Bacteriophages encoding human immune evasion factors adapt to livestock-associated MRSA through rounds of integration and excision

Helena Leinweber¹, Raphael Sieber², Jesper Larsen², Marc Stegger², Hanne Ingmer^{1*}.

¹Department of Veterinary and Animal Sciences, University of Copenhagen, Stigbøjlen 4, 1870 Copenhagen, Denmark

²Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Artillerivej 5, 2300 Copenhagen, Denmark.

*Hanne Ingmer.

Email: hi@sund.ku.dk

ORCID numbers:

Helena Leinweber 0000-0002-9949-819X

Raphael Sieber 0000-0002-0372-128X

Jesper Larsen 0000-0003-0582-0457

Marc Stegger 0000-0003-0321-1180

Hanne Ingmer 0000-0002-8350-5631

Author Contributions:

H.L. and H.I. designed the study, H.L. generated experimental data, did formal analysis, wrote the manuscript and visualized the data; R.S. supported bioinformatic analysis; J.L. provided strain material; M.S. conducted sequencing; H.L., H.I., R.S., M.S. J.L. conducted review and editing; H.I. provided funding acquisition and project administration.

Competing Interest Statement: The authors declare no conflict of interest.

Keywords: Livestock MRSA, CC398, Sa3Int, phi13, Φ13, excision, alternative integration, tyrosine recombinase, attP, attB

This file includes:

27 Main Text, Figures 1, 2 and 3, Tables 1 and 2

28

29 Abstract

30 In recent years there has been an increase in human infections with methicillin-resistant
 31 *Staphylococcus aureus* (MRSA) originating from livestock and strains carrying bacterial viruses of
 32 the Sa3int-family have disseminated into the community. Sa3int phages express immune evasion
 33 factors and are common in human staphylococcal strains. As the bacterial attachment site (*attB*) for
 34 Sa3int phages is mutated in livestock-associated strains, the integration frequency is low and a key
 35 question is how the phages are established. Here we show that Sa3int phages adapt to alternative
 36 bacterial integration sites by mutating the phage attachment sequence, *attP*, leading to enhanced
 37 integration at these sites. Using a model strain carrying the mutated *attB_{LA}* of livestock-associated
 38 strains we find that once established, the Sa3int phage, Φ13 is inducible with release of
 39 heterogenous phage populations carrying mutations in *attP* that in part increase homology to
 40 alternative integration sites or *attB_{LA}*. Compared to the original phage, the adaptive mutations
 41 increase phage integration in new rounds of infection. Also, Sa3int phages induced from livestock-
 42 associated outbreak strains reveal mutated *attP* sequences. We suspect that promiscuity of the
 43 phage-encoded recombinase allows this adaptation and propose it may explain how phages
 44 mediate "host jumps" that are regularly observed for staphylococcal lineages.

45

46

47 Introduction

48 *Staphylococcus aureus* colonizes both humans and animals and its preference is associated with
 49 the content of mobile genetic elements¹. One example is the prophages of the Sa3int family. These
 50 bacterial viruses are found in most human strains of *S. aureus* where they express one or more
 51 immune evasion factors believed to facilitate human colonization as well as to promote human-to-
 52 human transmission^{2,3}. In contrast, the methicillin-resistant *S. aureus* found in livestock (LA-MRSA)
 53 commonly lack Sa3int phages^{4,5}. In fact, LA-MRSA of the CC398 lineage have been derived from
 54 human-associated strains which, subsequent to a jump from humans to animals, lost the Sa3int
 55 prophage⁴.

56 Despite host preference, there is a growing number of human infections with LA-MRSA and in
57 2019, they accounted for 32% of all new MRSA cases in Denmark (DANMAP, 2019). People with
58 occupational livestock contact are most at risk ⁶⁻⁸ and the infections appear equally severe as those
59 caused by human-associated strains ⁹. Although human infections with LA-MRSA are considered to
60 be the result of spillovers from livestock, there have been examples of transmissions between
61 household members as well as into community and healthcare settings ^{3,7,8}. Importantly, such
62 transfer events were associated with LA-MRSA strains carrying prophages of the Sa3int family
63 ^{3,7,8,10}. Since 95% of tested Danish pig herds are positive for LA-MRSA (DANMAP, 2019),
64 establishment of Sa3int phages in these strains may pose an increased risk of community spread of
65 LA-MRSA strains.

66 Integration of Sa3int phages in *S. aureus* occurs through orientation-specific recombination
67 between identical 14bp phage and bacterial core attachment sequences (*attP* and *attB*,
68 respectively) and is mediated by the phage-encoded tyrosine recombinase, Int ^{11,12}. In livestock
69 strains, the sequence corresponding to *attB* has two nucleotide changes as underlined 5'-
70 TGTATCCGAAITGG-3' (*attB_{LA}*). These substitutions do not alter the amino acid sequence of the β-
71 hemolysin encoded by *hly* in which *attB* is located, but significantly decrease the ability of Sa3int
72 phages to insert at this location ¹³. Accordingly, in LA-MRSA strains Sa3int prophages are mostly
73 located at variable positions in the bacterial genome but occasionally also in *attB_{LA}* ^{7,13-16}.

74 The ability of *S. aureus* to change its preference for human or animal hosts has been observed
75 several times. Such "host jumps" are thought to arise from "spillover" events where infections of
76 less preferred hosts are followed by host adaptation ultimately leading to colonization ^{7,17}. Host
77 adaptation often involves acquisition or loss of mobile genetic elements such as prophages ¹ but
78 little is known of the molecular events involved. Using massive parallel sequencing we have
79 examined Sa3int phages excised from alternative integration sites and find phage populations with
80 variable *attP* sequences of which a greater part increase resemblance to the bacterial attachment
81 sequence. Infections of naïve strains carrying the *attB_{LA}* site with such phage pools result in
82 increased phage integration. Our results explain how Sa3int phages, by adapting to alternative
83 integration sites in LA-MRSA strains, can establish in these strains that ultimately may be more
84 successful at colonizing and infecting humans and to disseminate in the human population.

