δ*icaR*, FK300Δ*5bp*, and FK300Δ*rob*)⁴¹ and RN4220/*sadR*(chr) were grown overnight in TSB and then incubated at 37 °C for 3 h. Each culture (750 μL) was collected and washed once with 750 μL of CS2 medium⁴². The washed cells were transferred to 6-well plates and mixed with 250 μL of heat-killed SCC*mec* donor *S. aureus* (COL or N315). For preparation of heat-killed donor *S. aureus* cells, COL and N315 were cultured overnight in TSB at 37 °C with shaking at 200 rpm. Each overnight culture (150 μL) was then inoculated into 3 mL of fresh TSB medium and incubated at 37 °C for 3 h. Cells were heat-killed at 98 °C for 60 min, and complete killing was confirmed by plating on TSA without any antibiotics. The resulting heat-killed donor cells were used in the 6-well plate assay, in which plates were incubated at 37 °C and the supernatant was replaced every 24 h with an equal volume of fresh CS2 medium. Biofilm-associated cells were harvested by pipetting and plated on brain heart infusion agar supplemented with 4 μg/mL cefmetazole (Merck, Rahway, NJ, USA; Catalogue: C6048). Plates were incubated at 37 °C for 48 h.

Phage susceptibility testing

S. aureus RN4220 harbouring the relevant plasmid was grown overnight in TSB containing the appropriate antibiotics at 37 °C. Subsequently, 100 μL of the stationary-phase culture was mixed with 5 mL of molten soft agar (TSB containing 0.6% (w/v) agar; supplemented with chloramphenicol at 10 μg/mL when required) and overlaid onto TSA plates. Phage suspensions (2.5 μL) from a 10-fold dilution series were spotted onto the soft agar plates. Plates were incubated overnight at 37 °C before plaque formation was recorded.

Conservation phylogeny of Abi_C domain-containing protein

To detect SadR protein in other species, we collected all prokaryotic amino acid sequences and genome sequences using ncbi-genome-download (<https://github.com/kbclin/ncbi-genome-download>) v0.3.1 and ncbi datasets v16.37.0⁴³, focusing on *Staphylococcaceae* and including representative outgroup genera (total $n = 27,655$). The dataset comprised *Abyssicoccus* ($n = 2$), *Aliicoccus* ($n = 1$), *Gemella* ($n = 38$), *Jeotgalicoccus* ($n = 16$), *Macroccoccus* ($n = 30$), *Nosocomiicoccus* ($n = 11$), *Salinicoccus* ($n = 24$), *Pseudobacillus* ($n = 10$), *Staphylococcus* ($n = 20,499$), *Bacillus* ($n = 6,959$), and other genera ($n = 65$). All meta-data for the downloaded genomes were collected from https://ftp.ncbi.nlm.nih.gov/genomes/GENOME_REPORTS/. All accession numbers and species used for homologues of SadR proteins are listed in Supplementary Data 8. Collected sequences were analysed using local BLASTP (v.2.12) with the clade A SadR protein as the query, applying thresholds of $e\text{-value} < 1e-20$, query coverage $\geq 80\%$, and sequence identity $\geq 80\%$ to detect SadR homologues across all genomes. To account for increased sequence divergence in certain non-*Staphylococcus* taxa (e.g., *Macroccoccus*), we additionally applied a relaxed identity threshold ($\geq 60\%$) while keeping the $e\text{-value}$ and coverage thresholds unchanged. A phylogenetic tree was constructed using Clustal Omega (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>) based on multiple alignments of SadR candidate protein sequences. The SadR protein-based phylogenetic tree was visualised using iTOL.

Abi_C domain-containing proteins were detected using the NCBI Conserved Domain (CD) search tool (accessed in November 2023, Conserved Domain Database version v3.20) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), with an $e\text{-value}$ threshold of < 0.01 . Obtained sequences were further subjected to local BLASTp, employing an $e\text{-value}$ threshold of < 0.001 , and sequences with less than 50% query coverage were discarded. Sequences were aligned using MAFFT v.7.520⁴⁴. Approximately-maximum-likelihood phylogenetic trees were constructed using FastTree v2.1.11⁴⁵ with JTT + CAT model. The constructed tree was visualised using the iTOL software. HEPN-type RNase motif was defined as previously reported^{46,47}. For RNA motif detection, 120 amino acids from the C-terminal were extracted using Seqkit v2.10.0⁴⁸ with “subseq” subcommand. Subsequently, “locate” subcommand was used to detect the HEPN-RNase motif sequences. The extracted sequences were aligned using MAFFT software v.7.520. Amino acid sequence logos were constructed using WebLogo3 (<https://weblogo.threeplusone.com/create.cgi>), based on the aligned amino acid sequences.

