

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

CRISPR vs. Anti-CRISPR: How bacterial viruses fight CRISPR-Cas immunity

Permalink

<https://escholarship.org/uc/item/0s97t7px>

Author

Borges, Adair Louise

Publication Date

2020

Peer reviewed|Thesis/dissertation

CRISPR vs. Anti-CRISPR: How bacterial viruses fight CRISPR-Cas immunity

by
Adair Borges

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:



E2BBE6A63C2745B...


Alexander Johnson

Chair

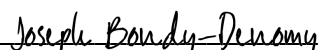
DocuSigned by:



DocuSigned by:



DocuSigned by:



D23522498F4247C...

Seemay Chou

Leor Weinberger

Joseph Bondy-Denomy

Committee Members

Copyright 2020

by

Adair L. Borges

Acknowledgements

I would like to thank my thesis advisor, Dr. Joseph Bondy-Denomy, for his support and excellent mentorship throughout my graduate work. The training I received in the Bondy-Denomy lab is of the highest quality, and I am extremely grateful for everything I learned there. I am very thankful for all the great friends that I have made in the Bondy-Denomy lab - it is a privilege to do science with such incredible people. I would especially like to thank my two Berkeley students, Jenny Yujie Zhang and Allyson Park, for being the best undergraduate researchers imaginable. I am very lucky to have had the opportunity to work with such brilliant young individuals. I am very grateful to all of my scientific collaborators at UCSF and beyond, I have learned so much from all of them and gotten to do lots of exciting and thought-provoking science. I would also like to thank my thesis committee, Drs. Seemay Chou, Sandy Johnson, and Leor Wienberger, for their valuable mentorship, guidance, and scientific insight. I would like to give an extra special thank you to Dr. Seemay Chou for being an amazing mentor and important part of my life at UCSF. Finally, thank you to my undergraduate advisor, Dr. Jon Boyle at University of Pittsburgh, for his excellent mentorship and for teaching me the foundations of scientific research.

Contributions

Chapter 1 contains work previously submitted for publication:

Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs. *Annu. Rev. of Virol.* **4**, 37–59 (2017).

Chapter 2 contains work previously submitted for publication:

Borges, A.L., Zhang JY, Rollins MF, Osuna BA, Wiedenheft B, Bondy-Denomy J. Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. *Cell* **174**, 917–925 (2018)

Chapter 3 contains work previously submitted for publication:

Marino ND*, Zhang JY*, **Borges AL***, Sousa AA, Leon LM, Rauch BJ, Walton RT, Berry JD, Joung JK, Kleinstiver BP, Bondy-Denomy J. Discovery of widespread type I and type V CRISPR-Cas inhibitors. *Science* **362**, 240–242 (2018).

*Equal contribution

Chapter 4 contains work previously submitted for publication:

Borges, A. L., Castro B, Govindarajan S, Solvik T, Escalante V, Bondy-Denomy J.. Bacterial alginate regulators and phage homologs repress CRISPR–Cas immunity. *Nat. Microbiol.* (2020). doi:10.1038/s41564-020-0691-3

CRISPR vs. Anti-CRISPR:

How bacterial viruses fight CRISPR-Cas immunity

Adair L. Borges

Abstract

Bacteria and the viruses that infect them (bacteriophages/phages) are locked in an ancient evolutionary arms race. Bacteria use CRISPR-Cas systems as a form of anti-phage immunity and phages encode “anti-CRISPR” proteins that antagonize CRISPR-Cas function. Anti-CRISPRs are high-affinity CRISPR-Cas inhibitors, but face a great challenge in rapidly neutralizing all the CRISPR-Cas complexes in the cell upon phage infection. I show that CRISPR defeats phages with anti-CRISPRs in >90% of infection events, and rare successes occur only when multiple phages cooperate to infect the same cell. Failed infections “immunosuppress” the bacterium by the production of anti-CRISPRs prior to phage destruction, increasing chances of survival for other co-infecting phages. This is the first description of altruism in viruses. This cooperative strategy for CRISPR-Cas neutralization is completely dependent on phage multiplicity of infection: if the concentration of phage is too low for coinfections to occur at an appreciable rate, the phage population will go extinct. While this strategy may be suitable for some classes of mobile elements, others may be unable to obtain cooperation thresholds, necessitating enzymatic or hyper-potent inhibitors. In searching for new mechanisms of CRISPR-Cas neutralization, we have now discovered anti-CRISPRs that function substoichiometrically, one of which has been shown to be a multi-turnover enzyme. Furthermore, I show that some phages have hijacked a transcriptional repressor of CRISPR-Cas immunity, and deploy it to limit CRISPR-Cas biogenesis. The CRISPR/anti-CRISPR arms race is a hotbed for evolutionary innovation, and is a source of novel and inventive mechanisms of immune neutralization.

Table of Contents

Chapter 1: Introduction.....	1
References.....	28
Chapter 2: Bacteriophage cooperation suppress CRISPR-Cas3 and Cas9 immunity....	43
References.....	56
Chapter 3: Discovery of widespread Type I and Type V CRISPR-Cas inhibitors.....	81
References.....	85
Chapter 4: Bacterial alginate regulators and phage homologs repress CRISPR-Cas immunity.....	103
References.....	113

List of Figures

Chapter 1

1.1. Anti-CRISPR (<i>acr</i>) locus organization.....	39
1.2. Mechanisms for anti-CRISPR protein function.....	41
1.3. Anti-Anti-CRISPR mechanisms.....	42

Chapter 2

2.1. Anti-CRISPRs are imperfect CRISPR-Cas inhibitors.....	70
2.2. Sensogram of AcrIF4 binding the Csy complex.....	71
2.3. Anti-CRISPR success requires cooperative infections during lytic growth.....	72
2.4. Output phages from liquid growth experiments remain CRISPR sensitive.....	73
2.5. Generating and validating Hybrid _{<i>acr</i>} phages.....	74
2.6. Immunosuppression facilitates acquisition of a marked prophage.....	75
2.7. Prophage content of immunosuppressed lysogens.....	76
2.8. Cas9 anti-CRISPR AcrIIA4 requires cooperative infection to neutralize Type II-A CRISPR immunity.....	77

Chapter 3

3.1. The discovery of a widespread Type I inhibitor.....	92
3.2. AcrIF11 phylogenetic tree.....	93
3.3. Guilt-by-association in self-targeting strains leads to Type V-A and Type I-C anti-CRISPR proteins.....	94

Chapter 4

4.2. Characterization of <i>kinB::Tn</i> mutants.....	127
4.1. A forward genetic screen identifies a role for an alginate-activating pathway in repressing CRISPR-Cas immunity.....	128
4.3. Double knockouts of pathway members.....	129
4.4. The KinB/AlgB pathway modulates Cas3 and Csy protein and RNA levels.....	130
4.5. AmrZ is a surface-activated repressor of CRISPR-Cas immunity.....	131
4.6. AmrZ activity in liquid growth.....	132
4.7. Phage-derived AmrZ homologs control CRISPR-Cas immunity.....	133
4.8. Cas and Csy RNA and protein levels across growth conditions.....	135

List of Tables

Chapter 2

2.1. DMS3m _{acr} phage genotypes.....	79
2.2. Plasmids used in this study.....	80

Chapter 3

3.1. A table of previously discovered <i>aca</i> genes (<i>aca1-3</i>) and <i>aca</i> genes found in this study (<i>aca4-7</i>).....	95
3.2. Protein sequences and accession numbers of Acr and Aca proteins found in this study.....	96
3.3 List of all accessions of AcrIF11 homologs.....	99
3.4. List of Aca accessions.....	101
3.5. Plasmids used for expression in <i>Pseudomonas aeruginosa</i>	102

Chapter 4

4.1. Mapped insertions from transposon mutagenesis screen.....	136
4.2. Mobile AmrZ homologs.....	137
4.3. AmrZ copy number analysis of two <i>Pseudomonas aeruginosa</i> strains.....	138
4.4. Strains and phages used in this study.....	139
4.5. Plasmids and primers used in this study.....	140

Chapter 1: Introduction

The genetic diversity of microbes ensures their widespread colonization of the planet. In addition to surviving in wide-ranging hostile niches, such as the human body and the depths of the ocean, microbes face a constant onslaught from viruses ¹. Bacteriophages (phages) are bacterial viruses that have earned the reputation of being the most abundant biological entities on the planet, with estimates of $\sim 10^{31}$ particles on earth ². Phages are intrinsically specific for the bacterial host that they infect, typically being restricted to a single bacterial species and even a subset of strains within that species. This specificity has enabled the careful dissection of the molecular determinants of phage-host interactions in many model bacterial systems, leading to an array of fundamental biological discoveries and ground breaking biotechnologies.

Anti-phage defence

Along with the widespread presence of viruses on the planet, anti-viral immune pathways are ubiquitous across the tree of life. Bacteria achieve resistance to viral infection through diverse mechanisms that can be broadly classified into those that act before phage genome injection and those that manifest after the phage nucleic acid is in the cell. Prior to phage injection, receptors on the cell surface are required for successful phage adsorption to the cell surface. These receptors can be absent, mutated, or masked through specific modifications as an anti-phage mechanism ³. Additionally, the poorly understood process of phage genome injection can be inhibited by proteins localizing to the cytoplasmic membrane or periplasm ⁴⁻⁶. Remarkably, many of these anti-phage mechanisms are encoded by integrated phages (prophages) and operate through their host as a phage superinfection exclusion mechanism.

Once inside the cell, phages that are entering the lytic cycle hijack host processes to convert the cell into a viral factory. Before phage replication proceeds to completion, the phage nucleic acid (often DNA) may be degraded by bacterial immune systems that target foreign DNA, such as restriction enzymes or CRISPR-Cas nucleases. Many of these anti-phage mechanisms have been described in detail in excellent reviews on the subject (See ^{7,8}). Other intracellular immune systems such as bacteriophage exclusion (BREX) in *Bacillus subtilis* ⁹ and phage inducible chromosomal island-like elements (PLE) in *Vibrio cholerae* ¹⁰ have recently been discovered although their mechanisms of action remain obscure.

Despite the numerous powerful systems that bacteria employ to block phage entry and replication, the abundance of phages on the planet shows that these mechanisms have not driven phages to extinction. This can be explained, in part, by a plethora of phage-encoded mechanisms that inhibit these bacterial defences. Phages can degrade restrictive outer membrane modifications, modify tail proteins to utilize alternate receptors, modify their DNA to avoid restriction endonucleases, and encode protein inhibitors of various bacterial processes ¹¹. Here, we discuss the mechanisms by which phages evade CRISPR-Cas function.

CRISPR-Cas Systems

Clustered regularly interspaced short palindromic repeats and CRISPR associated genes (CRISPR-Cas) comprise a bacterial adaptive immune system that utilizes RNA-guided nucleases to cleave foreign nucleic acids. By acquiring small fragments of DNA from a foreign element during an initial exposure, the CRISPR array forms a chronological record of past genomic transgressors ¹²⁻¹⁴. The repetitive elements in the CRISPR array provide semi-palindromic functional elements for both the construction of the CRISPR array and the process of interfering with foreign DNA, while the 'spacer' elements between the repeats specify the

sequence of the target. In fact, the first hints that CRISPR-Cas might comprise an adaptive immune system against phages was the identification of spacer sequences that are identical to phage genomes¹⁵⁻¹⁷. To function, the CRISPR array is transcribed and processed to generate mature CRISPR RNAs (crRNAs) which often possess repeat derived regions at the 5'- and 3'-ends, with the spacer-encoded sequence in the middle^{18,19}. This crRNA is assembled with 1-6 Cas proteins, depending on the type of system and this complex will surveil the cell²⁰. Upon recognition of invading complementary nucleic acid, nuclease activity of at least one enzyme is activated and mediates the destruction of that target²¹⁻²⁴. There are currently six distinct CRISPR-Cas types, which possess completely distinct sets of proteins that enable function. The details of the CRISPR-Cas types and mechanisms of action have been the subject of excellent reviews (see^{25,26}). Here, we will focus on the first two systems for which anti-CRISPRs were discovered, Type I (Cas3) and II (Cas9).

Type I CRISPR-Cas systems utilize an RNA-guided protein complex consisting of 3-5 proteins that process and guide the crRNA to a complementary target and signal for the recruitment of the *trans*-acting nuclease known as Cas3^{18,27,28}. Type I CRISPR-Cas systems are further categorized into multiple subtypes with distinct RNA-guided protein complexes (I-A through I-F), with all utilizing the Cas3 signature protein for DNA degradation²⁰. In contrast, Type II systems possess a single effector protein, Cas9. Cas9 participates in spacer acquisition, crRNA processing (together with a trans-encoded small RNA tracrRNA and RNase III), target identification, and cleavage^{19,29-32}. Type II CRISPR-Cas systems are also broken down into three subtypes (II-A through II-C), possessing distinct Cas9 homologs. Due to a reliance on a single protein for function, Cas9 homologs derived from different subtypes and species have been utilized for numerous far-reaching gene editing applications in recent years³³. In common between the type I and II CRISPR-Cas systems, they possess a reliance on near-perfect complementarity between the crRNA and a DNA target, and a subtype-specific protospacer

adjacent motif (PAM) ³⁴⁻³⁶. Point mutations in the PAM or the PAM-proximal region of the protospacer (denoted as the 'seed') can result in phages or plasmids that can escape CRISPR targeting and proceed to replicate despite a near perfect or perfect spacer match ³⁷.