85

86

87 **Results**

88 **Sa3int phages are adapting to alternative *attB* sites of LA-MRSA CC398**

89 In a recent study, 20 LA-MRSA CC398 strains from pigs and humans in Denmark were isolated and
 90 found to contain Sa3int prophages. In these strains the prophages were located at one of five
 91 different genomic locations (variant I-VI) ⁷ and the 14bp bacterial integration site carried two
 92 nucleotide mismatches (designated *attB_{LA}*) compared to the one found in human strains in other
 93 studies of LA-MRSA strains ^{13,15,16}. In the LA-MRSA CC398 genomes we determined the
 94 sequences flanking the prophage (*attL* and *attR*) and through comparisons with strains that lack the
 95 prophage, we deduced the corresponding *attB* sequences (Table 1). In all cases except one
 96 (variant V), the *attL* sequences differed from *attR*. This indicates non-matching *attB* and *attP* sites,
 97 as otherwise *attR* and *attL* would be identical, as seen with the original *attB*-site in *hlyB* of *S. aureus*
 98 8325-4.

99

100 To examine if mismatches between *attL* and *attR* affected excision of the prophage, we induced the
 101 lysogens with mitomycin and observed that in all strains the phages could be excised. From the
 102 resulting phages we determined the *attP* sequences using PCR amplification and Sanger
 103 sequencing (Table 1). For eight phages (one isolate of variant II, variant IV, variant V and five
 104 isolates of variant VI), the *attP* sequences were identical to that of the model Sa3int phage Φ13 ¹¹,
 105 showing that in these cases integration in the variant *attB* sites did not affect the *attP* sequence of
 106 the excised phage. In the remaining 12 phages however, mutations had arisen in the phage *attP*
 107 sequences. Importantly, in all cases the changes increased the sequence similarity between *attP*
 108 and the alternative *attB* site of the livestock-associated strains, as indicated in Table 1. These
 109 results indicate that Sa3int phages may be promiscuous with respect to both integration and
 110 excision and that integration of prophages at alternative bacterial attachment sites may alter the
 111 phage in such a way that its *attP* sequence bears greater resemblance to alternative *attB*
 112 sequences.

113

114 **Phage integration at multiple locations in a model strain carrying *attB_{LA}***

115 To examine the interactions between Sa3int phages and LA-MRSA strains in greater detail, we
 116 employed a derivative of *S. aureus* NCTC8325-4, designated *S. aureus* 8325-4attBmut, which
 117 contains 2bp point-mutations in *hly* to create the *attB_{LA}* of the LA-MRSA CC398 lineage¹³. With this
 118 strain we performed liquid infection with Φ 13kan^R, a derivative of the Sa3int phage, Φ 13 that
 119 encodes the staphylokinase (*sak*) but in which the immune evasion virulence genes *scn* and *chp*
 120 are replaced by the kanamycin resistance cassette *aphA3*¹³.

121 From eight independent lysogenization experiments we selected 22 lysogens being resistant to
 122 kanamycin. Alternative integration sites were confirmed for 20 of the lysogens by PCR (*hly*+, *sak*+) and two lysogens harbored the phage in the mutated *hly* site (*hly*-, *sak*+) (Supplementary Figure
 123 S1). The 22 isolates were whole-genome sequenced and analysis revealed 17 different integration
 124 sites for Φ 13kan^R in *S. aureus* 8325-4attBmut that were widely distributed across the bacterial
 125 chromosome (Supplementary Figure S2) and with the *attB* sequences listed in Figure 1. The
 126 integrations occurred in both non-coding and coding regions and were independent of
 127 transcriptional orientation.

129 When comparing the 14bp sequences of all alternative *attB* sites (Figure 1), they showed 29-86%
 130 homology compared to the original *attB* core sequence in the *hly* gene. However, the last three
 131 base pairs (5'-TGG-3') were highly conserved, being present in 20 out of 22 *attB* sites with lysogens
 132 6 and 20 being the exceptions. The nucleotides G at position 8 and T at position 11 signifying the
 133 *attB_{LA}* compared to *attB*, were not found in the same combination in any of the 17 *attB* sequences.
 134 Based on the conserved base pairs between the alternative *attB*-sites, we searched the
 135 chromosome of *S. aureus* NCTC8325 for the presence of 5'-NNNNNNCWNNTGG-3' (where W =
 136 A/T) and obtained more than 700 hits. Thus, there appears to be a multitude of potential integration
 137 sites in the staphylococcal genome.

138 Three of the alternative *attB* locations were observed as integration sites in lysogens obtained in
 139 independent lysogenization rounds, i.e. SAOUHSC_01067 CDS conserved hypothetical protein
 140 (lysogens 1,14 and 18), the intergenic region between open reading frames encoding hypothetical
 141 proteins SAOUHSC_01301 and SAOUHSC_01304 (lysogens 5 and 13) and SAOUHSC_00125

142 *cap5L* protein/glycosyltransferase (lysogens 10 and 21). As clonality can be excluded, these
 143 integration events show that there is some preference in selection of integration site when the bona
 144 fide *attB* sequence is mutated. However, when we screened the 300bp flanking regions of the
 145 alternative *attB* sites we found no common patterns in terms of sequence composition or distance of
 146 inverted repeats relative to the alternative *attB* core sequences (Supplementary Figure S3 and S4).
 147 Thus, it is still unclear why some integration sites are preferred over others.

148

149 **Phage evolution following excision from alternative integration sites**

150 Similar to what we had observed for Sa3int phages in livestock-associated strains, we found that
 151 mitomycin C induced $\Phi 13\text{kan}^R$ from all lysogens established in the 8325-4attBmut strain with the
 152 number of phage particles varying between 5×10^3 plaque-forming units (pfu)/ml and 4×10^6 pfu/ml
 153 (Supplementary Figure S5). This represents up to 1000-fold decrease in induction efficacy
 154 compared to the 6×10^6 pfu/ml obtained when the phage was induced from its integration site in the
 155 non-mutated *attB* of *S. aureus* 8325-4 (8325-4phi13kan^R control). Spontaneous phage release was
 156 also detected for many of the lysogens ranging from 2×10^1 to 3×10^3 pfu/ml compared to $1,0 \times 10^4$
 157 pfu/ml for the 8325-4phi13kan^R control (Supplementary Figure S5).

