



Phage gene expression and host responses lead to infection-dependent costs of CRISPR immunity

Sean Meaden¹ · Loris Capria¹ · Ellinor Alseth¹ · Sylvain Gandon² · Ambarish Biswas³ · Luca Lenzi⁴ · Stineke van Houte¹ · Edze R. Westra¹

Received: 17 March 2020 / Revised: 20 September 2020 / Accepted: 24 September 2020 / Published online: 3 October 2020
© The Author(s) 2020. This article is published with open access

Abstract

CRISPR-Cas immune systems are widespread in bacteria and archaea, but not ubiquitous. Previous work has demonstrated that CRISPR immunity is associated with an infection-induced fitness cost, which may help explain the patchy distribution observed. However, the mechanistic basis of this cost has remained unclear. Using *Pseudomonas aeruginosa* PA14 and its phage DMS3vir as a model, we perform a 30-day evolution experiment under phage mediated selection. We demonstrate that although CRISPR is initially selected for, bacteria carrying mutations in the phage receptor rapidly invade the population following subsequent reinfections. We then test three potential mechanisms for the observed cost of CRISPR: (1) autoimmunity from the acquisition of self-targeting spacers, (2) immunopathology or energetic costs from increased *cas* gene expression and (3) toxicity caused by phage gene expression prior to CRISPR-mediated cleavage. We find that phages can express genes before the immune system clears the infection and that expression of these genes can have a negative effect on host fitness. While infection does not lead to increased expression of *cas* genes, it does cause differential expression of multiple other host processes that may further contribute to the cost of CRISPR immunity. In contrast, we found little support for infection-induced autoimmunological and immunopathological effects. Phage gene expression prior to cleavage of the genome by the CRISPR-Cas immune system is therefore the most parsimonious explanation for the observed phage-induced fitness cost.

Introduction

CRISPR-Cas immune systems exhibit a widespread, but patchy, distribution in bacteria and archaea [1, 2]. This distribution is indicative of extensive gain and the loss of

such adaptive immune systems with high predicted rates of horizontal gene transfer [3–5]. Such gain or loss is likely the result of the balance between potential benefit, i.e., the targeting of phages and deleterious mobile genetic elements (MGEs) against the cost of carriage and expression [6, 7]. Put differently, fitness trade-offs associated with CRISPR immunity may select for alternative resistance strategies and the subsequent loss of CRISPR [8]. Previous work has demonstrated that the cost of CRISPR immunity against lytic phage can be inducible, where the cost is only realised upon phage infection [9, 10]. As a consequence, CRISPR immunity is only favoured over loss or mutation of the phage receptor at a low frequency of infection. In contrast, when infection risk is high, a constitutive defence, mutation of the phage receptor, is selected for since this is associated with a fixed cost that does not change, regardless of the number of encounters with phages [9].

Despite this understanding of the ecological conditions under which different strategies of resistance will be selected for, the mechanistic basis of the inducible cost of CRISPR immunity remains unclear [9, 11]. Using *Pseudomonas aeruginosa* strain PA14 and its lytic phage

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-020-00794-w>) contains supplementary material, which is available to authorised users.

✉ Sean Meaden
s.meaden@exeter.ac.uk

✉ Edze R. Westra
e.r.westra@exeter.ac.uk

- ¹ Biosciences, University of Exeter, Exeter TR10 9EZ, UK
- ² CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul Valéry Montpellier 3, Montpellier, France
- ³ Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand
- ⁴ Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK

DMS3vir, we perform a 30-day evolution experiment that supports the previously reported inducible cost of CRISPR immunity and then test hypotheses that have been proposed to explain this cost. Firstly, a cost of autoimmunity, where bacteria acquire self-targeting spacers during infections has been demonstrated in a number of systems and provides a plausible mechanistic basis for a cost of CRISPR immunity [12–15], although diverse mechanisms exist to avoid self-targeting spacer acquisition [16]. Ultimately the relative importance of such self-targeting events at the population level is unknown [17].

Secondly, a cost of enhanced expression of the CRISPR-Cas immune system following infection could also explain the induced fitness cost. Amongst a range of different cues that regulate CRISPR expression (reviewed in [18]), infection can trigger upregulation of *cas* genes and CRISPR arrays [19–21]. Enhanced expression presumably elevates the levels of protection but would likely be associated with a metabolic cost to the host.

The final proposed mechanism behind an inducible cost of CRISPR immunity is that expression of phage-encoded genes may take place prior to clearance of the infection, whereby the phage genome no longer has a deleterious effect on the host, which could be harmful to the host [22]. For example, phage can encode anti-CRISPR proteins which must be expressed before clearance by the immune system in order to function. Anti-CRISPRs are widespread, show a diversity of functions and are expressed rapidly upon infection [23–25]. It is therefore likely that other phage genes may also be expressed prior to clearance. Here we employ a combination of selection experiments, deep sequencing of CRISPR amplicons following phage infection and RNA sequencing approaches to explore the importance of each of these three non-mutually exclusive mechanisms in determining the induced fitness cost of CRISPR-Cas immune systems. Here we use extended evolution experiments that span 30 days under phage selection to examine the frequency of different resistance mechanisms that evolve.

Materials and methods

Bacterial strains and phages

The previously described *P. aeruginosa* strains UCBPP-PA14, isogenic mutant $\Delta cas1$ and phage DMS3vir have been previously described [26] and were used throughout this study. DMS3vir is a modified version of DMS3 that carries a truncated C-repressor, preventing lysogeny [26]. A variant of DMS3vir, known as DMS3mvir, which carries a modified protospacer such that it is targeted by WT PA14, was used for the $\Delta cas1$ competition experiments

and has been described previously [26]. Additionally, a variant of PA14 was used that carries two spacers targeting DMS3vir (BIM2 herein) has been described previously [9] and was used for competition experiments and gene expression profiling. A mutant of DMS3vir with the full *acr* operon, including the *acrIE3* and *aca1* genes and their promoter sequence, removed has previously been described in [27].

Evolution experiment

The evolution experiment was performed with six replicates by inoculating 6 mL M9 supplemented with 0.2% glucose with $\sim 10^6$ bacteria from an overnight cultures of the WT strain and adding 10^7 PFU of DMS3vir, followed by incubation at 37 °C while shaking at 180 rpm. Cultures were transferred daily 1:100 for 30 days. The ancestral phage (10^7 PFU) was added daily to maintain selection for resistance and to constantly induce the cost of CRISPR immunity.