Anti-CRISPR Genes in *Pseudomonas aeruginosa*

In the human pathogen, *Pseudomonas aeruginosa*, prophages have been implicated in phenotypes such as toxin production and overall virulence ^{38,39}. However, the mechanisms by which *P. aeruginosa* prophages modulate the physiology of their hosts are poorly understood. An effort to discover and characterize novel prophage-mediated phenotypes in this organism led to the serendipitous identification of the first phage-encoded inhibitors of CRISPR-Cas function. A survey of 30 distinct *P. aeruginosa* prophages revealed many examples of superinfection exclusion ⁴⁰. Surprisingly, examples were observed where a subset of integrated prophages *licensed* infection by a superinfecting phage, allowing a phage to infect the lysogenized host. This observation was highlighted by a $>10^6$ fold change in the efficiency of plating for phages that did not form plaques on the wild-type, unlysogenized strain, but were able to infect and replicate in the lysogenic strain ⁴¹. The same phages that could only infect the lysogenized host had been previously shown to be targeted by the natural type I-F CRISPR spacers in the very same wild-type strain ⁴², leading the authors to speculate that prophages were inactivating CRISPR-Cas function. By comparing the genomes of phages that were sensitive to the action of CRISPR and those that were inactivating it, an 'anti-CRISPR locus' was identified. Many related phages from a single phage family possessed genes in this locus, that were small (i.e. 150-450 base pairs) and of unknown function. Despite overall synteny and broad conservation of gene sequences throughout the rest of these phages, the anti-CRISPR locus was quite diverse (**Fig. 1.1a**). By testing these genes in isolation, five were attributed anti-CRISPR function, based on their ability to allow infection of a CRISPR-Cas targeted phage ⁴¹. These genes are now known

as *acrF1-F5*. In addition to being encoded by this closely related family of ‘Mu-like’ phages (i.e. phages that utilize transposition to replicate), homologs of these genes were also identified in conjugative islands and plasmids, suggesting a broad role in enhancing horizontal gene transfer in *Pseudomonas* ^{41,43}.

The *P. aeruginosa* type I-F CRISPR-Cas system also possesses non-canonical function, via an interaction with a prophage possessing a target sequence with five mismatches. This interaction leads to the inhibition of biofilm formation and swarming motility ^{44,45}. Furthermore, it was shown that phages that should have been targeted by the system (i.e. they possessed perfect matches to spacer sequences) were unhindered in their ability to replicate ⁴⁶. The inability of this system to block phage replication led to the conclusion that this CRISPR-Cas system did not perform canonical functions. In hindsight, it is now clear that the relevant phages being tested in this study possessed *acr* homologs that prevented the detection of CRISPR-Cas activity and the correction of the five mismatches to four or zero mismatches in phage DMS3 (which lacks an *acrF* gene) caused it to be targeted through canonical CRISPR-Cas activity ⁴². Recent work has revealed that the DMS3 prophage with five mismatches triggers an SOS response as a result of a self-targeting genome cleavage event, which causes death upon the initiation of group behaviors ⁴⁷. Together, these results highlight that *acrF* genes are important both during lytic and lysogenic growth, to protect a phage with perfect or mismatched protospacer targets. During lysogeny, the constitutive expression of an *acr* gene generates an immunocompromised host, which is now sensitive to other phages that CRISPR would have previously blocked. While this seems maladaptive for the prophage and lysogen, the inhibition of CRISPR-Cas function is an obligate part of lysogenic survival as genome cleavage that would result from self-targeting of the prophage would be lethal in the absence of an anti-CRISPR.

In addition to the type I-F system, type I-E systems had also been identified in many *P. aeruginosa* genomes⁴⁶. While type I-F function was serendipitously identified in a widely-used lab strain (PA14), a strain with a functional type I-E system had to be actively searched for among sequenced strains possessing this system. Strains with active systems were identified by designing plasmids possessing protospacers and assessing transformation efficiencies. Ultimately, two *P. aeruginosa* strains with functional type I-E systems were found. Using one of these strains, it was then possible to discover four distinct type I-E anti-CRISPR (*acrE1-4*) genes⁴⁸. The *acrE* genes were found as genomic neighbors to the *acrF* genes in the same family of Mu-like prophages. In many cases, individual phages were identified that encode both an *acrF* and an *acrE* gene.

The Discovery of anti-CRISPRs in Diverse Bacterial Species

Anti-CRISPR loci in the *P. aeruginosa* Mu-like phages possess a stereotypical genomic architecture (**Fig 1.1a**), with 1-3 *acrE/F* genes followed by a highly conserved gene that is referred to as anti-CRISPR associated gene 1 (*aca1*). While the anti-CRISPR genes possess no significant shared sequence identity between them, *aca1* homologs encode proteins with 95% sequence identity in this family of phages and only occur in phages that possess *acr* genes. While homologs of *acrE/F* genes have been found in diverse mobile elements within *Pseudomonas* species, homology searches did not identify any hits outside of this genus, making it difficult to predict whether ACRs are widespread. The conservation of *aca1*, however, provided a robust bioinformatics tool to identify novel *acr* genes both in and outside of *Pseudomonas* (**Fig. 1.1b**). Utilizing this conserved gene as a query, two new *acrF* genes (*acrF6*, *acrF7*) were discovered in *Pseudomonas* mobile elements. Excitingly, *acrF6* homologs were discovered in diverse Gammaproteobacteria and some of these homologues proved to be active against the *P. aeruginosa* I-F system, representing the first anti-CRISPRs found outside

of the *Pseudomonas* genus. Also important, one *acrF6* homologue was found next to a gene encoding an HTH-motif containing protein that was distinct from *aca1*, and was given the name *aca2*. Homologs of *aca2* led to the discovery of *acrF8-F10* in diverse organisms, as well as many other candidates that did not possess anti-CRISPR activity in *P. aeruginosa*. Together, these anti-CRISPR genes were identified broadly across the *Proteobacteria* phylum⁴⁹. Notably, some members of each new *acrF* genes discovered in this manner displayed a broad host range, inactivating the I-F systems of *P. aeruginosa* and *Pectobacterium atrosepticum*. The Cas proteins of the *Pectobacterium* system range from 40 to 60% sequence identity with their *P. aeruginosa* orthologues. This broad host range was a feature of only *acrF1* and *acrF2* from the original group. Furthermore, this study yielded the first dual specificity anti-CRISPR protein (*acrF6_{Pae}*) which could also inactivate the type I-E CRISPR-Cas system⁴⁹. This work demonstrated the power of utilizing a “guilt by association” bioinformatics approach to discover small, novel genes of unknown function next to *aca1/2* homologs as a method for the discovery of new anti-CRISPRs.

The Discovery of Anti-CRISPRs Inhibiting Cas9

Performing further BLAST searches with *Aca2* led to a gene encoding a putative Acr in a strain of *Brackiella oedipodus*⁵⁰. This strain did not encode a type I CRISPR-Cas system, but did encode a type II-C (Cas9) system, leading to the hypothesis that this might be a Cas9 inhibitor. A homologue of this putative Acr was found in *Neisseria meningitidis*, which also possesses a type II-C CRISPR-Cas system (**Fig. 1.1b**). Subsequent experiments showed that these proteins did inhibit the *N. meningitidis* Cas9 system in its natural context, proving the existence of anti-CRISPRs against a Cas9-based system. Further bioinformatic investigation uncovered two more families of anti-CRISPRs functioning to inhibit *N. meningitidis* Cas9 (*NmeCas9*). Excitingly, these anti-CRISPRs were also found to function in human cells to inhibit genome

editing mediated by NmeCas9. These studies were important in showing that Acrs were not limited to type I CRISPR-Cas systems.

Upon acquisition of foreign DNA encoding an *acr* gene, one outcome is that a strain may now possess an element in the genome with a target of the CRISPR-Cas system. For example, a temperate phage with a targeted protospacer and PAM can avoid CRISPR targeting by deploying an Acr protein, allowing integration and stable lysogeny due to continued production of the Acr protein. This results in a situation described as “self-targeting.” The continued expression of an Acr protein is now an essential process in this cell as loss of Acr expression will result in lethal genomic cleavage. This premise was utilized as a bioinformatics strategy to identify strains that possessed the first Acr proteins encoded in a Gram positive microbe⁵¹. Four distinct inhibitors of the Type II-A CRISPR-Cas9 system in *Listeria monocytogenes* were identified (*acrIIA1-4*), guided by examples of genomic self-targeting (**Fig 1.1c**). Most *acr* genes in this system were encoded by prophages in *L. monocytogenes* genomes, with some *acr* homologs being found in distant phages and plasmids of *L. monocytogenes* and other Firmicutes. Two of these newly discovered Acr proteins (AcrIIA2 and AcrIIA4) were able to block function of the widely used *S. pyogenes* Cas9 both in an *E. coli* test system and in a genome editing assay in human cells. *L. monocytogenes* and *S. pyogenes* Cas9 are 53% identical, showing that AcrIIA proteins can also function against distinctive systems.

The type II-A and type II-C Acrs together present important new additions to the Cas9 engineering toolkit, derived from the phage-bacterium arms race. Much work remains to be done for the AcrIIA/C proteins to understand how widespread they are, how many distinct proteins perform this task, and what the evolutionary implications are for their presence.

Anti-CRISPRs are widespread

While ~50% of bacteria possess CRISPR-Cas systems, an open question is whether any given system is active and able to respond to foreign DNA invasion⁵². Although it is impossible to experimentally interrogate every microbe possessing a type I-F system, for example, one can utilize bioinformatics to predict whether a given system may be capable of being inactivated by known *acr* genes. In the I-F system, where ten *acrF* sequences are available, it appears that nearly every known I-F system it is possessed by a genus that also has a known *acr* gene or a is closely related one that fits those criteria⁴⁹. This suggests widespread inactivation of I-F CRISPR-Cas systems across its entire distribution. As the best-studied group of *acr* genes, this provides a hypothesis going forward, that every CRISPR-Cas system may possess a similar and concomitant abundance of *acr* genes throughout its distribution.

The *acrIIC* genes were also found beyond the organism they were discovered in (*N. meningitidis*), suggesting the potential for widespread Type II-C CRISPR-Cas inactivation as well⁵⁰. The coverage was not as striking as the *acrF* genes, however, suggesting that there are likely *acrIIC* genes to discover. The *acrIIA* genes told a slightly different story, while homologs of *acrIIA2-4* were found only in *Listeria* and *Streptococcus* prophages and plasmids, *acrIIA1* homologs were found broadly across the Firmicutes. This distribution included many species encoding Type II-A CRISPR-Cas systems, suggesting widespread inactivation of Cas9 in these organisms. As the discovery of new *acr* genes continues, it will be exciting to track where their homologs are encoded to determine what percentage of CRISPR-Cas systems are likely “inhibitable” and what this will mean for bacteria and phages on an evolutionary timescale.

Anti-CRISPR Gene Organization

To defend against commonly encountered Type I-E and Type I-F CRISPR-Cas systems, many *P. aeruginosa* phages maintain an *acrE* gene alongside an *acrF* gene in their anti-CRISPR loci. The most interesting example of phage response to dual I-E and I-F CRISPR-Cas targeting is the evolution of the *P. aeruginosa* phage allele of *acrF6* (*acrF6_{Pae}*), which is a single protein possessing dual I-E and I-F inhibitory activity. This dual activity is unique to the *P. aeruginosa* phage allele, as *acrF6* homologs from 5 other diverse bacteria did not inhibit the I-E system of *P. aeruginosa*. Interestingly, this anti-I-E activity could be abolished by a C-terminal truncation of the final 2 residues of the AcrF6_{Pae} protein, all while leaving the anti-I-F activity of the protein unaffected. In contrast to the pervasive co-association of heterotypic Acrs in *P. aeruginosa* phages, examples of two *acrF* or *acrE* genes appearing together in the same genome is much rarer. The singular locus architecture with *acrF* co-association is the co-occurrence of *acrF3* and *acrF5*. Interestingly, *acrF3* often occurs in the absence of *acrF5*, but *acrF5* is never found without *acrF3*. The functional significance of this unique genetic interaction is unknown and remains to be investigated. In contrast to I-F and I-E inhibitors, multiple II-A inhibitors are often encoded together in the same ACR locus. An estimated 75% of *acrIIA* loci encode more than one AcrIIA protein, whereas only approximately 7% of *acrF* loci have both *acrF3* and *acrF5* and there are no examples of *acrE* genes occurring in tandem. Dominating the *acrIIA* landscape is *acrIIA1*, which pervasively co-occurs with *acrIIA2-4*, demonstrating a potential multi-pronged attack on the *L. monocytogenes* CRISPR-Cas system.

Mechanisms of Acr Function

A notable feature of each family of Acr proteins is their lack of sequence similarity to any proteins of known function. Furthermore, besides being small (~50-150 amino acids), there are no common features among them. For these reasons, no insight into the mechanisms of Acr

function could be gained until experimental studies on individual Acrs were undertaken. The first such study provide an in vitro mechanistic characterization of AcrF1, F2 and F3 ⁵³. It was shown that each of the Acrs functions through a different mechanism. AcrF1 and AcrF2 both bound to the type I-F CRISPR-Cas (Csy) complex, but they did this by binding to different subunits of the complex (**Fig 1.2**). AcrF1 bound with a stoichiometry of 2 or 3 to the Cas7f (Csy3) subunit, which is present in 6 copies in the Csy complex. By contrast, AcrF2 bound to the Csy complex with a stoichiometry of 1, and interacted with the Cas8f (Csy2):Cas5f (Csy1) heterodimer. Both of these Acrs inhibited the DNA-binding activity of the Csy complex. However, AcrF2 directly competed with DNA for a site on the Csy complex, while AcrF1 interacted with a site removed from the DNA interaction site. Interestingly, AcrF1 could still form a complex with the DNA-bound Csy complex if the DNA was added first. AcrF3 directly bound to the Cas3 helicase-nuclease protein and prevented its recruitment to the Csy-DNA complex. AcrF4 bound to the Csy complex like AcrF1 and AcrF2 but specific details were not obtained ⁵³.

To gain structural resolution of AcrF3 interacting with Cas3, co-crystal ⁵⁴ and cryo-EM ⁵⁵ structures were recently published. These structures revealed an AcrF3 dimer, where each monomer makes multiple asymmetric contacts with many residues and domains of Cas3. This effectively covers an entire face of the Cas3 protein, approximately 2,500 Å² in surface area ⁵⁴. The large size of this interaction interface suggests that individual mutations in Cas3 are unlikely to successfully evade AcrF3 function. In contrast to these results with AcrF3, an NMR solution structure of AcrF1 coupled with extensive mutagenesis revealed a small patch of the protein was required for function ⁵⁶. A single tyrosine to alanine mutation at position 6 of AcrF1 was sufficient to inactivate anti-CRISPR function *in vivo* and *in vitro*. These two Acr protein structures highlight the diversity of structure and mechanism of these inhibitor proteins.