158 To examine the integration and excision process of $\Phi 13\text{kan}^R$ at the alternative integration sites, we
 159 determined the *attL* and *attR* from the genome sequences of the lysogens and deduced the
 160 alternative *attB* sites by comparing with sequences prior to integration of the phage. In addition, we
 161 determined the *attP* sequences by induction of the lysogens and amplicon sequencing of PCR
 162 products obtained on phage lysate with primers spanning *attP* (sequencing depth range 10.000-
 163 180.000, average 100.000).

164 For the majority of the lysogens (Table 2, part A), *attL* was identical to *attB*, and *attR* was identical
 165 to *attP* as can be observed by the bold red nucleotides marking the nucleotide differences in the
 166 alternative *attB* site sequences compared to the original *attB*. For these lysogens, the integration
 167 cross-over likely occurred at the 5'-TGG-3' (Supplementary Figure S6a). For the remaining
 168 lysogens (Table 2, part B), both *attL* and *attR* displayed sequences matching the alternative *attB*
 169 site with *attL* matching the 5'-end and *attR* the 3'-end. In these cases, the integration cross-over

170 events may have occurred at variable positions within the core sequences (Supplementary Figure
171 S6b).

172 When assessing *attP* by amplicon sequencing we observed remarkable sequence variation at
173 single nucleotide positions in more than 40% of the phage populations obtained from 9 of the
174 lysogens (Figure 2). When comparing these changes to the sequence of the bacterial integration
175 site from which the phage was derived, we saw that in five instances (lysogens 3, 10, 12, 17 and
176 21) the excised phages displayed adaptation to the alternative *attB* site by adopting a nucleotide of
177 the alternative *attB* sequence (Figure 2). Phages from lysogens 6, 7, 15, and 23 also displayed
178 single nucleotide substitutions in *attP* but without matching the alternative *attB* sequences. These
179 may result from mismatch repair during excision.

180 The adaptability of the phage to the alternative integration sites was even more pronounced when
181 all sequence variation >1% was scored (Figure 2). Importantly most of the excised phage pools
182 contained variants with sequence changes adopting the nucleotides of the alternative *attB*
183 sequences and multiple sequence variations occurred within the individual pools (Figure 2, green).
184 Notable exceptions were lysogens 1, 14 and 18, where no variants >1% were observed. In these
185 lysogens, $\Phi 13\text{kan}^R$ had independently integrated in the same *attB* site and despite 7 mismatches
186 with the 14 bp *attB* from 8325-4, resolution to the original *attP* sequence occurred with the same
187 precision as seen when $\Phi 13\text{kan}^R$ excised from *attB* of 8325-4 $\Phi 13\text{kan}^R$. In summary, our results
188 demonstrate that excision of $\Phi 13\text{kan}^R$ from alternative integration sites leads to evolutionary
189 adaptation of the phage to the bacterium by increasing the number of *attP* nucleotides matching the
190 alternative *attB* sequences.

191

192 **Phage adaptation to alternative *attB* site**

193 After observing that induction of phages at alternative integration sites led to mutated phage
194 populations with increased base pair matches between *attP* and the alternative *attB* sites or *attB_{LA}*,
195 we wondered whether these phages, in comparison to the original $\Phi 13\text{kan}^R$, had increased
196 preference for such sites in a new infection cycle. To address this we quantified integration by
197 qPCR with primer pairs covering *attR*. We examined phage pools obtained from lysogen 2 and 7
198 (designated $\Phi\text{lys}2$ and $\Phi\text{lys}7$) excised from the *attB_{LA}* and compared them to the original $\Phi 13\text{kan}^R$

with respect to integration in either 8325-4 or 8325-4attBmut (Figure 3). As expected, we found that for the wildtype, homogeneous $\Phi 13\text{kan}^R$ there were much less integration in *attB_{LA}* compared to *attB* that matches the *attP* sequence. In contrast, this difference was essentially eliminated for the Φlys2 and Φlys7 phage pools. Further, the mutations in these pools significantly increased the integration frequency in 8325-4attBmut when compared to $\Phi 13\text{kan}^R$ with the original *attP* site. Our results show that a single round of integration and excision dramatically increases the preference of the phage for an alternative or mutated attachment site.

Discussion

Sa3int prophages encode immune evasion factors and are found in most human strains of *S. aureus*^{18,19}. In contrast, LA-MRSA commonly lack Sa3int phages⁴ but when present, they increase the risk of transmission between household members and into the community^{3,7}. The integration site for Sa3int phages is naturally mutated in livestock-associated strains and so integration is infrequent and occurs at alternative sites often leading to mismatches between the *attL* and *attR* sequences. Intriguingly, we show that induction of these lysogens results in phage populations that are heterogeneous with respect to their *attP* sequences and with mutations that increase overall identity to the alternative bacterial integration site (Table 1,2 and Figure 1). Importantly, these *attP* changes increase phage integration into the naïve 8325-4attBmut in a new round of infection (Figure 3). Further we find that Sa3int prophages are spontaneously released from alternative integration sites highlighting that environmental triggers are not necessary for dissemination of these phages. Thus, rounds of excision and integration are possible with the potential for phage adaptation in each round.

When examining Sa3int prophages from outbreak strains of LA-MRSA⁷ we observe a greater number of adaptive changes in the *attP* sites of the excised phages than from our model 8325-4attBmut strain. This suggests that adapted phages are circulating in the LA-MRSA CC398 population, a notion that is supported by a study of the Sa3int phage P282 from a *S. aureus* CC398 strain, where the *attP* sequence is identical to *attB_{LA}*¹⁵. Also, re-analysis of Sa3Int-prophages in

MRSA CC398 isolates from hospital patients in Germany¹⁶ revealed that in 10 out of 15 lysogens, the *attL* and *attR* sequences were identical to *attB_{LA}* (Supplementary Table S3) indicating that the prophages have adapted to the livestock-associated strains. This raises the question where phage adaptations may occur. Since about one in three humans is colonized with *S. aureus* of which the majority contains Sa3int phages, transfer can occur when humans, naturally colonized with *S. aureus*, are exposed to livestock-associated strains. Once established as a prophage in a livestock-associated strain, Sa3int phages will be released and, if adapted, will integrate more effectively than the original phage in the LA-MRSA population.

