Supplementary Materials for

Genome hypermobility by lateral transduction

John Chen, Nuria Quiles-Puchalt, Yin Ning Chiang, Rodrigo Bacigalupe, Alfred Fillol-Salom, Melissa Su Juan Chee, J. Ross Fitzgerald, José R Penadés.

Correspondence to: John Chen: miccjy@nus.edu.sg
José R Penadés: JoseR.Penades@glasgow.ac.uk

This PDF file includes:

Materials and Methods
Figs. S1 to S18
Tables S1 to S4

Materials and Methods

Bacterial strains and growth conditions

Bacterial strains used in these studies are listed in Table S1. *S. aureus* strains were grown in tryptic soy broth (TSB) and TSB agar. *E. coli* strains were grown in Luria-Bertani (LB) broth or on LB agar plates. Antibiotic resistant *S. aureus* were selected and maintained on 5 μg ml^{-1} erythromycin, 10 μg ml^{-1} chloramphenicol, 300 μg ml^{-1} streptomycin, 5 μg ml^{-1} tetracycline, 1.0 mM $NaAsO_2$, or 0.1 mM $CdCl_2$. Antibiotic resistant *E. coli* were selected and maintained on ampicillin, 100 μg ml^{-1}.

DNA methods

The oligonucleotides and plasmids used in this study are listed in Tables S2 and S3, respectively. The sequences of all phages and strains in this article were previously available in GenBank. Plasmid DNA preparation and agarose gel purification kits were purchased from Qiagen. Oligonucleotides were purchased from Integrated DNA Technologies. 1st Base (Singapore) or the University of Glasgow core facilities performed DNA sequencing. Phusion polymerase, restriction enzymes, and ligase were purchased from New England Biolabs.

Whole genome sequencing

Bacterial cultures were grown and at indicated time points after MC-treatment 4 ml of sample was taken for DNA extraction using GenElute Bacterial Genomic DNA kit (Sigma) following the manufacturer's instructions. DNA samples were quality control tested using Agilent Bioanalyzer 2100 at the University of Glasgow, Polyomics Facility. Whole genome sequencing (WGS) was performed at the University of Glasgow Polyomics Facility using Illumina NextSeq 500 obtaining 75 bp pair end reads with no initial PCR amplification step. Trimmed reads were mapped to the appropriate phage genome: 80α (NC_009526.1), ϕ11

(NC_004615.1), ϕNM1 (DQ530359.1), ϕNM2 (DQ530360.1), ϕNM3 (NC_008617.1), ϕNM4 (DQ530362.1), ϕ52a (NC_007062.1).

Total RNA extraction and mRNA enrichment

Bacterial cultures were grown and at indicated time points after MC-treatment samples for RNA extraction were taken and mixed with two volumes of RNA protect reagent (QIAgen) incubated for 5 minutes at room temperature and cell pellets harvested by centrifugation. The bacterial pellet was resuspended in 1 ml of TRIzol reagent (Ambion) and was lysed in a FastPrep-24 homogenizer (MP Biomedicals) using two cycles of 60 s at 6.5 m s^{-2} interrupted by a 5 minutes incubation on ice. Total RNA was extracted using the Ambion PureLink Trizol Plus kit according to the manufacturer's instructions. Genomic DNA was removed using on-column DNase digestion step using the RNase-free DNase kit (Qiagen) and residual DNA was removed by a second DNase treatment using RQ1 DNase (Promega). Total RNA samples were enriched for mRNA using MICROBexpress mRNA enrichment kit (Ambion). Experiments were performed in triplicate. Samples for RNA-seq analysis were quality control tested for mRNA enrichment using Agilent Bioanalyzer 2100 at the University of Glasgow, Polyomics Facility.

RNA-seq transcriptome analysis

cDNA synthesis and sequencing was performed at the University of Glasgow Polyomics Facility using Illumina NextSeq 500 obtaining 75 bp single end reads. Sequencing reads for each of the libraries were mapped to the appropriate reference genomes: 80α (NC_009526.1), ϕ11 (NC_004615.1), ϕNM1 (DQ530359.1), ϕNM2 (DQ530360.1), ϕ52a (NC_007062.1) using the RNA-Seq pipeline READemption (*27*). The coverage subcommand of this software was used to produce strand specific coverage counts normalized by 100,000 division. Strands coverage values of the phage spanning regions for different replicates were plotted with values >1 log10 corrected.

Allelic Exchange

For the generation of insertions in the *S. aureus* chromosome, allelic exchange was performed as previously described (*28-30*).

Phage transductions

Preparation of phage lysates, transduction, and titrations were performed as previously described (*28*, *29*, *31*, *32*). Briefly, for phage lysates, lysogens were grown to mid log in TSB, normalized by OD_{600}, and adjusted to 2 µg ml^{-1} Mitomycin C (Sigma) until complete lysis; lysates by infection were made by infecting the same density of cells with an MOI of 0.1 until complete lysis. Lysates were then adjusted to 1 µg ml^{-1} DNAse I and 1 µg ml^{-1} RNAase and filter sterilized (0.2 µm pore) before use. Phage titers were determined by plaque formation on RN450 using phage agar (*31*). For transductions, cells were infected for 30 minutes and then adjusted to 100 mM sodium citrate, mixed with 3 ml of top agar, and plated on selective agar. The results are reported as the transduction units (TrU) ml^{-1}, divided by the corresponding plaque forming units (PFU) ml^{-1}, and represented as the TrU of an average phage titer (1E+9 PFU).

Marker co-transduction analysis

Lysates from phage infection or lysogen induction of strains containing both a Cd^R and a Tet^R marker at varying distances apart were used to determine marker co-transduction frequencies. A Tet^R marker at +15kb was not evaluated because marker expansion would push the marker too close to the headful demarcation, and it is predicted there would not be sufficient flanking DNA to support homologous recombination in the recipient. The Cd^R transductants of three independent lysates (100 transductants each) were tested for Tet^R and the frequency represented as the (Tet^R/Cd^R) x 100%.

Spontaneous transduction analysis

Strains were grown to mid log in TSB and normalized by OD_{600}. Equal numbers (~1E+9 cells) of donor and recipient were adjusted to a final 1:1 mixture of TSB and phage buffer. For plating, a 500 µl equivalent of OD_{600} = 1.0 (to normalize for growth) was adjusted to 100 mM sodium citrate, mixed with 3 ml of top agar, and plated on TSA + $CdCl_2$ + Streptomycin. The results are reported as the transduction units (TrU) ml^{-1} equivalent of 1 ml of OD_{600} = 1.0 to adjust for minor growth.

Determination of the presence of defective 80α prophages in 80α lateral transductants

Transductants from lysates of 80α prophage induction were tested for defective 80α prophages by growing 100 Cd^R transductants and challenging them with wt 80α phage. Lysis by 80α suggests a lack of 80α lysogeny, and lack of lysis suggests the strain is protected from superinfection by 80α lysogeny. In addition, the same 100 Cd^R transductants were further tested for 80α lysogeny by PCR analysis using primer pairs 80alpha-1m/80alpha-2c and 80alpha-3m/80alpha-4c specific to 80α (Table S2).

Inducible complementation of the ϕ52a-fs mutant

Lysogens of wt ϕ52a and ϕ52a-fs (*rep*$_{\phi 52a}$ frame-shift) were grown at 32°C, normalized by OD, and MC-induced (t=0). Anhydrotetracycline (31.25 ng/ml) was added 0, 30, 60, 90, and 120 minutes for complementation by P_{tet}-*rep*$_{\phi 52a}$. The cells were mechanically lysed 2 hours (32°C, 100 rpm) after the addition of anhydrotetracycline in a Precellys 24 (Bertin) with 0.1mm Beadbug silica beads (Sigma-Aldrich). Physical disruption was done to harvest all lysates at fixed times because the ϕ52a-fs strains do not lyse on their own. Also, mechanical agitation with 0.1mm silica beads does not detectably degrade phage titers. The lysates were then filter sterilized and treated as above. Plaque forming units were determined on a *S. aureus* host that constitutively expresses *rep*$_{\phi 52a}$ because the ϕ52a-fs mutant is unable to form plaques on RN450 alone.

Analysis of gene content, gene synteny and recombination

Complete annotated genome sequences of 100 *S. aureus* strains representative of the species diversity were downloaded from the NCBI database (Table S4). A core genome alignment of all strains was obtained using Roary with default parameters (*33*) and FastTree (*34*) was used to construct a Maximum Likelihood phylogenetic tree with a general time-reversible (GTR) model of nucleotide substitution. Next, we used BLASTn to identify 1kb sequences that included the gene most proximal to the Sa2 and Sa6 *attB* sites in the genome of *S. aureus* strain JKD6159, and extracted 140 Kb of nucleotide sequence upstream and downstream of this locus in selected strains using Bedtools (*35*). Genomic coordinates were manually inspected to ensure homologous regions were being analyzed before gene annotation using Prokka (*36*). Clusters of orthologous genes were identified using the OrthoMCL algorithm (*37*) from the open-source get_homologues software (*38*) with parameters -C 50%, - S 30%, -F 1, before removal of paralogues. Easyfig (*39*) was used to visualize the genomic regions, and genes were coloured according to their conservation in the clusters identified. In order to analyze the impact of recombination on the flanking region of the Sa2 and Sa6 *attB* sites, the sequences of conserved clusters of genes (defined as being present in over 90% of the isolates) were aligned using MUSCLE (*40*). Subsequently, the individual alignments were concatenated into a super alignment using the software tool catfasta2phyml.pl, according to the gene order in the reference *S. aureus* strain JKD6159, before identification of recombinant regions with ClonalFrameML (*41*) using default parameters.

