Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces the fitness of lysogenic hosts

Gregory W. Goldberg¹, Elizabeth A. McMillan², Andrew Varble³, Joshua W. Modell¹, Poulami Samai¹, Wenyan Jiang¹ & Luciano A. Marraffini¹

CRISPR-Cas systems offer an immune mechanism through which prokaryotic hosts can acquire heritable resistance to genetic parasites, including temperate phages. Cotranscriptional DNA and RNA targeting by type III-A CRISPR-Cas systems restricts temperate phage lytic infections while allowing lysogenic infections to be tolerated under conditions where the prophage targets are transcriptionally repressed. However, long-term consequences of this phenomenon have not been explored. Here we show that maintenance of conditionally tolerant type III-A systems can produce fitness costs within populations of Staphylococcus aureus lysogens. The fitness costs depend on the activity of prophage-internal promoters and type III-A Cas nucleases implicated in targeting, can be more severe in double lysogens, and are alleviated by spacer-target mismatches which do not abrogate immunity during the lytic cycle. These findings suggest that persistence of type III-A systems that target endogenous prophages could be enhanced by spacer-target mismatches, particularly among populations that are prone to polylysogenization.
Prokaryotic organisms can limit the spread of mobile genetic elements (MGEs) with the help of various defense systems, including clustered, regularly interspaced, short palindromic repeat (CRISPR) loci, and CRISPR-associated (cas) genes that function together as adaptive immune systems. CRISPR-Cas systems provide immunity to invasive MGEs, such as bacteriophages (phages) and plasmids, by acquiring short “spacer” sequences from their nucleic acids and incorporating them between repeat sequences of CRISPR locus DNA. Transcription of a CRISPR locus, followed by cleavage and processing of its transcripts, generates mature CRISPR RNA (crRNA) guides containing individual spacer sequences. Ultimately, crRNA-guided Cas proteins are employed to locate and degrade the nucleic acids of genetic elements that bear a matching target sequence for one or more spacers. Four of the six (I–VI) types of CRISPR-Cas systems which have been classified to date include systems that cleave the DNA of their target elements.

Given that MGEs can facilitate the spread of beneficial genes within prokaryotic populations, it was proposed that resistance to foreign DNA elements by CRISPR-Cas systems could jeopardize the survival of bacteria which rely heavily on MGE-mediated horizontal gene transfer (HGT), and in turn promote the evolution of strains that do not harbor CRISPR-Cas systems. CRISPR-Cas systems are indeed absent from ~50% of sequenced bacterial genomes, despite evidence that they can be transferred horizontally. Moreover, the potential for genetic loss of CRISPR-Cas systems has been clearly demonstrated in laboratory settings where both beneficial plasmids and temperate phages were targeted. During the anti-plasmid CRISPR-Cas immune response, population bottlenecks imposed by antibiotics were found to select for mutant or deleted CRISPR-Cas systems when a target plasmid carrying resistance to the antibiotic was introduced prior to treatment. Similar outcomes were observed when type I and type II CRISPR-Cas systems were engineered to target temperate phages that produce antibiotic-resistant lysogens. In this manner, indiscriminate targeting of beneficial elements might impede the distribution of CRISPR-Cas systems in natural populations.

Fitness costs associated with the upkeep of CRISPR-Cas systems, including costs which result from basal expression and the occasional acquisition of toxic spacers that target their host’s chromosome, have also been proposed to influence the distribution of these systems. In certain cases, toxic chromosomal targeting may result from CRISPR-Cas immunity directed at temperate phages which integrate into the host chromosome as prophage DNA during lysogenic infections. We previously demonstrated that the type III-A system from Staphylococcus epidermidis RP62a can avert targeting of integrated lambda-like (lambdoid) prophages and give rise to “conditionally tolerant” lysogens in which the CRISPR-Cas system and prophage target have not been genetically altered. In other words, we observed a stable co-existence between functional type III CRISPR-Cas systems and prophage target DNA in lysogenic hosts. Type III systems are distinct from other DNA-cleaving CRISPR-Cas systems where targeting occurs only at sequences which are transcribed, and requires crRNAs which are complementary to the nascent transcripts. The type III-A system of S. epidermidis RP62a was further shown to encode a targeting complex with nucleases that license cleavage of RNA in addition to DNA, as well as an auxiliary RNase (Csm6) that can assist in the degradation of phage transcripts during lytic infection. Collectively, these type III-A nucleases allow the system to prevent temperate phage propagation when its targets are transcribed during lytic infection or prophage induction, and yet tolerate prophages while their targets are sufficiently repressed in the chromosome.

However, type-III A spacers displaying a perfect match with an endogenous prophage are yet to be found in the completed staphylococcal genomes. Thus, it remains to be determined whether conditional tolerance can influence the long-term persistence of type III-A systems in natural populations of lysogenic Staphylococci, including S. aureus lineages that are polylysogenized with more than one prophage.

In this work, we use laboratory derivatives of a model clinical isolate, S. aureus Newman, and show that maintenance of conditionally tolerant type III-A systems can become a fitness liability for lysogenic hosts. We also provide evidence that the costs result from transcription-dependent targeting at prophage loci, and investigate the effects of different spacers within populations of single or double lysogens.

**Results**

**Incomplete prophage tolerance by type III-A systems.** In a previous work, the CRISPR-Cas system of S. epidermidis RP62a was re-engineered to contain spacers targeting the temperate phages of S. aureus Newman, and introduced into heterologous RN42203 or TB4 hosts on pC194-based plasmid. TB4 is a prophage-cured derivative of Newman that can be re-infected with the native phages, including ΦN1 and ΦN4M. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132.

**Incomplete prophage tolerance by type III-A systems.** In a previous work, the CRISPR-Cas system of S. epidermidis RP62a was re-engineered to contain spacers targeting the temperate phages of S. aureus Newman, and introduced into heterologous RN42203 or TB4 hosts on pC194-based plasmid. TB4 is a prophage-cured derivative of Newman that can be re-infected with the native phages, including ΦN1 and ΦN4M. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132. One of our previously tested spacer sequences was modeled from a spacer occurring naturally in the type III-A system of S. argenteus MSHR1132.

Previous studies have demonstrated that chromosomal targeting by CRISPR-Cas systems in bacteria is detrimental to growth
and potentially lethal\(^{27,30}\), even in cases where prophage loci are targeted\(^{28,29,43}\). Type III-A systems can tolerate intact chromosomal targets, but only while they are sufficiently repressed. In light of these findings, we hypothesized that the spacer-dependent fitness costs could be derived from low levels of type III-A chromosomal targeting, licensed by transcription at prophage loci. Transcription of prophage loci can occur in a subset of cells during spontaneous prophage induction\(^{44}\). However, spontaneous prophage induction is lethal in the absence of targeting spacers, and it is therefore unexpected that non-targeting lysogens would be more fit than lysogens with targeting spacers, as we observed in our competition experiments (Fig. 1). Furthermore, we determined that all of these targeting spacers can disrupt phage lytic propagation via spontaneous (Supplementary Fig. 2a) or MMC-stimulated (Supplementary Fig. 2b) prophage induction. Therefore, if the fitness costs were derived from targeting effects which occur during prophage induction, each spacer might be expected to produce similar fitness costs. Instead, we observed significantly different fitness costs for at least one of the spacers we tested across six transfers (Fig. 1b–d; one-way ANOVA, \(p = 0.0499\)). To rule out the possibility that prophage induction is required for the fitness defects observed in competition assays, we disrupted \(\Phi NM1\)'s inducibility by introducing a serine to alanine (S124A) missense mutation in the catalytic domain of its CI-like repressor, generating \(\Phi NM1^{\text{ind-}}\) (Supplementary Fig. 2c). This designation is further supported by research on \(S.\ aureus\) lambdoid phages with regulated regulatory architecture\(^{77-79}\). Pairwise competition as in b–d, except that targeting and control lysogens both contain the \(\Phi NM1^{\text{ind-}}\) prophage mutations. All \(s_{\text{av}}\) values are derived from Supplementary Table 1 and provide mean ± standard deviation (s.d.)