Integration at secondary sites has also been observed for other phages when the primary site is absent or mutated^{20–23}. Excision of phage λ from such as site resulted in substitutions in *attP*^{24,25} and in P2, the authors stated that the new *attP* region contained DNA from *attR*^{26,27}. Similar to Φ 13, these phages encode tyrosine recombinases^{12,22}. This family of recombinases catalyzes recombination between substrates with limited sequence identity²⁸. We propose that the adaptive behavior of Sa3int phages is depending on this promiscuity. As tyrosine-type recombinases are employed also by other *S. aureus* phages encoding virulence factors²⁹, the results presented here may provide a broader explanation for how phages adapt to new bacterial strains and thereby enable the host jumps that are regularly observed for staphylococci¹.

Materials and Methods

Strains and media. Phage-cured *S. aureus* 8325-4³⁰ and its mutant 8325-4 Φ 13attBmut¹³ (here termed 8325-4attBmut) containing the 2bp variation in *hly* were used as recipients and indicator strains for Φ 13kan^R. Twenty LA-*S. aureus* strains harboring a Sa3Int-phages were analyzed for their *attR* and *attL* composition⁷. Sequencing data available at <https://www.ebi.ac.uk/ena/browser/home> with identifiers listed in Supplementary Table S4. *S. aureus* S0385 (GenBank accession no. NC_017333) were used as a reference strain for analysis of sequencing data of the LA-strains. The prophage Φ 13kan^R carries the kanamycin resistance

257 cassette *aphA3*, which replaces the virulence genes *scn* and *chp* and was obtained by induction of
258 8325-4phi13kan^R ¹³. A full strain list is provided in Supplementary Table S1. Strains were grown in
259 tryptone soy broth (TSB, CM0876, Oxoid) and tryptone soy agar (TSA, CM0131, Oxoid). Top agar
260 for the overlay assays was 0,2 ml TSA/ml TSB. Kanamycin (30 µg/ml) and sheep blood agar (5%)
261 were used to select for lysogens.

262

263 **Lysogenization assay.** To obtain the phage stock, 8325-4phi13kan^R was grown to late exponential
264 phase (37°C, 200 rpm, OD₆₀₀=0,8), mixed with 2 µl/ml mitomycin C and incubated for another 2-4
265 hours. Phages were harvested by centrifugation for 5 min at 8150 x g and filtering the supernatant
266 with a 0,2 µm membrane filter. The lysogens were obtained as described previously, with slight
267 adjustments ³¹. In brief, Φ13kan^R was added at a multiplicity of infection MOI=1 to the respective
268 recipients, incubated 30 min on ice to allow phage attachment, the non-attached phages were
269 washed off and after another incubation for 30 min at 37°C allowing phage infection, the culture was
270 diluted and plated on TSA with 5% blood and 30 µg/ml kanamycin. After overnight incubation at
271 37°C, 20 colonies showing β-hemolysis and two colonies without β-hemolysis were isolated and
272 used for further analysis. Lysogens were derived from eight independent lysogenization
273 experiments resulting in lysogens 1-5 (experiment 1); lysogens 6 and 7 (experiment 2), lysogen 8
274 (experiment 3), 10 and 11 (experiment 4), 12 and 13 (experiment 5), 14 and 15 (experiment 6), 16-
275 19 (experiment 7) and lysogens 20-23 (experiment 8).

276

277 **Spot assay and phage propagation.** Phage lysates were serially diluted in SM-buffer (100 mM
278 NaCl, 50 mM Tris (pH=7,8), 1 mM MgSO₄, 4 mM CaCl₂) and spotted on recipient lawn of either *S.*
279 *aureus* 8325-4 for pfu determination. To obtain an even lawn, 100 µl of fresh culture (OD=1) were
280 added to 3 ml top agar and poured on a TSA-plate supplemented with 10mM CaCl₂. After solidifying
281 of the top agar, drops of 3 x 10 µl of each dilution were spotted on the lawn.

282

283 **Induction assay.** To determine the different levels of phage release, the 8325-4attBmut-lysogens
284 were grown to OD₆₀₀=0,8, and centrifuged after adding 2 µg/ml mitomycin C and further incubation

for 2 hours. The sterile-filtered supernatant was diluted and spotted on an overlay of 8325-4 consisting of 100 µl culture mixed with 3 ml top agar.

287

Whole-genome sequencing and bioinformatics analysis. Genomic DNA was extracted by using DNeasy Blood and Tissue Kit (Qiagen) and whole genome sequences were obtained by 251bp paired-end sequencing (MiSeq, Illumina) as described previously³². Raw data can be accessed at <https://www.ebi.ac.uk/ena/browser/home> with identifiers listed in Supplementary Table S5.

Genomes were assembled using SPAdes³³. Geneious Prime 2020.1.1 was used to determine phage integration sites (www.geneious.com). The locations and core sequences were determined by extracting short sequences from the assembled draft genomes of the lysogens lying adjacent to the prophage and mapping it to the annotated genome of *S. aureus* 8325 (GenBank accession no. NC_007795). Reads obtained by sequencing the PCR amplicons spanning *attP*, were mapped to the Φ13 reference genome (GenBank accession no. NC_004617) and SNPs were called applying a variant frequency threshold of 50%. WebLogo3 was applied to detect gapped motifs in the flanking regions of the alternative *attB* sites³⁴.

300

PCR and amplicon sequencing. Direct colony PCR was used to determine (i) the presence of the phage using *sak*-primers, (ii) the integrity of the *hly* gene using *hly*-primers and (iii) *attP* using *attP*-primers³⁵ if the phage had spontaneously excised and was present in its circular form. Primer sequences and cycling conditions are listed in Supplementary Table S2. For each reaction, a well-isolated colony was picked, suspended in 50 µl MilliQ-water, heat-lysed for 5 min at 99°C and briefly centrifuged. One µl was used as template. To determine *attP* of induced phages in lysates, 1 µl of a 1:10 dilution of phage lysate was used as template. Each single reaction mix was composed of 20,375 µl water, 2,5ml Taq polymerase buffer, 1 µl each of forward and reverse primers (10 µM), 0,5 µl dNTPs and 0,125 µl Taq polymerase (Thermo Fisher). PCR products were purified with GeneJET PCR purification kit (Thermo Fisher) and sequenced either using Sanger Sequencing (Mix2Seq, Eurofins Genomics) for the Sa3Int-phages deriving from the LA-MRSA strains or using Illumina MiSeq (sequencing depth varied from 10.000-180.000 (average 100.000)).