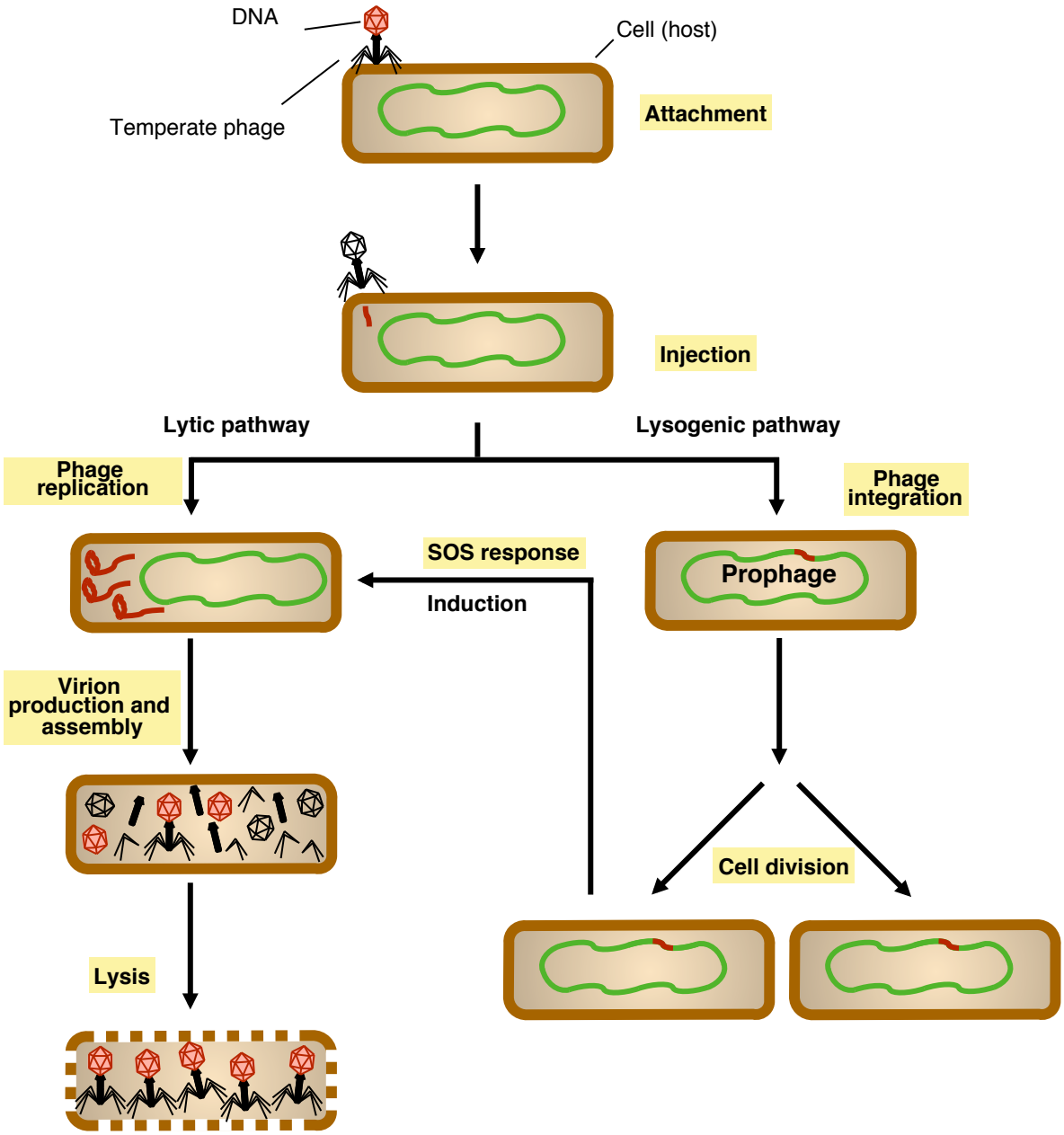


Figure S1. Lytic and lysogenic phage life cycles.

A

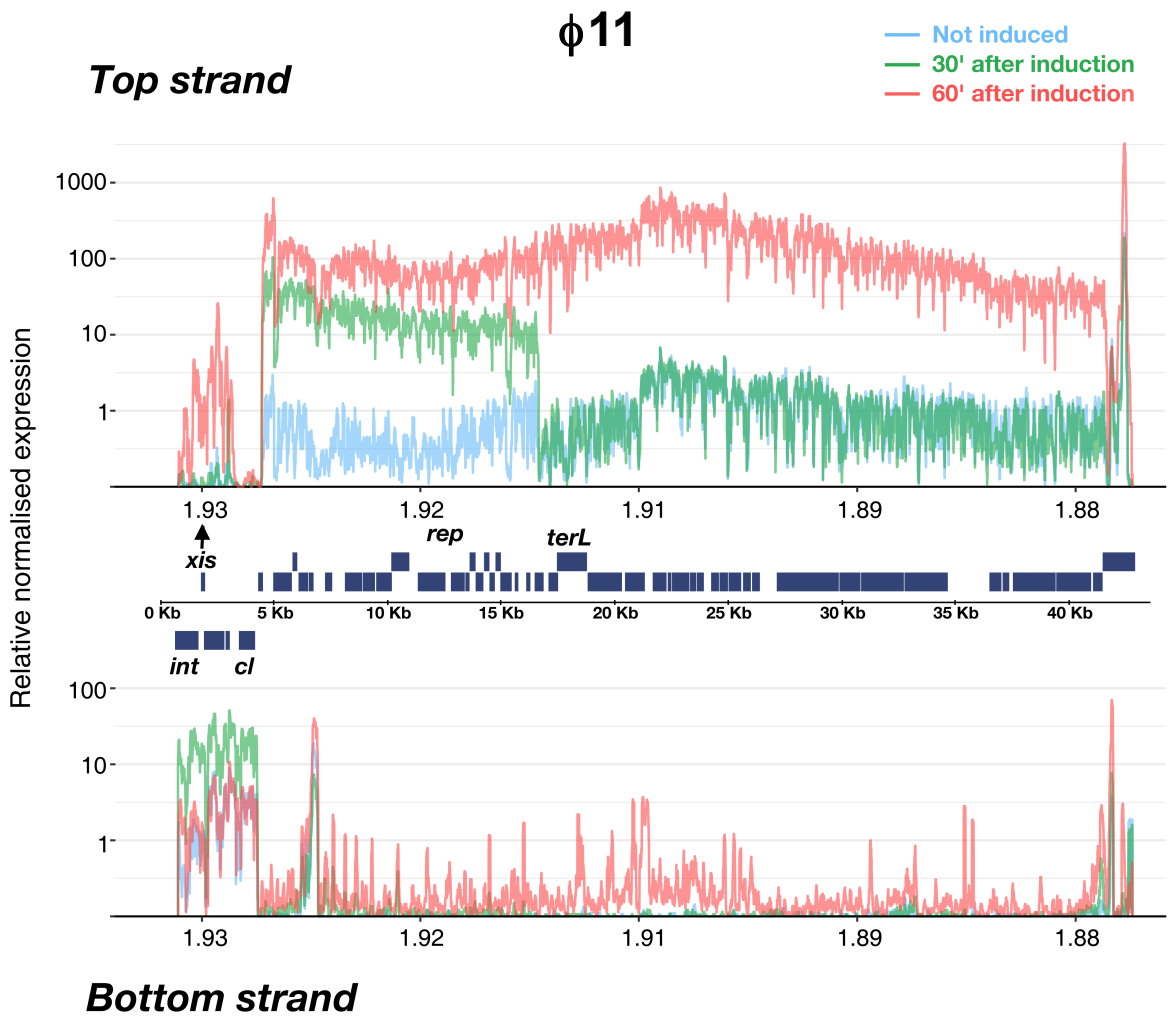
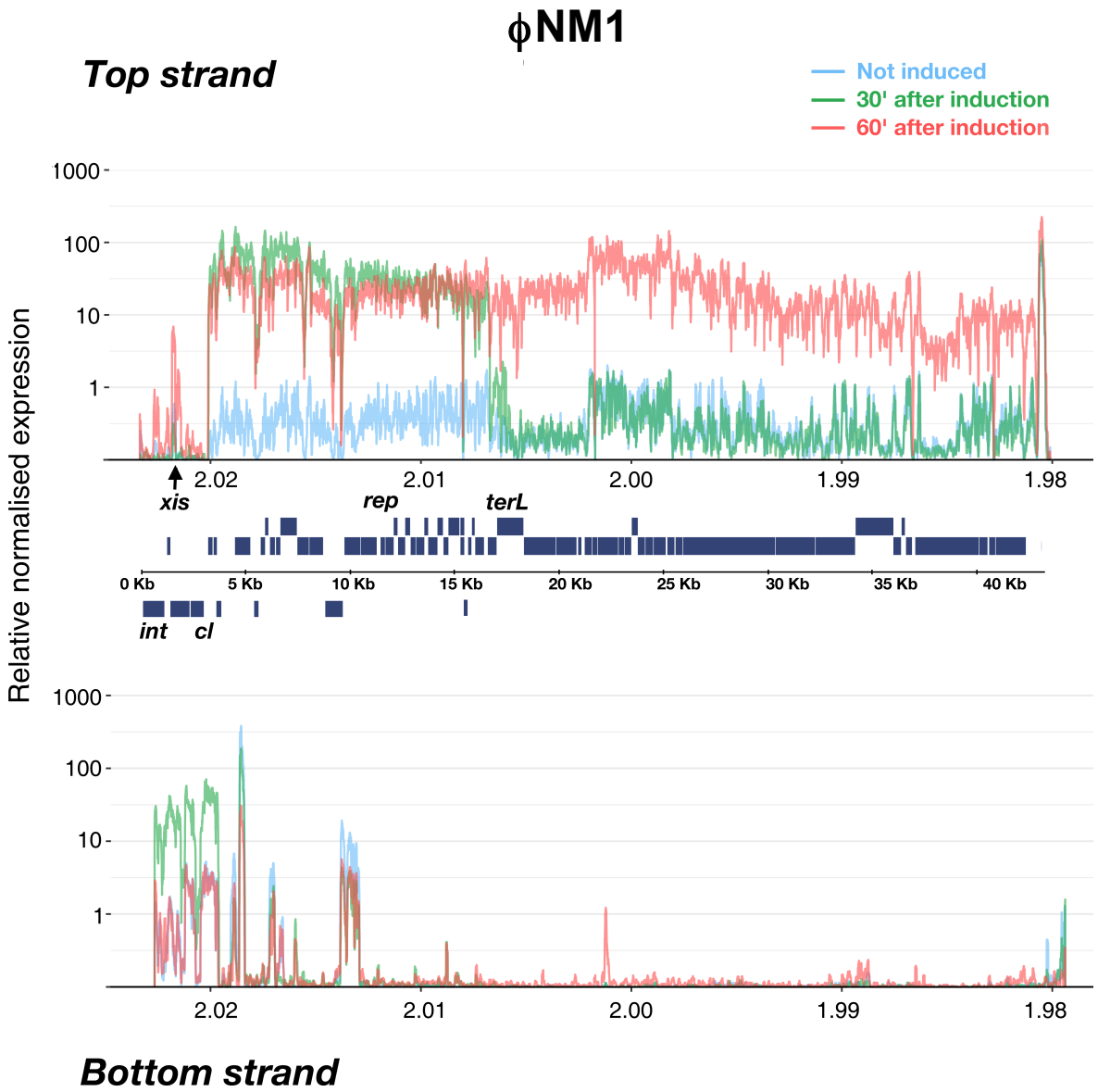