Other Possible anti-CRISPR Mechanisms

In contrast to newly discovered CRISPR/anti-CRISPR antagonism, other forms of bacterial immunity/counter-immunity are better studied and the specific details of counter-immunity evolution are better understood⁵⁷. Across diverse systems, bio-mimicry is employed as a mechanism to inhibit immune activity. This leads us to hypothesize that Acr proteins could have evolved by similar mechanisms of bio-mimicry and bacterial gene hi-jack.

Nucleic acid mimics: Restriction enzymes comprise the bacterial “innate” immune system and have been studied for many years⁵⁸. Diverse inhibitors of restriction enzyme immunity have been discovered, many which ultimately function by shielding the phage DNA from enzymatic attack using base modification⁵⁹. However, other inhibitors work by mimicking phage DNA and tightly sequestering restriction enzymes⁶⁰. The T7 *ocr* gene, which is an immediate-early gene that T7 uses to inhibit restriction activity in its *E. coli* host is highly acidic and structurally resembles 24 bp of bent B-form DNA⁶¹. Similarly, the *ardA* gene, a widely distributed inhibitor of Type I restriction system, functions as a homo-dimer that mimics a 42 bp stretch of B-form DNA⁶². Viral bio-mimicry of DNA is also seen in eukaryotic systems, where a virally encoded DNA mimic binds histones and is hypothesized to disrupt nucleosome assembly and prevent repair of DNA breaks⁶³.

Like restriction enzymes, CRISPR-Cas systems bind DNA and in principle, should be susceptible to inhibition by DNA mimics. DNA-binding activities that are independent of the sequence of the spacer-derived crRNA (i.e. the PAM site) ascribed to both type I and II CRISPR-Cas systems could provide a weakness for Acr DNA mimics to exploit. While inhibitors of RNA-binding CRISPR-Cas systems⁶⁴ have not yet been reported, RNA bio-mimicry could similarly function as an anti-CRISPR strategy. Furthermore, though RNA-based mechanisms of

anti-CRISPR activity have not been discovered, we hypothesize that virally encoded small RNAs could mimic crRNAs and interfere with CRISPR activity. Specifically, crRNA mimics could function by outcompeting bona-fide crRNAs for Cas proteins during CRISPR-complex assembly, or by directly displacing crRNAs in pre-loaded complexes. Interestingly, some *Clostridium* phages carry CRISPR arrays, the biological function of which is unknown⁶⁵.

Cas genes as proto-Acrs: Horizontal gene transfer between phages and host bacterial species is pervasive, and CRISPR-Cas elements have been found in phage genomes previously. In a striking example of horizontal Cas gene transfer, *Vibrio cholerae* phages acquired a I-F CRISPR Cas system which they deploy to inhibit a novel *V. cholerae* DNA-based anti-phage immune system¹⁰. Since Cas proteins interact in complex with each other, an Acr that mimics a Cas protein or Cas protein motifs could compete with or disrupt these bona fide Cas-Cas interactions. Despite these predictions, the structures of *P. aeruginosa* Acr proteins, AcrF1⁵⁶ and AcrF3^{54,55} bear no obvious resemblances or topology similarities to any of the *P. aeruginosa* Cas proteins for which there are structures: Cas1, Cas2-3, or Cas6^{18,54,66}. Further structural characterization of Cas:Acr protein interactions is urgently needed, especially for the multi-protein complexes utilized in type I systems as the structural intricacies of this complex may be absent without the interaction partners and crRNA present. This information will help illuminate the currently obscure evolutionary history of Acr proteins.

Anti-CRISPRs as Modifiers of CRISPR-Cas Function

The nuances of AcrF function go far beyond simple inhibition of interference, as these proteins have the ability to enable or disable new functions that were not initially predicted. For example, the inhibition of Cas3 recruitment mediated by AcrF3 converted the CRISPR system into a sequence-specific transcriptional repressor (CRISPRi) when the system was targeted to a

promoter region. This repression presumably occurred because the crRNA-guided complex could bind DNA and block RNA polymerase recruitment, but DNA cleavage did not occur⁵³. Type I CRISPRi had previously been demonstrated in the type I-E system via the deletion of the Cas3 nuclease^{67,68} and in the Type II system by catalytic inactivation of Cas9^{69,70}. This demonstrates the ability of an Acr protein to leave CRISPR-Cas function partially intact and may therefore enable new functionalities such as “natural CRISPRi.” Additionally, the discovery of priming acquisition (a mechanism of spacer acquisition that requires all components of the CRISPR-Cas system), connected the spacer acquisition and interference pathways, which were previously thought to be separate¹⁴. With this connection, it became clear that while AcrF proteins bind to the interference factors in the Type I-F system (Csy complex, Cas3), this also functions to block new spacer acquisition⁷¹.

The recent discovery of Cas9 inhibitors (AcrIIA and AcrIIC proteins) that were both able to interfere with Cas9 gene editing activity in human cells^{50,51} suggests that these also directly bind to Cas9. Indeed, AcrIIC1, 2 and 3 proteins all form direct physical interactions with type II-C Cas9 from *N. meningitidis* but not type II-A from *S. pyogenes*⁵⁰. These data broadly show the utility for Acr proteins to function in heterologous hosts with potential benefits such as enabling CRISPRi (Cas3 inhibition) and providing an off-switch for gene editing applications and dCas9-based CRISPRi applications.

While direct interactions with Cas proteins present a logical solution for phages to CRISPR-Cas based immunity, we envision many future strategies to achieve this end result. For example, base modifications have been previously shown to block Type II⁷² and Type I⁷³ immunity, although the Type II results seem to depend on the guide RNA design⁷⁴. Certain mechanisms of phage injection and replication may also be recalcitrant to CRISPR targeting, as was recently shown for phage T5, which injects its genome gradually and therefore only ~10% of the DNA is

a substrate for effective CRISPR-Cas immunity ⁷³. Additionally, Acr proteins that conduct enzymatic inactivation or destruction of Cas proteins or the crRNA, or transcriptional repression of any component of the CRISPR-Cas system might sufficiently shift the balance in favor of the phage during infection.

Anti-CRISPRs are phage accessory genes

Phage genomes are highly mosaic, possessing distinct functional “modules” with unique evolutionary histories ^{1,75,76}. Individual modules are assorted into phage genomes through diverse mechanisms of DNA recombination and/or ligation, and high fitness combinations being selected for, whereas low fitness assemblages are purged from the phage population. The frequency with which modules are moved in or out of genomes of related phages creates a conservation pattern that allows for the designation of core and accessory genes across a population of related phages. The core genome contains genes that are essential for lytic or lysogenic replication under all conditions and genetic backgrounds, such as genes encoding the phage virion components, lysis proteins, or repressor proteins ^{77,78}. While core genes are broadly conserved among groups of related phages, accessory genes will often be conserved in only subsets of phages, and may also be observed sporadically in diverse groups of phages ⁷⁹. Accessory genes may be essential under some conditions or provide a fitness advantage to the phage or its host (in the case of a prophage) under only certain conditions. In some cases, accessory genes have been referred to as “morons” and this term may be used to specifically refer to accessory genes of phages ^{5,80,81} as core and accessory genes are also a feature of bacterial genomes. The specific combination of accessory genes in a given phage genome likely reflects its adaptation to a specific host or niche, meaning that deletion of accessory genes often will not result in phenotypic change in standard laboratory growth conditions on a permissive host. Indeed, *acr* genes are conditionally essential; they can be deleted or disrupted

without phenotypic consequence, when infecting a bacterial host lacking a CRISPR-Cas system or CRISPR spacers targeting that specific phage.

The best-studied phage accessory genes increase the fitness of the bacterial host during lysogeny, participating in adaptive “lysogenic conversion”^{82,83}. Historically studied for their role in bacterial pathogenesis, diphtheria toxin, cholera toxin, and Shiga toxin are famous examples of prophage encoded accessory genes that dramatically alter the behavior of their hosts. Other conditionally essential phage accessory genes are involved in inter-phage warfare, preventing superinfection by a competitor phage. A recently published paper from the Hatfull group highlights the diverse roles that phage accessory genes play in during inter-phage antagonism⁸⁴. The authors first discovered a mycobacteriophage toxin/anti-toxin (TA) module that inhibited superinfection by competitor mycobacteriophages. The TA module, composed of toxic pp(p)Gpp synthetase and associated anti-toxin, inhibits replication of incoming superinfecting phages by promoting pp(p)Gpp synthesis, leading to cessation of cell growth and preventing lytic replication. The group then discovered accessory gene alleles encoded in a competitor phage, which had evolved to counter this prophage encoded defense system. This accessory gene specifically inhibited the anti-phage activity of the TA module, presumably by preventing TA dissociation. This impressive example of phage-host interactions stands out as an example of Red Queen selection dynamics, which predict counter-adaptation as a requirement for survival in “arms races” such as these⁸⁵. Previous examples of these dynamics have been demonstrated in both phages and eukaryotic viruses^{86,87}.

Anti-CRISPRs are another clear example of Red Queen dynamics at play in the phage accessory genome. These anti-immunity genes were first discovered in the genomes of a group of highly syntenic Mu-like phages with only a few pockets of genomic diversity – their accessory gene loci^{40,88}. Interestingly, accessory gene loci appear in conserved locations across the

genomes of these Mu-like phages, despite the sequences of the genes in the accessory locus being distinct. The anti-CRISPR accessory locus exemplifies the “grab-bag” nature of these loci – many diverse inhibitor proteins are encoded at the same location in the phage genome (**Fig 1.1**). It is interesting to broadly consider these loci as functional modules themselves – are genes in other syntenic accessory loci also inhibiting the same bacterial process through different mechanisms in these phages? In striking similarity to the *P. aeruginosa* Mu-like phages, the *L. monocytogenes* phages encoding *acrIIA* genes have highly syntenic genomes with conserved functional modules interspersed with accessory gene pockets. The diverse *acrIIA* genes are similarly anchored by a conserved gene (*acrIIA1*) encoding predicted HTH protein ⁵¹. In contrast to the *P. aeruginosa* Mu-like phage *acr* locus, not all of the accessory genes encoded in the *L. monocytogenes* “*acr* locus” have been demonstrated to have ACR function. It is worthy of follow-up to determine the functions of these genes. Similarly, many candidates from the 2016 survey of *aca1*-possessing loci across diverse prophages and mobile genetic elements in Proteobacteria discovered genes that did not exhibit anti-CRISPR activity when tested against the *P. aeruginosa* I-E and I-F CRISPR-Cas systems ⁴⁹. These proteins of unknown function are strong candidate inhibitors for other types of bacterial anti-phage immunity.

The striking diversity of *acr* genes across even closely related phages generates several questions- where were these diverse genes acquired from, and how did they evolve? Currently, no known “proto-anti-CRISPRs” have been discovered, and the evolutionary path of these novel proteins is mysterious. Analyses of the primary anti-CRISPR amino acid sequences have yielded have no recognizable domains or motifs, and likewise structural characterization of AcrF1 and AcrF3 has provided little insight into the origins of these inhibitors ⁵⁴⁻⁵⁶.

Why are anti-CRISPR Genes so diverse?

Many distinct *acr* genes have been identified thus far for I-E, I-F, II-A, and II-C systems. While the evolutionary history of anti-CRISPRs may be currently enigmatic, *acr* diversity presents an intriguing question: why there so many *acr* genes? We propose two non-mutually exclusive hypotheses to explain *acr* diversity:

I) Distinct *acrs* confer distinct, niche-specific fitness advantages to their host phage

II) *acr* diversity is a form of “distributed anti-immunity”

Hypothesis I: Like other phage accessory genes, the specific assemblage of *acr* genes on a given phage represents a snapshot of a unique set of fitness challenges experienced by that phage. As an obvious example, the combination of *acrE* and *acrF* genes found in *P. aeruginosa* phages reflects that they have likely cycled through hosts with both I-E and I-F CRISPR-Cas systems. It is less clear, however, what specific fitness advantages might be associated with using one particular AcrF protein over another. One immune inhibitor may impact phage fitness differently than another, and likewise the same Acr on a distinct type of genetic parasite (like a plasmid or mobile island) likely also has different fitness outcomes for the parasite. Non-lytic conjugative elements impose different selective pressures on their hosts and thus would experience a different set of fitness costs and benefits associated with Acr deployment.

Dependent on the host Cas protein sequences and expression levels, the potential for weakness in an Acr protein’s mode of action provokes the hypothesis that incomplete inhibition of CRISPR-Cas immunity may result. Although counterintuitive, this could benefit a phage by maintaining a population of infection-susceptible hosts and reducing selection for the evolution of alternative forms of anti-phage immunity such as phage receptor loss. Indeed, it has been shown that under the presence of high phage burden, surface modifications are favored over

CRISPR-based immunity⁸⁹. Furthermore, we also imagine the potential for Acr proteins to synergize with each other when two *acr*-encoding phages infect the same bacterial cell. By targeting different steps in CRISPR interference, infections with heterotypic *acr* genes could lead to more viral replication than homotypic infections, selecting for the maintenance of diverse *acr* genes in a viral population and facilitating accessory gene exchanges between closely related phages.

Hypothesis II: Diverse *acr* genes limit evolution of anti-anti-CRISPR (anti-Acr) mechanisms. An important facet of CRISPR immune function is the paradigm of distributed immunity, which is selection for the co-existence of many, equally fit, immune alleles in a population. This theory of CRISPR immunity was proposed first by the Whitaker group and tested using modeling approaches and experimentally evolved microbial populations of *Streptococcus thermophilus*⁹⁰. The distributed CRISPR immunity hypothesis is that viral predators select for the maintenance of a diverse spacer repertoire distributed across a microbial population. It is simple for a virus to escape targeting of one CRISPR spacer: a single point mutation can fully disable CRISPR immunity³⁷. However, distribution of many targeting spacers across a microbial population prevents individual viral escaper genotypes from emerging. Likewise, no single spacer will dominate the CRISPR landscape because immunity functions on the level of microbial populations not individual microbial genotypes.