313

314 **qPCR assay.** DNA for use in the qPCR assay (LightCycler 96, Roche) was extracted using the
 315 GenElute Bacterial Genomic DNA kit (Sigma). The samples of interest were obtained by
 316 lysogenizing *S. aureus* 8325-4 and 8325-4attBmut with the respective phage (Φ 13kanR, Φ lys2 or
 317 Φ lys7) and plating 2 x 100 μ l of the culture on TSA supplemented with 30 μ g/ml kanamycin. After
 318 overnight incubation, the colonies were scraped off (approx. 10.000 colonies) and re-suspended in
 319 1 ml saline. Of this, 100 μ l were used directly in the first lysis step of the kit. DNA concentration was
 320 measured using a QubitTM (Invitrogen) and diluted to 1 ng/ml of which 5 μ l were used in the qPCR
 321 reaction, consisting of 3 μ l water, 10 μ l FastStart Essential DNA Green Master 2x, 1 μ l of each
 322 forward and reverse primers (10 μ M). Primer sequences and cycling conditions can be found in
 323 Supplementary Table S2.

324

325 **Data availability**

326 All genomic data used or produced in this study is deposited at the European Nucleotide Archive
 327 (<https://www.ebi.ac.uk/ena/browser/home>). Accession numbers and identifiers are listed in
 328 Supplementary Table S4 and S5.

329 Source data for the qPCR-assay and Sanger amplicon sequencing are provided with this paper.

330

331 **Acknowledgments**

332

333 We thank Henrike Zschach for her contribution to the bioinformatic analysis and the staff of the
 334 Danish reference laboratory for Staphylococci at Statens Serum Institut for typing and handling of
 335 study isolates.

336 This project has received funding from the European Union's Horizon 2020 research No. 765147

337

338

339

340

341 **References**

342 1. Richardson, E. J. *et al.* Gene exchange drives the ecological success of a multi-host

- 343 bacterial pathogen. *Nat. Ecol. Evol.* **2**, 1468–1478 (2018).
- 344 2. Sieber RN, Urth TR, Petersen A, Møller CH, Price LB, Skov, R. Phage-mediated immune
345 evasion and transmission of livestock-associated methicillin-resistant *Staphylococcus*
346 *aureus* in humans. *Emerg. Infect. Dis.* (2020).
- 347 3. McCarthy, A. J. *et al.* *Staphylococcus aureus* CC398 clade associated with human-to-human
348 transmission. *Appl. Environ. Microbiol.* **78**, 8845–8848 (2012).
- 349 4. Price, L. B. *et al.* *Staphylococcus aureus* CC398: Host adaptation and emergence of
350 methicillin resistance in livestock. *MBio* **3**, 1–6 (2012).
- 351 5. Cuny, C., Abdelbary, M., Layer, F., Werner, G. & Witte, W. Prevalence of the immune
352 evasion gene cluster in *Staphylococcus aureus* CC398. *Vet. Microbiol.* **177**, 219–223 (2015).
- 353 6. Goerge, T. *et al.* MRSA colonization and infection among persons with occupational
354 livestock exposure in Europe: Prevalence, preventive options and evidence. *Vet. Microbiol.*
355 **200**, 6–12 (2015).
- 356 7. Sieber, R. N. *et al.* Phage-mediated immune evasion and transmission of livestock-
357 associated methicillin-resistant *Staphylococcus aureus* in humans. *Emerg. Infect. Dis.* **26**,
358 2578–2585 (2020).
- 359 8. Sieber, R. N. *et al.* Genome investigations show host adaptation and transmission of LA-
360 MRSA CC398 from pigs into Danish healthcare institutions. *Sci. Rep.* **9**, 1–10 (2019).
- 361 9. Becker, K., Ballhausen, B., Kahl, B. C. & Köck, R. The clinical impact of livestock-associated
362 methicillin-resistant *Staphylococcus aureus* of the clonal complex 398 for humans. *Vet.*
363 *Microbiol.* **200**, 33–38 (2017).
- 364 10. Larsen, J. *et al.* Methicillin-resistant *Staphylococcus aureus* CC398 is an increasing cause of
365 disease in people with no livestock contact in Denmark from 1999-2011. *Euro Surveill* **20**,
366 (2015).
- 367 11. Coleman, D. *et al.* Insertional inactivation of the *Staphylococcus aureus* beta-toxin by

- 368 bacteriophage phi 13 occurs by site- and orientation-specific integration of the phi 13
369 genome. *Mol. Microbiol.* **5**, 933–939 (1991).
- 370 12. Esposito, D. & Scocca, J. J. The integrase family of tyrosine recombinases: Evolution of a
371 conserved active site domain. *Nucleic Acids Res.* **25**, 3605–3614 (1997).
- 372 13. Tang, Y. *et al.* Commercial biocides induce transfer of prophage Φ 13 from human strains of
373 *Staphylococcus aureus* to livestock CC398. *Front. Microbiol.* **8**, 1–11 (2017).
- 374 14. Kashif, A. *et al.* *Staphylococcus aureus* ST398 Virulence Is Associated With Factors Carried
375 on Prophage $\square$ Sa3. *Frontiers in Microbiology* vol. 10 (2019).
- 376 15. Kraushaar, B. *et al.* Acquisition of virulence factors in livestock-associated MRSA: Lysogenic
377 conversion of CC398 strains by virulence gene-containing phages. *Sci. Rep.* **7**, 1–13 (2017).
- 378 16. Alen, S. Van, Ballhausen, B., Kaspar, U. & Köck, R. Prevalence and Genomic Structure of
379 Bacteriophage phi3 in Human-Derived Livestock-Associated Methicillin-Resistant
380 *Staphylococcus aureus* Isolates from 2000 to 2015. 1–11 (2018).
- 381 17. Matuszewska, M., Murray, G. G. R., Harrison, E. M., Holmes, M. A. & Weinert, L. A. The
382 Evolutionary Genomics of Host Specificity in *Staphylococcus aureus*. *Trends Microbiol.* **28**,
383 465–477 (2020).
- 384 18. Sung, J. M. L., Lloyd, D. H. & Lindsay, J. A. *Staphylococcus aureus* host specificity:
385 Comparative genomics of human versus animal isolates by multi-strain microarray.
386 *Microbiology* **154**, 1949–1959 (2008).
- 387 19. Wamel, W. J. B. Van, Rooijackers, S. H. M., Kessel, K. P. M. Van, Strijp, J. a G. Van &
388 Ruyken, M. The Innate Immune Modulators Staphylococcal Complement Inhibitor and
389 Chemotaxis Inhibitory Protein of *Staphylococcus aureus* Are Located on β -hemolysin
390 converting Bacteriophages. *J. Bacteriol.* **188**, 1310–1315 (2006).
- 391 20. Shimada, K., Weisberg, R. A. & Gottesman, M. E. Prophage lambda at unusual
392 chromosomal locations. II. Mutations induced by bacteriophage lambda in *Escherichia coli*
393 K12. *J. Mol. Biol.* **80**, 297–314 (1973).