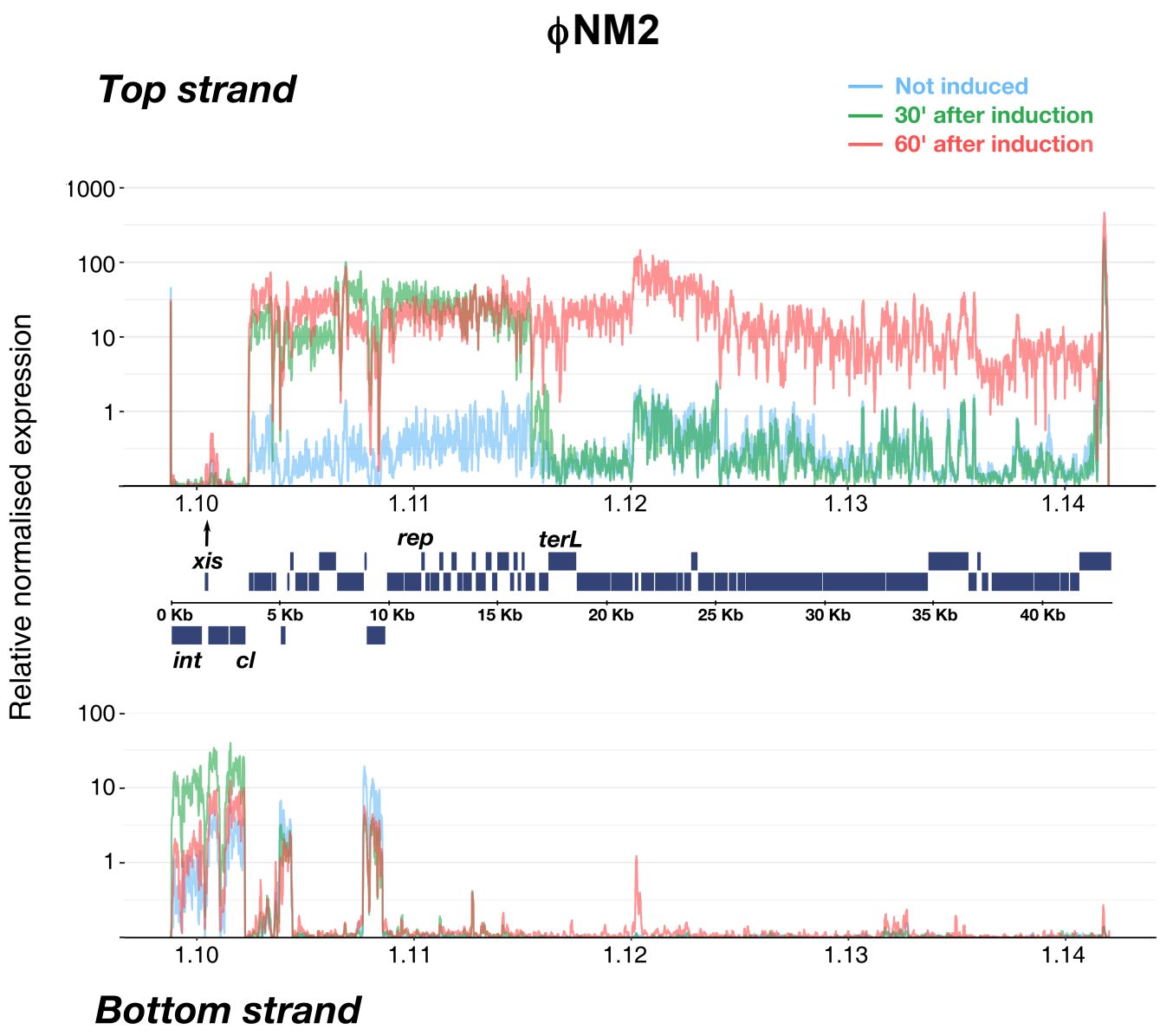
Figure S2. Transcriptomic analyses for phages ϕ11, ϕNM1, and ϕNM2.

Transcriptional analysis showing expression of the early and late genes from phages **(A)** ϕ11, **(B)** ϕNM1, and **(C)** ϕNM2. Strains RN451 (ϕ11) or Newman (ϕNM1 and ϕNM2) were induced with mitomycin C and samples were analyzed without induction (light blue) or at 30 (early genes, green) and 60 minutes (late genes, red) after induction.

B


ϕNM1
Top strand
Not induced
30' after induction
60' after induction
Relative normalised expression
1000
100
10
1
2.02
2.01
2.00
1.99
1.98
xis
rep
terL
0 Kb
5 Kb
10 Kb
15 Kb
20 Kb
25 Kb
30 Kb
35 Kb
40 Kb
int
cI
Bottom strand

C


ϕNM2
Top strand
Not induced
30' after induction
60' after induction
Relative normalised expression
1000
100
10
1
1.10
1.11
1.12
1.13
1.14
xis
rep
terL
0 Kb
5 Kb
10 Kb
15 Kb
20 Kb
25 Kb
30 Kb
35 Kb
40 Kb
int
cl
100
10
1
1.10
1.11
1.12
1.13
1.14
Bottom strand

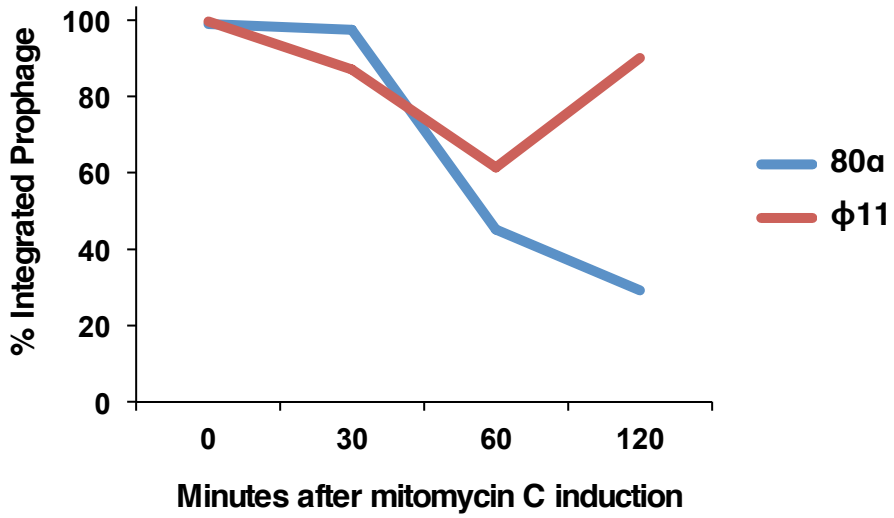


Figure S3. Prophages excise late from the bacterial chromosome.

Percentage of reads spanning the chromosome-*att*L phage region (integrated phage) out of the total (chromosome-*att*L phage plus chromosome-chromosome).

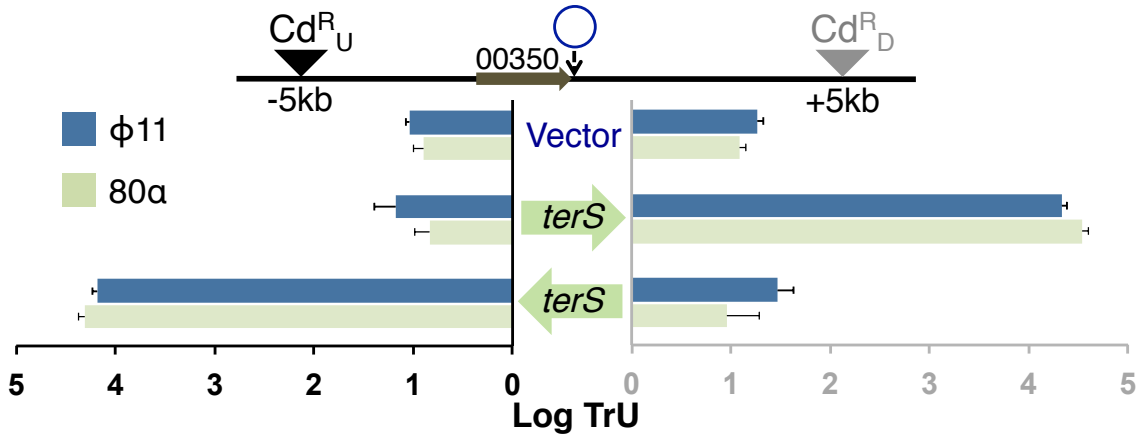


Figure S4. Determination of a *pac* site in *terS*$_{,11}$ and its directionality of packaging.