Immunity and resistance profiling

Bacterial immunity against the ancestral phage was determined as described before [9, 11, 22] by streaking individual clones (24 clones per sample) through ancestral phage DMS3vir and phage DMS3vir carrying the anti-CRISPR F1 (*acrIF1*) gene. Bacterial clones sensitive to both phages were scored as “sensitive”, those resistant to DMS3vir but sensitive to DMS3vir + AcrIF1 were scored as “CRISPR immune”, and bacterial clones resistant to both phages were scored as “surface mutants”. CRISPR-Cas-mediated immunity was further confirmed by PCR (see Table S1 for primers). Surface modification was further confirmed on the basis of colony morphology (since phage DMS3vir is pilus-specific, surface mutants have motility defects, resulting in a modified colony morphology) and a lack of new CRISPR spacers. From these analyses fractions of each phenotype (sensitive, CRISPR immune, surface mutant) were calculated for each replicate experiment.

$\Delta cas1$ competition assays

Relative fitness of bacterial populations with CRISPR-mediated immunity was determined by direct competition with a surface mutant for 3 days. Competitions were carried out using blue–white colony screening between a *lacZ* marked surface mutant and either the WT PA14 or a PA14 mutant that lacks the *cas1* gene. We performed eight replicate experiments of each competition and used a gradient of phage titres (0, 10^6 , 10^7 and 10^8 PFU/mL, $n = 64$). The phage used (DMS3mvir) is a mutant of DMS3vir that is targeted by both the WT PA14 and the $\Delta cas1$ strain.

For competition assays, an overnight culture was grown of each strain and a marked surface mutant. Mixtures were made at a 50:50 ratio and plated at T0. These mixtures were added to 6 mL M9 media supplemented with 0.2% glucose and serially transferred with a 1:100 dilution. Populations were plated after 3 days on LB agar supplemented with 30 µg of X-gal to determine the relative frequencies of the competing strains. Selection coefficients were determined as described in [11].

Amplicon sequencing

In order to track the changes in genetic diversity of the CRISPR loci, we deep sequenced both CRISPR arrays (CRISPR1 and CRISPR2) from each population described above (see “Evolution experiment”) at 2, 4, 6, 10 and 12 days. Full bacterial genomic DNA was isolated using the Qiagen QIAmp DNA mini kit as per the manufacturer’s protocols. A PCR amplification was performed for both CRISPR arrays (CRISPR1 and CRISPR2, see Tables S1 and S2). PCR reactions contained 5 µl DreamTaq master mix (Thermo-Scientific, UK), 0.5 µl forward primer, 0.5 µl reverse primer, 1.5 µl MiliQ water, 0.5 µl DMSO, 2 µl template DNA. Sample purity was determined by NanoDrop and DNA concentrations were quantified using a Qubit fluorometer (Thermo-Fisher, UK). Two separate CRISPR primer (CRISPR1 and CRISPR2 locus) pairs were designed for two first-round PCRs. Two microliter of DNA was used in a first-round of PCR. The primer design incorporates a recognition sequence to allow a secondary nested PCR process (Table S2). Samples were first purified with Ampure SPRI Beads before entering the second PCR performed to incorporate Illumina adaptor sequences. Samples were purified using Ampure SPRI Beads before being quantified using Qubit and assessed using the Fragment Analyzer. Successfully generated amplicon libraries were taken forward and pooled in equimolar amounts, then size selected on a Pippin prep using a range of 180–600 bps. The quantity and quality of each pool was assessed by Bioanalyzer and subsequently by qPCR using the Illumina Library Quantification Kit from Kapa on a Roche Light Cycler LC480II according to manufacturer’s instructions. Template DNA was denatured according to the protocol described in the Illumina cBot User guide and loaded at 12.5 pM concentration. To help balance the complexity of the amplicon library 15% PhiX was spiked in. The sequencing of each pool was carried out on one lane of an Illumina MiSeq, at 2 × 250 bp paired-end sequencing with v2 chemistry.

Bioinformatics analysis

Sequence quality control

Base-calling and de-multiplexing of indexed reads was performed by CASAVA (version 1.8.2) (Illumina) to

produce 30 samples. FASTQ files were trimmed to remove Illumina adaptor sequences using Cutadapt version 1.2.1 [28]. The option “-O 3” was set, so the 3’ end of any reads which matched the adaptor sequence over at least 3 bp was trimmed off. The reads were further trimmed to remove low quality bases, using Sickle version 1.200 with a minimum window quality score of 20. After trimming, reads shorter than 10 bp were removed. The raw reads were subjected to a Cutadapt trimming step to remove PCR primer sequences that could potentially introduce an artificial level of complexity in the samples. To improve base quality in both read pairs, sequencing errors were corrected in both forward and reverse reads using the error-correct module within SPAdes sequence assembler, version 3.1.0 [29]. Read pairs were aligned to produce a single sequence for each pair that would entirely span the amplicon using PEAR (version 0.9.10 [30]). Additionally, sequences with uncalled bases (Ns) were removed. To remove sequences originating from potential PCR primer dimers or from any spurious amplification events, a size selection was applied to each merged sequence set, respectively between 30–140 bp for CRISPR1 and 70–500 bp for CRISPR2. Fragmented PhiX phage genome was added to the sequence library in order to increase the sequence complexity. To remove any “bleed through” of PhiX sequences, each sample was compared with the complete PhiX sequence (GenBank gi9626372) using BLASTN [31]. Sequences matching PhiX (*E* value < 10⁻⁵) were filtered out of the dataset.

Clustering and diversity metrics

For each dataset, any sequences passing the filters (from any sample) were merged into a single file. This final sequence file, plus its own metadata file describing each sample, was used for the analysis by using a custom pipeline based on QIIME 1.9.0 [32]. Clusters were defined using SWARM [33], using the strictest (default) parameters. This tool aggregates a sequence to a cluster if the sequence shows similarity with any of the sequences already present in that cluster. Importantly, the similarity threshold is not fixed but defined within the dataset. A minimum cluster size filter is applied to retain clusters containing at least two sequences and potential chimeric sequences due to PCR events were discarded as well. To calculate the abundance of each cluster, sequences were then aligned on the centroid sequence identified for each cluster, using a minimum similarity threshold of 99% for the entire length of the sequence using the “usearch_global” function in VSEARCH.