RNA extraction

S. aureus strains were incubated overnight in TSB at 37 °C. The culture was incubated until OD₆₀₀ reached 0.5, and the 80α phage was added at a multiplicity of infection of 0.1 and incubated for 30 min. The culture was centrifuged at 8000 × g for 3 min, and the supernatant was discarded. Pellets were suspended using 700 μL of Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA). Cells were homogenised with glass beads for 45 s at 5-min intervals. The suspension was centrifuged at 11,363 × g for 15 min. Next, 500 μL of the supernatant was transferred to a new tube, and the same volume of 70% (v/v) ethanol was added. Total RNA was extracted from RN4220 cells using the miRNeasy Mini kit (Qiagen, Venlo, Netherlands), according to the manufacturer's instructions. Total RNA concentration was measured at a wavelength of 260 nm using a NanoDrop spectrophotometer.

RT-qPCR

Total extracted RNA (1.5 μg) was subjected to reverse transcription to synthesise cDNA using the ReverTrace qPCR RT master mix with gDNA remover (TOYOBO, Osaka, Japan; Catalogue: FSQ-301) according to the manufacturer's instructions. Quantitative reverse transcription PCR was performed using SYBR Green master mix (Thermo Fisher Scientific, Waltham, MA, USA, Catalogue: A66732), which was mixed with 2 μL of cDNA. Primers were used to amplify the *sadR* intergenic region. As internal controls, the transcription levels of *dnaA*, *dnaK*, and *rpoB* genes were also measured. qPCR was performed using Applied Biosystems StepOne (Applied Biosystems, Waltham, Massachusetts, USA) under the following cycling conditions: 95 °C for 1 min, followed by 30 cycles of 95 °C for 15 s and 60 °C for 30 s. The relative expression of *sadR* was determined using the 2^{−ΔΔC_t} method, normalised to the transcription level of housekeeping gene *rpoB*. All primers used for qPCR are listed in Supplementary Data 10.

Temperate phage induction

S. aureus strains were incubated overnight at 37 °C in 3 mL of TSB. Subsequently, 200 μL of overnight culture was inoculated into 20 mL of fresh TSB and incubated at 37 °C until OD₆₀₀ reached 0.2–0.3. Mitomycin C was added to a final concentration of 1 μg/mL, the cultures were further incubated for 5–6 h, and then centrifuged at 10,000 × g for 15 min. The supernatants were filtered through a 0.22 μm filter. The filtered supernatant was spotted onto soft agar containing *S. aureus* RN4220 to assess the presence of temperate phages, using the same media and plating conditions as described in the “Phage susceptibility testing” section.

Phage DNA sequencing

Genomic DNA was extracted from phages using the phenol/chloroform extraction method. Briefly, 1 mL of phage lysate was treated with 2 μ L of 10 mg/mL RNase A (Thermo, Catalogue; R1253) (final concentration, 20 μ g/mL, corresponding to ≥ 100 U/mL) and 2 μ L (10 U) DNase I (Takara Bio, Shiga, Japan, Catalogue; 2270 A), followed by incubation at 37 °C for 1 h, and then 15 μ L of 0.5 M EDTA was added to halt the reaction of DNase I. Thereafter, 4 μ L of 20% (w/v) SDS and 2 μ L of Proteinase K (Kanto chemical, Tokyo, Japan, Catalogue; 34060-96) were added to the reaction and then incubated at 65 °C for 15 min. The DNA was purified by sequential extraction with Phenol:chloroform:isoamyl alcohol (25:24:1, v/v). After each extraction, the sample was centrifuged at $17,750 \times g$ for 3 min, and the upper aqueous phase was transferred to a new tube. DNA was precipitated by adding 1/10 volume of 3 M sodium acetate and 2.5 volumes of ice-cold 100% (v/v) ethanol, followed by incubation at -20 °C for at least 30 min. The precipitated DNA was collected by centrifugation at $14,000 \times g$ for 30 min. The pellet was washed twice with 70% (v/v) ethanol, air-dried completely, and resuspended in 30 μ L of TE buffer or sterile water. DNA concentration and purity were assessed using a NanoDrop spectrophotometer. Each phage DNA library was constructed using a QIAseq FX DNA library kit (Qiagen, Venlo, Netherlands, Catalogue; 180473), and sequencing was performed using the Illumina MiSeq platform, generating paired-end reads (2×300 bp). Raw reads were assessed with FastQC (<https://github.com/s-andrews/FastQC>) v0.12.1. Genome assembly was performed using Unicycler v0.4.8⁴⁹ (parameters --mode normal). The coverage and the number of reads were calculated using samtools v1.14⁴⁸ (Supplementary Data 9). Annotation was performed using Pharokka v1.7.3⁴⁵ (parameters --eval 0.1).