To test the importance of distributed immunity, *P. aeruginosa* Mu-like phage DMS3 and artificially assembled populations of *P. aeruginosa* with varying degrees of spacer diversity distributed across the bacterial population⁹¹. They found that low diversity populations of *P. aeruginosa* with 1, 6, or 12 spacer genotypes routinely selected for the emergence of escaper phages that had presumably accumulated point mutations across protospacer regions. In contrast, high diversity populations with 24 or 48 spacer genotypes drove the DMS3 phage to

extinction. In the case of high-diversity populations, only Acr deployment could protect the phage from CRISPR immunity. We invoke a hypothesis of distributed anti-immunity to describe anti-CRISPR diversity. By maintaining a diverse repertoire of *acr* genes, viral populations limit the emergence bacterial anti-Acr mechanisms, such as point mutations in the target Cas protein. We next consider potential mechanisms for the emergence of anti-Acr strategies.

Putative anti-anti-CRISPR mechanisms

Acquisition of new CRISPR systems: CRISPR-Cas systems are diverse immune systems. Currently, there are six broad types of CRISPR-Cas systems which can be further subdivided into many subtypes^{20,92}. One of the simplest mechanisms by which bacteria could evolve to overcome phages with subtype specific *acr* genes is to accumulate multiple types of CRISPR-Cas systems. In order to survive, a phage would need to inhibit all systems. There are many examples of bacteria that have accumulated multiple types of CRISPR-Cas systems. *Streptococcus thermophilus*, the first organism in which CRISPR-Cas activity was demonstrated, has 3-4 different CRISPR-Cas systems: two Type II-A systems, a Type III-A system, and sometimes a Type I-E system^{93,94}. It is unknown if *acr* genes have selected for this CRISPR diversity in *S. thermophilus*, however it is notable that thus far *acr* genes inhibiting both II-A and I-E CRISPR systems have been characterized. In contrast, no Type III anti-CRISPR has been discovered. Similar to *S. thermophilus*, *Serratia* sp. ATCC39006 carries an active I-E, I-F, and III-A CRISPR systems, and these diverse CRISPR systems were recently discovered to be regulated coordinately by quorum sensing⁹⁵. Again, it is unknown if Acr proteins have driven selection for *Serratia* to carry multiple CRISPR-Cas systems, however at least one *acrF* gene (*acrF8*) is found in *Serratia marcescens* genomes⁴⁹. Finally, in *P. aeruginosa*, acquisition of multiple CRISPR subtypes may also be driven by CRISPR/Acr warfare. *P. aeruginosa* has both Type I-F and Type I-E CRISPR-Cas systems, which often co-occur in the same genome. Less

frequently, *P. aeruginosa* genomes contain Type I-C CRISPR-Cas systems that are mobilized on an integrative and conjugative element (ICE). I-C and I-F also co-occur in *P. aeruginosa* genomes, but there are no examples yet of genomes carrying all three⁴³. Currently, no anti-I-C anti-CRISPRs have been described.

Mutational escape: The anti-CRISPRs that have been biochemically characterized bind specific surfaces on Cas proteins⁵³⁻⁵⁵. By mutating these surfaces, bacteria could hypothetically evolve Acr resistant CRISPR systems. By employing diverse inhibitors that bind to unique surfaces on CRISPR-Cas proteins, a population of viruses will limit accumulation of such CRISPR-Cas escape mutations that could allow anti-CRISPR escape. Interestingly, in the co-crystal of the AcrF3 dimer bound to its Cas3 target, the AcrF3 dimer makes many contacts across the face of the Cas3 protein, suggesting that many Cas3 mutations would be required to disrupt the AcrF3/Cas3 interaction⁵⁴. More information about the residue-specific interactions between Acrs and Cas proteins will be critical to identify Acr resistant CRISPR-Cas systems.

Regulatory changes: The biochemically characterized anti-CRISPRs bind stoichiometrically to their Cas protein targets. Interestingly, overexpression of the type I-F CRISPR-Cas complex subunits in *P. aeruginosa* functions as an anti-ACR mechanism against the phages that use Acr proteins that bind this complex. In contrast, AcrF3, which targets the recruited Cas3 effector nuclease is not affected by increasing the intracellular concentration of proteins that it does not bind⁵³. This shows that Acr proteins can be overwhelmed by shifting intracellular Cas protein concentrations, and suggests the possibility for bacteria to overcome Acrs by overexpressing components of CRISPR-Cas systems. Multiple papers have reported different pathways involved in regulation of CRISPR-Cas systems in diverse bacteria⁹⁵⁻⁹⁸. In each case, CRISPR-Cas is dynamically regulated, suggesting a cost to constitutive CRISPR expression. We hypothesize that Acr proteins that target CRISPR-Cas subunits more toxic to overexpress would

have a selective advantage in this scenario. For instance, both Cas3 and Cas9 nucleases have the potential for genomic DNA cleavage, so Acr proteins that target these Cas proteins may be less susceptible to inhibition by CRISPR regulatory changes. Furthermore, some CRISPR systems are strongly induced during phage infection^{99,100}. At first glance, this can be interpreted as enhanced immune activity, but this could also present a mechanism to overwhelm inhibitor proteins deployed by the phage.

Dedicated anti-CRISPR inhibitors: Bacteria may possess dedicated inhibitors of Acr function, which prevent target binding or cause Acr protein degradation. Alternatively, blocking *acr* expression may also be possible, despite *acrs* themselves being diverse in sequence and mechanism, a commonality amongst them is a shared regulatory environment. By targeting conserved, cis-acting DNA elements such as promoters, operators, and terminators required for *acr* expression, the bacterial cell could shift the balance in favor of CRISPR. For example, I-E, I-F, and II-C anti-CRISPRs have conserved associated proteins *aca1*, *aca2*, and *aca3* (of unknown function), whereas Type II-A anti-CRISPR loci often carry *acrIIA1*. Though the functional relevance of these associated proteins is currently unknown, they are strikingly conserved relative to their associated *acr* genes and could potentially represent the Achilles heel of an otherwise rapidly evolving system. A summary of these putative mechanisms for anti-CRISPR evasion is provided in **Fig.1.3**.

CRISPR meets anti-CRISPR in lysogeny

Anti-CRISPRs are widespread across bacterial genomes. A recent report estimates that 64% of 449 *P. aeruginosa* I-F systems are inhibited by chromosomally encoded *acrF* genes⁴⁹. The same study concluded that the full diversity of I-F systems across the phylum Proteobacteria is potentially able to be inhibited by known anti-CRISPRs. A separate analysis of *P. aeruginosa* I-

E systems estimates that 53% of 81 I-E systems are inhibited by *acrE* genes⁴³. Similarly, >50% of II-A systems in *L. monocytogenes* are estimated to be inhibited by the recently discovered *acrIIA* genes. Though CRISPR-Cas systems are commonly inhibited by Acr proteins, the consequences of CRISPR/Acr co-occurrence are relatively unexplored.

Self-targeting: The most striking sign of inhibition of CRISPR-Cas activity is the stable co-existence of a CRISPR spacer and its target in the same cell. *AcrIIA* proteins were discovered using this strong genomic signature of “self-targeting” and this will likely be a useful strategy for the discovery of novel Acr proteins⁵¹. In this scenario, an *acr* is now an essential gene of the lysogen, as anti-CRISPR loss will trigger direct CRISPR autoimmunity. Type I and II CRISPR-Cas cannot distinguish between chromosomal “self” from incoming phage. In contrast, type III CRISPR-Cas systems, in which CRISPR activity is dependent on target transcription, have been demonstrated to conditionally tolerate their prophages with perfect protospacer matches¹⁰¹. Self-targeting is likely prevalent in bacterial genomes, and the phenotypic consequences of this potential autoimmune scenario are unknown.

CRISPR-Cas alternative functions: There is increasing evidence pointing toward CRISPR-Cas components (protein or RNA) performing alternative non-immunity-related functions¹⁰². In the pathogen *Francisella novicida*, non-canonical activity of CRISPR-Cas effector protein Cas9 in association with a small CRISPR-Cas associated RNA (*scaRNA*) and the *tracrRNA* directly regulate levels of a virulence-associated transcript through base pairing with the RNA target¹⁰³. It is currently unknown if *F. novicida* genetic parasites employ *acr* genes, but our current knowledge of Type II inhibitors suggests the potential for undiscovered *acrIIB* genes to impact virulence regulation in *F. novicida*. Furthermore, a recent publication has shown RNA targeting in the *P. aeruginosa* Type I-F CRISPR-Cas system, where a mismatched crRNA guides degradation of the *lasR* transcript, a master transcriptional regulator in *P. aeruginosa*¹⁰⁴. This

non-canonical, RNA-directed CRISPR-Cas activity is dependent on both the I-F CRISPR-Cas complex and recruited effector nuclease Cas3. Do endogenous Acr proteins disrupt non-canonical CRISPR activity as well as canonical immunity functions? Given the widespread distribution of *acrF* genes in *P. aeruginosa* and beyond ⁴⁹, *acr* genes have the potential to profoundly impact the biology of their bacterial hosts. It will be very interesting to see if *acr* genes can inhibit, alter, or enable alternative functions, and what the evolutionary consequences of these interactions may be.

Horizontal gene transfer

CRISPR-Cas immune systems, which destroy foreign DNA, can act as barriers to horizontal gene transfer (HGT). While inhibition of viral parasites is an obvious adaptive function of a CRISPR-Cas system ^{27,94}, the exclusion of potentially beneficial foreign DNA, such as a prophage ⁸³, can render a CRISPR-Cas system disadvantageous, and selection for CRISPR-Cas loss or inhibition can occur. By inhibiting CRISPR-Cas activity, chromosomally encoded *acrs* should enable foreign DNA acquisition in their hosts. HGT is pervasive in bacteria and has had a profound impact on shaping bacterial genomes, suggesting a strong potential cost to CRISPR-Cas activity and large potential benefit to anti-CRISPR acquisition.

While individual examples of CRISPR excluding HGT mediated by plasmids, prophages, and through natural transformation have been shown ¹⁰⁵⁻¹⁰⁸, it has been difficult to extrapolate these individual examples to broad principles of bacterial genome evolution. In 2015, the Koonin group performed a bioinformatics study, analyzing CRISPR-Cas activity (using CRISPR array length as a proxy for activity) and HGT across 1399 microbial genomes ¹⁰⁹. The authors found no evidence that CRISPR-Cas activity inhibited HGT on evolutionary time scales. Instead, they found that the best predictor of HGT was growth temperature, with lower genetic diversity at hotter temperatures. This counterintuitive finding suggests that propensity for HGT is an intrinsic

property of an organism and its ecological niche, and that CRISPR-Cas may exert its fitness impacts on the short-term population level rather than on long-term evolutionary timescales.

Emphasizing the population level importance of CRISPR-activity on HGT, a 2015 comparison of CRISPR-Cas distribution and HGT across a population of *P. aeruginosa* isolates demonstrated that CRISPR-Cas activity significantly restricted genome size⁴³. Importantly, each *P. aeruginosa* strain was only considered to be immune-competent if it had a CRISPR array, Cas genes, and lacked chromosomally encoded *acr* genes. The group showed that *P. aeruginosa* strains with active I-E and I-F CRISPR-Cas systems had genomes that were on average 300 kbp smaller than *P. aeruginosa* strains with no CRISPR-Cas systems. Fascinatingly, the authors also showed that *P. aeruginosa* strains with *acrE* or *acrF* genes had genome sizes that, on average, were not different in size compared to strains with no CRISPR-Cas system. Despite CRISPR inhibition likely being a relatively recent event in the evolutionary history of these bacterial strains, their HGT profile was similar to that of a strain that had presumably been without CRISPR for much longer. This result demonstrates the short-term, population-level impacts of CRISPR-Cas activity on bacterial genomes and emphasizes the rapid impacts that *acrs* can have on the biology of their host bacteria.

The rapid acquisition of additional mobile genetic elements (MGEs) after CRISPR-Cas inhibition facilitate interactions amongst multiple MGEs. With multiple MGEs being more likely to be stable in the same cell, this could increase the horizontal transfer of new genes, including *acrs* themselves. It is interesting to consider the strong impacts that *acr* genes could have on shaping the accessory genomes of their host phage by “opening the door” to downstream infection. CRISPR inhibition of bacterial immunity could also have negative fitness impacts for the prophage, as the immune-compromised bacterial host could be infected and killed by a superinfecting competitor phage.

Not all phages have *acr* genes, suggesting that there may also be fitness costs to Acr action, such as licensing superinfecting phages. Interestingly, many of the Mu-like Acr phages in *P. aeruginosa* utilize diverse mechanisms to inhibit superinfection of other phages^{40,110}, which likely ameliorates some costs of host-immune compromise. As phage accessory genomes become better defined, it will be interesting to correlate the presence of superinfection exclusion genes with the presence of *acr* genes. Such complex genetic interactions in the phage accessory genome have likely profoundly shaped phage evolution, and may in part control anti-CRISPR distribution across phage populations.

Conclusion

CRISPR-Cas immune systems are a relatively recent discovery in the arms race between phages and their hosts, but are likely ancient players in this battle. This new field has had a massive impact on our understanding of microbial evolution, phage biology, and horizontal gene transfer. Also remarkable is the elegance of many distinct, adaptive, sequence-specific RNA-guided nuclease systems possessed by bacteria, with some of them currently revolutionizing human gene editing and therapy. Anti-CRISPRs are an even more recent addition to the CRISPR story and are fascinating for many reasons, providing new insights into how CRISPR-Cas systems work, and how CRISPR systems and bacterial genomes have co-evolved with the moving target of mobile DNA. While it is still early, we have already seen examples of both CRISPR and anti-CRISPRs shaping bacterial population by dictating the horizontal DNA that is acquired versus what is excluded. Furthermore, as CRISPR has revolutionized gene editing, anti-CRISPRs have provided new biotechnological resources in our efforts to precisely edit the human genome and develop new tools to probe it. Future work should focus on the discovery of new anti-CRISPRs that inhibit distinct CRISPR-Cas systems, deciphering their mechanisms of

action and studying the counter response from CRISPR-Cas systems to combat anti-CRISPR emergence.