Fig. 1 Incomplete prophage tolerance by type III-A systems can be licensed by prophage-internal promoter activity. a Schematic representation of the integrated \(\Phi NM1\) prophage genome. The gp5, gp16, gp32, and gp43 open reading frames are colored to denote inclusion of a target sequence for one or more spacers tested in this work. Gray bent arrow indicates the position of \(\Phi NM1\)'s early lytic promoter (P\(_{\text{cro}}\)). b–d Pairwise competition experiments for three perfectly matching spacers that license conditional tolerance of \(\Phi NM1\). In each case, an Erm\(^R\)-marked control lysogen harboring the parent vector with non-targeting spacers was competed against a TetR-marked conditionally tolerant lysogen. Relative frequencies (y-axis) are plotted against the number of transfers (x-axis), with one transfer per day. Individual values from each biological replicate (\(n = 3\)) are depicted in black as a triangle, circle, or rhombus. Solid lines represent the average change in relative frequency across the three replicates, and share color coding with the target genes outlined in a. e Control co-culture experiment. Pairwise competition was performed as in b–d above, except that differentially marked TB4::\(\Phi NM1\) lysogens both harbor the same parent vector with non-targeting spacers. f Sequences of the wild-type \(\Phi NM1\) early lytic promoter and its inactive \(\Phi NM1^{\text{Pcro-}}\)-variant with point mutations (red lettering) in the −10 element. The location of this promoter, along with the approximate position of its transcriptional start site (+1), was inferred from RNA-sequencing analysis of \(\Phi NM1\) during lytic infections\(^{29}\). This designation is further supported by research on \(S.\ aureus\) lambdoid phages with related regulatory architecture\(^{77-79}\). g–i Pairwise competition as in b–d, except that targeting and control lysogens both contain the \(\Phi NM1^{\text{Pcro-}}\) prophage mutations.
each of the gp16- and gp43-targeting plasmids were not significantly reduced in this prophage-mutant background (Supplementary Fig. 2e–f; lme, $p = 0.9786$ and $p = 0.5431$, respectively), when selection coefficients were compared to those obtained in the wild-type background (Fig. 1c, d).

The results above confirm that de-repression during prophage induction is not required for fitness costs associated with conditionally tolerant type III-A systems in lysogens. We therefore postulated that the costs could be licensed by unregulated transcription from one or more promoter within the target prophage. To explore this possibility, we introduced inactivating mutations into the rightward promoter of ΦNM1 (Fig. 1a, f) that is predicted to be involved in the lysis/lysogeny decision of lambdoid phages49, generating ΦNM1Δprc38. As expected, mutation of this promoter was sufficient to abolish the prophage’s lytic cycle in two assays (Supplementary Figs. 2a and g), so we proceeded to test these mutant lysogens in competition co-cultures. The Pprc mutation virtually abolished the fitness cost associated with the gp5 spacer, which targets a sequence immediately downstream of the promoter, such that selection coefficients were no longer significantly higher than those obtained in control competitions (Fig. 1g; lme, $p = 0.2721$).

In contrast, the costs associated with the gp16 and gp43 spacers were only partially reduced or not significantly reduced at all, respectively, when compared to their costs in the wild-type ΦNM1 background (Fig. 1h, i; lme, $p = 0.0321$ and $p = 0.6711$, respectively). Potentially, these fitness costs are licensed by leaky transcription from other sequences in the prophage genome. Taken together, these results demonstrate that conditionally tolerant type III-A systems can impose spacer-dependent fitness costs even if their target prophages are not inducible, and rather suggest that the costs arise from leaky transcription of prophage targets within otherwise stably lysogenic cells.

**Contributions of targeting nucleases to fitness costs.** Having demonstrated that lysogenic hosts which maintain type III-A systems can incur spacer-dependent fitness costs, we wanted to determine whether this phenomenon results from CRISPR–Cas targeting. To this end, we measured fitness costs associated with the gp16 spacer in the presence of mutations in Cas nuclease active sites previously implicated in targeting by the *S. epidermidis* type III-A system (Fig. 2a). Mutation of Csm3 (D32A) was previously shown to abolish crRNA-guided transcript cleavage, while mutation of Cas10’s palm polymerase domain (D586A, D587A) was shown to abolish the complex’s capacity for transcription-dependent cleavage of DNA35. In addition, mutations in the HEPN domain of the Csm6 RNase (R364A, H369A) were shown to abolish its transcript degradation activity and eliminate its ability to target prophages by type III-A CRISPR–Cas systems.

Consequences of conditional targeting in polylysogens. Clinical isolates of *S. aureus* Newman, for example, does not harbor an endogenous CRISPR–Cas system but harbors four heterimmune lambdoid prophages38. However, among the staphylococcal type III-A CRISPR–Cas systems identified so far some spacers have been found to match conserved temperate phage lytic genes39-41, and presumably offer the potential to target multiple phages. These observations prompted us to investigate the consequences of maintaining type III-A systems in a polylysogenic scenario where more than one prophage can be targeted. This was accomplished using ΦNM1 and ΦNM4, which integrate into distinct chromosomal loci but share extensive regions of sequence homology, including a perfectly conserved target for the gp16 spacer (Fig. 3a). Provided that each target sequence falls within lytic gene clusters which are sufficiently repressed during lysogeny, we anticipated that polylysogeny would be tolerated even in the presence of perfectly matching spacers. To confirm this, TB4·ΦNM4 lysogens harboring CRISPR–Cas plasmids with spacers matching to ΦNM1’s gp5, gp16, or gp43 targets were infected with an erythromycin resistance-conferring derivative of ΦNM1 (ΦNM1-ErmR2) that allows quantification of newly formed lysogens. When compared with control lysogens harboring a fully mismatched (35 mm) spacer, we found no significant differences in the frequency of ΦNM1-ErmR2 secondary lysogenization for lysogens with the singly (gp5, gp43) and doubly (gp16) targeting spacers (Fig. 3b; one-way ANOVA, $p = 0.1906$), while plaque formation on the same hosts was strongly reduced in the presence of targeting spacers (Fig. 3c). However, differences in colony size phenotype were clearly discernible among the resulting double lysogens. Relative to double lysogens with the control spacer, double lysogens with the gp5 or gp43 spacers which do not target ΦNM4 had a less severe colony size reduction than lysogens with the gp16 spacer targeting both prophages (Fig. 3d). To determine whether these phenotypes reflect discrepancies in relative fitness, we mixed unmarked double lysogens harboring each of the three spacers into a single batch and competed them in liquid culture for 3 days. At each transfer interval, spacer abundances were measured by deep sequencing, and the frequency of each targeting spacer was calculated relative to the control spacer’s frequency in the sample. These pairwise relative frequencies were then used to calculate selection coefficients for each strain (Supplementary Table 1). Our analysis indicates that the fitness cost associated with the dual-targeting gp16 spacer in double lysogens was significantly greater than that obtained with the gp5 and gp43 spacers targeting only ΦNM1 (Fig. 3e; one-sided t tests, $p = 0.0095$ and $p = 0.0049$, respectively). Importantly, although the gp16-associated cost in double lysogens was significantly higher than its average cost in pairwise competitions with either of the marked TB4·ΦNM1 (Fig. 1c; one-sided t test, $p = 0.0056$) or TB4·ΦNM4 single lysogens (Supplementary Fig. 3; one-sided t test, $p = 0.0095$) across the first three transfers, we did not observe a significant difference when comparing the gp5- or gp43-associated costs in double lysogens (Fig. 3e) to their average costs in pairwise competitions with marked TB4·ΦNM1 single lysogens across the first three transfers (Fig. 1b, d; two-sided t tests, $p = 0.6322$ and $p = 0.5643$). These results demonstrate that type III-A CRISPR–Cas systems can become a greater fitness liability if they target multiple temperate phages, despite their capacity for conditional tolerance.