SadR plasmid construction for expression and protein purification

Wild-type *sadR* and site-directed variants were cloned into the His-tag vector pET-16b (Merck) with the NEBuilder HiFi DNA Assembly kit (New England Biolabs, Ipswich, MA, USA, Catalogue; E2621) following the manufacturer's instructions. For construction of the three mutants (R225A, N226A, H232A), inverse PCR was performed on the wild-type SadR-pET-16b plasmid as the template using Tks Gflex DNA Polymerase (Takara Bio, Shiga, Japan, Catalogue; R060). All constructs were transformed into *Escherichia coli* Rosetta 2 (DE3) (Merck, Catalogue; 71397) and transformants were selected on the LB + 100 μ g/mL ampicillin. Transformants were cultured in 3 mL overnight. The overnight culture was inoculated into 1 L LB medium and grown at 37 °C, 200 rpm to OD₆₀₀ $\approx$ 0.5. Protein expression was induced by adding IPTG (Wako, Osaka, Japan, Catalogue; 096-05861) to a final concentration of 0.2 mM, and the culture was incubated for an additional 3 h. Cells were harvested ($10,000 \times g$, 10 min, 4 °C) and resuspended in 60 mL of HisTALON xTractor Buffer (Takara Bio, Shiga, Japan); cells were disrupted by ultrasonication on ice. The lysate was clarified ($10,000 \times g$, 90 min, 4 °C) and subsequently, the supernatant was filtered through a 0.22 μ m membrane (Merck, Catalogue; SLGVR33RS). His-tagged SadR proteins were purified on a HisTALON (Takara Bio, Shiga, Japan) cobalt column according to the manufacturer's protocol. Protein purity and the SadR complex were assessed by 12.5 % gels for SDS-PAGE and 4-20% gradient gels for native-PAGE electrophoresis using pageRUN Ace (ATTO, Tokyo, Japan, WSE-1150P), respectively.

In vitro synthesis of retron-Ec83 msRNA

Retron-Ec83 msRNA was synthesised in vitro from the noncoding region of retron-Ec83. A synthetic DNA fragment containing the target sequence was obtained from Genewiz (Azenta Life Sciences, South Plainfield, USA) and used as PCR template. The DNA template for T7 transcription was generated using Phusion High-Fidelity DNA Polymerase (New England Biolabs, M0530) with primers introducing a T7

promoter upstream of the msRNA sequence. In vitro transcription was performed using the MEGAscript™ T7 Transcription Kit (Invitrogen, AM1333) in a 20- μ L reaction containing 2 μ L each of ATP, CTP, GTP, and UTP, 2 μ L 10 $\times$ reaction buffer, 2 μ L enzyme mix, and 200 ng DNA template. Reactions were incubated overnight at 37 °C, followed by Turbo DNase treatment (1 U/ μ L) for 15 min at 37 °C. RNA was then mixed with 15 μ L ammonium acetate stop solution (5 M ammonium acetate, 100 mM EDTA) and 65 μ L H₂O, extracted with 100 μ L PCI, and the aqueous phase was purified using a G25 spin column. The eluate was mixed with 300 μ L of 30:1 EtOH:3 M NaOAc (pH 6.5) and precipitated overnight at -20 °C. RNA was pelleted by centrifugation at $13,000 \times g$ for 30 min at 4 °C, washed with 100 μ L 75% EtOH, and centrifuged again at $13,000 \times g$ for 10 min at 4 °C. After removal of the supernatant, the pellet was air-dried and resuspended in 30 μ L nuclease-free water by shaking for 5 min at 65 °C.