Phage *pac* sites are commonly embedded in structural *terS* genes, and to determine if the same follows for the staphylococcal phages we inserted the ϕ11 *terS* gene (*terS*$_{\phi 11}$) at an unlinked site (SaPI 4 *attB*) in RN450, using a single-copy integration vector flanked by either a cadmium marker 5kb upstream (Cd^R_U) or downstream (Cd^R_D). Since 80α and ϕ11 carry nearly identical *terS* genes, *terS*$_{\phi 11}$ was used for both phages. These strains were infected with phage and the resulting phage lysates were tested as donors of cadmium resistance to *S. aureus*. Fig. S4 shows that the transfer frequencies were three orders of magnitude greater than vector alone, only when the *terS*$_{\phi 11}$ gene was oriented such that the 3' end was closer to the marker, showing that the *terS*$_{\phi 11}$ gene encodes the bona fide ϕ11 *pac* site that directs unidirectional packaging toward the 3' end of the gene.

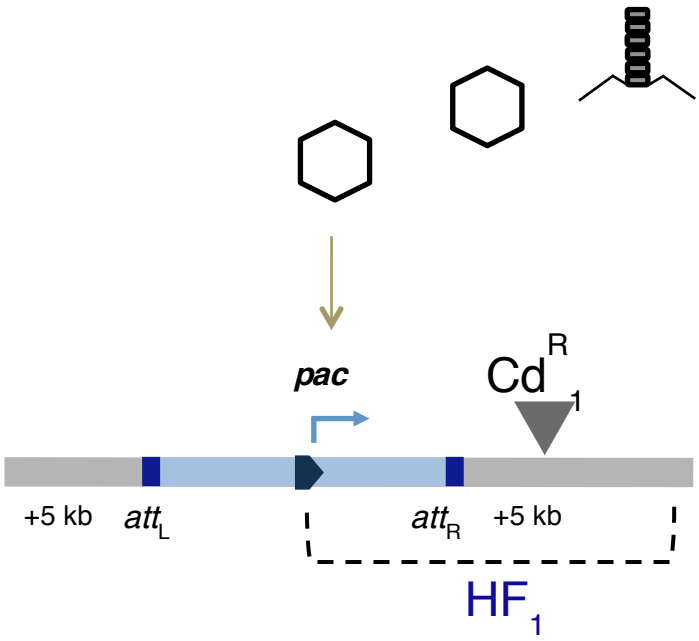


Figure S5. Model for gene transfer resulting from *in situ* packaging after prophage induction.

Location of the Cd^R_1 adjacent to the phage ϕ11 *attB* site.

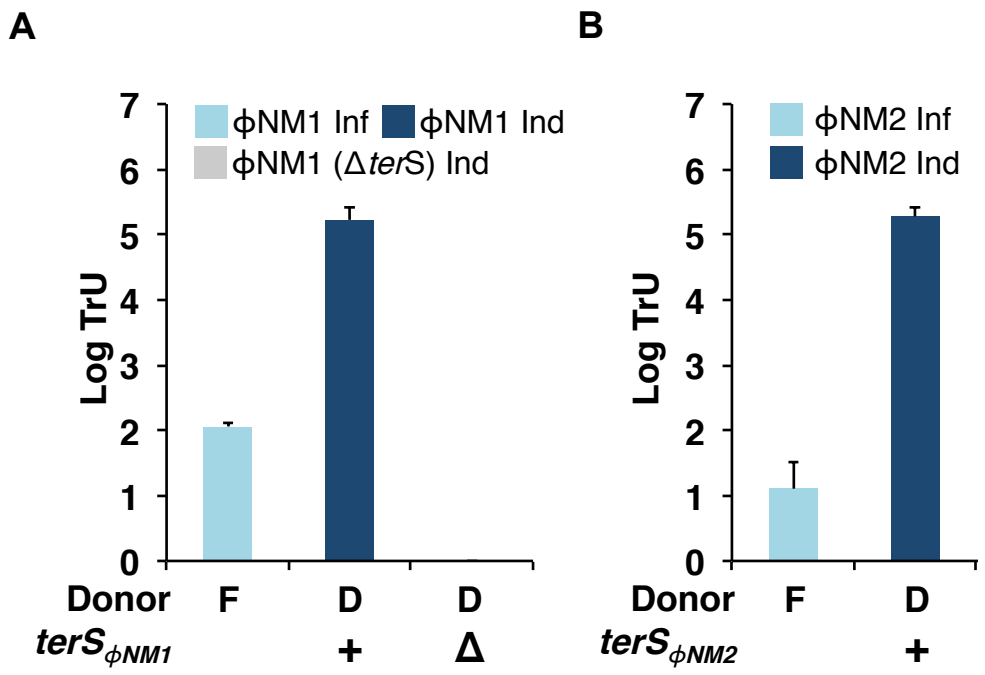


Figure S6. Staphylococcal phages and lateral transduction.

(A) ϕNM1 infection or lysogen induction tested for transfer of a Cd^R_1 5kb downstream of the *attB*$_{\phi NM1}$ (same as *attB*$_{\phi 11}$). **(B)** ϕNM2 infection or lysogen induction tested for transfer of a Cd^R 4kb downstream of the *attB*$_{\phi NM2}$ (same as *attB*$_{80\alpha}$). Non-lysogenic strains (light blue) were infected (F) or the lysogenic wt (+, dark blue) and *terS* deletion (Δ, gray) strains were induced (D) with mitomycin C and the lysates tested for transduction into *S. aureus*. Transduction units (TrU) ml^{-1} were normalized by plaque forming units (PFU) ml^{-1} and represented as the TrU of an average phage titer (1E+9 PFU). TrU ml^{-1} for Δ*terS* were <10. Values are means ± SD (*n* = 3 independent samples).

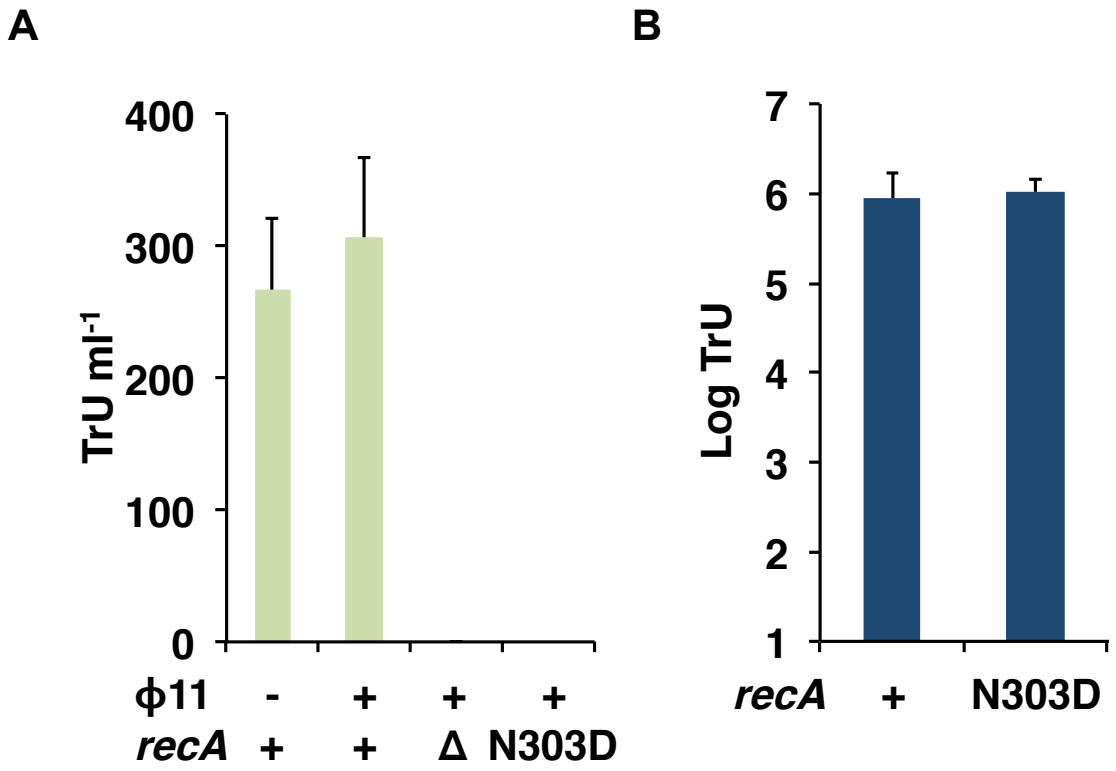


Figure S7. Recombination does not play a role in the generation of transducing particles by lateral transduction.

(A) Mutants of *recA* tested for recombination competence. Strains were used as recipients for an 80α lysate containing *agr*::*arsABC* for recombination. The results are represented as transduction units (TrU) ml^{-1}. **(B)** ϕ11 lysogen induction tested for transfer of a Cd^R 5kb downstream of the $attB_{\Phi 11}$ to *S. aureus*. Strains lysogenic (+) or non-lysogenic (-) for ϕ11 are indicated. The *recA* genotypes are wild type (+), deletion (Δ), and *recA* (N303D). TrU ml^{-1} were normalized by plaque forming units (PFU) ml^{-1} and represented as the TrU of an average phage titer (1E+9 PFU). Values are means ± SD (*n* = 3 independent samples).

A *recA* (N303D) mutant is defective for recombination but not for LexA cleavage and the SOS response. This mutant was necessary because a Δ*recA* null mutant is defective for the SOS response and unable to induce the prophage and produce a phage lysate. Shown in Fig. S7A, ϕ11 lysogens with Δ*recA* null or a *recA*(N303D) allele were both completely defective for recombination when used as recipients for GT of an *agr* loci (accessory gene regulator) marked with an arsenite resistance cassette. However, lysates from ϕ11 induction in a wt or *recA*(N303D) background transferred the $Cd^R{}_1$ marker at comparable frequencies (Fig. S7B), showing that recombination is not involved in $Cd^R{}_1$ packaging.

