Recombination between phages and CRISPR–cas loci facilitates horizontal gene transfer in staphylococci

Andrew Varble1, Sean Meaden2, Rodolphe Barrangou3, Edze R. Westra2 and Luciano A. Marraffini1,4*

CRISPR (clustered regularly interspaced short palindromic repeats) loci and their associated (cas) genes encode an adaptive immune system that protects prokaryotes from viral and plasmid invaders. Following viral (phage) infection, a small fraction of the prokaryotic cells are able to integrate a small sequence of the invader’s genome into the CRISPR array. These sequences, known as spacers, are transcribed and processed into small CRISPR RNA guides that associate with Cas nucleases to specify a viral target for destruction. Although CRISPR–cas loci are widely distributed throughout microbial genomes and often display hallmarks of horizontal gene transfer, the drivers of CRISPR dissemination remain unclear. Here, we show that spacers can recombine with phage target sequences to mediate a form of specialized transduction of CRISPR elements. Phage targets in phage 85, ΦNM1, ΦNM4 and Φ12 can recombine with spacers in either chromosomal or plasmid-borne CRISPR loci in Staphylococcus, leading to either the transfer of CRISPR-adjacent genes or the propagation of acquired immunity to other bacteria in the population, respectively. Our data demonstrate that spacer sequences not only specify the targets of Cas nucleases but also can promote horizontal gene transfer.

Bioinformatic analysis of CRISPR (clustered regularly interspaced short palindromic repeats)–cas (encoding CRISPR-associated protein) loci have uncovered hallmarks of horizontal gene transfer and CRISPR–Cas modularity. Phylogenies based on either CRISPR repeats or the universal Cas1 protein revealed poor correlations between bacterial species trees, suggesting evidence of HGT of CRISPR–cas loci between distantly related bacterial species. Furthermore, genomic studies have suggested that CRISPR systems evolved from a common ancestor and co-opted a diverse set of effector modules, potentially via HGT and CRISPR–Cas modularity. Phylogenies based on either CRISPR repeats or the universal Cas1 protein revealed poor correlations between bacterial species trees, suggesting evidence of HGT of CRISPR–cas loci between distantly related bacterial species. Furthermore, genomic studies have suggested that CRISPR systems evolved from a common ancestor and co-opted a diverse set of effector modules, potentially via HGT and CRISPR–Cas modularity.

We explored whether recombination between newly acquired spacers and their targeted phage could mediate transduction. It is well established that even short sequences with homology to a phage direct recombination events that integrate the homologous DNA into the viral genome, leading to a drastic enhancement of the transduction rates of plasmids, for example. The acquisition of a 30–40 base pair spacer sequence from the infecting virus during the CRISPR–Cas adaptive immune response would introduce homology between the phage genome and the CRISPR locus, and could facilitate recombination and elevated rates of transduction. Such a mechanism would lead to the transduction of genes adjacent to the CRISPR locus (Fig. 1a). To test this, we added an erythromycin resistance gene adjacent to the chromosomal type III-A system within the methicillin resistance cassette (SCCmec) of Staphylococcus aureus 08BA02176 (ref. 28), and inserted a target site for the first spacer of this CRISPR locus at two locations (P1 and P2) of the staphylococcal phage 85 genome or a new spacer (C1) matching orf28 (Supplementary Fig. 1A). This erythromycin-resistant strain was infected with each of the three phages or a non-targeted, wild-type phage as a control, and the lysates were used to transduce the marker into the wild-type strain. We observed that spacers P1 and C1 enhanced transduction of the antibiotic resistance cassette by one order of magnitude over the non-targeting control (Fig. 1b). As there are no differences in the viability of the recipient cells (Supplementary Fig. 1B), these results suggest that recombination between spacers P1 and C1 and the phages harboring their targets can direct the transfer of genomic locations adjacent to the CRISPR locus at rates that exceed those observed for generalized transduction (mediated by the non-targeted control phage). We also investigated the possibility of spacer-mediated transduction of entire chromosomal CRISPR–cas loci to CRISPR-lacking strains. Using two different empirical systems, the type I-F CRISPR–Cas system of Pseudomonas aeruginosa and the DMS3vir phage and the type II-A CRISPR–Cas system of Streptococcus thermophilus and the 2972 phage, with adjacent chromosomal markers to track transduction (Supplementary Fig. 1C), we observed generalized but not spacer-mediated transduction of the entire CRISPR–Cas system (Supplementary Fig. 1D,E). Most likely, this is a result of the presence of only one region of homology for integration (Supplementary Fig. 1C).
Although most CRISPR–Cas systems reside on chromosomes, an important fraction of CRISPR loci has been reported to be carried on plasmids. Plasmid-borne CRISPR loci offer unique advantages for their lateral transfer via spacer recombination: the increased copy number elevates the probability of recombination and the circular nature allows the insertion of the entire CRISPR–cas locus into the plasmid genome, facilitating its packaging and recircularization into the recipient host (Fig. 1c). To explore this, we first tested whether the transduction of CRISPR-carrying plasmids can be mediated by the newly acquired spacers. We infected *S. aureus* RN4220 (ref. 31) cells (which lack endogenous CRISPR–cas loci) carrying the type II-A locus of *Streptococcus pyogenes* (Supplementary Fig. 2A) into the 2.9-kb staphylococcal plasmid pCI94 (pCRISPR, conferring chloramphenicol resistance), with a staphylococcal *pac* phage carrying a virulent mutation, *ΦNM4y4* (ref. 31). Staphylococcal *pac* phage carrying pCRISPR, but not an empty vector control, recovered at 12 h (Fig. 1d) through the acquisition of new spacers (Supplementary Fig. 2B), at the same time as the phage titres began to decline (Fig. 1d). We then checked for the presence of pCRISPR-transducing particles in phage filtrates (which also contain infective phages) by infecting *S. aureus* RN4220 recipients and selecting for chloramphenicol-resistant colonies. We detected an increase of transduction starting at 16 h post-infection, with a peak frequency of 1 transduced colony per 10^4 plaque-forming units (p.f.u.) in the filtrates (Fig. 1d and Supplementary Fig. 2C–F). To determine whether transduction could transfer expanded pCRISPR adaption were determined by mixing cells at a 1 CRISPR*+* (naive, without a targeting spacer) to 5 CRISPR*–* ratio and infecting the mix with *ΦNM4y4* at an MOI of 1. Cultures were collected 20 h post-infection, and survivors resulting from the acquisition of new spacers or the transduction of the adapted pCRISPR were measured by enumerating colonies on plates containing different antibiotic combinations. As a control, the experiment was repeated after mixing non-CRISPR resistant cells with naive CRISPR*+* staphylococci. Means ± s.d. of three biological replicates are reported. ND, not detected.