References

1. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci USA* **96**, 2192–2197 (1999).
2. Cobián Güemes, A. G. *et al.* Viruses as Winners in the Game of Life. *Annual Review of Virology* **3**, 197–214 (2016).
3. Seed, K. D. Battling Phages: How Bacteria Defend against Viral Attack. *PLoS Pathog* **11**, e1004847 (2015).
4. Bebeacua, C. *et al.* X-ray structure of a superinfection exclusion lipoprotein from phage TP-J34 and identification of the tape measure protein as its target. *Molecular Microbiology* **89**, 152–165 (2013).
5. Cumby, N., Edwards, A. M., Davidson, A. R. & Maxwell, K. L. The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. **194**, 5012–5019 (2012).
6. Cumby, N., Reimer, K., Mengin-Lecreulx, D., Davidson, A. R. & Maxwell, K. L. The Phage Tail Tape Measure Protein, an Inner Membrane Protein, and a Periplasmic Chaperone Play Connected Roles in the Genome Injection Process of E. coli Phage HK97. *Molecular Microbiology* (2014). doi:10.1111/mmi.12918
7. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat Rev Micro* **8**, 317–327 (2010).
8. Westra, E. R. *et al.* The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annu. Rev. Genet.* **46**, 311–339 (2012).
9. Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* **34**, 169–183 (2015).

10. Seed, K. D., Lazinski, D. W., Calderwood, S. B. & Camilli, A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. **494**, 489–491 (2013).
11. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat Rev Micro* **11**, 675–687 (2013).
12. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Research* **40**, 5569–5576 (2012).
13. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. **520**, 505–510 (2015).
14. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Communications* **3**, 945 (2012).
15. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading, Engl)* **151**, 653–663 (2005).
16. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, Engl)* **151**, 2551–2561 (2005).
17. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).
18. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science* **329**, 1355–1358 (2010).
19. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).

20. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Micro* **13**, 722–736 (2015).
21. Westra, E. R. *et al.* CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. *Mol Cell* **46**, 595–605 (2012).
22. Mulepati, S. & Bailey, S. In vitro Reconstitution of an Escherichia coli RNA-guided Immune System Reveals Unidirectional, ATP-dependent degradation of DNA target. *J Biol Chem* (2013). doi:10.1074/jbc.M113.472233
23. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* **527**, 110–113 (2015).
24. Jiang, W., Samai, P. & Marraffini, L. A. Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* **164**, 710–721 (2016).
25. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
26. Mohanraju, P. *et al.* Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
27. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
28. Westra, E. R. *et al.* Cascade-mediated binding and bending of negatively supercoiled DNA. *RNA Biol* **9**, 1134–1138 (2012).
29. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* **109**, E2579–86 (2012).
30. Jinek, M. *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).
31. Heler, R. *et al.* Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* **519**, 199–202 (2015).

32. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
33. Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nature Biotechnology* **34**, 933–941 (2016).
34. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).
35. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. **190**, 1401–1412 (2008).
36. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390–1400 (2008).
37. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences* **108**, 10098–10103 (2011).
38. Winstanley, C. *et al.* Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa*. *Genome Research* **19**, 12–23 (2009).
39. Nakayama, K., Kanaya, S., Ohnishi, M., Terawaki, Y. & Hayashi, T. The complete nucleotide sequence of phi CTX, a cytotoxin-converting phage of *Pseudomonas aeruginosa*: implications for phage evolution and horizontal gene transfer via bacteriophages. *Molecular Microbiology* **31**, 399–419 (1999).
40. Bondy-Denomy, J. *et al.* Prophages mediate defense against phage infection through diverse mechanisms. *The ISME Journal* **10**, 2854–2866 (2016).
41. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).

42. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas Adaptive Immune System of *Pseudomonas aeruginosa* Mediates Resistance to Naturally Occurring and Engineered Phages. 1–40 (2012).
43. van Belkum, A. *et al.* Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant *Pseudomonas aeruginosa*. *mBio* **6**, e01796–15 (2015).
44. Zegans, M. E. *et al.* Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* **191**, 210–219 (2009).
45. 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).
46. Cady, K. C. *et al.* Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* **157**, 430–437 (2011).
47. 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**, e00129–15 (2015).
48. Pawluk, A., Bondy-Denomy, J., Cheung, V. H. W., Maxwell, K. L. & Davidson, A. R. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* **5**, e00896–e00896–14 (2014).
49. Pawluk, A. *et al.* Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology* **1**, 1–6 (2016).
50. Pawluk, A. *et al.* Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* **167**, 1829–1838.e9 (2016).
51. Rauch, B. J. *et al.* Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150–158.e10 (2017).

52. Bondy-Denomy, J. & Davidson, A. R. To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends Microbiol* **22**, 218–225 (2014).
53. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
54. Wang, X. *et al.* Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Struct. Mol. Biol.* **23**, 868–870 (2016).
55. Wang, J. *et al.* A CRISPR evolutionary arms race: structural insights into viral anti-CRISPR/Cas responses. *Cell Res.* **26**, 1165–1168 (2016).
56. Maxwell, K. L. *et al.* The solution structure of an anti-CRISPR protein. *Nature Communications* **7**, 13134 (2016).
57. Stern, A. & Sorek, R. The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43–51 (2011).
58. Arber, W. & Dussoix, D. Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda. *Journal of Molecular Biology* **5**, 18–36 (1962).
59. Krüger, D. H. & Bickle, T. A. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol. Rev.* **47**, 345–360 (1983).
60. Dryden, D. T. F. & Tock, M. R. DNA mimicry by proteins. *Biochem Soc Trans* **34**, 317–319 (2006).
61. Walkinshaw, M. D. *et al.* Structure of Ocr from bacteriophage T7, a protein that mimics B-form DNA. *Mol Cell* **9**, 187–194 (2002).
62. McMahon, S. A. *et al.* Extensive DNA mimicry by the ArdA anti-restriction protein and its role in the spread of antibiotic resistance. *Nucleic Acids Research* **37**, 4887–4897 (2009).

63. Wang, H.-C. *et al.* White spot syndrome virus protein ICP11: A histone-binding DNA mimic that disrupts nucleosome assembly. *Proceedings of the National Academy of Sciences* **105**, 20758–20763 (2008).
64. Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
65. Soutourina, O. A. *et al.* Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. **9**, e1003493 (2013).
66. Wiedenheft, B. *et al.* Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* **17**, 904–912 (2009).
67. Luo, M. L., Mullis, A. S., Leenay, R. T. & Beisel, C. L. Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Research* (2014). doi:10.1093/nar/gku971
68. Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P. R. & Lundgren, M. Efficient programmable gene silencing by Cascade. *Nucleic Acids Research* **43**, 237–246 (2015).
69. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
70. Gilbert, L. A. *et al.* CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154**, 442–451 (2013).
71. Vorontsova, D. *et al.* Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. *Nucleic Acids Research* **43**, 10848–10860 (2015).
72. Bryson, A. L. *et al.* Covalent Modification of Bacteriophage T4 DNA Inhibits CRISPR-Cas9. *mBio* **6**, e00648–15 (2015).

73. Strotskaya, A. *et al.* The action of Escherichia coli CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Research* gkx042 (2017). doi:10.1093/nar/gkx042
74. Yaung, S. J., Esvelt, K. M. & Church, G. M. CRISPR/Cas9-mediated phage resistance is not impeded by the DNA modifications of phage T4. *PLoS ONE* **9**, e98811 (2014).
75. Hatfull, G. F. & Hendrix, R. W. Bacteriophages and their genomes. *Curr Opin Virol* **1**, 298–303 (2011).
76. Botstein, D. A theory of modular evolution for bacteriophages. *Ann. N. Y. Acad. Sci.* **354**, 484–490 (1980).
77. Comeau, A. M., Bertrand, C., Letarov, A., Tétart, F. & Krisch, H. M. Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. *Virology* **362**, 384–396 (2007).
78. Hatfull, G. F. Bacteriophage genomics. *Curr Opin Microbiol* **11**, 447–453 (2008).
79. Hatfull, G. F. Dark Matter of the Biosphere: the Amazing World of Bacteriophage Diversity. *Journal of Virology* **89**, 8107–8110 (2015).
80. Juhala, R. J. *et al.* Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *Journal of Molecular Biology* **299**, 27–51 (2000).
81. Cumby, N., Davidson, A. R. & Maxwell, K. L. The moron comes of age. *Bacteriophage* **2**, 225–228 (2012).
82. Bondy-Denomy, J. & Davidson, A. R. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *J Microbiol* **52**, 235–242 (2014).
83. Brüßow, H., Canchaya, C. & Hardt, W.-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**, 560–602– table of contents (2004).

84. Dedrick, R. M. *et al.* Prophage-mediated defence against viral attack and viral counter-defence. *Nature Microbiology* **2**, 16251 (2017).
85. Dawkins, R. & Krebs, J. R. Arms races between and within species. *Proc. R. Soc. Lond., B, Biol. Sci.* **205**, 489–511 (1979).
86. Betts, A., Kaltz, O. & Hochberg, M. E. Contrasted coevolutionary dynamics between a bacterial pathogen and its bacteriophages. *Proceedings of the National Academy of Sciences* **111**, 11109–11114 (2014).
87. Elde, N. C. *et al.* Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* **150**, 831–841 (2012).
88. Cazares, A., Mendoza-Hernández, G. & Guarneros, G. Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics* **15**, 1146 (2014).
89. Westra, E. R. *et al.* Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. *Curr. Biol.* **25**, 1043–1049 (2015).
90. Childs, L. M., England, W. E., Young, M. J., Weitz, J. S. & Whitaker, R. J. CRISPR-induced distributed immunity in microbial populations. *PLoS ONE* **9**, e101710 (2014).
91. van Houte, S. *et al.* The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* **532**, 385–388 (2016).
92. Shmakov, S. *et al.* Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Micro* **15**, 169–182 (2017).
93. Carte, J. *et al.* The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Molecular Microbiology* **93**, 98–112 (2014).
94. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

95. Patterson, A. G. *et al.* Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. *Mol Cell* **64**, 1102–1108 (2016).
96. Høyland-Kroghsbo, N. M. *et al.* Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proceedings of the National Academy of Sciences* **114**, 201617415–135 (2016).
97. Medina-Aparicio, L. *et al.* The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. **193**, 2396–2407 (2011).
98. Patterson, A. G., Chang, J. T., Taylor, C. & Fineran, P. C. Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Research* **43**, gkv517–6048 (2015).
99. Agari, Y. *et al.* Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *Journal of Molecular Biology* **395**, 270–281 (2010).
100. Young, J. C. *et al.* Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS ONE* **7**, e38077 (2012).
101. Goldberg, G. W., Jiang, W., Bikard, D. & Marraffini, L. A. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* **514**, 633–637 (2014).
102. Westra, E. R., Buckling, A. & Fineran, P. C. CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Micro* **12**, 317–326 (2014).
103. Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y.-L. & Weiss, D. S. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. **497**, 254–257 (2013).
104. Li, R. *et al.* Type I CRISPR-Cas targets endogenous genes and regulates virulence to evade mammalian host immunity. *Cell Res.* **26**, 1273–1287 (2016).

105. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).
106. Bikard, D., Hatoum-Aslan, A., Mucida, D. & Marraffini, L. A. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12**, 177–186 (2012).
107. Palmer, K. L. & Gilmore, M. S. Multidrug-resistant enterococci lack CRISPR-cas. *mBio* **1**, (2010).
108. Edgar, R. & Qimron, U. The Escherichia coli CRISPR system protects from λ lysogenization, lysogens, and prophage induction. *J. Bacteriol.* **192**, 6291–6294 (2010).
109. Gophna, U. *et al.* No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *The ISME Journal* **9**, 2021–2027 (2015).
110. Chung, I. Y., Jang, H. J., Bae, H. W. & Cho, Y.-H. A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. *Proceedings of the National Academy of Sciences* (2014). doi:10.1073/pnas.1403537111

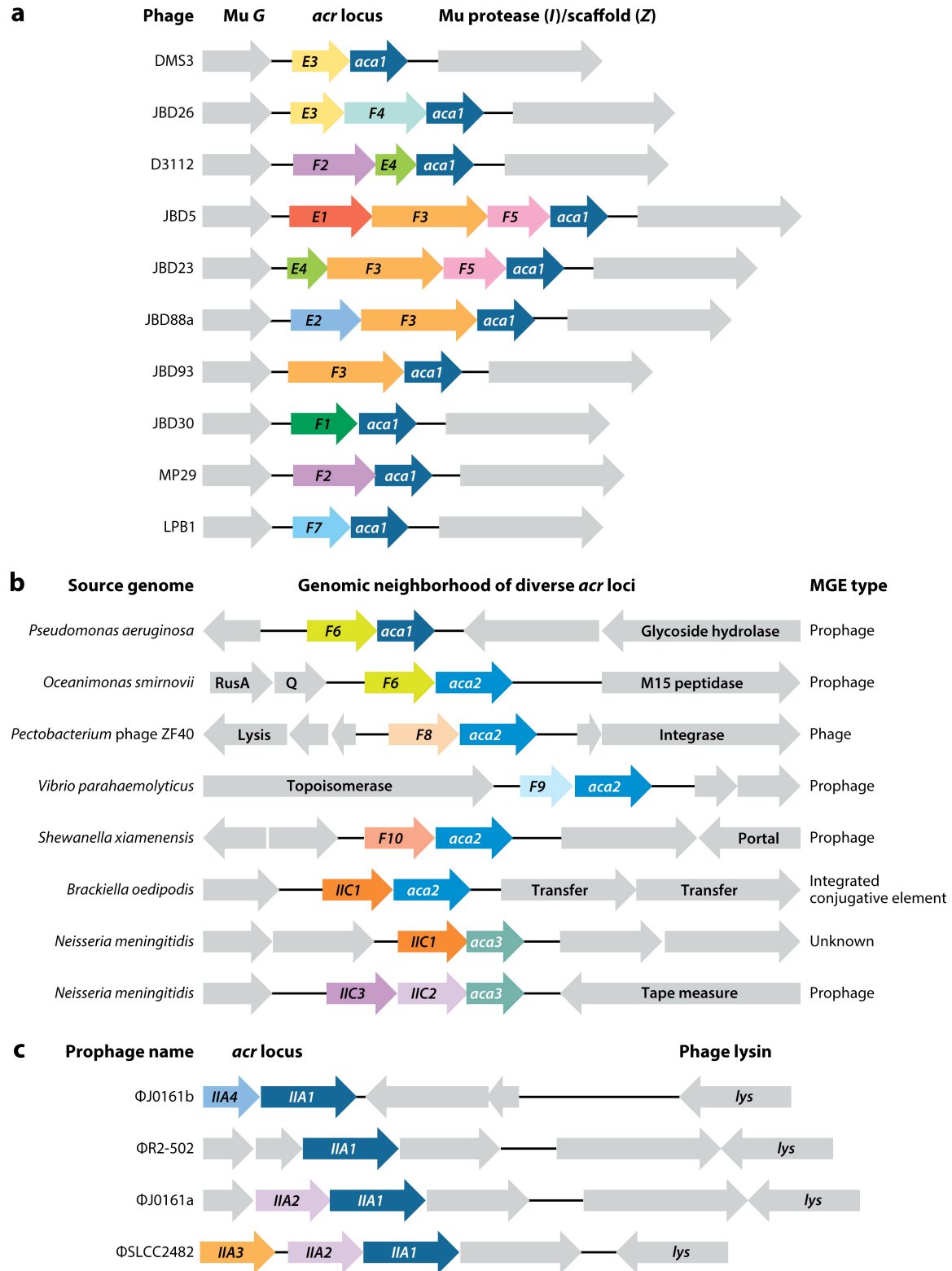


Figure 1.1 Anti-CRISPR (*acr*) locus organization.