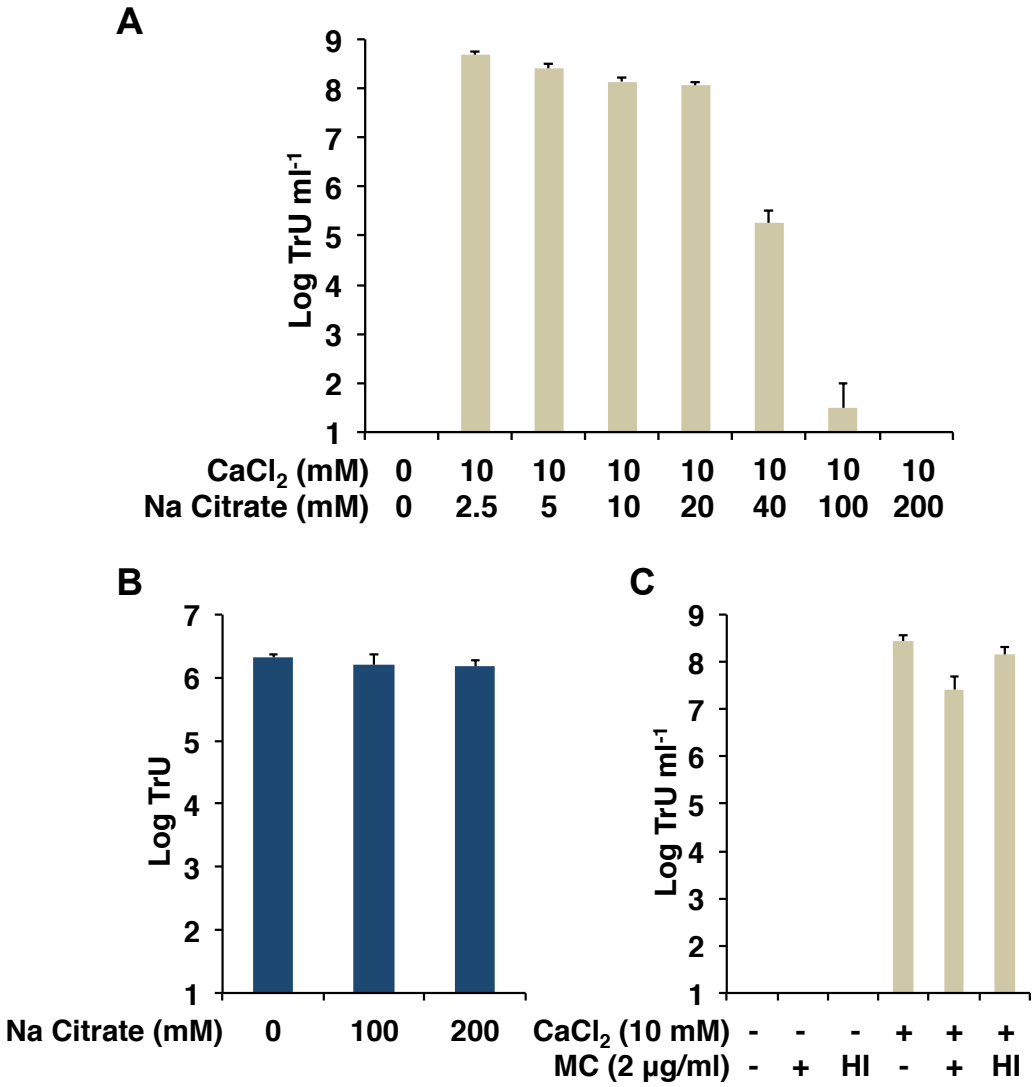


Figure S8. Super-infection does not play a role in lateral transduction.

(A) Determination of optimal concentration of chelating agent, Na citrate, for complete inhibition of phage adsorption. $CaCl_2$ is required for staphylococcal phage adsorption and 10mM $CaCl_2$ is standardly used for transductions. Increasing concentrations of Na citrate were added to *S. aureus* cells in 10mM $CaCl_2$ and infected with an 80α-SaPI 2 (*tsst1*::*tetM*) lysate. The results are represented as transduction units (TrU) ml^{-1}. **(B)** ϕ11 lysogens were induced with mitomycin C (MC) in the presence of 100mM or 200mM Na citrate and tested for transfer of a Cd^R 5kb downstream of the $attB_{\phi 11}$ to *S. aureus*. TrU ml^{-1} were normalized by plaque forming units (PFU) ml^{-1} and represented as the TrU of an average phage titer (1E+9 PFU). **(C)** *S. aureus* cells with or without 10mM $CaCl_2$ and mitomycin C (+) or heat inactivated MC (HI) were infected with an 80α-SaPI2 (*tsst1*::*tetM*) lysate. The results are represented as TrU ml^{-1}. Values are means ± SD (*n* = 3 independent samples).

Staphylococcal phages require divalent cations for adsorption to host cells, and the induction media contains only tryptic soy broth (TSB) and mitomycin C. To confirm that TSB is not able to support phage adsorption, we made use of a *S. aureus* pathogenicity island (SaPI) marked with a tetracycline resistance marker that is specially packaged by 80α and transferred at high frequencies. Fig. S8A shows that in the presence of 10mM $CaCl_2$, SaPI2 transferred at a frequency between 1E+8 to 1E+9 TrU, but was not detectable without $CaCl_2$, showing that phage adsorption is unlikely under the conditions of prophage induction. Furthermore, we induced ϕ11 in the presence of sodium citrate at concentrations (100 or 200mM) sufficient to chelate 10mM $CaCl_2$ and block phage adsorption (Fig. S8B), and showed that the packaging and transfer frequencies of the $Cd^R{}_1$ marker were unaffected. Lastly, we tested mitomycin C and heat-inactivated mitomycin C to determine if there was a contamination of divalent cation and found that SaPI2 transfer was not detectable with mitomycin C and heat-inactivated mitomycin C unless $CaCl_2$ was added (Fig. S8C).

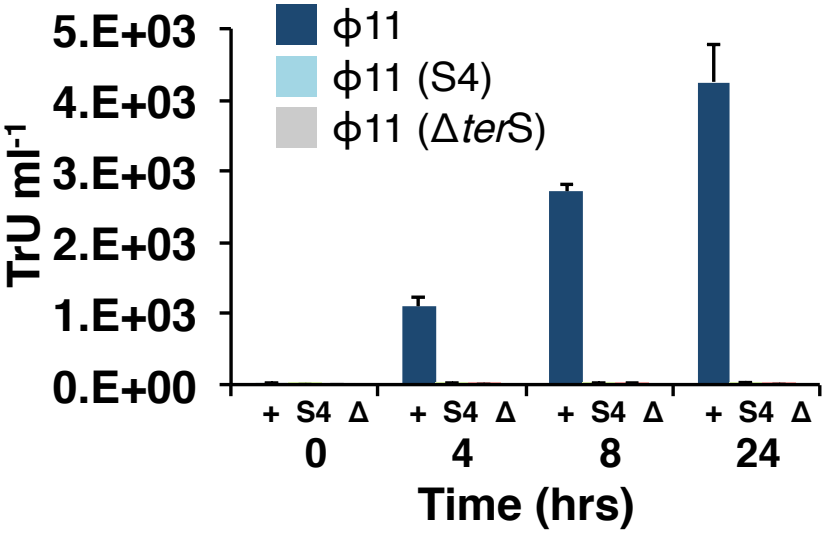


Figure S9. Spontaneous lateral transduction.

ϕ11 lysogenic strains were mixed with a streptomycin-resistant *S. aureus* recipient and tested for transfer of the Cd^R_2 marker. Strains were left to induce spontaneously, without UV or chemical induction. Genotypes of lysogens are indicated as wild type (+), ϕ11 prophage integrated at the SaPI 4 *attB* (S4), and wild type ϕ11 prophage carrying a deletion in the *ter*S gene (Δ). Transduction units (TrU) ml^{-1} are represented as the equivalent of 1ml of OD_{600} = 1.0 to adjust for minor growth. Values are means ± SD (*n* = 3 independent samples).

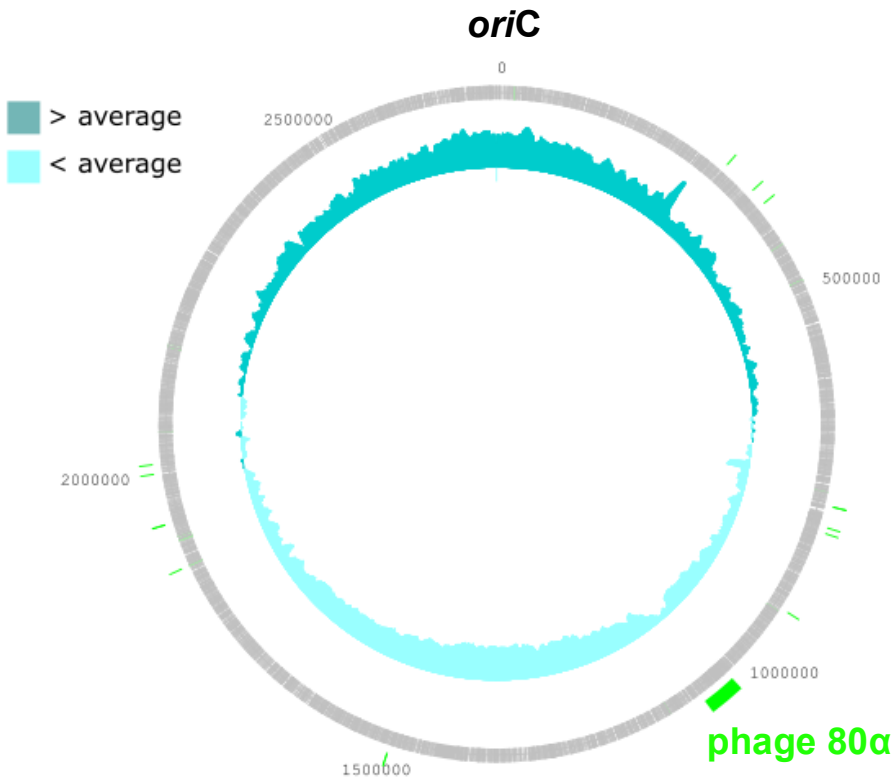


Figure S10. Whole genome sequencing and DNA quantification after induction of an 80α lysogen.