Effects of spacer-target mismatches on fitness costs. The above results confirm that type III-A systems that target multiple temperate phages are not a barrier to polylysogenization per se, even if they possess perfectly matching spacers. However, when searching the fully sequenced staphylococcal genomes which were
previously found to contain type III-A systems \cite{49}, we were unable to identify cases where perfectly matching targets co-existed with spacers within the same genome. Meanwhile, as previously reported, one of these genomes possesses a spacer with two partially mismatched targets (each within a separate prophage), and we were able to confirm this using modified search parameters. Having found, also, that our partially mismatched gp32 spacer is still capable of targeting \Phi NM1 during the lytic cycle (Supplementary Fig. 2) but displayed a fitness cost which appeared to be lower (although this reduction was not statistically significant) than that of its perfectly matched \textit{gp32} variant in single lysogens (Supplementary Fig. 1), we wondered if spacer-target mismatches could likewise help to minimize fitness tradeoffs in polylysogenic populations. To investigate this, we constructed a series of mismatched spacers based around our \textit{gp16} spacer, which targets an identical sequence found in both structured a series of mismatched spacers based around our.

Fig. 2 Effect of targeting nuclease active site mutations on fitness costs associated with type III-A systems. a Schematic diagram summarizing the arrangement of CRISPR-Cas loci used throughout the work, with position 2 of the CRISPR array occupied by the \textit{gp16} spacer (purple rectangle) in this example. The naturally occurring \textit{res} spacer (white rectangle) from \textit{S. epidermidis RP62a} is maintained at position 1 in all cases. Open reading frames (ORFs) encoding Cas nucleases implicated in targeting are highlighted in blue (\textit{cas10}), red (\textit{csm3}), or yellow (\textit{csm6}), and the positions of their active site mutations are labeled to scale within each ORF. Other features of the diagram are not drawn to scale. b-e Pairwise competition experiments for mutant CRISPR-Cas plasmids harboring the \textit{gp16} spacer with perfect matches to \Phi NM1. In each case, an \textit{erm}\textsuperscript{R}-marked control lysogen harboring the wild-type parent vector with non-targeting spacers was competed against a TetR-marked lysogen harboring a mutant CRISPR-Cas plasmid with the \textit{gp16} spacer. Relative frequencies (y-axis) are plotted against the number of transfers (x-axis), with one transfer per day. Individual values from each biological replicate (n = 3) are depicted as a triangle, circle, or rhombus, and are colored to match the ORFs highlighted in a for each mutant, or orange for the \textit{csm3}/\textit{csm6} double mutant. Solid lines represent the average change in relative frequency across the three replicates, and are colored purple to match the \textit{gp16} spacer’s target gene as outlined in Fig. 1. The $s_{av}$ values are derived from Supplementary Table 1 and provide mean ± s.d.

$$s_{av} = 0.0745 ± 0.05$$

$$s_{av} = 0.0555 ± 0.02$$

$$s_{av} = 0.0113 ± 0.02$$

$$s_{av} = 0.0055 ± 0.02$$

$$s_{av} = 0.0258 ± 0.02$$

The naturally occurring \textit{res} spacer (white rectangle) from \textit{S. epidermidis RP62a} is maintained at position 1 in all cases. Open reading frames (ORFs) encoding Cas nucleases implicated in targeting are highlighted in blue (\textit{cas10}), red (\textit{csm3}), or yellow (\textit{csm6}), and the positions of their active site mutations are labeled to scale within each ORF. Other features of the diagram are not drawn to scale. b-e Pairwise competition experiments for mutant CRISPR-Cas plasmids harboring the \textit{gp16} spacer with perfect matches to \Phi NM1. In each case, an \textit{erm}\textsuperscript{R}-marked control lysogen harboring the wild-type parent vector with non-targeting spacers was competed against a TetR-marked lysogen harboring a mutant CRISPR-Cas plasmid with the \textit{gp16} spacer. Relative frequencies (y-axis) are plotted against the number of transfers (x-axis), with one transfer per day. Individual values from each biological replicate (n = 3) are depicted as a triangle, circle, or rhombus, and are colored to match the ORFs highlighted in a for each mutant, or orange for the \textit{csm3}/\textit{csm6} double mutant. Solid lines represent the average change in relative frequency across the three replicates, and are colored purple to match the \textit{gp16} spacer’s target gene as outlined in Fig. 1. The $s_{av}$ values are derived from Supplementary Table 1 and provide mean ± s.d.
mismatched control spacer’s frequency in the sample. Our analysis indicates that the fitness costs were significantly reduced for spacers with four or six mismatches when compared to those of the perfectly matching gp16 spacer within their respective populations of single (Fig. 4c; one-sided t tests, $p = 0.0280$ and $p = 0.0290$, respectively) or double (Fig. 4d; one-sided t tests, $p = 0.0436$ and $p = 0.0281$, respectively) lysogens. As expected from the observed lack of targeting during the phage’s lytic cycle, no significant costs were measured for any of the spacers with eight or more mismatches in either population (one-way ANOVAs, $p = 0.7408$ and $p = 0.6265$, respectively), when selection coefficients were compared with that of the average obtained in control TB4::ΦNM1 pairwise competitions across their first three transfers (Fig. 1e). Importantly, we observed no significant difference between the cost associated with the perfectly matching gp16 spacer in competitions with marked and unmarked TB4::ΦNM1 single lysogens (Fig. 1c and 4c; two-sided t test, $p = 0.3862$), and this served to validate our usage of deep sequencing to assess relative fitness. We also found that the cost associated with the perfectly matching gp5 spacer targeting only ΦNM1 (Fig. 3e) was not significantly different from the costs measured for the dual-targeting spacers in double lysogens harboring the different CRISPR-Cas plasmids tested in b, as indicated by the labels. Efficiency ratios were calculated for each plasmid relative to plaques quantified on a TB4 lawn harboring the non-targeting parent vector. Dotted line represents the limit of detection under these assay conditions. Error bars, mean ± s.d. ($n = 3$, biological replicates).