---

**Fig. 1** Transfer of CRISPR–Cas elements through spacer-mediated transduction. **a**, Transduction of CRISPR-adjacent loci (blue arrow) after phage recombination with a chromosomal CRISPR–cas locus. The red arrows indicate the common adjacent locus. PS, protospacer; S, CRISPR spacer; R, repeat. **b**, Transducing particle production from *S. aureus* strain 08BA02176 tagged with an erythromycin resistance cassette. Liquid cultures were infected at an MOI of 50. Wild-type phage (ctl) or phage with the target site (trgt) of the first spacer of the CRISPR array inserted at 15 kb (P1) or 20 kb (P2) positions on the phage genome were used. Wild-type phage was also used to infect a strain with a chromosomally inserted phage-targeting spacer (C1). Means ± s.d. of three biological replicates are reported. Two-tailed unpaired t-test was used to calculate P values: *P* = 0.023, ****P* = 0.00058. **c**, Transduction of a plasmid containing a CRISPR-cas locus after recombination with the phage genome. J1 and J2, recombination junctions. **d**, Growth and phage titres of infected cultures of bacteria containing plasmids with either the type II-A CRISPR–Cas system (CRISPR*–*) or the empty vector control (CRISPR*+*). Liquid cultures were infected at an MOI of 1 with *ΦNM4y4*. The growth of cultures was determined by measurement of absorbance at 600 nm (A_{600}). Titres, p.f.u/ml^−1, and levels of transducing-immune phage particles, colony-forming units per ml (c.f.u. ml^−1), were determined at every time point. No transducing-immune particles were detected using a vector control. Means ± s.d. of three biological replicates are reported. **e**, Levels of transduction during CRISPR adaption were determined by mixing cells at a 1 CRISPR*+* (naive, without a targeting spacer) to 5 CRISPR*–* ratio and infecting the mix with *ΦNM4y4* at an MOI of 1. Cultures were collected 20 h post-infection, and survivors resulting from the acquisition of new spacers or the transduction of the adapted pCRISPR were measured by enumerating colonies on plates containing different antibiotic combinations. As a control, the experiment was repeated after mixing non-CRISPR resistant cells with naive CRISPR*+* staphylococci. Means ± s.d. of three biological replicates are reported. ND, not detected.
compared the spacer repertoires of CRISPR-resistant and CRISPR-transduced cells, obtained 20h post-infection, using next-genera-
tion sequencing. Spacer sequences from four biological replicates
were mapped onto the ΦNM4γ4 genome and plotted against their
average number of reads (Fig. 2a and Supplementary Data File 1).
The relative frequency of transduced spacers was consistent in each
experiment (Supplementary Fig. 3A) and we did not detect a corre-
lation between the frequency of spacer acquisition and transduction
efficiency, measured as the ratio of transduced c.f.u. to the total number of p.f.u. of phage in the lysates obtained after infection of staphylococci harbouring pCas9 carrying the L, I, H1 and H2 spacers, or no targeting spacer as a control. Means ± s.d. of three biological replicates are reported. Two-tailed unpaired t-test was used to calculate P values: NS, not significant (P = 0.1554), *P (H1) = 0.024, *P (H2) = 0.024. c. Same as b, but measuring the transduction of pSpacer plasmids, that is, not carrying cas9. Means ± s.d. of three biological replicates are reported. One-way analysis of variance (ANOVA) was used to calculate P values: NS (P = 0.59). d, PCR products after amplification of J1 and J2 junctions (Fig. 1c) either from lysates (phage DNA) or infected cells (cell DNA) obtained after infection of cells harbouring pCas9 (+Cas9) or pSpacer (−Cas9) plasmids with L, I, H1 or H2 spacers. These results are representative of three independent experiments. e, Next-generation sequencing of phage DNA harvested after infection of cells containing pCas9(H1). Reads were aligned to the putative J1 junction. The RPM for each nucleotide within this region are shown. The dotted lines indicate a 75-nucleotide sequence that is unique to the junction. Results from a single experiment are shown. f. Same as e, but for the J2 junction. Results from a single experiment are shown.