In vitro nucleotide degradation assay

Purified SadR protein (about 10 ng) and the following nucleotides: ssRNA (CGCUCUUCUAAUUCUGUAAUUCGUUGUCACUUGCACUUGGUAGU), dsRNA (CGCUCUUCUAAUUCUGUAAUUCGUUGUCA), ssDNA (ATCTCGCCCTTTTCTTGAGATAAGTTCACGAAGCTGATTGATACGCTCTCTAATTCTGTAATATCGTTGTCACCTTGCACTTGGTAGTTCCTCGCGTTGTAGA), and dsDNA (ATCTCGCCCTTTTCTTGAGATAAGTTCACGAAGCTGATTGATACGCTCTCTAATTCTGTAATATCGTTGTCACCTTGCTAGTTCCTCGCGTTGTAGA) were mixed with PBS or NEbuffer r3.1 (NEB, Catalogue; B6003S) and incubated at 37 °C for 1 h. For total RNA, 200 ng RNA extracted from *S. aureus* USA300 was mixed with the SadR protein. Retron-Ec83 msRNA was referenced as previously described⁵⁰. Electrophoresis was then performed using Xcell SureLock (Invitrogen, Carlsbad, CA, USA) and Novex TBE-urea PAGE (10 % TBE-Urea Gel) at 200 V. Subsequently, the gel was stained with Sybr-gold (Invitrogen, Catalogue; S11494) and gel bands were visualised under LED illumination using a FAS-Digi PRO imaging system (Nippon Genetics, Tokyo, Japan).

Statistical analysis and plotting

Statistical analyses were performed using Python version 3.9.7 and SciPy Module⁵¹ v1.13.1 or pingouin v0.5.5 (<https://pingouin-stats.org/build/html/index.html>). All statistical tests were two-tailed unless otherwise specified. Data are presented as mean $\pm$ SD or, where appropriate, 95% confidence intervals (CI). *P*-values are provided in the figures or legends. Data visualisation was performed using Seaborn⁵², the ComplexHeatmap package⁵³ in R, and RAWGraphs (<https://www.rawgraphs.io/>).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequencing data for *S. aureus* generated in this study have been deposited in the DDBJ/ENA/GenBank databases under BioProject accession [PRJDB10779](https://www.ncbi.nlm.nih.gov/bioproject/PRJDB10779) and [PRJDB13052](https://www.ncbi.nlm.nih.gov/bioproject/PRJDB13052); each accession number was provided in Supplementary Data 3. Phage genome sequences have been deposited under BioProject accession (<https://www.ncbi.nlm.nih.gov/bioproject/PRJDB40234/>); accession numbers are provided in Supplementary Data 9. The escape phage sequencing data are associated with accession numbers DRR1048348 to DRR1048350 (SAM01914136 to SAM01914138). ST188 strains from Vietnam are available in Bioproject accession [PRJNA646358](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA646358) (SRR12233367 to SRR12233563). Source data are provided with this paper.

Code availability

The custom code used in this study is available at Zenodo (<https://doi.org/10.5281/zenodo.20077086>).