Stereotypical organizations of *acr* loci encoded by phages and mobile genetic elements (MGEs) are shown. Unique *acr* genes are named and shown in color, whereas non-*acr* genes are

shown in gray and are annotated with predicted functions when possible. (a) *Pseudomonas aeruginosa* Mu-like phage *acr* locus. The *acr* genes are all integrated at the same locus between two highly conserved structural genes (gray) that are homologous to Mu phage gene G and Mu phage protease (I)/scaffold (Z) genes. Many loci encode both type I-E (AcrE1–4) and I-F (AcrF1–5) Acr proteins, all adjacent to the conserved anti-CRISPR-associated gene 1 (*aca1*). A representative phage is indicated for each unique locus architecture. Panel adapted from Reference 48. (b) *acr* loci in diverse Proteobacteria are shown. These *acr* loci do not share a common “genomic neighborhood,” but all are anchored by HTH-encoding anti-CRISPR-associated genes (*aca1–3*). Representatives of each *acr-aca* association are shown in the indicated species. Panel adapted from Reference 49. (c) Listeriophage *acrIIA* locus. The listeriophage locus is near the left end of the integrated prophage genome and a highly conserved endolysin gene (*lys*). All listeriophage *acr* loci are anchored by the HTH-encoding gene *acrIIA1*. Panel adapted from Reference 51.

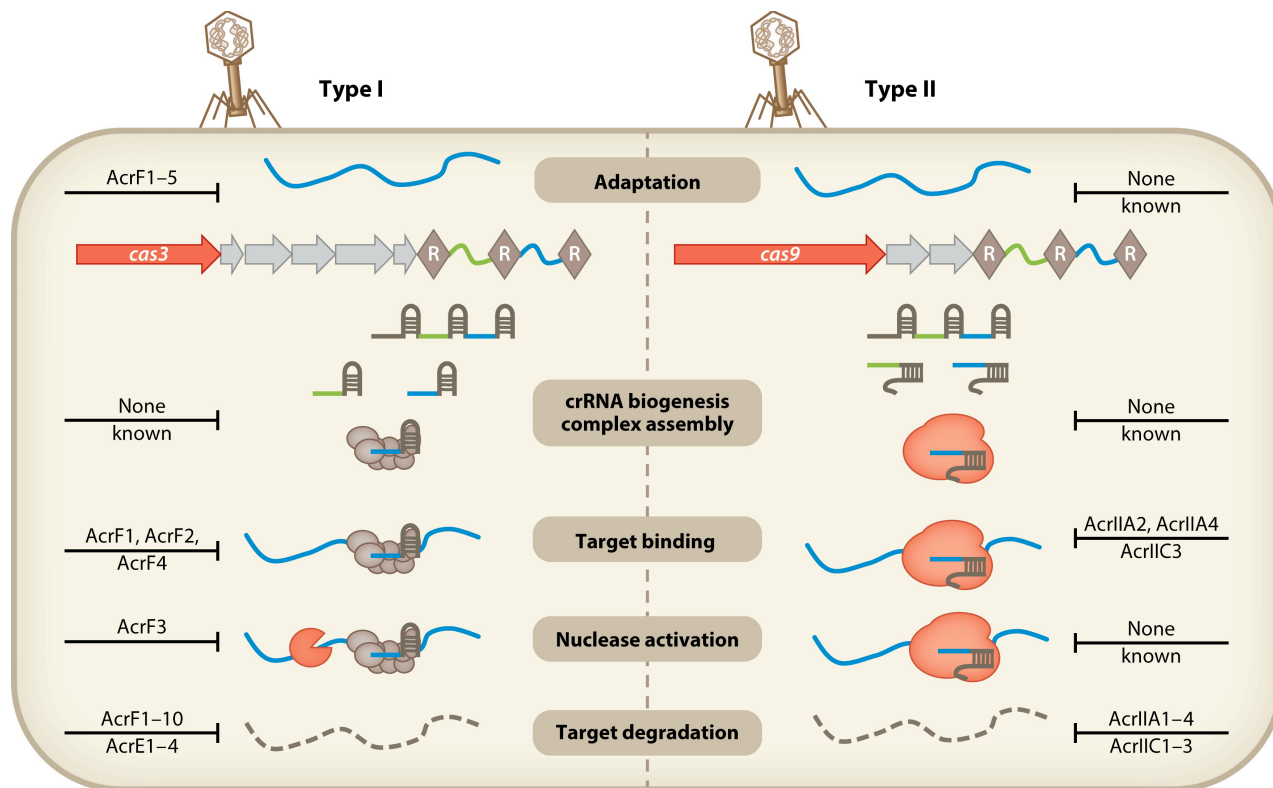


Figure 1.2 Mechanisms for anti-CRISPR protein function

CRISPR-Cas immune function is broken down into five distinct processes, shown in brown boxes. Acr proteins that inhibit these processes are shown for both type I and type II CRISPR-Cas systems. All characterized type I-F Acr proteins (AcrF1–5) have been demonstrated to inhibit both adaptation and immunity by preventing either foreign DNA recognition (AcrF1, AcrF2, and AcrF4) or Cas3 nuclease recruitment (AcrF3). AcrIIA2, AcrIIA4, and AcrIIC3 prevent DNA target binding by Cas9. All anti-CRISPRs are defined by their ability to ultimately prevent foreign DNA destruction, though the mechanisms by which most of them accomplish this task are still unknown. Abbreviations: crRNA, CRISPR RNA; R, repeat.

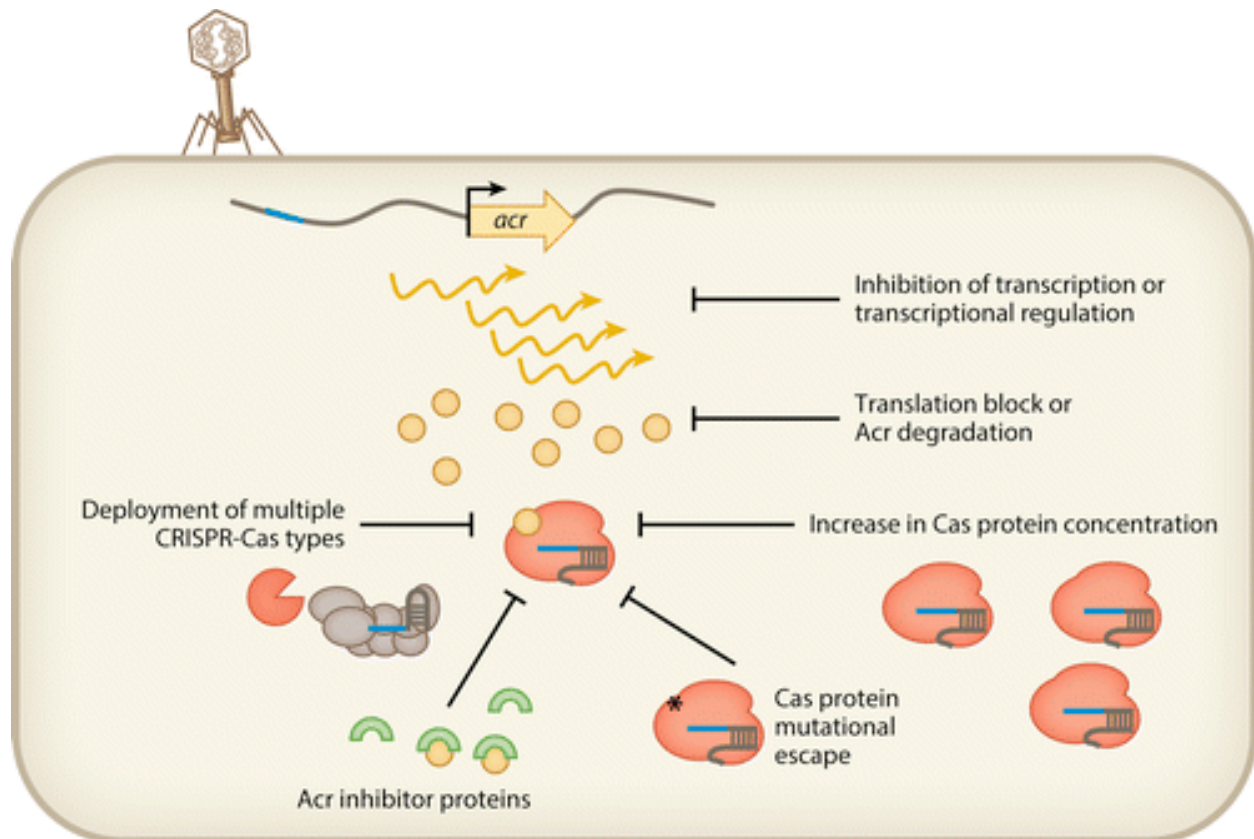


Figure 1.3 Anti-anti-CRISPR mechanisms

To inhibit CRISPR-Cas immunity, *acr* genes need to be transcribed and translated inside a host cell. Currently, there are no described mechanisms by which bacterial hosts perturb anti-CRISPR transcript or protein levels, but AcrF proteins can lose efficacy when the intracellular concentration of Cas protein targets is increased. Cas mutations that lower or abolish Acr binding affinity for the Cas target could also serve to shift the balance in favor of the CRISPR-Cas system, as could protein inhibitors that sequester Acr proteins and prevent them from binding their Cas targets. Lastly, deployment of multiple types of CRISPR-Cas systems is a mechanism by which cells can protect themselves from subtype-specific Acr proteins and may in part explain the accumulation of multiple CRISPR-Cas system types and subtypes in diverse bacteria.

Chapter 2: Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity

ABSTRACT

Bacteria utilize CRISPR-Cas adaptive immune systems for protection from bacteriophages (phages), and some phages produce anti-CRISPR (Acr) proteins that inhibit immune function. Despite thorough mechanistic and structural information for some Acr proteins, how they are deployed and utilized by a phage during infection is unknown. Here, we show that Acr production does not guarantee phage replication when faced with CRISPR-Cas immunity, but instead, infections fail when phage population numbers fall below a critical threshold. Infections succeed only if a sufficient Acr dose is contributed to a single cell by multiple phage genomes. The production of Acr proteins by phage genomes that fail to replicate leave the cell immunosuppressed, which predisposes the cell for successful infection by other phages in the population. This altruistic mechanism for CRISPR-Cas inhibition demonstrates inter-virus cooperation that may also manifest in other host-parasite interactions.

INTRODUCTION

Bacteria and the viruses that infect them (phages) are engaged in an ancient evolutionary arms race, which has resulted in the emergence of a diversity of CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) adaptive immune systems ¹. CRISPR-Cas immunity is powered by the acquisition of small fragments of phage genomes into the bacterial CRISPR array, the subsequent transcription and processing of these arrays to generate small CRISPR RNAs, and the RNA-guided destruction of the phage genome ²⁻⁵. The destruction of foreign DNA by CRISPR-Cas has been shown to prevent the acquisition of plasmids, DNA from the environment, phage lytic replication, and prophage integration ^{2,3,6-9}. In bacterial populations, these systems provide a fitness advantage to their host microbe when phage are present in the environment ^{10,11}.

To combat the potent action of RNA-guided CRISPR-Cas nucleases, phages have developed inhibitor proteins called anti-CRISPRs (Acrs). Acr proteins have been discovered in phages, prophages, mobile islands, and core genomes across many distinct bacteria and archaea¹²⁻¹⁴. Specific Acr proteins that inhibit Type I-F, I-E, and I-D CRISPR-Cas3 systems have been identified¹⁴⁻¹⁷, as well as proteins that inhibit Type II-A and II-C CRISPR-Cas9 systems¹⁸⁻²⁰. Phylogenetic studies indicate that these proteins are likely ubiquitous in coevolving populations of bacteria and phages¹³ and provide a significant replicative advantage to phages in the presence of CRISPR-Cas immunity¹¹.