Fig. 2 | Spacer sequences determine the frequency of pCRISPR transduction. a. Cultures containing pCRISPR were infected with ΦNM4γ4 at an MOI of 1. Expanded CRISPR arrays were analysed by next-generation sequencing. The reads per million (RPM) values of the acquired spacers were plotted against the ΦNM4γ4 genome (blue line, CRISPR resistant). Lysates containing ΦNM4γ4 as well as pCRISPR-transducing particles were collected at 20 h post-

infection and were used to infect cells without CRISPR-cas at an MOI of 1. Cells were plated to collect pCRISPR transductants and their spacer content was analysed by next-generation sequencing (green line, CRISPR transduced). The ratio of transduced spacers over resistant spacers was also plotted for each acquired spacer (solid black line). The horizontal dashed black line indicates a 1/1 transduced/resistant ratio. Select spacers with L and high (H1 and H2) ratios, along with the phage poc site, are indicated. The mean of four replicates is reported. b, Transduction efficiency of immunity, measured as the ratio of transduced spacer-repeat unit, without Cas9 targeting. If pCRISPR transduction occurs through the formation of...
recombinants between the acquired spacer and the phage protospacer sequence, these recombinants will maintain a full target in one of the recombination junctions that could be cleaved by Cas9. To investigate this, staphylococci carrying the pSpacer or pCas9 plasmids were infected with \( \Phi NM4\gamma4 \), and DNA was isolated from bacterial pellets (containing infected cells) or culture supernatants (containing virions) for PCR amplification of both recombination junctions (Fig. 2d) as well as chromosomal and viral genes as controls for the fractionation (Supplementary Fig. 4C). pSpacer–phage recombinants were detected at both junctions, for all spacer sequences, both in infected cells and in virions, a result that explains the equal transduction levels of these plasmids (Fig. 2c). By contrast, pCas9–phage recombinants were also detected, but PCR products were much less abundant for the targeted junction (J2) in infected cells (Fig. 2d). In virions, we only detected the non-targeted junction (J1) for the constructs containing H1 and H2 spacers, along with a faint PCR product for the I spacer construct; however, we were unable to detect the protospacer adjacent motif (PAM)-flanked junction (J2) (Fig. 2d). These results were corroborated by next-generation sequencing of DNA extracted from the virion fraction after infection of cells containing pCas9(H1). We found abundant reads spanning the non-targeted phage–CRISPR junction (J1) (Fig. 2c), whereas the targeted junction (J2) had relatively fewer reads (Fig. 2f). Altogether, these experiments demonstrate two important aspects of spacer-mediated recombination. First, recombination between the spacer sequence in pCRISPR and the protospacer sequence in \( \Phi NM4\gamma4 \) results in the formation of hybrid DNA molecules, which can be encapsidated into virions during infection. This does not depend on the host RecA (Supplementary Fig. 5A) and can also occur via the staphylococcal cos phage \( \Phi 12 \) (Supplementary Fig. 5B). Second, Cas9 targeting of the PAM-flanked spacer–protospacer junction within these molecules reduces the efficacy of their packaging into viral capsids and therefore the efficiency of transduction.

The presence of phage particles containing spacer–phage recombinants from infected CRISPR-immune cells suggests that incomplete protection of the host allows for the formation and release of the CRISPR-containing virions. Inefficient CRISPR immunity can result from at least two scenarios. One possibility is that the immunity provided by the acquired spacer can be bypassed by phages containing target mutations, known as ‘escapers’; in this case, the spacers that target regions with a high rate of mutation in the phage genome will be more prone to transduction. However, the experiments described in Supplementary Fig. 6 ruled out this scenario. A second possibility is that the acquired spacer provides only partial immunity, that is, a proportion of the adapted cells can be lysed by the phage and produce CRISPR-transducing particles. To test this, we measured immunity using an assay that reproduces the high multiplicity of infection (MOI) faced by cells that acquire new spacers\(^\text{31}\). In these conditions, CRISPR–Cas systems programmed with the L and I spacers enabled the complete recovery of the host, and cells containing H1 and H2 spacers showed only a partial recovery of the infected staphylococci (Fig. 3a). To determine the strength of the immunity mediated by all of the spacers present in the CRISPR-transducing particles (not just H1 and H2), we followed the survival of naive staphylococci upon infection with phages collected during the CRISPR–Cas immune response (Fig. 1d, 22 h time point), which contains both \( \Phi NM4\gamma4 \) as well as transducing particles that can provide immunity against the phage. We obtained similar partial survival curves to those provided by the H1-containing and H2-containing CRISPR–Cas systems (Fig. 3b). These results suggest that the complete destruction of the virus by the I-transducing and L-transducing spacers limits the formation of CRISPR-transducing particles. If true, these spacers should have a dominant effect on the H1 and H2 spacers, that is, they will reduce their frequency of transduction. To test this, we combined different pairs of pCRISPR plasmids (with different antibiotic-resistance markers) in the same cell, infected them with \( \Phi NM4\gamma4 \) and counted the number of transductants (Fig. 3c). We found that the combination of CRISPR–Cas systems harbouring I or L spacers with either H1 or H2 spacers resulted in a low transduction frequency, that is, the effect of I and L spacers predominates over H1 and H2 spacers. Similarly, a reduction in the transduction of a second plasmid (pE194, 2.9 kb) is observed in fully protected cells containing I and L spacers when compared to partially immune cells with H1 and H2 spacers (Supplementary Fig. 7A). Finally, we looked at the transduction of pCRISPR plasmids harbouring an inactivating mutation in Cas9 (dCas9)\(^\text{3}\). Corroborating our hypothesis, the reduction in immunity caused by this mutation enhanced the transduction rate for spacers I, H1 and H2 (Supplementary Fig. 7B,C). Altogether, these results demonstrate that spacer–phage recombination is the primary driver of spacer-enhanced transduction and that spacer sequences mediating highly efficient CRISPR immunity prevent the transduction of the CRISPR–cas locus.