The sequencing depth of all samples was explored using the “Chao 1” [34] richness index plotted as a rarefaction curve. Counts in the cluster abundance tables were repeatedly sub-sampled (rarefied; 33 repetitions) at sampling

depths of 1000, 12,000, 22,000, ... 150,000. The average Chao 1 value obtained by repeating the test 33 times is assigned as alpha-diversity at that specific number of reads for that sample implemented in Qiime. Because all samples reached a clear asymptote, i.e., no samples were under-sampled with regards to spacer diversity, rarefaction was not applied. An abundance table for each locus was used to estimate the richness and evenness of the samples using the following estimators: total observed sequence variants, Shannon, Simpson, Simpson evenness again conducted with Qiime.

The resulting spacers were extracted from the clustered contigs via CRISPRdetect [35] and mapped to the phage and host genome. CRISPRtarget [36] was also used to identify spacers that mapped to the host genome.

Gene expression profiling

Ten replicate cultures of PA14 BIM2 and five of WT were grown overnight in shaken glass vials (180 RPM) of 6 mL LB media at 37 C then standardised to an optical density of 0.5 OD₆₀₀ (~2 × 10⁹ CFU/mL). Five cultures of BIM2 and five cultures of WT were inoculated with 8 × 10⁹ PFU (in 500 µL) DMS3vir (MOI 0.5). Five cultures of BIM2 remained uninfected as controls. All replicates were vortexed then incubated statically. After 35 min, 1 and 2 h, 1.5 mL of cells were pelleted and snap-frozen at – 80 C. For RNA extraction, 1 mL of TRIzol reagent was added immediately after removal from the freezer. The PureLink RNA mini extraction kit (Invitrogen) was used following instructions for use with TRIzol and the on-column PureLink DNase treatment (Invitrogen). All protocols followed the manufacturer's instructions except for a substitution of 100 µL BCP instead of chloroform for phase separation. Sample quantities, purities and size distribution were quantified by Qubit, nanodrop and TapeStation respectively. In total, 150 bp libraries were prepared with the TruSeq directional kit, following manufacturer's instructions, and sequenced on an Illumina NovaSeq at Exeter University sequencing centre.

The resulting reads were quality filtered using Sickle ([37] default settings, version 1.33) with ~6 million read pairs per sample remaining. These were subsequently mapped to the PA14 and DMS3vir reference genomes with bwa ([38] default settings, version 0.7.17). The resulting read alignments were assigned to genomic features and counted using HTSeq (version 0.11.2 [39], using "union" mode, non-unique setting set to "none" and the reverse stranded option). Differential expression analysis was conducted in R using the DESeq2 package [40].

Phage gene expression programme was determined by normalising each gene by maximum and minimum expression levels then averaged across the five replicates of

WT PA14 infected by DMS3vir. These profiles were clustered using the "hclust" function in the R "stats" package. A dendrogram was made with the "dendextend" [41]. Sequence data are available from the European Nucleotide Archive under accession number PRJEB31514.

Competition assays in the presence of Acr(–) phage

In order to determine the effect of *acrIE3* and *acal* gene expression on the fitness of immune hosts we competed a WT PA14 strain carrying two spacers that target DMS3vir (BIM2) against a LacZ marked surface mutant in the presence of a phage lacking the full Acr operon and a WT control. While both of hosts are immune to DMS3vir, the surface mutant prevents phage genome injection and therefore any subsequent phage gene expression. Competition experiments were carried out as described above for 24 h under with the addition of 10⁵, 10⁷ and 10⁹ PFU DMS3vir.

Prol cloning and assay

The protease I encoding gene from DMS3vir was cloned into the pHERD30T vector under control of an arabinose inducible promoter. The protease I encoding gene was amplified from DMS3vir via PCR and introduced into pHerD30T between an EcoRI and an XbaI site, downstream of an arabinose inducible promoter. Primers used for cloning are available in Table S1. Colonies were selected using a gentamycin marker and blue/white colony screening in a background of DH5-alpha cells (New England Biolabs, UK). The resulting construct was transformed into PA14 and the clone used for assays was verified by Sanger sequencing (University of Sheffield, UK). Optical density was recorded during a 24-h growth curve at 37 C in LB supplemented with 1% arabinose. Growth curves of the clone expressing the protease I encoding gene and an empty vector control were recorded. Growth rate during log-phase was extracted from the growth curves in R (version 3.5.2).

Results

Fitness costs of CRISPR can select for alternative resistance strategies, and the subsequent invasion of bacteria carrying surface-based resistance [9]. However, it has remained unclear what the mechanistic basis of this cost is. We envisaged three possible reasons for this inducible cost: (1) a cost of autoimmunity, where bacteria acquire self-targeting spacers during infections (2) a fitness cost due to enhanced expression of the CRISPR-Cas immune system following infection or (3) due to expression of phage-encoded genes in the period between phage genome

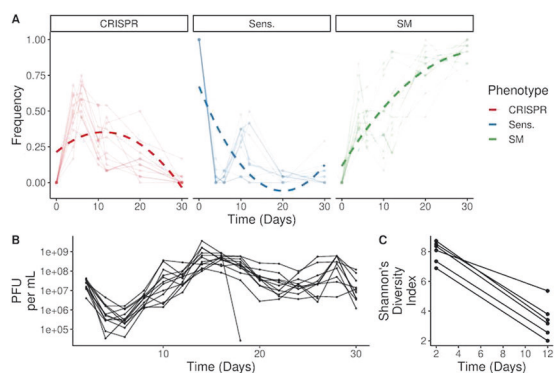


Fig. 1 Population dynamics of each bacterial genotype and phage titres. **a** Each point represents the proportion of bacterial phenotypes ($n = 24$ clones per population) that exist in each population. Shaded lines represent individual populations. Dashed lines represent model fit from a GLM that includes a polynomial spline (knots = 2). **b** Lines represent the phage titres present in each population throughout the evolution experiment. **c** Genetic diversity (Shannon's diversity index) of the CRISPR2 array during the first 12 days of the experiment from amplicon sequence data.