DNA coverage decreases linearly from the *ori*C origin of replication towards the terminus. The degree coverage greater than the average of the entire chromosome is represented in teal. The degree coverage lesser than the average of the entire chromosome is represented in turquoise. Hence, the region furthest from the *ori*C displays the least DNA coverage and the greatest teal. Induction of the 80α prophage results in a large amplification of the viral genome, hence the reduction in lesser than the average turquoise at the 80α prophage site.

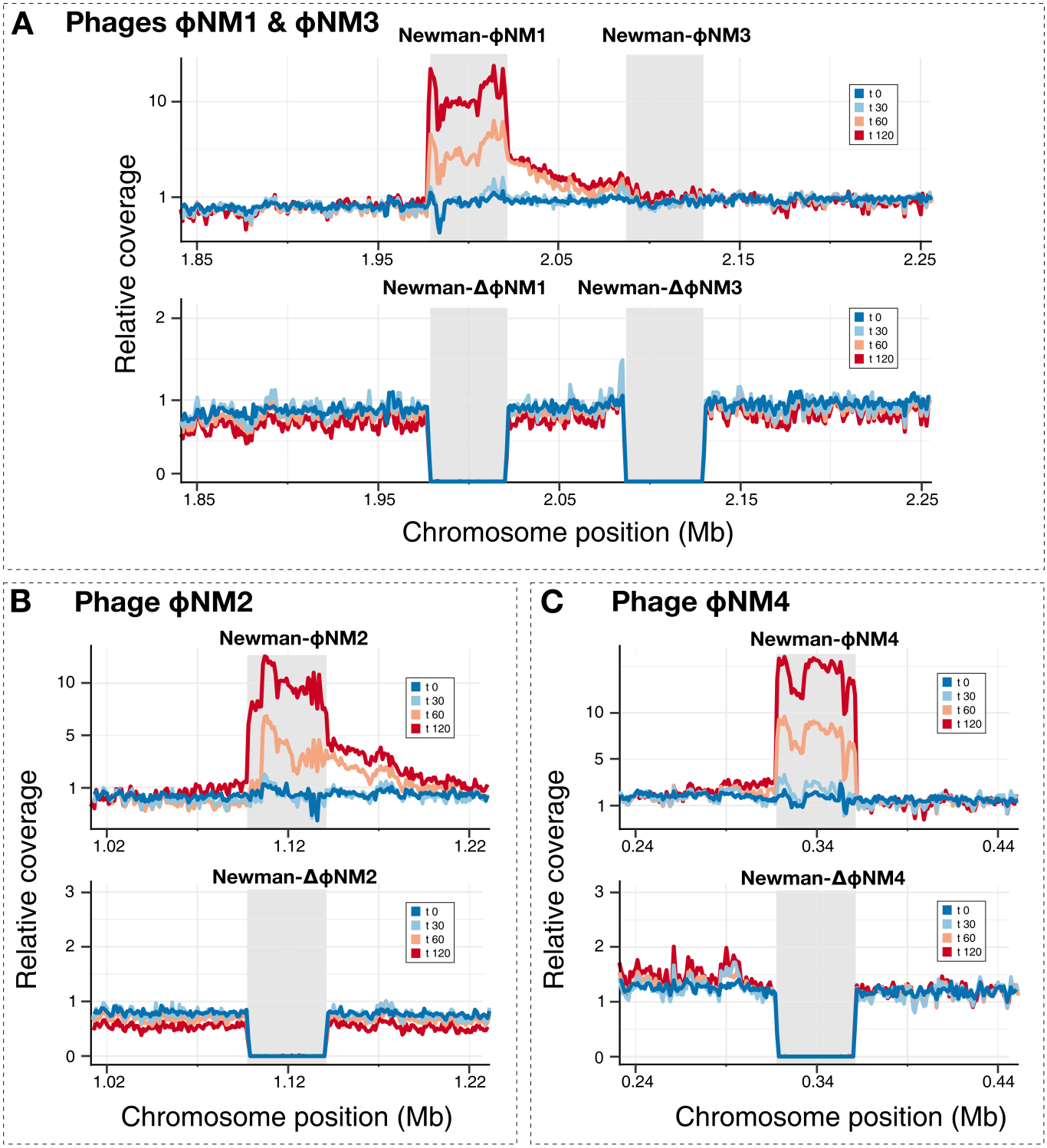


Figure S11. Staphylococcal phages replicate *in situ* before excision.

Relative abundance of phage genomic DNA and the chromosomal regions proximal to where they integrate for **(A)** ϕNM1 and ϕNM3, **(B)** ϕNM2, and **(C)** ϕNM4. Samples were analysed at 0 (blue), 30 (light blue), 60 (pink) and 120 minutes (red) post-induction with mitomycin C. Shaded grey rectangles represent the location of the prophages in the *S. aureus* chromosome.

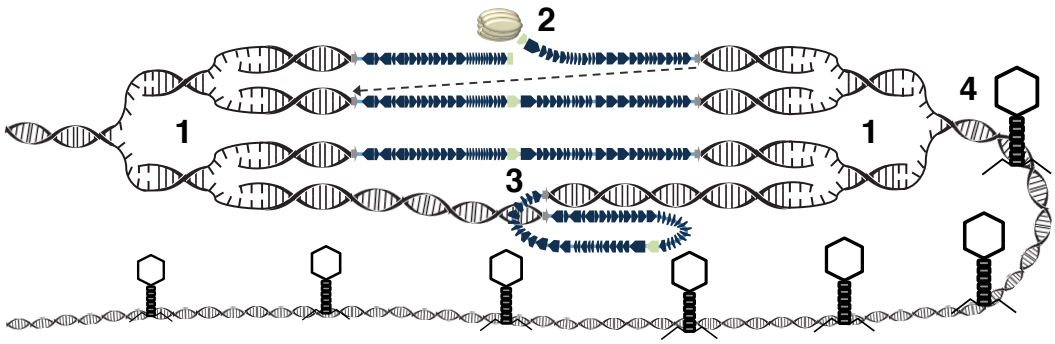


Figure S12. Model for lateral transduction.

The induced prophage genome (blue) replicates *in situ* bidirectionally (1) and amplifies the adjacent bacterial DNA (gray). Phage terminase initiates *in situ* DNA packaging from a *pac* site (2), while prophage excision occurs (3) in parallel. Prophage excision produces infectious phage particles. Lateral transducing particles are completed (4) when a headful capacity is reached. The DNA packaging machinery then proceeds to package additional headfuls of bacterial chromosome for several hundred kilobases. An arrow indicates a possible intramolecular integration joining two prophages that could result in an infectious particle by *in situ* packaging.

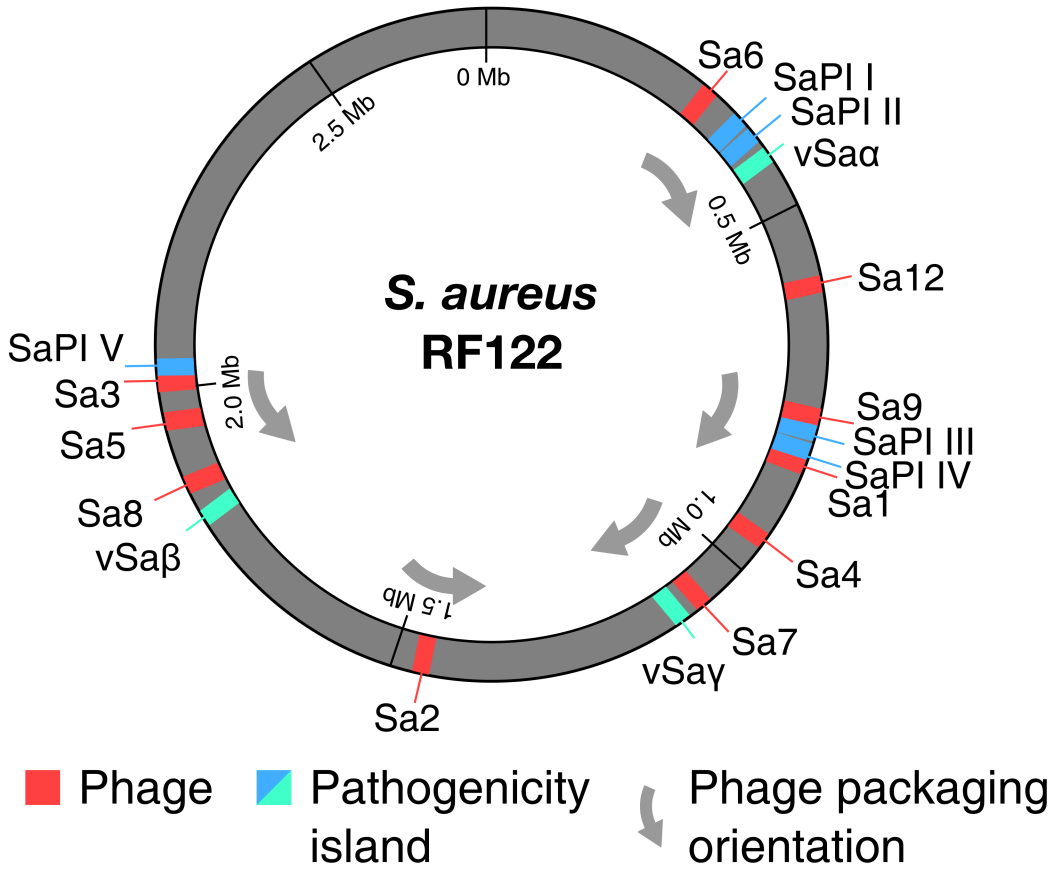


Figure S13. Lateral transduction influences genomic architecture.

The common *attB* sites of staphylococcal phages and pathogenicity islands mapped on to the genome of *S. aureus* strain RF122. Phage packaging directionality is determined by the orientation of the *terS* gene.