Here, we show that spacers acquired by CRISPR–Cas systems can perform a form of specialized transduction that requires their recombination with the phage target as well as incomplete CRISPR immunity (Fig. 3d). The recombination between the spacer and its viral target connects the locus with the packaging sequences on the phage genome, enabling foreign DNA uptake at much higher rates than observed during typical generalized transduction in our experimental system. If the CRISPR locus resides in the host chromosome, this recombination can mediate the transfer of genes adjacent to the CRISPR locus and thus facilitate the dissemination and exchange of cas modules. If the CRISPR–Cas system resides in a circular genetic element, spacer-mediated recombination leads to the spread of CRISPR immunity among naive CRISPR-negative hosts. Interestingly, the host RecA is not required for protospacer–spacer recombination, a result suggesting that this phenomenon is mediated by the phage’s own machinery, which can significantly elevate recombination rates\(^\text{12,13}\). Our analysis of four different spacers showed that transduction rates are higher for spacers that mediate poor cleavage of the pCRISPR–\( \Phi NM4\gamma4 \) recombinants. We believe that the next-generation sequencing experiment shown in Fig. 2a, which includes data for all transduced spacers, supports this correlation: the sequences of the spacers mediating the lowest transduction rates (5–20 kb of the \( \Phi NM4\gamma4 \) genome) are located around the \( \text{pac} \) site, which a previous study showed to be one of the regions of this phage best targeted by Cas9 (ref. \(^\text{39}\)).

Our data show that both types of spacer-enhanced transduction events that we describe here occur at low frequencies, and in nature, could be happening at even lower frequencies. However, as it is the case with most situations involving horizontal transfer of genetic material, the importance of these events relies not so much in their rate of occurrence, but in their capacity to increase the genetic pool of the recipients\(^\text{31}\); given the appropriate environmental conditions, the genes and plasmids transferred though spacer-mediated transductions could provide a crucial selective advantage to the population. For example, the exchange of CRISPR–adjacent modules could expand the repertoire of cas genes of a CRISPR locus and generate the genetic diversity\(^\text{14–16}\) required to stay ahead in the arms race with different phages and their anti-CRISPR inhibitors\(^\text{36}\). Plasmids and potentially excisable (circular) genomic islands harbouring CRISPR–cas loci are relatively common\(^\text{17,37}\), and their spread though spacer-mediated transduction could provide critical spacers and/or full defence cassettes for phage defence. Even if the transduced CRISPR–cas locus does not harbour the most efficient spacers, as our data indicate, their spread could help to increase the spacer diversity necessary to prevent the raise in the number of phage escapers\(^\text{40}\) and/or provide partial defence to enable the acquisition of more potent sequences. Finally, it is worth noting that CRISPR–Cas systems have been identified within phage genomes\(^\text{11–12}\). Although
we do not know their origin, it is possible that these arose by the type of spacer-mediated recombination that we demonstrated in this study.

It is interesting to consider that acquired spacers could have a dual role during CRISPR immunity: a major one in the generation of CRISPR RNA (crRNA) guides and a minor one in mediating HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT. In support of this idea, recent work has found that Tn7-like transposons that harbour CRISPR arrays without effecting HGT.

Methods

Bacterial strains and growth conditions. Culture of S. aureus RN4220 (ref. 33) was carried out in BHI medium at 37 °C with agitation at 220 r.p.m. Liquid experiments were carried out in 3 ml medium in 15 ml conical tubes. BHI media were supplemented with 10 μg ml−1 chloramphenicol, 10 μg ml−1 erythromycin or 25 μg ml−1 kanamycin for plasmid maintenance and/or chromosomal marker selection.

Culture of S. aureus RN4220 (ref. 33) was carried out in BHI medium at 37 °C with agitation at 220 r.p.m. Liquid experiments were carried out in 3 ml medium in 15 ml conical tubes. BHI media were supplemented with 10 μg ml−1 chloramphenicol, 10 μg ml−1 erythromycin or 25 μg ml−1 kanamycin for plasmid maintenance and/or chromosomal marker selection.