injection and CRISPR-mediated clearance of the phage. To explore the first hypothesis, we first carried out an experiment where we infected WT PA14 daily with a high titre of phage DMS3vir, followed by deep sequencing of CRISPR amplicons to measure the frequency of self-targeting spacers. As expected, bacteria initially evolved high levels of CRISPR immunity with corresponding spacer acquisition, which was followed by a consistent increase in the frequency of surface mutants due to the induced fitness costs of CRISPR immunity in the presence of phage (Fig. 1). Deep sequencing of the CRISPR arrays from these bacterial populations revealed a corresponding decline in genetic diversity of CRISPR arrays (Fig. 1). Because of primed spacer acquisition, most spacers target the phage DMS3vir genome in the area proximal to the priming site (position 27847), and the distribution of target sites on the phage genome is similar to that reported in previous studies on this empirical system (Fig. S1 [9, 27]) as well of that of other empirical systems where primed spacer acquisition was observed [42–44]. Arrays carrying self-targeting spacers with either a canonical PAM site ($n = 3$) or a non-canonical PAM site ($n = 3$) to the original host–genome were rare. Moreover, the majority of the spacers with a canonical PAM were only observed in a single sample, consistent with the idea that there is strong selection against a self-targeting spacer. These results therefore suggest that self-targeting is either deleterious or rare. By comparing the number of spacers acquired from the phage with either a canonical or non-canonical PAM site we estimated the background ratio between these acquisition events. By identifying the number of reads that mapped to arrays carrying self-targeting spacers with a non-canonical PAM, we can estimate the proportion of cells that self-

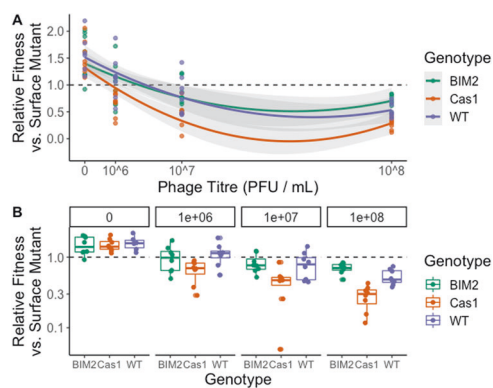


Fig. 2 Relative fitness of bacterial populations with CRISPR-mediated immunity at 3 d.p.i. when competing with a surface mutant. Competitions were carried out between a marked surface resistant strain and either the WT PA14 (purple), a mutant PA14 lacking a functional *cas1* (orange) or a WT-derived strain carrying two additional spacers targeting phage DMS3mvir (BIM2; green). Experiments were carried out across a gradient of phage titres, as indicated. The phage used (DMS3mvir) is a mutant of DMS3vir that is targeted by WT PA14. $n = 8$ replicates per genotype per phage titre. **a** Data presented with phage titre as a continuous variable. **b** Data grouped by phage titre to better visualise differences within titres.

target. We found that 0.0068% of cells likely self-target (see Supplementary information for Methods). As PAM sites are a key component of self-target avoidance, this further suggests the phenomenon is rare.

To explore whether these self-targeting events have a measurable negative impact on the fitness of the bacteria, we performed a competition experiment between a surface mutant and either WT bacteria with CRISPR immunity, which can acquire self-targeting spacers, or a *cas1* deletion mutant with CRISPR immunity that is unable to acquire novel spacers. This competition was performed across a gradient of titres of DMS3mvir (a mutant of DMS3vir that is targeted by spacer1 in CRISPR array 2 of the PA14Δ*cas1* strain; hence both the surface mutant and competing bacterial genotypes are insensitive to this phage). If the induced fitness cost of CRISPR-Cas immunity was due to the acquisition of self-targeting spacers, we would expect the *cas1* deletion mutant to have a higher relative fitness compared to the WT strain. However, this was not observed; the *cas1* deletion mutant had a lower relative fitness than the WT strain during competition with the surface mutant (GLM, $F_{1,70} = 17.85$, $p < 0.001$, Fig. 2). This is likely due to the evolution of escape phages that carry mutations in their target site, which the *cas1* mutant cannot overcome. Indeed we found a higher number of escape phages evolving in the *cas1* mutant compared to the WT strain at the high titre (10^8 PFU per mL) (GLM, $F_{1,14} = 15.5$, $p < 0.01$), but this was not observed at a lower phage titre (10^6 PFU per mL) (GLM, $F_{1,15} = 2.2$, $p = 0.16$). In summary these data show that, even at low phage titres when escape phages are likely to be extremely rare, the

benefits of spacer acquisition outweigh any autoimmune effects. Taken together with the deep sequencing data, these results do not imply that the acquisition of self-targeting spacers is the primary driver of the observed phage-induced fitness cost that is associated with CRISPR immunity.

To explore the second hypothesis, that the induced fitness cost is caused by induction of CRISPR-Cas gene expression, we performed RNAseq of uninfected controls as well as CRISPR-immune bacteria infected with DMS3vir. To avoid evolution of “escape phage”, which carry mutations in the protospacer, we infected a CRISPR-immune clone that carries two spacers targeting DMS3vir (BIM2), and extracted RNA at $t=0$, 35, 60 and 120 min post infection. BIM2 shows complete immunity to DMS3vir and no phage particles are produced following infection (limit of detection = 200 PFU/mL, data not shown; also see [11]). Differential expression analysis found no evidence that CRISPR-Cas expression is enhanced following infection (Fig. 3). Of all *cas* genes, only *cas1* showed a significant difference, with slightly lower expression in infected BIM2 populations relative to controls (FDR adjusted p value = 0.017, log₂ fold change = -0.375). These data therefore do not support the idea that the induced cost of CRISPR immunity observed in this experimental system is due to the upregulation of CRISPR-Cas expression.

Interestingly, we identified a number of other host genes that showed significantly different gene expression in the infected BIM2 strain relative to the uninfected controls (Table S3).