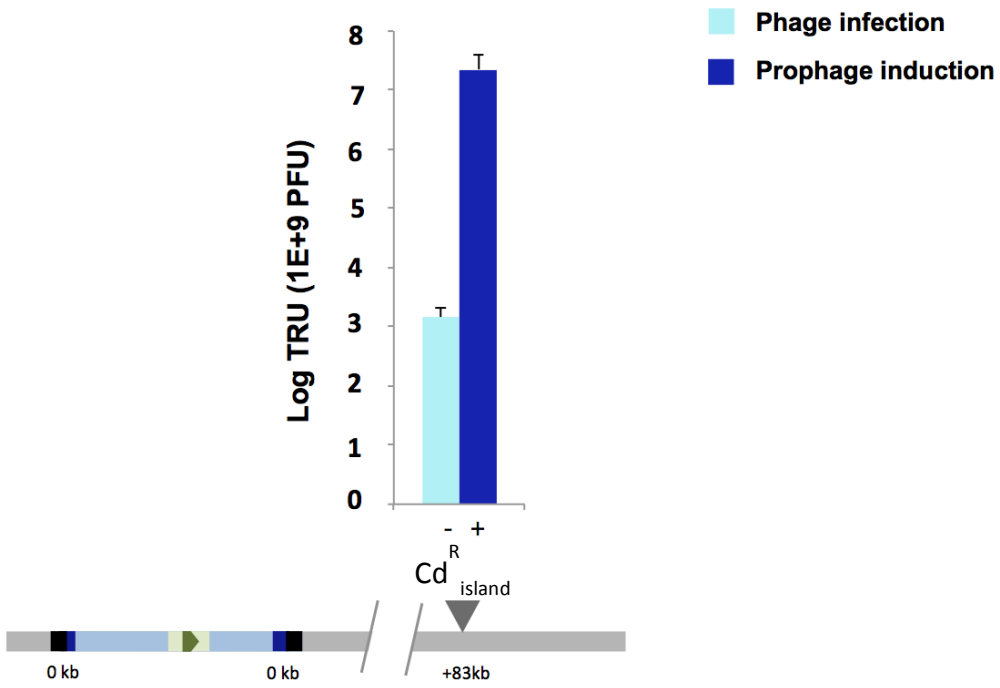


Figure S14. High frequency transfer of vSaα by lateral transduction.

Transfer of the chromosomal vSaα island after ϕ52A infection or lysogen induction. Non-lysogenic strains (light blue) were infected with ϕ52A or the lysogenic ϕ52A wt (dark blue) strains were induced with mitomycin C and the lysates tested for transduction into *S. aureus*. The Cd^R marker is shown inserted into the vSaα island. Transduction units (TrU) ml^{-1} were normalized by plaque forming (PFU) units ml^{-1} and represented as the TrU of an average phage titer (1E+9 PFU). Values are means ± SD (*n* = 3 independent samples).

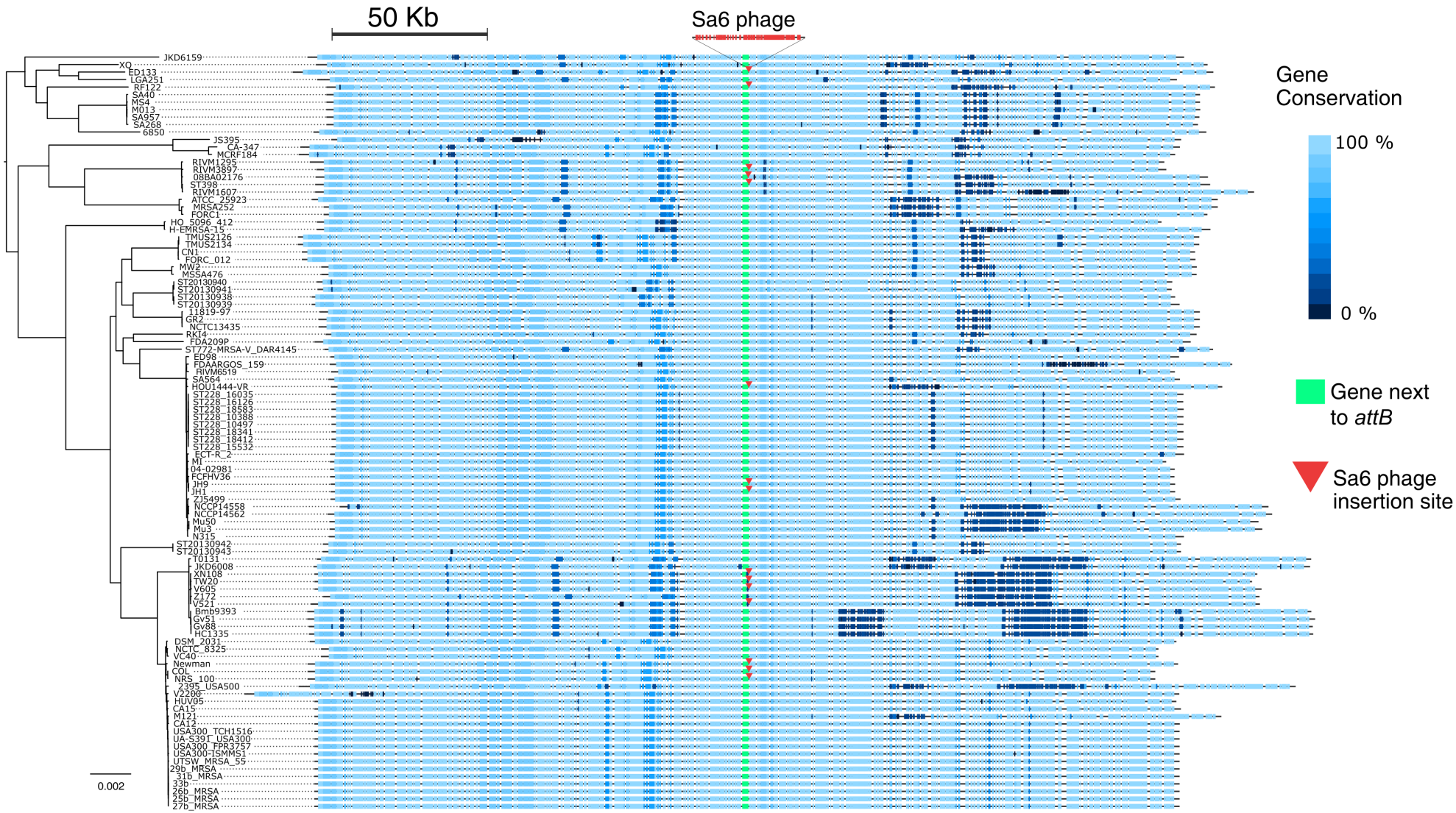


Figure S15. Genomic comparisons of the Sa6 phage insertion site flanking region.

Analysis of gene conservation by classifying clusters of homologous gene families using the algorithm OrthoMCL, marking the genes based on their degree of conservation. Genes located on the left side of the *attB* site are more conserved than those at the right. The dark blue genes found downstream the *attB* site correspond with the SaPI *attB* site type I, SaPI site type II and vSaα islands.

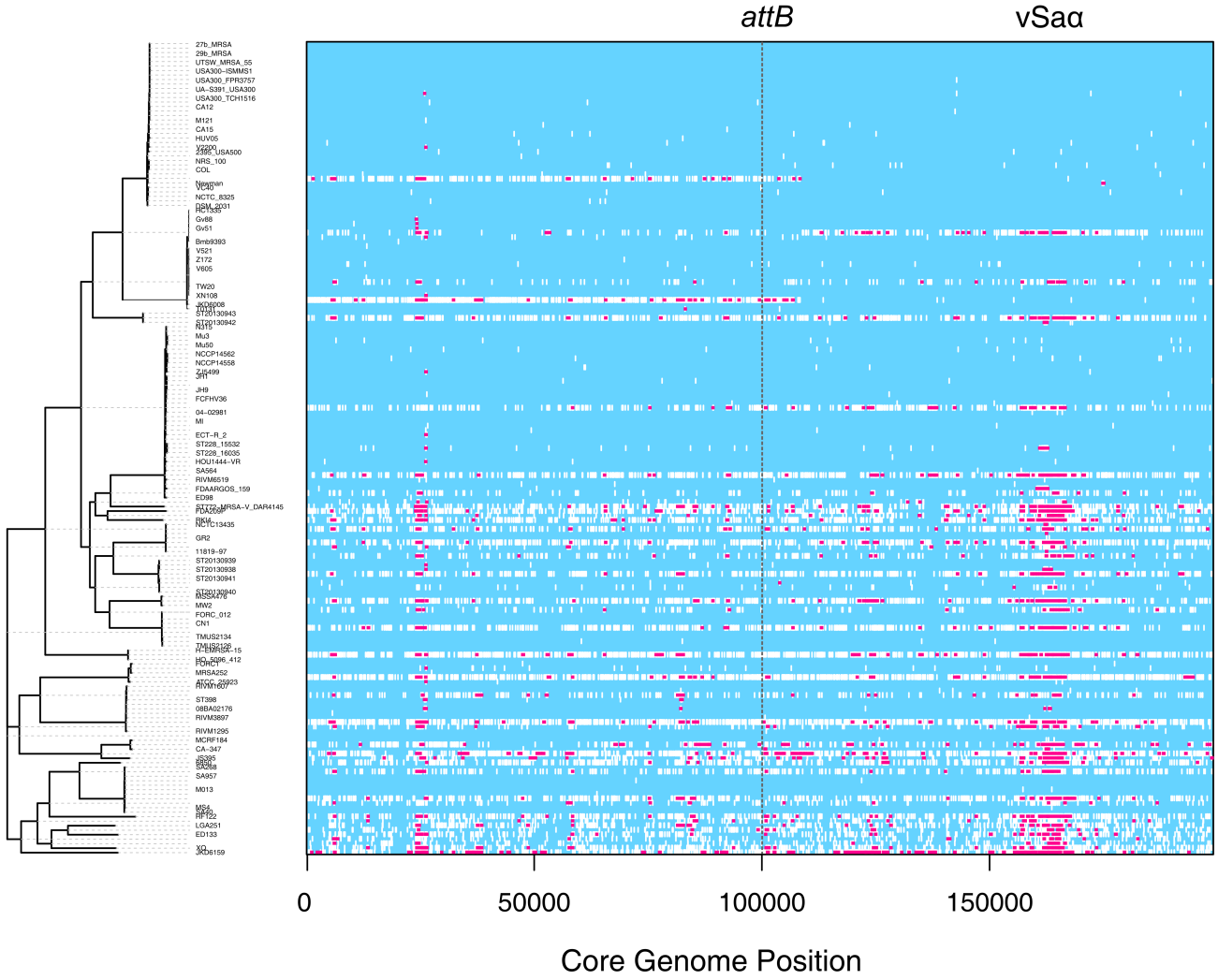


Figure S16. ClonalFrameML analysis of recombination of the Sa6 *attB* flanking region on the 200 Kb conserved across 101 *S. aureus* genomes.