Culture of S. aureus RN4220 (ref. 33) was carried out in BHI medium at 37 °C with agitation at 220 r.p.m. Liquid experiments were carried out in 3 ml medium in 15 ml conical tubes. BHI media were supplemented with 10 μg ml−1 chloramphenicol, 10 μg ml−1 erythromycin or 25 μg ml−1 kanamycin for plasmid maintenance and/or chromosomal marker selection.

Quantification of CRISPR–Cas-transducing particles. In S. aureus, overnight cultures of pWJ40 (ref. 34) or pC194 (ref. 33) were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with...
SNMP4y (ref. 2) at an MOI of 1. Phage was collected at indicated time points and filtered with 0.45-μm syringe filters (Acrodisc). Harvested phage was then used to infect lawns of S. aureus strain OS2 (ref. 2) suspended in 50% BHI supplemented with 5 mM CaCl₂, at an MOI of 1 on a BHI base supplemented with erythromycin and chloramphenicol to select for recipient cells and CRISPR–Cas transduction. For quantification of transducing particles produced from strains already containing CRISPR immunity, overnight cultures were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with φSNM4y. At A₆₀₀nm = 0.4 at an MOI of 1. Twenty minutes post-infection, phage was collected and filtered with 0.45-μm syringe filters (Acrodisc). Harvested phage was then used to infect lawns of OS2 suspended in 50% BHI supplemented with 5 mM CaCl₂ at an MOI of 1 on a BHI base supplemented with erythromycin and chloramphenicol to select for recipient cells and CRISPR–Cas transduction. Phages that were not of sufficient titres to infect at an MOI of 1 were supplemented and chloramphenicol to select for recipient cells and CRISPR–Cas transduction.

Detection of spacer acquisition. To check for spacer acquisition in S. aureus, transduced colonies were resuspended in colony lysis buffer (250 mM KCl, 5 mM CaCl₂. At 600 nm = 0.4, cultures were infected with φSNM4y at an MOI of 50 for targeting strains or 1 for non-targeting strains. Phages were collected from indicated strains 60 min post-infection. Supernatants were filtered with a 0.45-μm filter and then concentrated with an AMicon Ultra-4 100k centrifugal 50-ml spin columns (Amicon). Concentrates were resuspended with ΔNase I buffer, 20 mM Tris-HCl, pH 8.0, and 2 mM MgCl₂ and recontaminated two times. The suspension was then treated with 25 units of DNase I (Sigma) for 1 h. Following DNAse treatment, the enzyme was inactivated by heating at 70°C for 10 min and the addition of 5 mM EDTA. Phages were then incubated with 8 units of proteinase K (NEB) and 0.5% SDS at 37°C for 1 h. Phage DNA was isolated using a phenol/chloroform/isoamyl alcohol extraction (Fisher). Cellular DNA was collected andused in Quantification of transduction. Overnight cultures of PW40 were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with φSNM4y at an MOI of 1. Twenty hours post-infection, DNA was collected from recovered cells (CRISPR resistant). Phages were also collected and filtered with 0.45-μm syringe filters (Acrodisc). Overnight cultures of OS2 were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with φSNM4y at an MOI of 50. Phages were collected post-infection, 40 mM sodium citrate was added to the cultures. For erythromycin transduction, the cells were incubated for an additional 40 min then pelleted and washed twice with fresh BHI supplemented with 40 mM sodium citrate, whereas for chloramphenicol transduction, cells were washed immediately. Cells were then plated on BHI plates supplemented with the antibiotics selecting for the recipient strain and transduced marker along with 20 mM sodium citrate for type II-A plasmids and no sodium citrate for type III-A plasmids. For S. thermophilus, we isolated bacterial colonies that had acquired spacers in the erythromycin-tagged CRISPR1 locus of JAV28 following infection by phage 2972 using procedures previously described.28 Genomic DNA from strain JAV33 was amplified and sequenced with AV638–AV724 and found to have a spacer targeting the top strand beginning at position 26,553 of phage 2972. Phage 2972 was passaged on the strain in liquid culture on agar. Single plaques were picked and repassaged to single plaques on JAV33. Individual ‘escape’ phage clones were isolated, followed by sequencing of the amplicon containing the protospacer and PAM sequences. A single ‘PAM-escape’ mutant was used in the transduction assays (G–G, position 25,926) along with the wild-type DSMS3vir phage.

CRISPR adaptation and escape phage generation. For *P. aeruginosa*, to monitor the development of spacers targeting the phage, following procedures previously described.27 Next, we isolated DSMS3vir ‘escape’ mutants by inoculating a 96-well plate with 200 μl of 20% CRISPR-resistant cells and allowing the strains to transduce colonies on JAV33. Phage DNA was extracted by boiling the phage, and 2972Δvir DNA was amplified and sequenced with AV688–AV689. 2972Δvir contained a mutation in the PAM region (A>G, 26,588).