Anti-CRISPRs were first identified in phages that neutralize the *Pseudomonas aeruginosa* Type I-F system (anti-CRISPR type I-E, AcrIF1-5)¹⁵, and five more I-F anti-CRISPRs (AcrIF6-10) were subsequently identified in various mobile genetic elements¹⁶. The I-F Csy surveillance complex (also called I-F Cascade) is comprised of an unequal stoichiometry of four proteins (Csy1-4) that assemble with a 60 nt CRISPR RNA (crRNA) guide²¹⁻²⁷. The Csy complex locates and binds foreign dsDNA targets complementary to the crRNA, then recruits a *trans*-acting nuclease/helicase protein called Cas2/3 to degrade the target²⁸⁻³⁰. Anti-CRISPR proteins function by interacting directly with the Csy complex and inhibiting DNA binding, or bind to Cas2/3 and prevent nuclease-mediated degradation³¹. The structures of Type I-F Acr proteins AcrIF1, AcrIF2, AcrIF3, and AcrIF10 have been solved in complex with their target proteins, revealing mechanistically distinct inhibitors that bind tightly to their targets^{25-27,29,30,32}. Together with the recent identification and characterization of proteins that inhibit Cas9, all characterized Acr proteins block phage DNA binding or cleavage^{18,20,33-36}.

All AcrIF proteins are robust inhibitors of CRISPR-Cas activity when expressed from high copy plasmids prior to phage challenge, however this method of CRISPR-inactivation is not reflective

of anti-CRISPR deployment by phages in nature. When phage DNA cleavage has been assessed *in vivo*, it occurs in as little as 2 minutes³, suggesting that phage genome degradation may outpace *de novo* Acr synthesis and function. We therefore hypothesized that successful inhibition of CRISPR-Cas immunity by Acr proteins during phage infection would be challenging, as all components of the *P. aeruginosa* immune system are expressed prior to phage infection^{6,15}.

Here, we demonstrate that complete CRISPR-Cas inactivation by a phage-produced Acr protein is challenging, and that the concentration of Acr proteins required to inactivate CRISPR-Cas is contributed by multiple phage genomes. While initial phage infections fail due to rapid genome degradation by the CRISPR-Cas system, Acr deposition prior to phage destruction causes cellular immunosuppression. If the cell is re-infected, Acr proteins from the initial phage infection enhance the likelihood of subsequent phage replication. We propose that pathogens can contribute to the “remodeling” of their host cell via rapid protein production, even if the initial infecting genomes are cleared, opening the door for their clones.

RESULTS

Anti-CRISPR proteins are imperfect CRISPR-Cas inhibitors

We utilized the diversity of *acr* genes encoded by phages infecting *P. aeruginosa* to determine the mechanism of CRISPR-Cas neutralization during infection. Five natural phages, each encoding a single *acrIF* gene, were selected to represent *acrIF1-IF4* and *acrIF7* (*acrIF5* does not exist as the sole *acrIF* gene on any phage, *acrIF6*, *F8-F10* are not encoded by this phage family). Three of the five phages exhibited reduced efficiency of plaquing (EOP) on *P. aeruginosa* strain PA14, which possesses a naturally active Type I-F CRISPR-Cas system with 1 or 2 spacers targeting these phages (**Fig. 2.1A**, WT:pEmpty normalized to plaquing on Δ CRISPR). Overexpression of a targeting crRNA (WT:pSp1) exacerbated anti-CRISPR

inefficiency, limiting the replication of all phages by at least one order of magnitude. This suggests that Acr proteins are unable to fully protect their associated phage genome.

To assess anti-CRISPR strength directly, an isogenic phage panel was generated by replacing the *acrIE3* gene in the anti-CRISPR locus of phage DMS3m with single *acrIF* genes *F1-F7* (DMS3m_{*acrIF1*}-DMS3m_{*acrIF7*}). *acrIF1-F5*, and *acrIF7* are all encoded by DMS3m-like phages in syntenic anti-CRISPR loci, while *acrIF6* was discovered in a distinct type of *P. aeruginosa* phage. WT PA14 (1 spacer targeting DMS3m, “1sp”) and a PA14 derivative which acquired 4 more spacers against DMS3m through laboratory evolution (“5sp”) were challenged with this panel of recombinant phages. For phages encoding *acrIF1*, *F2*, *F3*, *F6* or *F7*, >90% of phage in the population failed to replicate (EOP=10⁻¹) when faced with 5 targeting spacers (**Fig. 2.1B**). *acrIF4* and *acrIF5* were very weak, with 99.0-99.99% of phages failing to replicate, depending on the CRISPR spacer content. Phages must rely on *acrIF* genes when infecting the 5sp strain, as the *acrIE3*-encoding phage is unable to escape CRISPR targeting via protospacer mutation alone. We conclude that phages encoding anti-CRISPRs remain sensitive to CRISPR-Cas immunity, suggesting that anti-CRISPR deployment and action is an imperfect process.

The observations above identified groups of “strong” and “weak” Acr proteins. We selected one representative from each group for downstream experiments, and a third Acr that does not target the I-F CRISPR system (i.e. AcrIE3), as a negative control. AcrIF1 was selected as a model strong inhibitor, as its mechanism and binding affinity are known (Csy complex binding, $K_D = 3 \times 10^{-11} \text{ M}$ ^{25,31}). In contrast, AcrIF4 is a weak inhibitor that also binds the Csy complex³¹, but with a significantly slower on-rate and faster off-rate compared to AcrIF1 (**Fig. 2.1C**, **Fig. 2.2**).

Lytic replication requires a critical Acr protein concentration

We next assessed the survival of bacterial populations when infected with phages that rely on apparently imperfect Acr proteins for survival. To assay the lytic cycle only, phages were prevented from entering lysogeny by knocking out the C repressor gene (*gp1*) in DMS3m_{acrIF1}, DMS3m_{acrIF4}, and DMS3m_{acrIE3}. The virulent (vir) phages were used to infect the 5sp strain in liquid culture, and bacterial growth measured. Given that AcrIF4 has a K_D for its binding partner that is >4 orders of magnitude weaker than AcrIF1 for its binding partner, we reasoned that a higher concentration of phages encoding AcrIF4 may be required to inactivate CRISPR-Cas function. In the presence of CRISPR-Cas immunity, bacterial death only occurred at multiplicities of infection (MOI, input plaque forming units per colony forming unit) greater than 0.02 ($\geq 10^5$ PFU) for *acrIF1* (**Fig. 2.3A**) and greater than an MOI of 2.0 ($\geq 10^7$ PFU) for *acrIF4* (**Fig. 2.3B**). Phage replication observed here was due to Acr function, and not a result of phage escape mutations, as output phages remained as sensitive to CRISPR-Cas immunity as the input phage population (**Fig. 2.4A-C**). Furthermore, the phage encoding *acrIE3* had no impact on bacterial survival when faced with CRISPR immunity (**Fig. 2.3C**), while in the absence of CRISPR, phages at all concentrations cleared bacterial cultures (**Fig. 2.3D-F**). These data demonstrate that Acr-mediated CRISPR-Cas inactivation requires a critical phage concentration that is inversely proportional to Acr strength.

We hypothesized that the phage concentration dependence that dictates Acr success is achieved by the contribution of Acr proteins from multiple phage genomes in a single cell, which is not achieved at low MOIs. To this end, we rendered a subset of phages in the population non-replicative Acr donors to test if Acr donation alone is sufficient to rescue a failing (i.e. low MOI) infection. The C repressor gene (*gp1*) and surrounding immunity region from a DMS3m-like phage (JBD30) was introduced into DMS3m phages, generating a hybrid phage. The replication of the hybrid phage could be specifically prevented by overexpression of the JBD30 C repressor

(*gp1*, **Fig. 2.5A**), a protein that does not interfere with DMS3m phage with wild-type immunity regions (**Fig. 2.5B**). This enabled the mixing of two independent phage populations: a sacrificial Acr “donor” that cannot replicate and a wild-type (replication competent) Acr “acceptor”.

In the presence of donor phages encoding AcrIF1 (10^6 PFU, MOI = 0.2), we observed a striking contribution to CRISPR-Cas neutralization, despite the inability of this phage to replicate (**Fig. 2.5C**). The acceptor phages DMS3m_{acrIF1} (**Fig. 2.3G**) and DMS3m_{acrIF4} (**Fig. 2.3H**) replicated robustly from input MOIs that are unsuccessful in the absence of an AcrIF1 donor phage (**Fig. 2.3G-H**, see “IE3” and “buffer”). The presence of AcrIF1 donor phages had a mildly protective effect on the DMS3m_{acrIE3} acceptor phage (**Fig. 2.3I**), though it was not able to reach high titers. Notably, the acceptor phage output from these experiments remain as sensitive to CRISPR-Cas targeting by the 5sp host as the original input phages, demonstrating escape mutations do not arise under these conditions (**Fig. 2.4D-G**). Additionally, any potential lysogens formed by the donor phage in this experiment would not have amplified the replicating phage, as these lysogens are resistant to superinfection (**Fig. 2.5D**). These data demonstrate that the determinant of phage replicative success is the concentration of Acr proteins reached in single cells, which is achievable by Acr production from independent phage genomes (**Fig. 2.3J**).

Lysogeny requires Acr proteins contributed by transient intracellular genomes

All phages encoding Acr proteins that infect *P. aeruginosa* are naturally temperate, and can form lysogens by integrating into the bacterial genome. We therefore measured the impact of CRISPR and Acr proteins on lysogeny establishment during a single round of infection. While previous experiments examined cumulative phage replication in the lytic cycle over many hours, assaying lysogen formation over a short time frame is ideal for understanding the initial events that determine phage genome survival or cleavage. Additionally, lysogeny provides a direct readout for phage genome survival (i.e. a cell with an integrated prophage), while in lytic

replication, phage survival leads to a dead cell that cannot be recovered. For these experiments, we selected the weak AcrIF4 protein as it provided the largest dynamic range of inefficiency in a single round of infection.

We generated derivatives of DMS3m_{acrIF4} and DMS3m_{acrIE3} marked with a gentamicin resistance cassette at the end of the genome, replacing a nonessential gene, *gp52*. This allowed the independent titration of two distinct replication-competent phage populations and the selection and analysis of stable lysogens after the experiment. These phages were used to infect Δ CRISPR cells (0sp) for a time span less than a single round of infection (50 minutes, data not shown), and the number of gentamicin resistant lysogens was assessed. In the absence of CRISPR selection, a linear increase in the number of lysogens with increasing MOI was observed, over ~4 orders of magnitude (**Figs. 2.6A-B**, circles). In the presence of spacers targeting DMS3m (5sp), CRISPR immunity reduced the number of lysogen forming units (LFUs) for the weak *acr* phage DMS3m_{acrIF4} (**Fig. 2.6A**, triangles). DMS3m_{acrIF4} demonstrated concentration dependence for successful lysogeny, with efficiency of lysogeny (EOL) values below or at the limit of detection for lower MOIs, increasing to EOL = 0.01 at higher MOIs (**Fig. 2.6C**). Phage DMS3m_{acrIE3} formed no lysogens at all input concentrations tested, demonstrating that Acr-mediated immune suppression is required to establish lysogeny (**Figs. 2.6B, 2.6D**).

We hypothesized that phage concentration dependence for CRISPR neutralization during lysogeny could also be explained by phage cooperation, and that below-threshold concentrations of DMS3m_{acrIF4 gp52::gent} could be rescued by the addition of wild-type (replication competent) Acr donor phages *in trans*. To test this hypothesis, we infected the 5sp strain with a mixture of 10³ LFU marked acceptor phage and 10⁷ PFU of unmarked Acr donor phages, and measured the EOL of the acceptor phage. The EOL of the acceptor phage DMS3m_{acrIF4 gp52::gent} increased by 2 orders of magnitude with Acr donor phage DMS3m_{acrIF1}, and by 1 order of

magnitude with the DMS3m_{acrIF4} donor phage (**Fig. 2.6E**). The addition of Acr donor phages DMS3m_{acrIE3}, or an escaper phage DMS3m_{acrIE3}^{*} had no effect on the EOL of the acceptor phage, demonstrating that the donor phage must be an Acr-producer. A marked acceptor phage lacking an *acrIF* gene (DMS3m_{acrIE3 gp52::gent}) only established rare lysogens in the presence of the AcrIF1 donor phage (**Fig. 2.6F**).

To determine the specific mechanism of anti-CRISPR donation leading to survival of the acceptor phages, we used the resulting lysogens as a genetic record of infection success for both the marked acceptor phage and the unmarked donor phage (**Fig. 2.6G**). This family of Mu-like phages integrates randomly into the host genome, allowing for the formation of strains with multiple prophages³⁷. We assayed the lysogens resulting from the experiment described above (**Fig. 2.6E-F**) for the presence of the donor prophage genome in addition to the acceptor prophage. All resulting DMS3m_{acrIF4 gp52::gent} lysogens (n=48) possessed only the marked acceptor prophage, with none possessing the Acr donor prophage (**Fig. 2.7A-B**). Furthermore, the DMS3m_{acrIF4 gp52::gent} acceptor phages induced from the lysogens that formed only in the presence of Acr donor phages remained sensitive to CRISPR-Cas targeting, indicating these lysogens did not arise due to phage protospacer mutation (**Fig. 2.7B**). Double lysogens only emerged when the marked *acrIE3* phage was used as an acceptor phage, which would be incapable of maintaining lysogeny alone due to CRISPR-Cas self-targeting (**Fig. 2.6F**). These results demonstrate that the transient presence (i.e. no lysogeny) of an Acr donor phage genome in a cell was sufficient to generate enough Acr protein to protect the marked acceptor phage, leading to the establishment of lysogens that would not exist if not for the Acr donor (**Fig. 2.6E**, compare “Buffer” to “IF1”). Collectively, these data demonstrate that the production of Acr proteins from a phage genome prior to its cleavage generates an immunosuppressed cell that can be successfully parasitized by another phage upon re- or co-infection(s).