The central line represents the *attB* related gene, the recombination events (red) were more frequent and much longer downstream the *attB* site (right side). 1169 recombination events were detected on 70 branches of the *S. aureus* phylogeny, with multiple importations of up to 3.7 Kb identified 60 Kb downstream the *attB* site (vSaα), suggesting this region is a recombination hotspot. Although some recombination was also identified upstream the *attB* site, we found that the number of recombination events and the number of recombinant bases downstream the *attB* site were significantly higher compared to the recombination identified upstream this site, suggesting it is due to the mobile genetic elements present in that area.

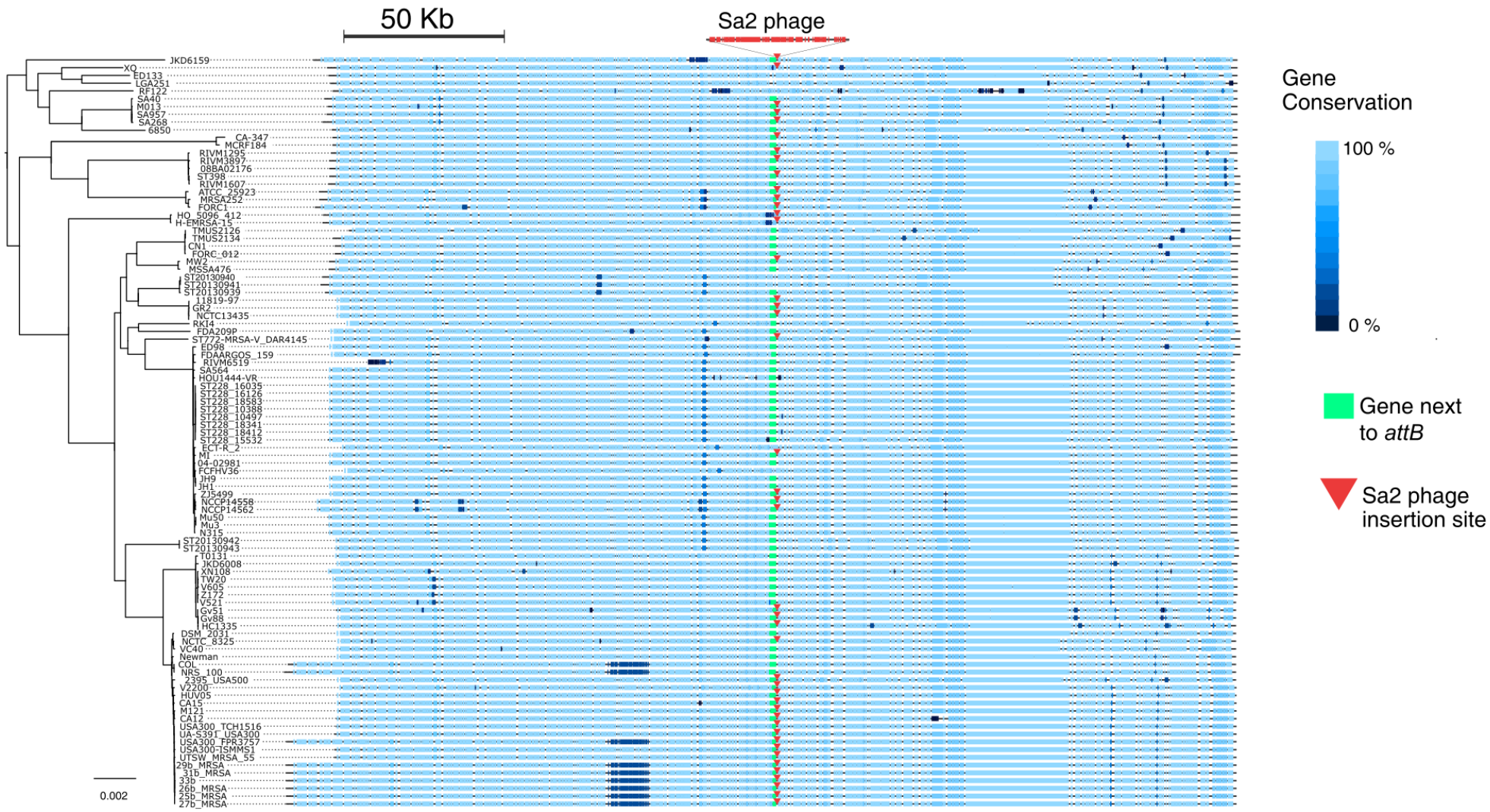


Figure S17. Genomic comparisons of the Sa2 phage insertion site flanking regions.

Analysis of gene conservation by classifying clusters of homologous gene families using the algorithm OrthoMCL, marking the genes based on their degree of conservation. Genes located on both the left and right side of the *attB* site are very conserved, but on the right side the highly recombinant *ebh* gene is present.

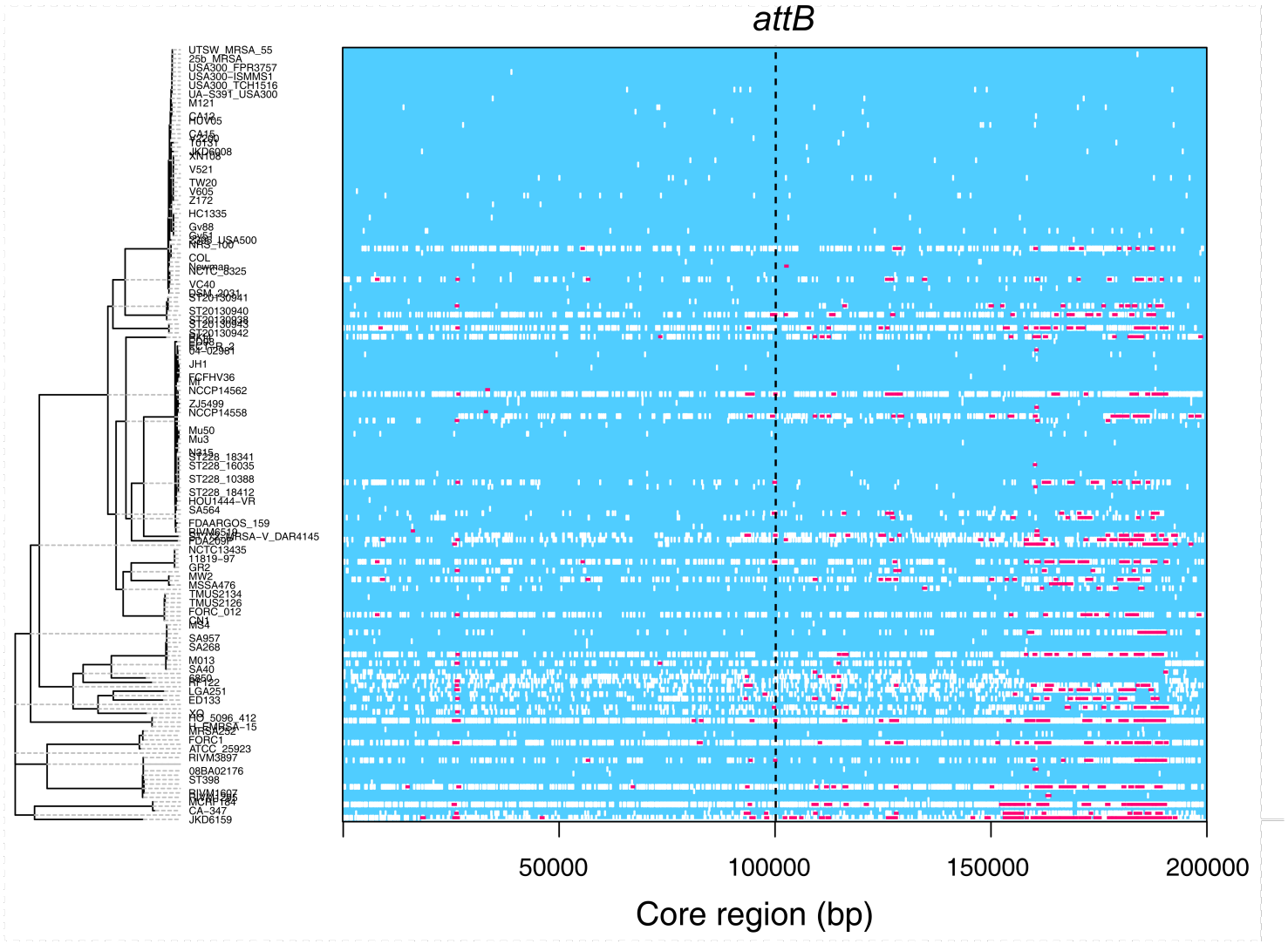


Figure S18. ClonalFrameML analysis of recombination of the Sa2 *attB* flanking region on the 200 Kb conserved across 101 *S. aureus* genomes.
With the central line representing the *attB* site, recombination events (red) were more frequent and much longer downstream the *attB* site (right side), suggesting that this region is a recombination hotspot (paired *Wilcoxon test* for recombinant events *p*-value = 1.308e-08, unpaired *Wilcoxon test* for recombinant nucleotides *p*-value = 5.514e-07). This core genome hotspot corresponds with the *ebh* gene, encoding the surface giant protein Ebh.

Table S1. Strains used in this study.

Table S2. Oligonucleotides used in this study.

Table S3. Plasmids used in this study.

Table S4. *S. aureus* strains used for the genomic analyses.