Quantification of transduction. For S. aureus, overnight donor cultures were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with either φSNM4y at an MOI of 1 or 85×1–3 and φNM1γ6 at an MOI of 50. Following lysis of the culture at 2 h, phages were collected and filtered with 0.45-μm syringe filters (Acrodisc). Overnight recipient cultures were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with φSNM4y at an MOI of 1. Twenty hours post-infection, DNA was collected from recovered cells (CRISPR resistant). Phages were also collected and filtered with 0.45-μm syringe filters (Acrodisc). Overnight cultures of OS2 were diluted 1:100 in fresh BHI with appropriate antibiotics and 5 mM CaCl₂. At A₆₀₀nm = 0.4, cultures were infected with φSNM4y collected from the PW40 culture at an MOI of 1. Twenty hours post-infection, DNA was collected from recovered cells (CRISPR transduced). Spacer were amplified with RH50 and JW655–JW662 for sample barcoding. The sequences of the oligonucleotide targets in this study are listed in Supplemental Table 2. Adapted bands were gel extracted and subjected to Illumina MiSeq sequencing. Data analysis was performed in Python. Spacer reads were extracted from the raw MiSeq FASTA files and aligned to the phage genome. The number of reads and PAM was designated for each spacer. Spacers were designated as reads per million and plotted against the φSNM4y genome in 2,000-base-pair bins. Deep sequence phage–CRISPR DNA junctions, cultures containing spacer H1, were infected with φSNM4y at an MOI of 50 and phages were collected 90 min post-infection. Phage DNA was isolated as described above. DNA was then prepped with the Illumina TruSeq Nano kit according to the manufacturer’s instructions. Prepped DNA was then subjected to NextSeq sequencing. BWA-MEM (arXiv:1303.3997v1) was used to align sequenced DNA to the PAM junction, which contains 200 base pairs of the upstream CRISPR sequence (leader and direct repeat) and 205 base pairs of the downstream phage sequence (spacer, PAM and phage genome), or the repeat junction, which contains 205 base pairs of the upstream phage sequence (phage genome and spacer) and 200 base pairs of the downstream CRISPR sequence (direct repeat and downstream plasmid sequence). A Python script was then used to sort and bin reads that spanned the full 75-nucleotide read length, allowing for one mismatch.

Phage titre assay. Phage titre assays were performed as previously described.20

Efficiency of plaquing assays. Efficiency of plaquing assays were performed as previously described.20

Simulation of CRISPR immunization. Simulation of CRISPR immunization was performed as previously described.20

Strain construction. To make the recA-knockout JAV9, the allelic replacement system developed by Wenyen Jiang using pPW244 was applied as previously described.20 Briefly, pAV4 was transformed into RN4220 and integrants were isolated. Double crossover events were selected for by a temperature-sensitive cat mutation. recA phage deletion were performed by introducing plasmid outside the homology arms, AV223 and AV224. To make JAV21, OS2 was infected with φSNM4y at an MOI of 1 to produce transducing particles carrying the genomic erythromycin cassette. These particles were used to infect JW263 (ref. 25) as described in quantification of transduction. Colonies that were resistant to kanamycin and erythromycin were struck out two times on plates supplemented
with 20 mM sodium citrate, kanamycin and erythromycin. JAV29 and JAV32 were constructed by transforming suicide vectors pAV253 and pAV282. Integration was confirmed using primers AV584 and AV812 for pAV253 and AV648 and AV525 for pAV282. JAV29 and JAV32 were passed on soft agar for 72 h. JAV29 was passaged on soft agar for 72 h. JAV29 was passaged on soft agar and then amplified and sequenced with oGG38–oGG40 to confirm target insertion.

To confirm target insertion, 2 were assembled using BsaI cloning γ homology arms from RN4220. AV149, pAV150, pAV153, pAV155, H-transducing cells as described elsewhere. AV149, pAV150, pAV153, pAV155, H-transducing cells as described elsewhere. The sequences of the oligonucleotides used in this system that targets the portion of the phage interrupted with the spacer 08BA02176 were sequenced. AV626–AV673, and a three piece Gibson assembly was used to create the final product for transformation. Also in JAV27, CRISPR3 was eliminated by amplifying homology arms with AV664–AV665 and AV683–AV686. The chloramphenicol resistance cassette was amplified from pC194 (ref. 29) with W1055–W1056, and a three-piece Gibson assembly was used to create the final product for transformation. JAV27 was made by first knocking out CRISPR1 and then repeating the procedure for CRISPR3. For JAV28, CRISPR1 was tagged with erythromycin resistance by amplifying homology arms with AV667–AV692 and AV693–AV694. The erythromycin cassette was amplified from pE194 (ref. 29) with AV177–AV695, and a three-piece Gibson assembly was used to create the final product for transformation. To make JAV28, the three-piece Gibson assembly fragments into cells, overnight culture was washed in tris-EDTA buffer and then diluted 1:100 in tris-EDTA buffer and then diluted 1:100 in 1 ml chemically defined medium, then diluted 1:100 in 1 ml chemically defined medium. Following 1.5 h of incubation at 37 °C, 10 µl Gibson product along with 1 µl ComSm–α peptide (LPPYFGACGL, Genescript) were added. Following a 4-h incubation, cells were plated with the appropriate antibiotic and incubated for 36 h.

Phage construction. To create phages to study transduction in S. aureus 08BA02176 (ref. 29), phage 85 (ref. 29) was used to infect this strain at a high MOI on soft agar. 8501 was isolated for its ability to form plaques on 08BA02176. To make 85α, the 08BA02176 type III-A target was inserted at an early-genome site in the phage. 85α was passaged on soft agar for two times and confirmed to be insensitive to S. aureus 08BA02176 infection. 4. After a 24-h incubation, a resistant colony was picked, restreaked two times and confirmed to be insensitive to 08BA02176 infection.