Across all three timepoints, genes involved in motility, polyamine breakdown and transport and metabolism were significantly upregulated. The two polyamines associated with phage infection were putrescine and spermidine which are linked to biofilm production (reviewed in [45]). In addition to these polyamines, we found many components

of flagellum production to also be upregulated subsequent to phage infection including flagellar rod proteins and the flagellar motor switch. This suggests that bacteria are producing and activating flagella in response to infection. In contrast, we found a significant downregulation of many genes implicated in chemotaxis as well as oxidoreductases and cytochrome *c* genes and metabolism associated genes. These results suggest a shift in both the metabolism and motility of the bacteria in response to phage infection. Although the communication molecule PQS has previously been shown to be produced during phage infection in order to repel sensitive bacteria from the area [46] we found no significant upregulation of PQS genes. We also observed no obvious effect of time on host gene expression, with the most upregulated genes all occurring at similar levels throughout the experiment.

Finally, to test whether phage gene expression prior to clearance of the infection by the immune system could be responsible for the induced fitness cost, we measured whether phage gene expression could be detected in the infected BIM2 populations. To this end, we mapped all the RNAseq reads to the DMS3vir genome. This revealed significant levels of genome-wide phage gene expression in the BIM2 populations, despite their CRISPR-based immunity. In total, phage expression in the BIM2 populations was around fivefold lower compared to infected WT populations (Figs. 4 and S2). Amongst the expressed phage genes, we identified very high expression levels of the anti-CRISPR (*acrIE3*) and its associated repressor (*acaI*) (Fig. S3), and to a lesser degree, many other genes (Fig S3). *AcrIE3* and *acaI* have previously been reported to be amongst the most strongly expressed genes on the phage DMS3vir genome [25]. This *acrIE3* protein is specific for type I-E CRISPR-Cas systems and does not impact the type I-F CRISPR-Cas system encoded by PA14. Collectively,

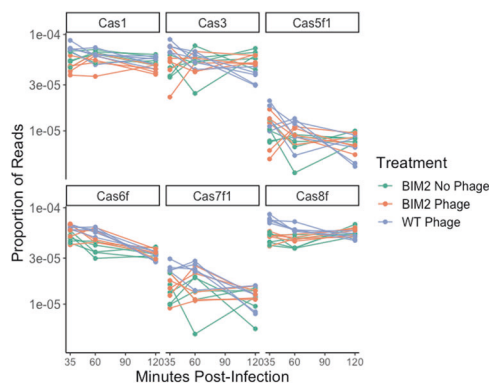


Fig. 3 Expression of CRISPR associated genes following infection with 8×10^9 PFU DMS3vir (MOI 0.5) across PA14 WT (purple, $n = 5$), PA14 BIM2 (orange, $n = 5$) and PA14 BIM2 uninfected controls (green, $n = 5$). Samples taken 35, 60 and 120 min post infection.

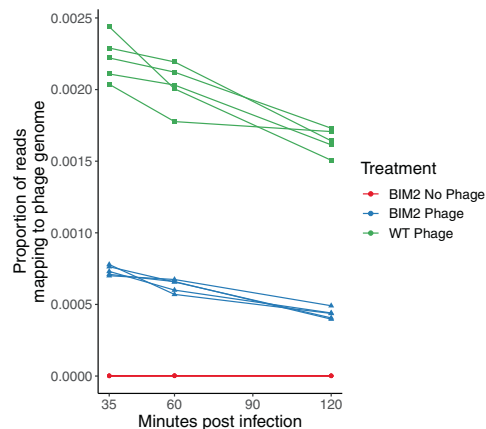


Fig. 4 Total phage gene expression following infection with 8×10^9 PFU DMS3vir (MOI 0.5) across PA14 WT (green, $n = 5$), PA14 BIM2 (blue, $n = 5$) and PA14 BIM2 uninfected controls (red, $n = 5$). Samples taken 35, 60 and 120 min post infection.

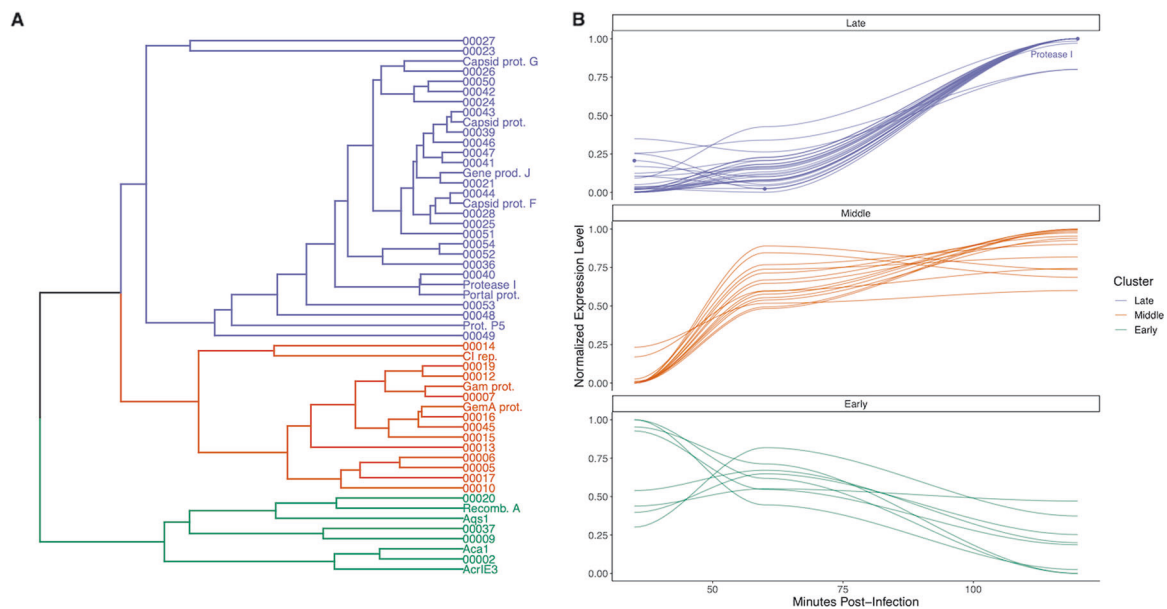


Fig. 5 Normalised phage gene expression of DMS3vir infecting WT PA14. Normalised gene expression was averaged across replicates ($n = 5$ per time point) and grouped by hierarchical clustering to identify distinct phases of the transcriptional programme. **a** Dendrogram showing discrete clusters with annotated genes and hypothetical

proteins (numbered). **b** Normalised gene expression of each gene through time grouped by each cluster. Lines show local-regression (loess) model fits. Points denote values for Protease I gene expression.

these data show that the phage is capable of expressing its genes prior to the completion of CRISPR-mediated cleavage.