Discussion

We report here that certain conditionally tolerant type III-A CRISPR–Cas systems can put their (poly)lysogenic hosts at a fitness disadvantage in mixed populations of S. aureus, despite the potential for these CRISPR–Cas systems and their prophage target(s) to co-exist in (poly)lysogenized clones. The fitness costs varied in severity with the position of targets within the prophage genome, and likely result from incomplete prophage tolerance by...
the system’s transcription-dependent targeting machinery. We previously showed that transcription-dependent chromosomal targeting by the type III-A system of S. epidermidis RP62a causes severe growth defects39,35, and similar effects have been observed with other types of CRISPR–Cas systems27,28,30. At least in part, this appears to result from toxicity associated with frequent nicking or cleavage of the host’s chromosomal DNA. Minor growth defects, in line with the fitness costs described in this work, might instead be expected if type III-A chromosomal targeting were infrequent; e.g., licensed by transcription in a small subpopulation of cells. When lambdoid lysogens are cultured during the non-inducing conditions of lysogeny, perhaps in part due to the terminus in this regulatory region. This cluster lies downstream of a regulatory region that was shown to contain a conserved transcriptional terminator55, and we identified such a region in the PHN1 between gp16 and gp43. In staphylococcal phages with related genomic architecture, this cluster lies downstream of a regulatory region that was shown to contain a conserved transcriptional terminator55, and we identified such a region in the PHN1 between gp16 and gp43. Conceivably, the magnitude of the gp43 spacer’s fitness cost is determined by one or more late promoters which are insulated from upstream transcription during the non-inducing conditions of lysogeny, perhaps in part due to the terminator in this regulatory region.

The potential for prophage lytic regions to be de-repressed independently of spontaneous prophage induction is not unprecedented for lambdoid temperate phages, and was previously demonstrated using a genetic reporter system46. In this previous study, the authors detected slight de-repression above background even when prophages carried a non-inducible cI allele similar to that of our PHN1ind– mutant. Leaky transcription arising in this manner could likewise be responsible for the fitness costs detected in our study. A model such as this would perhaps explain the exacerbated fitness cost observed with double lysogens where both prophages were targeted with the gp16 spacer (Fig. 3e), if we assume that transcription of each prophage is only well-coordinated in cells undergoing prophage induction56. In this scenario, if instances of leaky transcription in each prophage were poorly correlated within individual cells, the combined effect of having both targets might essentially be additive at the
population level. Consistent with this idea, the cost associated with the gpl6 spacer was found to be greater in double lysogens than in single lysogens of ΦNM1 and ΦNM4 alone (Fig. 1c and Supplementary Fig. 3a). Examples of uncorrelated leaky expression from distinct chromosomal loci in bacteria are also not unprecedented. It was previously demonstrated that cell–cell variability becomes elevated under transcriptionally repressed conditions, even when expression from two loci with identical promoters and identical repressors is compared. Single cell-approaches could be used to clarify whether expression of the two prophages used in this study is occasionally uncoupled.

Additional evidence supporting transcription-dependent chromosomal targeting as the source of the fitness costs was provided by experiments with mutant type III-A Cas nucleases. Ablation of csm3’s RNase activity did not reduce fitness costs (Fig. 2b). Given its role in specific cleavage of target RNAs by type III systems, this was perhaps expected; knockdown of prophage lytic transcripts should not be toxic to lysogenic cells per se. Meanwhile, fitness costs appeared to be completely abolished by mutation of cas10’s palm polymerase domain (Fig. 2c), in line with past reports indicating that transcription-dependent DNA degradation by this system requires an intact palm polymerase domain. Interestingly, mutation of csm6’s HEPN RNase domain strongly reduced the cost associated with the gpl6 spacer as well (Fig. 2d, e). Our previous study showed that Csm6’s RNase activity contributes to the degradation of phage lytic transcripts, and this should not necessarily be toxic for lysogens. However, given that the specificity of Csm6’s RNA knockdown effect has not been fully determined, it is possible that this activity could cause occasional off-target cleavage of essential host transcripts, and detectably impact fitness in our competition assays. Of relevance here, it was recently shown that the related HEPN domains of type VI effector RNases can mediate cleavage of non-target transcripts, or “collateral” RNA cleavage, in the presence of target RNA. Growth defects associated with type VI targeting in vivo were accordingly proposed to result from toxic off-target cleavage of cellular transcripts. More recently, binding of target RNAs by type III complexes was found to enable synthesis of a cyclic oligonucleotide that stimulates the RNase activity of Csm6 proteins, and this depended on the Cas10 subunit’s intact palm polymerase domain. The phenotypic similarity observed with our cas10 and csm6 mutants in fitness assays (Fig. 2) might therefore be explained, at least in part, by a dominant effect of the cas10 palm domain mutation.

The results presented in this work suggest that, among (poly)lysogenization-prone populations of S. aureus, natural selection could favor the persistence of conditionally tolerant type III-A systems with spacers that are partially mismatched relative to their temperate phage targets. In line with this possibility, a putative conditionally tolerant isolate of S. aureus was recently reported to contain five mismatches in each prophage target relative to a spacer in its type III-A CRISPR–Cas system. Conceivably there are various evolutionary routes that could lead to such co-existences between a type III-A system and one or more partially mismatched prophage targets. Although the spacers acquired upon phage infection during the adaptive stage of the CRISPR–Cas immune response are expected to match their original targets with perfect complementarity, conditionally tolerant type III-A spacers can allow for stable lysogenization even if they are perfectly matching. However, as our current work illustrates, the resulting conditionally tolerant (poly)lysogens can suffer fitness costs. In principle, mutations which ablate these fitness costs could progressively accumulate in the spacer or prophage target sequence(s) after (poly)lysogeny is established. This, however, is unlikely to happen given the low frequency of spontaneous mutation in bacteria. In addition, most of the cells that acquire a new spacer during phage infection will not become lysogens (depending on the phage and host, the lysogenization frequency in staphylococci varies between 10^-3 and 10^-4). We believe that these non-lysogens could provide an alternative route for the establishment of conditionally tolerant mismatched spacers with minimal fitness costs. Due to the tolerance of spacer-target mismatches that is characteristic of type III-A CRISPR–Cas immunity, the new spacer could provide efficient defense against related temperate phages with incomplete homology with their target sequence. If infection with such phage were to result in type III-A conditional tolerance and lysogenization, the spacer will contain mismatches with the prophage target that could lead to its persistence without a fitness cost. Interestingly, in Pseudomonas aeruginosa, it was previously found that five spacer-target mismatches could enable coexistence between a type I-F system and its prophage target, although these mismatches abolished immunity to lytic infections. Intriguingly, chromosomal targeting effects were not completely abrogated in that scenario, and were sufficient to drive a host-encoded response that inhibited group behavior in surface-grown populations. Whether or not type III-A chromosomal targeting could give rise to a similar environment-specific phenotype remains to be investigated. Finally, it is important to note that type III-A systems have also been found to display signatures of horizontal acquisition within S. aureus genomes. Hence, we further speculate that the genetic stability afforded by conditional tolerance in the presence of spacer-target mismatches, could be achieved through the horizontal dissemination of prophage-targeting type III-A systems into (poly)lysogenic populations harboring prophages containing mismatches with the targeting spacers. Additional investigation will be required to establish how these features of type III-A temperate phage targeting influence fitness in actively CRISPR-adapting populations, where spacer diversity is potentially much greater.