To create a P. aeruginosa PA14 strain carrying a streptomycin resistance cassette immediately adjacent to the type I-F CRISPR–Cas system in the genome (PA14-Sm, with the Sm gene inserted at position 2,937,360), we used homologous recombination. The streptomycin (Sm) resistance gene and its promoter were PCR amplified from pBAm1–Sm using primers pB_Sm_F and pB_Sm_R, and inserted into the Nhel restriction site of pHERD307, flanked by amplicons FL1 (flank1; generated using primer pairs FL1_F and FL1_R) and FL2 (flank2; generated using primer pairs FL2_F and FL2_R). To select for recombinants, a crRNA targeting the junction between two homology sequences with a gRNA repair primer was included. To create S. thermophilus strains, PCR products were generated with homology arms approximately 2-kb long that flank antibiotic resistant-cassette transfected and transformed into the wild-type strains. For JAV27, CRISPR2 was eliminated by amplifying homology arms with AV666–AV665 and AV686–AV667. The chloramphenicol resistance cassette was amplified from pC194 (ref. 29) with W1055–W1056, and a three-piece Gibson assembly was used to create the final product for transformation. Also in JAV27, CRISPR2 was eliminated by amplifying homology arms with AV668–AV669 and AV688–AV683. The chloramphenicol resistance cassette was amplified from pC194 (ref. 29) with W1055–W1056, and a three-piece Gibson assembly was used to create the final product for transformation. JAV27 was made by first knocking out CRISPR1 and then repeating the procedure for CRISPR2. For JAV28, CRISPR1 was tagged with erythromycin resistance by amplifying homology arms with AV667–AV692 and AV693–AV694. The erythromycin cassette was amplified from pE194 (ref. 29) with AV177–AV695, and a three-piece Gibson assembly was used to create the final product for transformation. To make JAV28, the three-piece Gibson assembly fragments into cells, overnight culture was washed in tris-EDTA buffer and then diluted 1:100 in tris-EDTA buffer and then diluted 1:100 in 1 ml chemically defined medium, then diluted 1:100 in 1 ml chemically defined medium. Following 1.5 h of incubation at 37 °C, 10 µl Gibson product along with 1 µl ComSm–α peptide (LPPYFGACGL, Genescript) were added. Following a 4-h incubation, cells were plated with the appropriate antibiotic and incubated for 36 h.

Phage construction. To create phages to study transduction in S. aureus 08BA02176 (ref. 29), phage 85 (ref. 29) was used to infect this strain at a high MOI on soft agar. 8501 was isolated for its ability to form plaques on 08BA02176. To make 85α, the 08BA02176 type III-A target was inserted at an early-genome site in the phage. 85α was passaged on soft agar for two times and confirmed to be insensitive to 08BA02176 infection. 4. After a 24-h incubation, a resistant colony was picked, restreaked two times and confirmed to be insensitive to 08BA02176 infection.

In JAV27, CRISPR2 was eliminated by amplifying homology arms with AV666–AV665 and AV686–AV667. The chloramphenicol resistance cassette was amplified from pC194 (ref. 29) with W1055–W1056, and a three-piece Gibson assembly was used to create the final product for transformation. Also in JAV27, CRISPR2 was eliminated by amplifying homology arms with AV668–AV669 and AV688–AV683. The chloramphenicol resistance cassette was amplified from pC194 (ref. 29) with W1055–W1056, and a three-piece Gibson assembly was used to create the final product for transformation. JAV27 was made by first knocking out CRISPR1 and then repeating the procedure for CRISPR2. For JAV28, CRISPR1 was tagged with erythromycin resistance by amplifying homology arms with AV667–AV692 and AV693–AV694. The erythromycin cassette was amplified from pE194 (ref. 29) with AV177–AV695, and a three-piece Gibson assembly was used to create the final product for transformation. To make JAV28, the three-piece Gibson assembly fragments into cells, overnight culture was washed in tris-EDTA buffer and then diluted 1:100 in tris-EDTA buffer and then diluted 1:100 in 1 ml chemically defined medium, then diluted 1:100 in 1 ml chemically defined medium. Following 1.5 h of incubation at 37 °C, 10 µl Gibson product along with 1 µl ComSm–α peptide (LPPYFGACGL, Genescript) were added. Following a 4-h incubation, cells were plated with the appropriate antibiotic and incubated for 36 h.

Phage construction. To create phages to study transduction in S. aureus 08BA02176 (ref. 29), phage 85 (ref. 29) was used to infect this strain at a high MOI on soft agar. 8501 was isolated for its ability to form plaques on 08BA02176. To make 85α, the 08BA02176 type III-A target was inserted at an early-genome site in the phage. 85α was passaged on soft agar for two times and confirmed to be insensitive to 08BA02176 infection. 4. After a 24-h incubation, a resistant colony was picked, restreaked two times and confirmed to be insensitive to 08BA02176 infection.
45. Hatoum-Aslan, A., Maniv, I. & Marrafiini, L. A. Nature clustered, regularly interspaced, short palindromic repeat (CRISPR) RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proc. Natl Acad. Sci. USA 108, 21218–21222 (2011).