To test whether expression of the *acrIE3* or *aca1* genes might contribute to the phage-induced fitness cost that is observed in CRISPR-immune bacteria, we next competed the BIM2 strain and the surface mutant in the presence of either WT DMS3vir or a mutant carrying a deletion of the entire *acr* operon. This revealed a similar fitness cost of CRISPR immunity regardless of the phage genotype (GLM, effect of genotype on relative fitness of BIM vs SM, $F_{1,34} = 2.00$, $p = 0.16$), suggesting that the high expression levels of the *acrIE3* and *aca1* genes is not causing the observed induced fitness cost in the host. However, given that many other phage genes are also expressed, albeit at lower levels, the induced cost may well be due to expression of one or more of these other phage genes.

We next clustered the phage genes using their normalised gene expression during infection of WT PA14 (i.e., carries an imperfect CRISPR match which enables priming but not immunity), following the methods of [47] (Fig. 5). We found that consistent with Stanley et al. [25] the *acrIE3* and *aca1* gene are expressed early. We also searched the annotations for potentially damaging genes and identified a protease I gene. Expression of the protease I gene might be particularly costly, since proteolytic activity could conceivably cause cytotoxicity in the cell. This gene is located immediately downstream of the *acrIE3* and *aca1* genes and was found to be expressed both in this experiment as well as

a preliminary, independent RNAseq experiment using nanopore sequencing (data not shown). Given that protease I is involved in virion assembly, a deletion mutant would likely not be viable, we therefore cloned and expressed the protease gene in WT *P. aeruginosa* to measure the cost of expression of this gene for the host. This showed that protease I expression reduced cell growth rates by ~13% relative to an empty vector control (Fig. S4, GLM, $F_{1,46} = 28.72$, $p < 0.001$). This suggests that expression of phage genes prior to phage genome cleavage by the CRISPR-Cas immune system can be costly for the host. Given that this is just one of the many genes expressed by the phage during infection, we argue that phage gene expression prior to clearance of the infection is an important source of the observed phage-induced cost of CRISPR-based immunity.

We developed a mathematical model to explore the competition between different resistant bacteria in the presence or in the absence of phages. Only the CRISPR-immune bacteria are potentially affected by the presence of phages because surface mutants cannot be infected. Our model shows that the dynamics we observe experimentally can be recreated even when self-targeting is absent, provided there is a cost associated with phage infection because of induced immunity or toxicity (see Supplementary information). Importantly, the model predicts that in the absence of a phage-induced cost, bacteria with CRISPR-Cas systems that experience high levels of self-targeting would decline in frequency (Fig. S5C). If the CRISPR-immune bacteria is deficient in spacer acquisition (i.e., $p = 0$) it would not

decline in frequency (Fig. S5A). However, this result is inconsistent with our experimental data, which shows that both bacteria with fully functional CRISPR immunity and those deficient in spacer acquisition lose the competition with surface mutants (Fig. 2). Therefore, this model suggests the observed population dynamics are best explained by toxicity effects.

Discussion

Previous work has demonstrated that CRISPR-immune populations experience infection-dependent costs [9, 10]. Here, we test three hypotheses that might explain the observed infection-induced cost of CRISPR immunity. We find that upregulation of *cas* genes is unlikely to be the cause of this cost. While we cannot rule out increased *cas* protein activity levels, previous work found that plasmid targeting was unaffected by phage pre-infection [48]. Additionally, we find that autoimmunity due to acquisition of self-targeting spacers, although present, is unlikely to be an important source of the observed fitness costs of CRISPR immunity. Although quantifying the true extent of auto-immune effects of self-targeting spacers is difficult as they, by definition, are rapidly lost from the populations, we also observe very low numbers of self-targeting spacers that target sequences with non-canonical PAM. As these spacers would be selected against less strongly it supports the notion that self-targeting spacer acquisitions are rare [12]. Moreover, we see that carrying CRISPR in the absence of phages is not costly [9], that an acquisition-deficient mutant is equally fit to the wild type in the absence of phages (Fig. 2), and there is no detectable difference in CRISPR-Cas expression in the presence or absence of phages (suggesting that rates of self-targeting spacer acquisition would be independent of the presence or absence of phage). This result is also consistent with previous studies that report higher rates of acquisition from MGEs than the host genome, even in the absence of priming [13, 49].

Our data demonstrate that there are high levels of phage gene expression following infection, even with highly effective CRISPR targeting. Phage gene expression likely creates cytotoxic effects for the infected bacterial cells and host gene expression is fundamentally altered. The finding that phages can rapidly express their genes prior to clearance by the CRISPR-Cas immune system is consistent with previous studies [25]. For example, phages that encode *acr* genes can successfully infect bacteria with CRISPR immunity, even though the Acr proteins are produced during the infection cycle. This has selected for extremely rapid and strong expression of the *acr* gene (*acrIE3*) and its repressor (*aca1*) [25], which we also observed in our RNAseq data. These high expression levels of the *acrIE3* gene, which