**Methods**

**Bacterial strains and growth conditions.** Unless otherwise noted, cultivation of S. aureus RN4220 or TB4 and their derivatives was carried out in sterile 15 or 50 ml conical-bottom Falcon tubes (Corning) containing Difco TSB liquid media (volume 5-14% tube capacity), or on sterile petri dish plates containing Difco TSA solid media. Media was supplemented with erythromycin (Erm) at 10 μg/ml or tetracycline (Tet) at 5 μg/ml only when streaking out single colonies of marked TB4 derivatives containing pe194 or pT181 plasmids, respectively, and during the construction of these strains as needed. All media was supplemented with chloramphenicol (Cm) at 10 μg/ml when cultivating strains with CRISPR–Cas plasmids, and/or CarGc; 5 mM if the media was intended to support phage infection and transduction. Unless otherwise noted, plates were incubated at 37°C for 12–18 h and then stored at 4°C for up to a week. Overnight liquid cultures were inoculated from single colonies (biological replicates), grown for 12–16 h at 37°C with shaking, and then stored at 4°C for 3 h or less before briefly vortexing and subculturing at a 1:100 dilution as needed. Where applicable, subcultures were treated with mitomycin C (AG Scientific) to a final concentration of either 1.0 or 2.0 μg/ml or the virulent phage ΦNM167 at a multiplicity of infection ~7.5, as indicated in the figure legends. When the overnight cultures were to be used in multi-day competition experiments, 24 h of growth was instead allowed, followed by immediate vortexing, mixing of aliquots, and passing at a 1:1000 dilution. When streaking out lysogens with targeting spacers, aberrantly large colonies were presumed to have lost conditional tolerance either via genetic inactivation of CRISPR–Cas plasmids or prophage deletion, and not studied further in this work.

**DNA preparation and cloning.** Plasmid DNA of E. coli DH5α was purified from 4 to 6 ml overnight cultures using plasmid miniprep reagents from Qiagen, according to the manufacturer’s protocol. DNA from S. aureus RN4220 or TB4 was purified similarly. For samples that 2 ml overnight cultures were used, cells were treated with 10–15 μl lysostaphin (1 mg ml^-1) at 37°C for 1.5 h immediately after resuspension in P1 buffer. Minipreps were carried out with either QiaGen or EconoSpin columns. For PCR-based cloning procedures, DNA was amplified with Phusion polymerase (Thermo) and purified using QiaGen reagents and EconoSpin columns. When generating amplicons for Sanger sequencing, TopTaq polymerase (Qiagen) was often used for amplicons smaller than 750 bp, and DNA was not necessarily purified after the PCRs.
Electrocortropment RN4220 cells (described previously) were used for cloning in S. aureus. The pGG79 parent vector (Supplementary Table 3) is a derivative of pGG3-Bsal20 with the Bsal placeholder sequence at position two trimmed at its 3’ end by one base pair, and a CRISPR repeat inserted immediately downstream. This was accomplished by inverse PCR21 using primers gOG281/gOG282 (Supplementary Table 4), followed by blunt end NotI digestion. The CRISPR array was sequenced and S. aureus was transformed previously21, and the expected plasmid size was verified via analytical digestion during subsequent manipulations. The immune functionality of this plasmid in S. aureus was later also confirmed by resistance to infection with \( \Phi NM4_{11} \), a derivative of \( \Phi NM4_{12} \) possessing a target for the \( \Phi NM4_1 \) spacer in position one of pGG79's CRISPR array. Except where noted, all other CRISPR-Cas plasmids with wild-type cas genes were constructed via scarless insertion of spacers between CRISPR repeats 2 and 3 of the parent vector, pGG79. This was accomplished by restriction digesting the two oppositely oriented Bsal sites and then ligating them with an annealed oligonucleotide pair possessing compatible overhangs and the desired spacer sequence (Supplementary Tables 3 and 4). CRISPR-Cas arrays were then verified as described previously22. The pGG139 vector contains the \( \Phi g 6_{16} \) spacer and one or more cas genes; mutations were also constructed by the annealed oligo cloning method in most cases, and checked similarly, except that mutation parent vectors were used as backbones for insertion of the spacer. The \( \text{cm}3_{5} \) (D32A) mutant vector, pS959, was constructed via 2-piece Gibson assembly of PCR products amplified from pGG79 using primers S153/S154, and the expected plasmid size was verified by Sanger sequencing with primers oGG191/W235. The pGG90 plasmid was isolated, and the presence of its recombinant target was confirmed by restriction digesting of the pAV71 vector containing \( \text{cm}3_{5} \). The \( \text{cm}3_{5} / \text{cm}6_{3} \) double mutant vector, pGG89, was also constructed via 2-piece Gibson assembly of PCR products amplified from pGG79 using primers W852/oGG458, and a backbone fragment amplified from pGG139 using primers oGG457/W614. The pGG139 vector contains \( cas10 \) mutations in both the HD (H14A, D15A) and palm domain (D586A, D587A) domains, and the same was done for pGG89, except that pW242 was used for templating the backbone with both \( \text{cm}3_{5} \) and \( \text{cm}6_{3} \) mutations. To construct the pGG167 \( cas10 \) (D328A, D587A) plasmid, the expected plasmid size was confirmed by restriction digest, and regions of interest were sequenced using Sanger sequencing with primers listed in Supplementary Table 4 (Gibson assembly junctions, CRISPR CRISPR-assisted editing of phages. Recombinant phage plasmids were isolated, and the presence of its recombinant target was confirmed by restriction digesting of the pAV71 vector containing \( \text{cm}3_{5} / \text{cm}6_{3} \).)

Preparation of transducing lysates. Overnight culture aliquots (100 μl) of RN4220 or TB4 derivatives harboring a plasmid of interest were infected with \( \Phi NM4_1 \) (MOI - 0.1 - 0.25) in HIA soft agar supplemented with CaCl2 at 5 mM, and incubated overnight at 37 °C to produce a lysed lawn. With 48 h of storage at 4 °C, soft agar was scraped and decanted into a 50 ml conical-bottom Falcon tube, supplemented with an additional 1.2 ml fresh HB, and then centrifuged at 4030 g for 8-10 min. The supernatants containing phage were collected and DNA was isolated by autoclaving 1.5 ml tubes (Eppendorf) at 4 °C. When the plasmid of interest carried a CRISPR-Cas system, care was taken to lyse its host with a suitable non-targeted phage, among the following: \( \Phi NM1_{18} \), \( \Phi NM2_{12} \), \( \Phi NM4_{4} \), or \( \Phi g 1_{22} \). Phages not described previously were procured via CRISPR-assisted editing as outlined below.

CRISPR-assisted editing of phages. Recombinant phage plasmids were isolated, and the presence of its recombinant target was confirmed by restriction digesting of the pAV71 vector containing \( \text{cm}3_{5} / \text{cm}6_{3} \). The pGG12, a single mutant (lytic) derivative described above was used in place of their growth medium, and prior to centrifugation. Following centrifugation, fresh TSB supplemented with sodium citrate (40 mM) was used to resuspend and dilute pellets for plating on Tryptone Soy Agar (TSA) lawns harboring the pGG12. A backbone fragment was ampliﬁed from pGG171 using primers JW443/JW444. The resultant lysates were used to infect soft agar lawns containing a suitable target sequence.