- 394 21. Shimada, K., Weisberg, R. A. & Gottesman, M. E. Prophage lambda at unusual
395 chromosomal locations. I. Location of the secondary attachment sites and the properties of
396 the lysogens. *J. Mol. Biol.* **63**, 483–503 (1972).
- 397 22. Serra-Moreno, R., Jofre, J. & Muniesa, M. Insertion site occupancy by stx2 bacteriophages
398 depends on the locus availability of the host strain chromosome. *J. Bacteriol.* **189**, 6645–
399 6654 (2007).
- 400 23. Bertani, G. & Six, E. Inheritance of prophage P2 in bacterial crosses. *Virology* **6**, 357–381
401 (1958).
- 402 24. Weisberg, R. A., Foeller, C. & Landy, A. Role for DNA homology in site-specific
403 recombination The isolation and characterization of a site affinity mutant of coliphage λ .
404 (1983).
- 405 25. Nash, H. A. Integration and Excision of Bacteriophage λ : The Mechanism of Conservative
406 Site Specific Recombination. *Annu. Rev. Genet.* **15**, 143–167 (1981).
- 407 26. Six, E. Specificity of P2 for prophage site. 1. On the chromosome of *Escherichia coli* strain
408 C2. *Virology* **29**, 106–125 (1966).
- 409 27. Yu, A., Bertani, L. E. & Haggård-Ljungquist, E. Control of prophage integration and excision
410 in bacteriophage P2: nucleotide sequences of the int gene and att sites. *Gene* **80**, 1–11
411 (1989).
- 412 28. Rajeev, L., Malanowska, K. & Gardner, J. F. Challenging a Paradigm: the Role of DNA
413 Homology in Tyrosine Recombinase Reactions. *Microbiol. Mol. Biol. Rev.* **73**, 300–309
414 (2009).
- 415 29. Oliveira, H. *et al.* Staphylococci phages display vast genomic diversity and evolutionary
416 relationships. *BMC Genomics* **20**, 1–14 (2019).
- 417 30. Novick, R. Properties of a cryptic high-frequency transducing phage in *Staphylococcus*
418 *aureus*. *Virology* **33**, 155–166 (1967).

- 419 31. Pleška, M., Lang, M., Refardt, D., Levin, B. R. & Guet, C. C. Phage-host population
420 dynamics promotes prophage acquisition in bacteria with innate immunity. *Nat. Ecol. Evol.* **2**,
421 359–366 (2018).
- 422 32. Leinweber, H. *et al.* Vancomycin resistance in *Enterococcus faecium* isolated from Danish
423 chicken meat is located on a pVEF4-like plasmid persisting in poultry for 18 years. *Int. J.*
424 *Antimicrob. Agents* **52**, (2018).
- 425 33. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes De
426 Novo Assembler. *Curr. Protoc. Bioinforma.* **70**, 1–29 (2020).
- 427 34. Crooks, G., Hon, G., Chandonia, J. & Brenner, S. WebLogo: a sequence logo generator.
428 *Genome Res* **14**, 1188–1190 (2004).
- 429 35. Goerke, C., Koller, J. & Wolz, C. Ciprofloxacin and trimethoprim cause phage induction and
430 virulence modulation in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **50**, 171–
431 177 (2006).
- 432

List of Figures and Tables

Table 1. Comparison of *attP*, *attB*, *attL*, and *attR* sites of Sa3int phages from 20 LA-MRSA CC398 isolates. Magenta underlined nucleotides indicate mismatches between the Φ 13kan^R *attP* core sequence (5'-TGTATCCAACTGG-3') and the *attB* site for each LA-MRSA CC398 isolate. Underlined nucleotides in green indicate putative adaptive changes in the *attP* site that mimic the *attB* site under the assumption that the phages contained the *attP* site core sequence upon integration into the different LA-MRSA CC398 genomes. Integration sites refer to annotated genes in *S. aureus* ST398 reference strain S0385 (GenBank accession no. NC_017333).

Table 2a and b. Comparison of *attP*, *attB*, *attL*, and *attR* sites of Φ 13kan^R in 8325-4attBmut. Magenta underlined nucleotides indicate mismatches between the *attB* site in 8325-4 and the *attB* site for each lysogen. Green underlined nucleotides indicate adaptive changes in the *attP* site that mimic the *attB* site. Blue underlined nucleotides indicate other changes in the *attP* site. The threshold for variant calling was set to 50%. Part A includes the lysogens, where *attL* matches *attB* and *attR* matches *attP*. Part B includes the lysogens where parts of *attL* and *attR* both match *attB* and *attP*.

Figure 1. Alternative integration sites of Φ 13kan^R in *S. aureus* 8325-4attBmut. The core *attB* sites are presented by color-coding of the different base pairs (A = yellow, C = dark green, T = light green, G = grey). The mutated base pairs in *hlyB* representing *attB_{LA}* in the recipient strain are indicated by a bold frame.^{a)} The percentages in the bottom row correspond to the proportions of conserved nucleotides in the 17 alternative *attB* sites found in the 22 lysogens with respect to the original *attB* in 8325-4.

Figure 2. Variant nucleotides and respective frequency (%) of the *attP* sequences after excision of the phage as determined by amplicon sequencing. The green shading of the percentages indicates adaptive changes in the *attP* site that mimic the respective alternative *attB* site. Blue shading of the percentages indicates other changes at this position. A hyphen (-) indicates that no nucleotide was detected at this respective position. Dots indicate conservation of the base pair compared to *attP* in Φ 13kan^R, the sequence of which is indicated in the first row. Note that for Φ 13kan^R, Φ lys1, Φ lys14 and Φ lys18 no variants with frequencies >1% were detected across the entire *attP* sequence.