References

- Xia, G. & Wolz, C. Phages of *Staphylococcus aureus* and their impact on host evolution. *Infect., Genet. Evolution* **21**, 593–601 (2014).
- Malachowa, N. & Deleo, F. R. Mobile genetic elements of *Staphylococcus aureus*. *Cell. Mol. Life Sci.* **67**, 3057–3071. (2010).
- Bitrus, A. A., Peter, O. M., Abbas, M. A. & Goni, M. D. *Staphylococcus aureus*: a review of antimicrobial resistance mechanisms. *Veterinary Sciences: Research and Reviews* **4**, 43–54 (2018).
- Touaitia, R. et al. *Staphylococcus aureus*: A Review of the Pathogenesis and Virulence Mechanisms. *Antibiotics* **14**, 470 (2025).
- Utsui, Y. & Yokota, T. Role of an Altered Penicillin-Binding Protein in Methicillin-and Cephem-Resistant *Staphylococcus Aureus*. *ANTI-MICROBIAL AGENTS AND CHEMOTHERAPY*. <https://journals.asm.org/journal/aac> (1985).
- Hayes, M. V., Curtis, N. A. C., Wyke, A. W. & Ward, J. B. Decreased affinity of a penicillin-binding protein for β -lactam antibiotics in a clinical isolate of *Staphylococcus aureus* resistant to methicillin. *FEMS Microbiol. Lett.* **10**, 119–122 (1981).
- Naghavi, M. et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. [https://doi.org/10.1016/S0140-6736\(24\)01867-1](https://doi.org/10.1016/S0140-6736(24)01867-1). (2024).
- Shinya Tsuzuki & Nobuaki matunaga. National trend of blood-stream infection attributable deaths caused by *Staphylococcus aureus* and *Escherichia coli* in Japan Publisher's Note. *J. Infect. Chemother.* **26**, i–ii (2020).
- Obata, S. et al. Comprehensive genomic characterization of *Staphylococcus aureus* isolated from atopic dermatitis patients in Japan: correlations with disease severity, eruption type, and anatomical site. *Microbiol. Spectr.* **11**, (2023).
- Thuy, D. B. et al. Colonization with *staphylococcus aureus* and *klebsiella pneumoniae* causes infections in a vietnamese intensive care unit. *Microb. Genom.* **7**, 1–14 (2021).
- Mašlanová, I. et al. Description and comparative genomics of *Macroccoccus caseolyticus* subsp. *hominis* subsp. nov., *Macroccoccus goetzii* sp. nov., *Macroccoccus epidermidis* sp. nov., and *Macroccoccus bohemicus* sp. nov., Novel *Macroccocci* from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Front. Microbiol.* **9**, 1178 (2018).
- Hossain, M., Aslan, B. & Hatoun-Aslan, A. Tandem mobilization of anti-phage defenses alongside SCCmec elements in staphylococci. *Nat. Commun.* **15**, 8820 (2024).
- Hochhauser, D., Millman, A. & Sorek, R. The defense island repertoire of the *Escherichia coli* pan-genome. *PLoS Genet.* **19**, e1010694 (2023).
- Wu, Y. et al. Bacterial defense systems exhibit synergistic anti-phage activity. *Cell Host Microbe* **32**, 557–572.e6 (2024).
- Rita C. A. et al. Accumulation of Defense Systems in Phage-Resistant Strains of *Pseudomonas Aeruginosa*. *Sci. Adv.* <https://www.science.org> (2024).
- Duan, N., Hand, E., Pheko, M., Sharma, S. & Emiola, A. Structure-guided discovery of anti-CRISPR and anti-phage defense proteins. *Nat. Commun.* **15**, 649 (2024).
- Jiang, X. et al. Type I BREX system defends against antibiotic-resistant plasmids in *Escherichia coli*. *Antimicrob. Agents Chemother.* **68** e01128–23 (2024).
- Rousset, F. et al. Phages and their satellites encode hotspots of antiviral systems. *Cell Host Microbe* **30**, 740–753.e5 (2022).
- Getz, L. J., Fairburn, S. R., Vivian Liu, Y., Qian, A. L. & Maxwell, K. L. Integrins are anti-phage defence libraries in *Vibrio parahaemolyticus*. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-025-01927-7> (2025).
- Nowrouzian, F. L. et al. *Staphylococcus aureus* sequence type (ST) 45, ST30, and ST15 in the gut microbiota of healthy infants—persistence and population counts in relation to ST and virulence gene carriage. *Eur. J. Clin. Microbiol. Infect. Dis.* **42**, 267–276 (2023).
- Chopin, M. C., Chopin, A. & Bidnenko, E. Phage abortive infection in lactococci: Variations on a theme. *Curr. Opin. Microbiol.* **8**, 473–479 (2005).
- Uchiyama, J. et al. In silico analysis of AHJD-like viruses, *Staphylococcus aureus* phages S24-1 and S13', and study of phage S24-1 adsorption. *Microbiologyopen* **3**, 257–270 (2014).
- Lee, C. Y., Buranen, S. L. & Ye, Z.-H. Construction of Single-Copy Integration Vectors for *Staphylococcus Ureus* (Recombinant DNA; Site-Specific Recombination; *Staphylococcal Bacteriophage*; Cloning Vector; Lysogenization). *Gene* **103** (1991).
- Monk, I. R., Shah, I. M., Xu, M., Tan, M. W. & Foster, T. J. Transforming the untransformable: Application of direct transformation to manipulate genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *mBio*. **3**, e00277-11 (2012).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Eddy, S. R. et al. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
- Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
- van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).
- Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
- Payne, L. J. et al. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Res* **50**, W541–W550 (2022).
- Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
- Tonkin-Hill, G. et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol.* **21**, 180 (2020).
- Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* **52**, W78–W82 (2024).
- Petit, R. A. & Read, T. D. *Staphylococcus aureus* viewed from the perspective of 40,000+ genomes. *PeerJ*. **2018**, e5261 (2018).
- Bortolaia, V. et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **75**, 3491–3500 (2020).
- Monecke, S. et al. Characterisation of *S. aureus*/MRSA CC1153 and review of mobile genetic elements carrying the fusidic acid resistance gene *fusC*. *Sci. Rep.* **11**, 8128 (2021).
- Ito, T. et al. Structural comparison of three types of staphylococcal cassette chromosome mec integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **45**, 1323–1336 (2001).
- Hiramatsu, K. et al. Genomic basis for methicillin resistance in *Staphylococcus aureus*. *Infection and Chemotherapy* vol. 45 117–136 <https://doi.org/10.3947/ic.2013.45.2.117> (2013).
- Sato'o, Y. et al. Optimized universal protocol for electroporation of both coagulase-positive and -negative *Staphylococci*. *J. Microbiol. Methods* **146**, 25–32 (2018).
- Maree, M. et al. Natural transformation allows transfer of SCCmec-mediated methicillin resistance in *Staphylococcus aureus* biofilms. *Nat. Commun.* **13**, 2477 (2022).
- Kutsuno, S. et al. Non-deacetylated poly-N-acetylglucosamine-hyperproducing *Staphylococcus aureus* undergoes immediate autoaggregation upon vortexing. *Front. Microbiol.* **13**, 1101545 (2023).
- Toledo-Arana, A. et al. *Staphylococcus aureus* develops an alternative, ica-independent biofilm in the absence of the *arlRS* two-component system. *J. Bacteriol.* **187**, 5318–5329 (2005).