Transduction. Subcultures were grown at 37 °C in TSB supplemented with CaCl2 at 5 mM for 1 h at 25 °C to 0.5-1.0 attendee, \( \text{Dou} \_\text{l} \), and 990 μl aliquots were subsequently mixed with 10 μl of transducing lysate containing the plasmid of interest in autoclaved 1.5 ml tubes (Eppendorf). After 15 min of growth at 37 °C, infected cultures were treated with filter-sterilized sodium citrate to a final concentration of 40 mM, and then pelleted by centrifugation at 16,100 x g for 2 min with refrigeration (4 °C). When transducing pE194 or pT181 plasmids, an additional 1 h 45 min of growth was allowed immediately after treatment with citrate and prior to centrifugation. Following centrifugation, fresh TSB supplemented with sodium citrate (40 mM) was used to resuspend and dilute pellets for plating on solid media supplemented with both sodium citrate at 20 mM and antibiotics to select for the transductants. To avoid residual free phosphate, an additional re-streak was performed on solid media in the presence of citrate (20 mM) and antibiotics to select for the strain's plasmid(s).

Phage-sensitivity assay. Throughout cloning and strain construction procedures, or otherwise necessary, clones were randomly checked for sensitivity to a phage of interest according to the strain method previously performed with \( \Phi NM2_{12} \), except that “clear” mutant (lytic) derivatives described above were used in place of their parental temperate phages to facilitate scoring, and chloramphenicol-supplemented TSA without added CaCl2 was usually used in place of HIA.

Strain construction. Single lysogens of TB4 were obtained by mixing 100 μl of a TB4 overnight culture with either \( \Phi NM1_{18} \) or \( \Phi NM4_1 \) (MOI - 0.1 - 1.0) in HIA soft agar supplemented with CaCl2 at 5 mM, and incubated overnight at 37 °C to produce a lysed lawn. With 48 h of storage at 4 °C, soft agar was scraped and decanted into a 50 ml conical-bottom Falcon tube, supplemented with an additional 1.2 ml fresh HB, and then centrifuged at 4030 g for 8-10 min. The supernatants containing phage were collected and DNA was isolated by autoclaving 1.5 ml tubes (Eppendorf) at 4 °C. When the plasmid of interest carried a CRISPR-Cas system, care was taken to lyse its host with a suitable non-targeted phage, among the following: \( \Phi NM1_{18} \), \( \Phi NM2_{12} \), \( \Phi NM4_{4} \), or \( \Phi g 1_{22} \). Phages not described previously were procured via CRISPR-assisted editing as outlined below.
by assaying for sensitivity to clear mutants (φNM1γ- or φNM4γ). Both clones were also typed via PCR using φNM1- or φNM4-specific primer pairs, oG66/oG67 (φNM1) and oG70/oG71 (φNM4), respectively. The identity of the φNM1 genome was confirmed and typed similarly, except that an overnight culture of TB4: φNM1 was mixed with φNM1 in soft agar to produce the turbid lysate lawn. φNM1 prophage mutants were constructed using CRISPR-assisted genome editing as outlined below. All plasmid-containing lysogenic derivatives were subsequently generated via transduction. Plasmid-containing non-lysogens were generated either via transformation of electrocompetent TB4 cells, or via transduction. Where applicable, strains were always marked via transduction with pE194 or pT181 plasmids first, followed by delivery of the CRISPR-Cas plasmid in a subsequent round of transduction.

CRISPR-assisted genome editing of *S. aureus*. Homology-directed allelic exchange coupled with CRISPR-Cas counter selection was performed essentially as described previously for the pW327 allelic replacement system38, with modifications. pW327-derived vectors containing homology to φNM1 and the prophage mutation of interest were used to transform electrocompetent RN4220 non-lysogens instead of RN4220:φ12, and a subsequent transduction step was used to transfer the plasmids from RN4220 to TB4:φNM1. Transduction was performed with φNM1γ as described above, except that transducing lysates were raised and injected at 28 °C, with 2 h 30 min of growth allowed for the recipient subculture. After isolating putative co-integrants via two consecutive re-streaks at 37 °C in the presence of chloramphenicol and verifying them with PCR using primers W1250/oGG7 (pGG170ts) or W1250/oGG32 (pGG171ts), clones were inoculated overnight in plain media at 37 °C instead of 28 °C. Subcultures were grown at 28 °C in plain media supplemented with CgCl to obtain logarithmic-phase cultures for treatment with the pW326 phagemid, but treated cultures were plated at 30 °C on media containing erythromycin after 3 h of growth at 30 °C rather than 1 h. An additional re-streak at 30 °C was performed in the presence of erythromycin before clones were checked by PCR and Sanger sequencing using primers oG455/oG462 (φNM1-φNM2 for the presence of the desired mutation (φNM1φφ or φNM1φφφ), respectively). Clones which had also lost the integrated pW326 amplicon (again determined by PCR with primers W1250/oGG7 or W1250/oGG32) were inoculated overnight in plain media but grown at 42 °C instead of 37 °C in order to cure the strain of pW326. Dilutions were plated on plain media at 37 °C, and single colonies were replica-plated at 37 °C on both plain and erythromycin-supplemented media to confirm plasmid loss. After re-streaking an erythromycin-sensitive clone on plain media, a PCR amplicon spanning the entire homology region was procured with primers oG453/oG47 (pGG170ts) or oG475/oG32 (pGG171ts), and sequenced by Sanger using additional primers listed in Supplementary Table 4. The clone was also re-streaked on media supplemented with chloramphenicol as a final check for sensitivity to the antibiotic; and, checked for sensitivity to φNM1γ to ensure that lysogenic immunity had not been lost. The pGG170ts- and pGG171ts-derived strains were renamed GWG7 and GWG9, respectively. TSB and TSA media were used in place of BHI throughout. The pW326 phagemid lysate used in this work was procured via an InM4- or InM4-derivative harboring the pGG79 parent vector.

Enumeration of plaque-forming units from lysogenic cultures. "Spontaneous" release from single lysogens was measured essentially as described previously except that filtered supernatants were collected immediately after 12 h of growth overnight. For quantification of particles released from MMC-induced cultures, overnight cultures were grown for 15 h and subcultures were grown for 1 h 15 min (~0–0.66 attenuation, D600nm) before treatment with MMC. Following 4 h of growth post treatment, subcultures were pelleted by centrifugation at 4300 x g for 6 min, and filtered supernatants were collected for serial dilution and spotting on lawns of TB4 harboring the CRISPR-Cas plasmid parent vector.

High-resolution growth curves. Overnight cultures grown for 14 h were subcultured essentially as described above, except that 200 µL culture volumes were used in 96-well microplates, and incubation at 37 °C with shaking was performed with an Infinite M200 PRO plate reader (TECAN) measuring attenuation (A600nm) every 10 min. In phage infection experiments, subcultures were treated with MMC to a final concentration of 2.0 µg/ml, following 70 min of growth. In phage infection experiments, subcultures were treated with φNM1γ at a multiplicity of infection of 4, followed by mixing of growth with additional subculture was performed in parallel to monitor growth in the absence of MMC or φNM1γ.