Figure 3. Normalized Cq-values after qPCR assay for detection of integration of evolved phages. The normalized Cq-value (calculated by $2^{Cq(pta)-Cq(hlyB)}$) normalizes the cycle number of the gene of interest to the reference gene *pta*. Primers are identifying integration in *hlyB* (*attB* or *attB_{LA}*) for infection of 8325-4 (dark green) or 8325-4attBmut (light green) either with Φ 13kan^R or the evolved phages Φ lys2 and Φ lys7, previously integrated at *attB_{LA}*. Statistical analysis was carried out in GraphPad Prism 9.1.0, using Two-way ANOVA. *P* values: ns > 0,1234, *=0,0332, **=0,0021, ***=0,0002, ****<0,0001. Error bars represent standard deviation of three biological replicates with three technical replicates.

Figures and Tables

Table 1. Comparison of *attP*, *attB*, *attL*, and *attR* sites of Sa3int phages from 20 LA-MRSA CC398 isolates. Magenta underlined nucleotides indicate mismatches between the Φ 13kan^R *attP* core sequence (5'-TGTATCCAAACTGG-3') and the *attB* site for each LA-MRSA CC398 isolate. Underlined nucleotides in green indicate putative adaptive changes in the *attP* site that mimic the *attB* site under the assumption that the phages contained the *attP* site core sequence upon integration into the different LA-MRSA CC398 genomes. Integration sites refer to annotated genes in *S. aureus* ST398 reference strain S0385 (GenBank accession no. NC_017333).

Variant	I	II		III		IV	V	VI	
Integration site (CDS, locus tag)	Intergenic, between SAPIG_RS03725 and SAPIG_RS03730	<i>sfaC</i> SAPIG_RS11745		<i>sph</i> SAPIG_RS10795		<i>cidA</i> SAPIG_RS13630	<i>sph</i> SAPIG_RS10795	<i>CocE/NonD</i> SAPIG_RS13905	
No. isolates	5	1	1	2	2	1	1	5	2
<i>attP</i> Φ 13kan ^R	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG
<i>attB</i>	CTAGTCC <u>TT</u> ACTG <u>T</u>	<u>GT</u> TATCCAA <u>T</u> CTGG	<u>GT</u> TATCCAA <u>T</u> CTGG	TGTATCC <u>GA</u> <u>T</u> TGG	TGTATCC <u>GA</u> <u>T</u> TGG	<u>TACCGCTAACT</u> TGG	TGTATCC <u>GA</u> <u>T</u> TGG	<u>TT</u> TATC <u>GTTT</u> CTGG	<u>TT</u> TATC <u>GTTT</u> CTGG
<i>attL</i>	CTAGTCC <u>TT</u> ACTG <u>T</u>	<u>GT</u> TATCCAA <u>T</u> CTGG	<u>GT</u> TATCCAA <u>T</u> CTGG	TGTATCC <u>GA</u> <u>T</u> TGG	TGTATCC <u>GA</u> <u>T</u> TGG	<u>TACCGCTAACT</u> TGG	TGTATCC <u>GA</u> <u>T</u> TGG	<u>TT</u> TATC <u>GTTT</u> CTGG	<u>TT</u> TATC <u>GTTT</u> CTGG
<i>attR</i>	TGTATCC <u>TT</u> ACTG <u>T</u>	TGTATCCAA <u>T</u> CTGG	TGTATCCAA <u>T</u> CTGG	TGTATCC <u>GA</u> <u>T</u> CTGG	TGTATCC <u>GA</u> <u>T</u> CTGG	TGTATCCAAACTGG	TGTATCC <u>GA</u> <u>T</u> TGG	TGTATCCAAACTGG	TGTATCCAAACTGG
<i>attP</i> after excision	TGTATCC <u>TT</u> ACTG <u>T</u>	TGTATCCAA <u>T</u> CTGG	TGTATCCAAACTGG	TGTATCC <u>GA</u> <u>T</u> TGG	TGTATCC <u>GA</u> <u>T</u> CTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCC <u>TT</u> ACTG <u>T</u>
Adaptive changes in <i>attP</i>	2/7	1/3	0/3	2/2	1/2	0/7	0/2	0/5	2/5
Other changes in <i>attP</i>	0	0	0	0	0	0	0	0	0

Table 2a and b. Comparison of *attP*, *attB*, *attL*, and *attR* sites of Φ 13kan^R in 8325-4attBmut. Magenta underlined nucleotides indicate mismatches between the *attB* site in 8325-4 and the *attB* site for each lysogen. Green underlined nucleotides indicate adaptive changes in the *attP* site that mimic the *attB* site. Blue underlined nucleotides indicate other changes in the *attP* site. The threshold for variant calling was set to 50%. Part A includes the lysogens, where *attL* matches *attB* and *attR* matches *attP*. Part B includes the lysogens where parts of *attL* and *attR* both match *attB* and *attP*.

a

	Lysogen									
	2,7	1, 14, 18	5, 13	8	10, 21	11	16	17	19	22
<i>attP</i> Φ 13, <i>attB</i>	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG
<i>attB</i> _{LA}	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG
<i>attB</i> alternative	TGTATCCGAA <u>T</u> TGG	<u>GTTTCTC</u> CACTGG	<u>TTTTC</u> CTTCTGG	<u>TTCTCA</u> CACTGG	TGATTCCTAACTGG	<u>GTTTCTC</u> CACTGG	<u>GCTCTATC</u> ACTGG	<u>TTTCTCT</u> AACTGG	<u>GGAGTCCA</u> CCTGG	<u>CGATAC</u> TACACTGG
<i>attL</i>	TGTATCCGAA <u>T</u> TGG	<u>GTTTCTC</u> CACTGG	<u>TTTTC</u> CTTCTGG	<u>TTCTCA</u> CACTGG	TGATTCCTAACTGG	<u>GTTTCTC</u> CACTGG	<u>GCTCTATC</u> ACTGG	<u>TTTCTCT</u> AACTGG	<u>GGAGTCCA</u> CCTGG	<u>CGATAC</u> TACACTGG
<i>attR</i>	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG
<i>attP</i> after excision	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG
Adaptive changes in <i>attP</i>	0/0	0/7	0/6	0/4	1/3	0/6	0/8	0/5	0/6	0/6
Other changes in <i>attP</i>	0	0	0	0	0	0	0	0	0	0