46. Senemova, E. T. et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl Acad. Sci. USA 108, 10098–10103 (2011).

47. Kreiswirth, B. N. et al. The toxic shock syndrome toxin 552 structural gene is not detectably transmitted by a prophage. Nature 305, 709–712 (1983).

48. Horinouchi, S. & Weisblum, B. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. 150, 815–825 (1982).

49. Schneewind, O., Podl, P. & Fischetti, V. A. Sorting of protein A to the staphylococcal cell wall. Cell 70, 267–281 (1992).

50. Hynes, A. P. et al. Detecting natural adaptation of the Streptococcus thermophilus CRISPR-Cas systems in research and classroom settings. Nat. Protoc. 12, 547–565 (2017).

51. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of Pseudomonas aeruginosa mediates resistance to naturally occurring and engineered phases. J. Bacteriol. 194, 5728–5738 (2012).

52. Goldberg, G. W., Jiang, W., Bikard, D. & Marrafiini, L. A. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature 514, 633–637 (2014).

53. McGinn, J. & Marrafiini, L. A. CRISPR-Cas systems optimize their immune response by specifying the site of spacer integration. Mol. Cell 64, 616–623 (2016).

54. Martinez-Garcia, E., Calles, B., Arevalo-Rodriguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. BMC Microbiol. 11, 38 (2011).

55. Husmann, L. K., Scott, J. R., Lindahl, G. & Stenberg, L. Expression of the Arp protein, a member of the M protein family, is not sufficient to inhibit phagocytosis of Streptococcus pneumoniae. Infect. Immun. 63, 345–348 (1995).

56. Horinouchi, S. & Weisblum, B. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macroline, lincosamide, and streptogramin type B antibiotics. J. Bacteriol. 150, 804–814 (1982).

57. Letort, C. & Julliard, V. Development of a minimal chemically-defined medium for the exponential growth of Streptococcus thermophilus. J. Appl. Microbiol. 91, 1023–1029 (2001).

58. Fontaine, L. et al. Development of a versatile procedure based on natural transformation for marker-free targeted genetic modifications in Streptococcus thermophilus. Appl. Environ. Microbiol. 76, 7870–7877 (2010).

59. Bikard, D. et al. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat. Biotechnol. 32, 1146–1150 (2014).

60. Wang, H., Claveau, D., Vaillancourt, J. P., Roemer, T. & Meredith, T. C. High-frequency transposition for determining antimicrobial mode of action. Nat. Chem. Biol. 7, 720–729 (2011).

Acknowledgements

We thank J. W. Modell, J. T. Rostol and A. M. Pham for helpful discussion and critical reading of the manuscript. J. W. Modell (The Rockefeller University) also provided the strain pAV293-296. A.V. is supported by the Arnold O. Beckman Postdoctoral Fellowship. L.A.M. is supported by the Rita Allen Scholars Program and an NIH Director’s Pioneer Award (DP1GM128184-01). The work carried out by E.R.W. and S.M. was supported by the Biotechnology and Biological Sciences Research Council (BB/N017412/1) and Natural Environment Research Council (NE/M018350/1).

Author contributions

A.V. and L.A.M. conceived the study. A.V., S.M., R.B., E.R.W. and L.A.M. designed the experiments. A.V. executed the experimental work. S.M. executed the experimental work with B. aerogenes. A.V., S.M., R.B., E.R.W. and L.A.M. wrote the paper.

Competing interests

L.A.M. is a co-founder and Scientific Advisory Board member of Intella Therapeutics and a co-founder of Eligo Biosciences. R.B. is a co-founder and Scientific Advisory Board of Intella Therapeutics, a co-founder of Locus Biosciences, an advisor to ICR and a shareholder of DuPont and Caribou Biosciences. A.V., E.R.W. and S.M. declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41464-019-0400-2.

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

Correspondence and requests for materials should be addressed to L.A.M.

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

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- □ confirmed
  - The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
  - An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
  - The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
  - A description of all covariates tested
  - A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
  - A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
  - For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
    - Give P values as exact values whenever suitable.
  - For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
  - For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
  - Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
  - Clearly defined error bars
  - State explicitly what error bars represent (e.g. SD, SE, CI)

- [ ] Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

- None.

Data analysis

- For statistical analysis: GraphPad Prism 7
- For spacer alignment to phage genome: Custom code in Python 2.7.9
- For alignment to phage/CRIOSPR junctions: BWA-MEM (arXiv:1303.3997v1) used to align reads to junctions, and custom code in Python 2.7.9 used to bin reads.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files). Raw sequencing data is available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences
☐ Behavioural & social sciences
☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size       | Sample sizes were not statistically predetermined. |
|-------------------|---------------------------------------------------|
| Data exclusions   | No data was excluded from the analysis.           |
| Replication       | All experiments were reliably replicated.          |
| Randomization     | Not relevant to this study.                       |
| Blinding          | Not relevant to this study.                       |

Reporting for specific materials, systems and methods

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a                              | n/a     |
| ☒ Involved in the study          | Involved in the study |
| ☒ Unique biological materials   | ☒ ChiP-seq |
| ☒ Antibodies                    | ☒ Flow cytometry |
| ☒ Eukaryotic cell lines         | ☒ MRI-based neuroimaging |
| ☒ Palaeontology                 |         |
| ☒ Animals and other organisms   |         |
| ☒ Human research participants   |         |