**Efficiency of plaquing.** Assays were performed by spotting high-titer lysate serial dilutions onto both CRISPR-containing and control lawns to enumerate pfu and calculate efficiency as pfu/cm² (φNM1γ). The φNM1 genome was constructed and typed similarly, except that an overnight culture of TB4: φNM1 was mixed with φNM1 in soft agar to produce the turbid lysate lawn. φNM1 prophage mutants were constructed using CRISPR-assisted genome editing as described previously29, except that the sensitive control lawns utilized a TB4 derivative harboring the pGG79 parent vector.

**Lysogenization with ermC-marked φNM1.** Lysogenization was performed essentially as described previously29, except that overnight cultures were always grown for 14 h, TS broth was used in place of TSB, multiplicity of infection was ~1, and incubation times were reduced to 10 min on ice and 20 min at 37 °C. Following the 20 min incubation, cultures were supplemented and plated with citrate as described for the transduction protocol above, except that chloramphenicol was maintained in minimal media and cultures were supplemented with erythromycin (5 µg/ml) to quantify acquisition of the marked prophage. The φNM1-ErmRβ phage used for infections is a variant of φNM1-ErmRβ with a modified ermC 5’ UTR that includes the SNP found in the tyc-1 allele, previously reported to improve constitutive expression of the cassette’s gene product in *B. subtilis*90. To estimate the concentration of total recipients harboring CRISPR-Cas systems, serial dilutions of the untreated overnight cultures were plated in the presence of chloramphenicol alone.

**Batch competition of unmarked lysogens for deep sequencing.** Aliquots of overnight cultures were mixed in equal proportions by volume and then diluted for passaging as described above. The remaining culture (~2 ml) was pelleted by centrifugation at 4300 x g for 6 min. Supernatants were discarded, and pellets were stored at −20 °C after drying for 10 min at benchtop. Passaging and storage of pellets was repeated in this manner every 24 h for 3 days.

**Preparation DNA for deep sequencing.** Frozen pellets from batch competition experiments were thawed on ice for 20 min and then miniprepped as described above. In total, 100 ng of miniprepped DNA from each sample was used as the template for barcoded PCRs to amplify a region of the CRISPR array that fully spans spacer position 2, using forward and reverse primers with identical barcode pairs. Barcoded PCR products were gel purified and then pooled at roughly equal proportions by normalizing to the least concentrated NanoDrop Spectrophotometer (Thermo) reading. Library preparation was subsequently carried out on the pooled sample using a TrueSeq Nano DNA LT kit (Illumina).

**Deep sequencing and data analysis.** Deep sequencing was performed on an Illumina MiSeq essentially as described in the manufacturer’s protocol for low-complexity ampiclon sequencing with the v3 reagent kit (~25% PhiX spike-in). Fastq output files were parsed using the Biopython package for python, and reads containing the expected sequences with 100% identity (either forward or reverse) were tallied for each barcode set. Relative frequency datapoints are plotted as the number of reads containing a particular barcode and the perfect or partially matching spacer of interest divided by the sum of that count plus reads containing the fully mismatched spacer with the same barcode.

**Pairwise competition assays with marked non-lysogens.** Aliquots of overnight cultures grown for 14 h were mixed 1:1 by volume and then subcultured for 45 min, prior to withdrawing an aliquot for selective plating on tetracycline (5 µg/ml) and chloramphenicol or erythromycin (5 µg/ml) and chloramphenicol, and subsequently infecting the remaining culture with φNM1γ at a multiplicity of infection of 4, followed by mixing of growth with additional subculture was performed in parallel to monitor growth in the absence of MMC or φNM1γ.

**Data availability.** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**References**
complexity regions was left unchecked. In order to compile the 70-spacer master list, the previously available collection of 39 unique spacers from S. aureus and S. argenteus type III-A systems was expanded to include spacers from the non-
aureus and non-argenteus Staphylococci in this cohort. The additional spacers were manually collected from sequences located between previously defined type III-A repeats, except for the downstream-most spacer in each CRISPR array, which was instead defined as the sequence located between a known repeat and a degenerate repeat. The sequence of each degenerate repeat was manually inferred based on its distal location in the CRISPR array, partial homology to the nearest upstream repeat, and length (~36 bp). Spacers identified in this manner were subsequently added to the master list, if not already present within the file.

Data availability: Relevant data supporting the findings of the study are available in this published article and its Supplementary Information files, and from the corresponding author upon request.

Received: 15 June 2017 Accepted: 11 December 2017
Published online: 04 January 2018

References

1. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. Nature 526, 55–61 (2015).
2. van Houwename, S., Buckling, A. & Westra, E. R. Evolutionary ecology of prokaryotic immunity mechanisms. Microbiol. Mol. Biol. Rev. 80, 745–763 (2016).
3. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712 (2007).
4. Marraffini, L. A. Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Nat. Rev. Microbiol. 15, 1–6 (2017).
5. Amitai, G. & Sorek, R. CRISPR Cas adaptation: insights into the mechanism of type III-A repeats, except for the downstream-most spacer in each CRISPR array, which was instead defined as the sequence located between a known repeat and a degenerate repeat. The sequence of each degenerate repeat was manually inferred based on its distal location in the CRISPR array, partial homology to the nearest upstream repeat, and length (~36 bp). Spacers identified in this manner were subsequently added to the master list, if not already present within the file.