targets type I-E CRISPR-Cas immune systems and not the type I-F system of the PA14 strain used here, was not associated with a detectable cost. However, we found genome-wide phage gene expression suggesting multiple opportunities for phage mediated damage prior to clearance (Fig. S4). We also found that plasmid-based expression of a phage gene (protease I) showed cytotoxicity and may be cleaving host proteins as well as performing its likely role in capsid packaging [50]. We suggest this may be representative of a widespread cost of phage gene expression and further work is needed to assess the point at which phage DNA no longer has a negative effect on the host cell. Given that, in some cases, the CRISPR-immune system preferentially targets the portion of the phage genome that is injected into the cell first, it appears that preventing phage gene expression as soon as possible is crucial for immunity [51]. Understanding the exact timing of CRISPR interference remains a vital question as the effects of targeting during different parts of the phage replication cycle may be important for the efficiency of immune response. Indeed similar effects may also occur during anti-plasmid immunity, when a plasmid carries genes that are toxic to the host [52]. In addition to the protease we identified a number of host processes that are induced upon phage infection, which may further contribute to the observed infection-induced fitness cost of CRISPR-immune bacteria. Notably the export of spermidine, breakdown of putrescine and construction and activation of flagella may all lead to intrinsic energetic costs. It is unclear why these host processes are induced and their adaptive significance. Given that *Pseudomonas* phages appear unable to replicate on hosts deficient for spermidine production [53], export of spermidine would likely interfere with phage propagation. The upregulation of degradation and export of these polyamines is interesting as polyamine production has been linked to biofilm production and protection from antibiotics—exogenous polyamines may also confer a protective barrier against phages. It is also possible that the change in host processes is a manipulation from the phage, analogous to some eukaryotic viruses that enhance host movement during infection in order to facilitate greater virus dispersal (reviewed in [54]). In this context the phage could induce flagella-based motility in order to disseminate virions to a wider area. An alternative is that flagella production is a host response to leave an area of high relatedness and therefore limit subsequent infections to kin. Such movement away from infected populations has been previously observed in *P. aeruginosa* [46] and motility has been shown to confer a fitness advantage in the presence of phages—although the precise mechanism is unclear [55].

In addition to the effects we observed, there may be additional unknown factors that arise from phage infection. These could include cellular stress from membrane piercing during phage injection, or the costs of producing phage

proteins even if they are not cytotoxic. Similarly, Mu-like phages replicate via replicative transposition, which involves integrated into the genome during the lytic cycle, and it is possible that CRISPR targeting of the phage during this step would be costly to the host. Such targeting during lysogenic infection has been observed to result in the upregulation of the SOS response [56] and can lead to the loss of a functional CRISPR system [8]. Although we do not see significant upregulation of the genes associated with this response, we cannot rule out this effect from happening in a sub-population of cells (<1% based on the effect sizes seen in Heussler et al. [56]). As more phage defences are discovered and described it will be fascinating to see how their specific costs and benefits will compare to CRISPR and other well described mechanisms, such as restriction modification. For example, the lag between infection and detection may represent a widespread cost intrinsic to all post-infection defence systems. Our results suggest that although the maintenance of a CRISPR system is likely to be low-cost, the infection-dependent cost is likely to be much higher. Therefore, if the frequency of phage challenge is high, constitutive mechanisms are selected for. Moreover, the interplay between costs and benefits may vary by environment [57]. For example, in our liquid media experiments, spatial structure is absent, or low, which may reduce the requirement for a pilus (the DMS3vir receptor). In a highly structured environment, such trade-offs may differ. More generally, the ecological context will influence how costly it is to lose the pilus, for example coexistence with other bacteria can amplify the costs [58]. In turn, these results may help explain the patchy distribution of CRISPR systems observed in nature as different niches will be associated with varying trade-offs and infection risks. Indeed it is increasingly understood that CRISPR systems may be transferred among bacteria leading to the frequent gain and loss of such systems. This may, in part, be driven by the relative cost–benefit relationship between inducible cost and infection frequency [59].

Acknowledgements ERW acknowledges the Natural Environment Research Council, and the Biotechnology and Biological Sciences Research Council for funding. ERW further acknowledges the Wellcome Trust and the European Research Council for funding. SVH acknowledges funding from BBSRC (grant number BB/R010781/1). We also thank Angus Buckling for valuable discussions and the anonymous reviewers for their constructive comments. RNA sequencing was conducted at the University of Exeter sequencing centre which is supported by the Wellcome Trust.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- van Belkum A, Soriaga LB, LaFave MC, Akella S, Veyrieras JB, Barbu EM, et al. Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. MBio. 2015;6:e01796–15.
- Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJ, et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol. 2019;18:1–17.
- Portillo MC, Gonzalez JM. CRISPR elements in the Thermococcales: evidence for associated horizontal gene transfer in *Pyrococcus furiosus*. J Appl Genet. 2009;50:421–30.
- Watson BN, Staals RH, Fineran, PC. CRISPR-Cas-mediated phage resistance enhances horizontal gene transfer by transduction. MBio. 2018;9:e02406–17.
- Varble A, Meaden S, Barrangou R, Westra ER, Marraffini LA. Recombination between phages and CRISPR-cas loci facilitates horizontal gene transfer in staphylococci. Nat Microbiol. 2019;4: 956–63.
- Levin BR. Nasty viruses, costly plasmids, population dynamics, and the conditions for establishing and maintaining CRISPR-mediated adaptive immunity in bacteria. PLoS Genet. 2010;6: e1001171.
- van Houte S, Buckling A, Westra ER. Evolutionary ecology of prokaryotic immune mechanisms. Microbiol Mol Biol Rev. 2016;80:745–63.
- Rollie C, Chevallereau A, Watson BN, Chyou TY, Fradet O, McLeod I, et al. Targeting of temperate phages drives loss of type I CRISPR-Cas systems. Nature. 2020;578:149–53.
- Westra ER, van Houte S, Oyesiku-Blakemore S, Makin B, Broniewski JM, Best A, et al. Parasite exposure drives selective evolution of constitutive versus inducible defense. Curr Biol. 2015;25:1043–9.
- Vale PF, Lafforgue G, Gatchitch F, Gardan R, Moineau S, Gandon S. Costs of CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. Proc R Soc B: Biol Sci. 2015;282:20151270.
- van Houte S, Ekroth AK, Broniewski JM, Chabas H, Ashby B, Bondy-Denomy J, et al. The diversity-generating benefits of a prokaryotic adaptive immune system. Nature. 2016;532:385–8.
- Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. Self-targeting by CRISPR: gene regulation or autoimmunity? Trends Genet. 2010;26:335–40.
- Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature. 2015;520:505–10.
- Vercoc RB, Chang JT, Dy RL, Taylor C, Gristwood T, Clulow JS, et al. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS Genet. 2013;9:e1003454.