b

	Lysogen						
	3	4	6	12	15	20	23
<i>attP</i> Φ 13, <i>attB</i> 8325-4	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAAACTGG
<i>attB</i> _{LA}	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG	TGTATCCGAA <u>T</u> TGG
<i>attB</i> alternative	<u>GGGGA</u> CTTAACTGG	<u>TCAAT</u> CTTAACTGG	<u>GCTTAT</u> CAAACTGA	<u>GTATCCAA</u> CTGG	<u>CCACTT</u> CACTACTGG	<u>ACATCAGAGT</u> CTGA	<u>TTTG</u> TCCAAAGTGG
<i>attL</i>	<u>GGGGA</u> CCAACTGG	<u>TCAAT</u> CAAACTGG	<u>GCTTAT</u> CAAACTGG	<u>GTATCCAA</u> CTGG	<u>CCACTT</u> CACTACTGG	<u>ACATCAGAGT</u> CTGG	<u>TTTG</u> TCCAAACTGG
<i>attR</i>	TGTATCC <u>T</u> AACTGG	TGTATCC <u>T</u> AACTGG	TGTATCCAACTGA	TGTATCCAACTGG	TGTATCCAACTGG	TGTATCCAACTGA	TGTATCCAAAGTGG
<i>attP</i> after excision	TGTATCCAAACTGG	TGTATCCAAACTGG	TGTATCCAACTGG	TGTATCCAACTGG	TGTATCCAACTGG	TGTATCCAACTGG	TGTATCCAACTGG
Adaptive changes in <i>attP</i>	0/5	0/5	0/6	1/3	0/6	0/10	0/4
Other changes in <i>attP</i>	0	0	1	0	1	0	1

Figure 1. Alternative integration sites of $\Phi 13\text{kan}^{\text{R}}$ in *S. aureus* 8325-4attBmut. The core attB sites are presented by color-coding of the different base pairs (A = yellow, C = dark green, T = light green, G = grey). The mutated base pairs in *hly* representing attB_{LA} in the recipient strain are indicated by a bold frame.^{a)} The percentages in the bottom row correspond to the proportions of conserved nucleotides in the 17 alternative attB sites found in the 22 lysogens with respect to the original attB in 8325-4.

<i>attP</i> Φ13kan ^R	T	G	T	A	T	C	C	A	A	A	C	T	G	G	Average sequencing depth
Φ13kan ^R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	150,000
Φlys1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	40,000
Φlys2	.	.	.	.	.	.	.	G (5)	T (2)	.	T (5)	.	.	.	150,000
Φlys3	.	.	.	.	.	.	.	T (48)	.	.	.	.	.	.	100,000
Φlys4	.	.	.	.	.	.	.	T (9)	.	C (8)	.	.	.	.	180,000
Φlys5	.	.	.	.	.	.	.	T (6)	T (5)	T (5)	.	.	.	.	70,000
Φlys6	.	.	.	.	.	.	.	T (55)	.	C (2)	.	.	.	.	90,000
Φlys7	.	.	.	.	.	.	.	G (2)	.	T (42)	T (2)	.	.	.	90,000
Φlys8	.	.	.	.	.	.	- (9)	- (9)	.	.	.	.	.	.	100,000
Φlys10	.	.	.	.	.	.	.	T (72)	.	.	.	.	.	.	150,000
Φlys11	.	.	.	.	.	T (2)	.	C (6)	.	T (37)	T (18) - (24)	G (11)	.	.	90,000
Φlys12	.	.	.	.	.	.	.	.	.	T (52)	.	.	.	.	160,000
Φlys13	.	.	.	.	.	.	.	T (4)	T (4)	T (4)	.	.	.	.	90,000
Φlys14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100,000
Φlys15	.	.	.	.	.	.	.	G (57)	.	C (21)	C (21)	.	.	.	160,000
Φlys16	.	.	.	.	.	.	A (13)	T (13)	C (12)	.	T (1)	.	.	.	100,000
Φlys17	.	.	.	.	.	.	.	T (46)	.	.	.	.	.	.	100,000
Φlys18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	90,000
Φlys19	.	.	.	.	.	.	.	C (1)	C (1)	T (1)	.	.	.	.	30,000
Φlys20	.	.	A (1)	T (1)	C (1)	.	G (3) A (3)	G (3)	G (4)	T (3)	T (2)	C (2)	.	.	10,000
Φlys21	.	.	.	.	.	.	.	T (50)	.	.	.	.	.	.	120,000
Φlys22	.	.	.	.	.	.	T (5)	.	C (5)	.	.	.	.	.	90,000
Φlys23	.	.	.	.	.	.	.	T (57)	.	G (21)	G (21)	.	.	.	150,000

Figure 2. Variant nucleotides and respective frequency (%) of the *attP* sequences after excision of the phage as determined by amplicon sequencing. The green shading of the percentages indicates adaptive changes in the *attP* site that mimic the respective alternative *attB* site. Blue shading of the percentages indicates other changes at this position. A hyphen (-) indicates that no nucleotide was detected at this respective position. Dots indicate conservation of the base pair compared to *attP* in Φ13kan^R, the sequence of which is indicated in the first row. Note that for Φ13kan^R, Φlys1, Φlys14 and Φlys18 no variants with frequencies >1% were detected across the entire *attP* sequence.

Figure 3. Normalized Cq-values after qPCR assay for detection of integration of evolved phages. The normalized Cq-value (calculated by $2^{Cq(pta)-Cq(hlb)}$) normalizes the cycle number of the gene of interest to the reference gene *pta*. Primers are identifying integration in *hlb* (*attB* or *attB_{LA}*) for infection of 8325-4 (dark green) or 8325-4attBmut (light green) either with $\phi 13kan^R$ or the evolved phages $\phi lys2$ and $\phi lys7$, previously integrated at *attB_{LA}*. Statistical analysis was carried out in GraphPad Prism 9.1.0, using Two-way ANOVA. *P* values: ns > 0,1234, * = 0,0332, ** = 0,0021, *** = 0,0002, **** < 0,0001. Error bars represent standard deviation of three biological replicates with three technical replicates.