43. O'Leary, N. A. et al. Exploring and retrieving sequence and meta-data for species across the tree of life with NCBI datasets. *Sci. Data* **11**, 732 (2024).
44. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2018).
45. Bouras, G. et al. Pharokka: a fast scalable bacteriophage annotation tool. *Bioinformatics* **39**, btac776 (2023).
46. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V. & Aravind, L. Comprehensive analysis of the HEPN superfamily: Identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct* **8**, 15 (2013).
47. Pillon, M. C., Gordon, J., Frazier, M. N. & Stanley, R. E. HEPN RNases—an emerging class of functionally distinct RNA processing and degradation enzymes. *Crit. Rev. Biochem. Mol. Biol.* **56**, 88–108 (2021).
48. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Giga-science* **10**, giab008 (2021).
49. 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).
50. Azam, A. H. et al. Evasion of antiviral bacterial immunity by phage tRNAs. *Nat. Commun.* **15**, 9586 (2024).
51. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
52. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
53. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

Acknowledgements

We thank Chika Arai of the Antimicrobial Resistance Research Centre for the support.

Author contributions

M.S. and J.H. conceptualised this study. K.Kondo provided funding, designed, and conducted the experiments, analysed data, and draughted the manuscript. A.H.A. supervised microbiological experiments and reviewed the manuscript. A.H.A., S.O., K.C., T.N., W.Y., and A.T. contributed to data collection. S.N. performed bioinformatic analysis. J.H., J.U., Y.S., M.A., T.Y., Y.I., and S.K. contributed to data collection and strain isolation and assisted in genomic analysis. K.W., Y.T., J.H., and K.Kiga supervised the study and provided funding.

M.S., J.H., K.Kiga contributed to writing and reviewing the manuscript. M.S. supervised the study, provided funding, and critically reviewed the manuscript.

Funding

This work was supported by the Japan Agency for Medical Research and Development (grant number JP25fk0108699 to M. S. and grant number JP23wm0325065 to A.H.A., K. Kondo, K.C., and K. Kiga) and JSPS KAKENHI (21K16331 and 25K20585 to K.Kondo). This work was also supported by the MHLW Research Programme for Promoting Food Safety, Grant Number JPMH24KA1005, to M.S.

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/s41467-026-75971-0>.

Correspondence and requests for materials should be addressed to Motoyuki Sugai.

Peer review information *Nature Communications* thanks Agnes Marie Figueiredo and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://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) 2026