15. Staals RH, Jackson SA, Biswas A, Brouns SJ, Brown CM, Fineran PC. Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR–Cas system. *Nat Commun.* 2016;7:1–13.
16. Weissman JL, Stoltzfus A, Westra ER, Johnson PL. Avoidance of Self during CRISPR immunization. *Trends Microbiol.* 2020;28:543–53.
17. Koonin EV. Open questions: CRISPR biology. *BMC Biol.* 2018;16:1–3.
18. Patterson AG, Yevstigneyeva MS, Fineran PC. Regulation of CRISPR–Cas adaptive immune systems. *Curr Opin Microbiol.* 2017;37:1–7.
19. Quax TE, Voet M, Sismeiro O, Dillies MA, Jagla B, Coppee JY, et al. Massive activation of archaeal defense genes during viral infection. *J Virol.* 2013;87:8419–28.
20. Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, Shah M, et al. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS ONE.* 2012;7:e38077.
21. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J Mol Biol.* 2010;395:270–81.
22. Chabas H, van Houte S, Høyland-Kroghsbo NM, Buckling A, Westra ER. Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. *Proc R Soc B: Biol Sci.* 2016;283:20160721.
23. Pawluk A, Staals RH, Taylor C, Watson BN, Saha S, Fineran PC, et al. Inactivation of CRISPR–Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat Microbiol.* 2016;1:1–6.
24. Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, et al. Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature.* 2015;526:136–9.
25. Stanley SY, Borges AL, Chen KH, Swaney DL, Krogan NJ, Bondy-Denomy J, et al. Anti-CRISPR-associated proteins are crucial repressors of anti-CRISPR transcription. *Cell.* 2019;178:1452–64.
26. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O’Toole GA. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J Bacteriol.* 2012;194:5728–38.
27. Chevallereau A, Meaden S, Fradet O, Landsberger M, Maestri A, Biswas A, et al. Exploitation of the cooperative behaviors of anti-CRISPR phages. *Cell Host Microbe.* 2020;27:189–98.
28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17:10–12.
29. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19:455–77.
30. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics.* 2014;30:614–20.
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–10.
32. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010;7:335–6.
33. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ.* 2015;3:e1420.
34. Chao A. Nonparametric estimation of the number of classes in a population. *Scand J Stat.* 1984;11:265–70.
35. Biswas A, Staals RH, Morales SE, Fineran PC, Brown CM. CRISPRDetect: a flexible algorithm to define CRISPR arrays. *BMC Genom.* 2016;17:1–14.
36. Biswas A, Gagnon JN, Brouns SJ, Fineran PC, Brown CM. CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol.* 2013;10:817–27.
37. Joshi N. & Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for fastQ files (version 1.33) [software]. 2011. <https://github.com/najoshi/sickle>.
38. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009;25:1754–60.
39. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31:166–9.
40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
41. Galili T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics.* 2015;31:3718–20.
42. Richter C, Dy RL, McKenzie RE, Watson BN, Taylor C, Chang JT, et al. Priming in the Type I CRISPR–Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.* 2014;42:8516–26.
43. Fineran PC, Gerritzen MJ, Suárez-Diez M, Künne T, Boekhorst J, van Hijum SA, et al. Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc Natl Acad Sci USA.* 2014;111:E1629–38.
44. Xue C, Seetharam AS, Musharova O, Severinov K, Brouns J, Severin SJ, et al. CRISPR interference and priming varies with individual spacer sequences. *Nucleic Acids Res.* 2015;43:10831–47.
45. Wortham BW, Oliveira MA, Patel CN. Polyamines in bacteria: pleiotropic effects yet specific mechanisms. In: (eds Perry RD, Fetherston JD). *The genus Yersinia*. New York, NY: Springer; 2007. p. 106–15.
46. Bru JL, Rawson B, Trinh C, Whiteson K, Høyland-Kroghsbo NM, Siryaporn A. PQS produced by the *Pseudomonas aeruginosa* stress response repels swarms away from bacteriophage and antibiotics. *J Bacteriol.* 2019;201:e00383–19.
47. Doron S, Fedida A, Hernández-Prieto MA, Sabehi G, Karunker I, Stazic D, et al. Transcriptome dynamics of a broad host-range cyanophage and its hosts. *ISME J.* 2016;10:1437–55.
48. Landsberger M, Gandon S, Meaden S, Rollie C, Chevallereau A, Chabas H, et al. Anti-CRISPR phages cooperate to overcome CRISPR–Cas immunity. *Cell.* 2018;174:908–16.
49. Wei Y, Terns RM, Terns MP. Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. *Genes Dev.* 2015;29:356–61.
50. Fokine A, Rossmann MG. Common evolutionary origin of pro-capsid proteases, phage tail tubes, and tubes of bacterial type VI secretion systems. *Structure.* 2016;24:1928–35.
51. Modell JW, Jiang W, Marraffini LA. CRISPR–Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature.* 2017;544:101–4.
52. van Sluijs L, van Houte S, van Der Oost J, Brouns SJ, Buckling A, Westra ER. Addiction systems antagonize bacterial adaptive immunity. *FEMS Microbiol Lett.* 2019;366:fnz047.
53. Wagemans J, Blasdel BG, Van den Bossche A, Uytterhoeven B, De Smet J, Paeshuysse J, et al. Functional elucidation of anti-bacterial phage ORFans targeting *Pseudomonas aeruginosa*. *Cell Microbiol.* 2014;16:1822–35.
54. van Houte S, Ros VI, van Oers MM. Walking with insects: molecular mechanisms behind parasitic manipulation of host behaviour. *Mol Ecol.* 2013;22:3458–75.
55. Taylor TB, Buckling A. Bacterial motility confers fitness advantage in the presence of phages. *J Evol Biol.* 2013;26:2154–60.
56. Heussler GE, Cady KC, Koepfen K, Bhujji S, Stanton BA, O’Toole GA. Clustered regularly interspaced short palindromic

- repeat-dependent, biofilm-specific death of *Pseudomonas aeruginosa* mediated by increased expression of phage-related genes. *MBio*. 2015;6:e00129–15.
57. Hernandez CA, Koskella B. Phage resistance evolution in vitro is not reflective of in vivo outcome in a plant-bacteria-phage system. *Evolution*. 2019;73:2461–75.
58. Alseth EO, Pursey E, Luján AM, McLeod I, Rollie C, Westra ER. Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature*. 2019;574:549–52.
59. Westra E, Levin BR. How important is CRISPR-Cas for protecting natural populations of bacteria against infections with badass DNAs? *BioRxiv*. 2020. <https://www.biorxiv.org/content/10.1101/2020.02.05.935965v2>.