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-017-02557-2 ARTICLE

25. Godberdgotit, S. et al. Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and prophages. Mol. Microbiol. 79, 35–49 (2011).
26. Fischer, S. et al. An archaeological immune system can detect multiple Protoplasmic Adjacent Motifs (PAMs) to target invader DNA. J. Biol. Chem. 287, 33351–33363 (2012).
27. Jiang, W. et al. Dealing with the evolutionary downside of CRISPR immunity: balancing and beneficial plasmids. PLoS Genet. 9, e1003844 (2013).
28. Edgar, R. & Qimron, U. The Escherichia coli CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. J. Bacteriol. 192, 6291–6294 (2010).
29. Goldberg, G. W. et al. A conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature 518, 633–637 (2014).
30. Vercoe, R. B. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Nat. Rev. Microbiol. 15, 1–6 (2017).
31. Marraffini, L. A. Resistance and tolerance to foreign elements by prokaryotic immune systems-curating the genome. Nat. Rev. Microbiol. 15, 717–724 (2015).
32. Heler, R. et al. Mutations in Cas9 enhance the rate of acquisition of viral spacer sequences during the CRISPR-Cas immune response. Mol. Cell 65, 168–175 (2017).
33. Gill, S. et al. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. J. Bacteriol. 187, 2426–2438 (2005).
34. Deng, L. et al. CRISPR-Cas immunity against phages: its effects on the evolution and survival of bacterial pathogens. J. Bacteriol. 190, 300–310 (2008).
35. Nair, D. et al. Whole-genome sequencing of Staphylococcus aureus strain RN4220, a key laboratory strain used in virulence research, identifies mutations that affect not only virulence factors but also the fitness of the strain. J. Bacteriol. 193, 2332–2335 (2011).
36. Kaye, T. B., Fischetti, V. A. & Fraser, P. W. Prophages in Staphylococcus aureus and their contribution to virulence. Mol. Microbiol. 62, 1035–1047 (2006).
37. Horinouchi, S. & Weisblum, B. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. 150, 815–825 (1982).
38. Holt, D. C. et al. A very early branching Staphylococcus aureus lineage lacking transposon/mobilization-related plasmid sequences. Gene 333, 881–895 (2003).
39. Horinouchi, S. et al. CRISPR sequence and functional map of pC194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin B type I antibiotics. J. Bacteriol. 170, 841–842 (1982).
40. Khan, S. A. & Novick, R. P. Complete nucleotide sequence of pT181, a tetracycline-resistance plasmid from Staphylococcus aureus. Plasmid 10, 251–259 (1983).
41. Cady, K. C., Bondy-Denomy, J., Heusssler, 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 phages. J. Bacteriol. 194, 5728–5738 (2012).
42. Nanda, A. M., Thomann, K. & Frountzi, J. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. J. Bacteriol. 197, 410–419 (2015).
43. Erdl, I., Campoy, S. & Barbe, J. A review of bacterial SOS response: ways to study bacterial SOS response. FEBS Mol. Microbiol. Rev. 31, 637–656 (2007).
44. Livny, J. & Friedman, D. I. Characterizing spontaneous induction of Stx encoding plasmids using a selectable reporter system. Mol. Microbiol. 51, 1691–1704 (2004).
45. Mo, C. Y., Birdwell, L. D. & Kohli, R. M. Specificity determinants for proteolysis of LexA, a key regulator of bacterial SOS mutagenesis. Biochemistry 53, 3158–3168 (2014).
46. Johnson, A. D. et al. Lambda repressor and c-cr components of an efficient molecular switch. Nature 294, 217–223 (1981).
47. Cao, L. et al. Characterization and functional study of type III-A CRISPR-Cas systems in clinical isolates of Staphylococcus aureus. Int. J. Med. Microbiol. 306, 686–696 (2016).
48. Anemans, D. M. et al. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive Staphylococcus aureus in Europe. MBio. 7, e00444-16 (2016).
49. Goerke, C. et al. Diversity of prophages in dominant Staphylococcus aureus clonal lineages. J. Bacteriol. 191, 3462–3468 (2009).
52. Manica, A., Zebec, Z., Steinkellner, J. & Schleper, C. Unexpectedly broad target recognition of the CRISPR-mediated virus defence system in the archaeon Sulfolobus solfataricus. *Nucleic Acids Res.* **41**, 10509–10517 (2013).

53. Maniv, I., Jiang, W., Bikard, D. & Marraffini, L. A. Impact of different target sequences on type III CRISPR-Cas immunity. *J. Bacteriol.* **198**, 941–950 (2016).

54. Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O. & Marraffini, L. A. Broad targeting specificity during bacterial type III CRISPR-Cas immunity constrains virus evolution. *Nature* **507**, 518–521 (2014).

55. Ferrer, M. D. et al. RinA controls phage-mediated packaging and transfer of virulence genes in Gram-positive bacteria. *Nucleic Acids Res.* **39**, 5866–5878 (2011).

56. Retardi, D. Within-host competition determines reproductive success of temperate bacteriophages. *ISME J.* **5**, 1451–1460 (2011).

57. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).

58. Staals, R. H. et al. RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol. Cell* **56**, 518–530 (2014).

59. Tamulaitis, G. et al. Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol. Cell* **56**, 506–517 (2014).

60. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5737 (2016).

61. Smargon, A. A. et al. Cas13b is a type VI-B CRISPR-associated RNA-guided RNAse differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630.e617 (2017).

62. Kazlauskienė, M., Tamulaitis, G., Kostiuk, G., Venclovas, C. & Siksnys, V. Spatiotemporal control of Type III-A CRISPR-Cas immunity: coupling DNA degradation with the target RNA recognition. *Mol. Cell* **62**, 295–306 (2016).

63. Niewoehner, O. et al. Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **543**, 543–548 (2017).

64. Drake, J. W. A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl Acad. Sci. USA* **87**, 7160–7164 (1991).

65. Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. Rates of spontaneous mutation. *Genetics* **148**, 1667–1688 (1998).

66. Cady, K. C. & O’Toole, G. A. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* **193**, 3433–3445 (2011).

67. Heussler, G. E. et al. Clustered regularly interspaced short palindromic repeat-dependent, biofilm-specific death of *Pseudomonas aeruginosa* mediated by increased expression of phage-related genes. *MBio* **6**, e00129–00115 (2015).

68. Golding, G. R. et al. Livestock-associated methicillin-resistant *Staphylococcus aureus* sequence type 398 in humans, Canada. *Emerg. Infect. Dis.* **16**, 587–594 (2010).

69. Kinnevey, P. M. et al. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome mec (SCC mec)-SCC-SCCIRISPR composite element in Irish hospitals. *Antimicrob. Agents Chemother.* **57**, 524–531 (2013).

70. Moore, S. D. & Prevelige, P. E. Jr. A P22 scaffold protein mutation increases the robustness of head assembly in the presence of excess portal protein. *J. Virol.* **76**, 10245–10255 (2002).

71. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

72. Hatoum-Aslan, A., Maniv, I., Samai, P. & Marraffini, L. A. Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J. Bacteriol.* **196**, 310–317 (2014).

73. Kiro, R., Shir dit, D. & Qimron, U. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol.* **11**, 42–44 (2014).

74. Martel, B. & Moineau, S. CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages. *Nucleic Acids Res.* **42**, 9504–9515 (2014).

75. Modell, J. W., Jiang, W. & Marraffini, L. A. CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature* **544**, 101–104 (2017).

76. Gryczan, T. J., Grandi, G., Hahn, J., Grandi, R. & Dubnau, D. Conformational alteration of mRNA structure and the posttranscriptional regulation of erythromycin-induced drug resistance. *Nucleic Acids Res.* **8**, 6081–6097 (1980).

77. Iandolo, J. J. et al. Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of *Staphylococcus aureus* 8325. *Gene* **299**, 109–118 (2002).

78. Ganguly, T. et al. Physicochemical properties and distinct DNA binding capacity of the repressor of temperate *Staphylococcus aureus* phage phi11. *FEBS J.* **276**, 1975–1985 (2009).

79. Quiles-Puchalt, N. et al. Staphylococcal pathogenicity island DNA packaging system involving cos-site packaging and phage-encoded HNH endonucleases. *Proc. Natl Acad. Sci. USA* **111**, 6016–6021 (2014).

Acknowledgements

We are grateful to Philip M. Nussenzweig, Robert Heler, and Pascal Maguin for additional advice on deep sequencing of low-complexity amplicon pools. We are also grateful to Robert Heler for instructions on extracting selection coefficient values from fitness data. We would like to thank Seth Darst for advice on promoter inactivation, as well as Charlie Mo for advice on site-directed mutations in proteins of the LexA superfamily. Finally, we would like to thank The Rockefeller University Genomics Resource Center for deep sequencing library preparation. L.A.M. is supported by the Burroughs Wellcome Fund PATH Award, an NIH Director’s Pioneer Award (DP1GM128184-01) and a HHMI-Simons Faculty Scholar Award.

Author contributions

G.W.G. and L.A.M. designed experiments. G.W.G. provided the initial conception of the study and conducted the experiments. E.A.M. advised on statistical reporting and carried out statistical tests. A.V., J.W.M., P.S. and W.J. contributed reagents and technical expertise. G.W.G. and L.A.M. wrote the paper.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-02557-2.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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

© The Author(s) 2017