Cas9 inhibitors require bacteriophage cooperation

The intrinsic inefficiency of stoichiometric inhibitors is likely due to the requirement for the rapid synthesis of a high concentration of inhibitors before phage genome cleavage. To determine whether this model generally applies to other stoichiometric inhibitors of bacterial immunity, we engineered a *P. aeruginosa* strain to express the Cas9 protein from *Streptococcus pyogenes* (SpyCas9) and a DMS3m phage to express a previously identified Cas9 inhibitor, AcrIIA4^{18,33}. With this entirely heterologous system, we again observed inefficiency for a phage relying on an Acr protein. Spot-titration of phage lysates on a strain expressing a single guide RNA (sgRNA) targeting DMS3m decreased the titer of DMS3m_{acrIE3} by >3 orders of magnitude, while DMS3m_{acrIIA4} was protected (**Fig. 2.8A**). However, EOP quantification again revealed that relying on an Acr protein for replication is imperfect, with an EOP = 0.4 (**Fig. 2.8B**). In lytic replication infection experiments, DMS3m_{acrIIA4} displayed concentration-dependent bacterial lysis in the presence of CRISPR targeting (**Fig. 2.8C**), while DMS3m_{acrIE3} did not affect bacterial growth (**Fig. 2.8D**). The replication of DMS3m_{acrIIA4} was not due to protospacer mutation leading to phage escape because the output phage population displayed the same EOP as the input (**Fig. 2.4H-I**). In the absence of CRISPR-Cas targeting, however, both phages killed their hosts at all phage concentrations tested (**Fig. 2.8E-F**).

To determine whether this concentration dependence for Cas9 inhibition was also a result of insufficient intracellular Acr dose, a non-replicative hybrid DMS3m_{acrIIA4} phage was generated and used as an Acr donor during infection. Indeed, increased delivery of AcrIIA4 to cells enhanced replication of the wild-type DMS3m_{acrIIA4} acceptor phage by 4 orders of magnitude (**Fig. 2.8G**), demonstrating phage cooperation neutralizes CRISPR-Cas9. AcrIIA4 donation was able to slightly rescue an acceptor phage without a II-A Acr, DMS3m_{acrIE3} (**Fig. 2.8H**), however this phage was unable to replicate to high titers. Furthermore, the effect of an AcrIIA4 donor rescuing either DMS3m_{acrIIA4} or DMS3m_{acrIE3} was not due to mutational escape of the acceptor

phage (**Fig. 2.4J-L**), demonstrating these phages had survived solely due to the immunosuppressive effect of AcrIIA4 donation (**Fig. 2.8I**). Collectively, these data demonstrate that phage-phage cooperation via cellular immunosuppression is a broadly useful strategy to overcome bacterial immunity.

DISCUSSION

Here we demonstrate that the necessary intracellular concentration of an anti-CRISPR protein to achieve inactivation of CRISPR-Cas immunity depends on the relative strengths of both the inhibitor and CRISPR immunity, which dictates the number of infecting viruses required in the population. We conclude that a single cell can become immunosuppressed by Acr protein contributions from independent infection events. In the absence of viral replication, these infection events serve to contribute to the inactivation of cellular immunity, thus enhancing the probability of successful infection events in the future. We expect that cooperation of this sort is necessary when the immune process acts rapidly and irreversibly on the infecting viral genome, as CRISPR-Cas immunity does.

Anti-CRISPR deployment and successful CRISPR-Cas inactivation requires a critical concentration of phage in the population to allow replication in the lytic or lysogenic cycle. We used three distinct genetic strategies to monitor phage-phage cooperation within an otherwise clonal population, allowing the independent titration and tracking of isogenic phages: **i)** non-replicative Acr donor phages, **ii)** marked and unmarked phages to follow the fate of only one phage, and **iii)** the prophage status of lysogens, as a genetic record of phage success. In the presence of non-replicative Acr donor phages, we observed the successful lytic amplification of a low-dose of wild-type phages, otherwise destined for replication failure (**Fig. 2.3G-H**). This provided an explanation for the observed phage inefficiencies during plaque assays (**Fig. 2.1A-B**) and population concentration thresholds in liquid infections (**Fig. 2.3A-B**). Next, the

acquisition of a marked prophage was monitored in the presence of wild-type Acr-donor phages. AcrIF proteins provided *in trans* caused cellular immunosuppression, enabling the formation of lysogens that were not established in their absence (**Fig. 2.6E-F**). The presence of only a single, marked prophage in the bacterial genome demonstrates that the donor phage neither entered the lytic cycle (this would kill the cell), nor lysogenized (prophage would be integrated), but had been present in the cell transiently.

The key result here is the observation that phages can remodel their host cell, even in the absence of a replicating or integrated genome. It has long been known that integrated prophages modulate host phenotypes via gene expression, including superinfection exclusion, toxin production, and the production of Acr proteins^{15,37-40}. Furthermore, the Imm protein produced by the lytic phage T4 prevents other phages in the environment from infecting the cell that one phage is currently replicating within⁴¹. This has been attributed to preventing sequential infections and the disruption of the carefully timed phage replication cycle. In contrast to these examples, we propose a new model of phage-induced host remodeling, whereby a transient, unsuccessful infection produces proteins that inactivate defense, enabling future infections.

Consistent with our observations of viral cooperation, beneficial virus-virus interactions in both eukaryotic and prokaryotic systems have been previously observed. Broadly, these phenomena can be separated into 2 categories: **i)** cooperative interactions between distinct viral genotypes and **ii)** group behaviors manifested in clonal viral populations: **i)** Similar to Acr proteins functioning as a public good, genetically distinct viruses can share protein products during coinfection⁴², even bypassing deleterious mutations *in cis* via functional complementation *in trans*⁴³⁻⁴⁵. Additionally, the direct exchange of viral genetic material can also increase viral fitness. The mosaic nature of phage genomes⁴⁶⁻⁴⁸ and the high abundance of chimeric viruses

in nature highlights the importance of coinfection and genetic exchange in viral evolution⁴⁹⁻⁵¹. In fact, CRISPR-targeted phages can evade CRISPR-Cas immunity via homologous recombination with genetically distinct phages, disrupting protospacers^{52,53}. **ii)** Group behaviors manifesting in clonal populations of virus is less frequently reported, likely because they leave no genetic signature. However, the lambda lytic/lysogeny switch is a famous example of phage group behavior: during lambda phage co-infection, high concentrations of the CII protein product derived from multiple infecting clones drives the cooperative decision to enter lysogeny⁵⁴⁻⁵⁶. In more recent literature, the discovery of the widespread *arbitrium* system as the first phage-phage communication mechanism demonstrates the potential of phages to act as a group and manifest cooperative behaviors⁵⁷. The immunosuppressive mechanism of anti-CRISPR function further exemplifies cooperation within clonal populations of phages, which may occur more often than was previously appreciated. The distinct aspect here is the altruistic nature of immunosuppression: to neutralize CRISPR-Cas immunity, many infections must fail such that a few can succeed. To our knowledge this is the first documented example of true viral altruism, which is evolutionary beneficial only through kin selection.

A distinct, but notable observation from this work is that not all Acr proteins operate at equivalent strengths. However, encoding even a weak inhibitor (e.g. AcrIF4) still provides a significant advantage to the phage, compared to lacking them entirely (**Fig. 2.1B**). We show that AcrIF4 binds the Csy complex with affinities that are orders of magnitude weaker (**Fig. 2.1C**) than Acr proteins like AcrIF1 and AcrIF2²⁵. We selected AcrIF1 as a model strong Acr protein because of its comparable mechanism of action to AcrIF4 (i.e. Csy complex binding), and consider it representative of other strong Acr proteins (AcrIF2, F3, F6, F7), based on EOP data. Going forward, we speculate that the strongest Acr proteins would be enzymatic in nature, allowing rapid and efficient inactivation of CRISPR complexes in a sub-stoichiometric manner, although no such Acr mechanism has been discovered. While not an enzyme, the recent

demonstration of the AcrIIC3 protein inactivating two Cas9 proteins at the same time would likely be a more efficient path towards CRISPR neutralization³⁴. It is also interesting to consider individual bacterial strains that encode multiple CRISPR-Cas system subtypes⁵⁸⁻⁶⁰, all which must be neutralized in order for a targeted phage to replicate. A dual-activity inhibitor is likely at a specific disadvantage in this scenario, as one protein would be tasked with inhibiting Cas proteins produced by two different systems. This may in part explain why DMS3m-like *Pseudomonas* phages often encode dedicated Type I-E and Type I-F Acrs in the same Acr locus^{15,17}, instead of employing dual I-F and I-E inhibitors such as AcrIF6¹⁶. Although encoding multiple Acrs comes with the burden of more genetic cargo in a phage's genome, this strategy could be advantageous on a biochemical level when infecting a bacterial strain with multiple CRISPR-Cas subtypes.

The challenge of neutralizing a pre-expressed CRISPR-Cas system likely explains why stoichiometric inhibitors like Acr proteins are imperfect, and phages relying on them are partially targeted by CRISPR. The sacrificial, population-level aspect of CRISPR inhibition is reminiscent of the manifestations of CRISPR adaptation in populations of bacterial cells. The majority of infected naïve host cells die, before a clone with a new spacer emerges^{2,61}. In the case of anti-immunity, many phages die in order to inhibit CRISPR on a single cell level, and this must happen at a sufficient frequency within a community for phage to prevail. We suspect that this mechanism of cellular immunosuppression and inter-parasite cooperation may have parallels in other host-pathogen interactions, where concentration dependence manifests at predictable levels due to the strengths of immune and anti-immune processes.

REFERENCES

1. Koonin, E. V., Makarova, K. S. & Wolf, Y. I. Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annu Rev Microbiol* **71**, annurev-micro-090816-093830 (2017).
2. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
3. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
4. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
5. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. **520**, 505–510 (2015).
6. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **194**, 5728–5738 (2012).
7. Bikard, D., Hatoum-Aslan, A., Mucida, D. & Marraffini, L. A. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12**, 177–186 (2012).
8. Edgar, R. & Qimron, U. The *Escherichia coli* CRISPR system protects from λ lysogenization, lysogens, and prophage induction. *J. Bacteriol.* **192**, 6291–6294 (2010).
9. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).
10. Westra, E. R. *et al.* Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. *Curr. Biol.* **25**, 1043–1049 (2015).
11. van Houte, S. *et al.* The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* **532**, 385–388 (2016).

12. Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs. *Annual Review of Virology* **4**, 37–59 (2017).
13. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: discovery, mechanism and function. *Nat Rev Micro* **16**, 12–17 (2018).
14. He, F. *et al.* Anti-CRISPR proteins encoded by archaeal lytic viruses inhibit subtype I-D immunity. *Nature Microbiology* **3**, 461–469 (2018).
15. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
16. Pawluk, A. *et al.* Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology* **1**, 1–6 (2016).
17. Pawluk, A., Bondy-Denomy, J., Cheung, V. H. W., Maxwell, K. L. & Davidson, A. R. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* **5**, e00896–e00896–14 (2014).
18. Rauch, B. J. *et al.* Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150–158.e10 (2017).
19. Hynes, A. P. *et al.* An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. *Nature Microbiology* **315**, 1 (2017).
20. Pawluk, A. *et al.* Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* **167**, 1829–1838.e9 (2016).
21. Wiedenheft, B. *et al.* RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences* **108**, 10092–10097 (2011).
22. Przybilski, R. *et al.* Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol* **8**, 517–528 (2011).

23. Richter, H. *et al.* Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. *Nucleic Acids Research* **40**, 9887–9896 (2012).
24. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science* **329**, 1355–1358 (2010).
25. Chowdhury, S. *et al.* Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. *Cell* **169**, 47–57.e11 (2017).
26. Peng, R. *et al.* Alternate binding modes of anti-CRISPR viral suppressors AcrF1/2 to Csy surveillance complex revealed by cryo-EM structures. *Cell Res.* **132**, 712 (2017).
27. Guo, T. W. *et al.* Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex. *Cell* **171**, 414–426.e12 (2017).
28. Rollins, M. F. *et al.* Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. *Proceedings of the National Academy of Sciences* **23**, 201616395–E5121 (2017).
29. Wang, J. *et al.* A CRISPR evolutionary arms race: structural insights into viral anti-CRISPR/Cas responses. *Cell Res.* **26**, 1165–1168 (2016).
30. Wang, X. *et al.* Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Struct. Mol. Biol.* **23**, 868–870 (2016).
31. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
32. Maxwell, K. L. *et al.* The solution structure of an anti-CRISPR protein. *Nature Communications* **7**, 13134 (2016).
33. Dong, D. *et al.* Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* **546**, 436–439 (2017).

34. Harrington, L. B. *et al.* A Broad-Spectrum Inhibitor of CRISPR-Cas9. *Cell* **170**, 1224–1233.e15 (2017).
35. Shin, J. *et al.* Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci Adv* **3**, e1701620 (2017).
36. Yang, H. & Patel, D. J. Inhibition Mechanism of an Anti-CRISPR Suppressor AcrIIA4 Targeting SpyCas9. *Mol Cell* **67**, 1–17 (2017).
37. Bondy-Denomy, J. *et al.* Prophages mediate defense against phage infection through diverse mechanisms. *The ISME Journal* **10**, 2854–2866 (2016).
38. Weigle, J. J. & Delbruck, M. Mutual exclusion between an infecting phage and a carried phage. *J. Bacteriol.* **62**, 301–318 (1951).
39. Waldor, M. K. & Mekalanos, J. J. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910–1914 (1996).
40. Bondy-Denomy, J. & Davidson, A. R. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *J Microbiol* **52**, 235–242 (2014).
41. Lu, M. J. & Henning, U. The immunity (imm) gene of Escherichia coli bacteriophage T4. *Journal of Virology* **63**, 3472–3478 (1989).
42. Xue, K. S., Hooper, K. A., Ollodart, A. R., Dingens, A. S. & Bloom, J. D. Cooperation between distinct viral variants promotes growth of H3N2 influenza in cell culture. *Elife* **5**, e13974 (2016).
43. Cicin-Sain, L., Podlech, J., Messerle, M., Reddehase, M. J. & Koszinowski, U. H. Frequent Coinfection of Cells Explains Functional In Vivo Complementation between Cytomegalovirus Variants in the Multiply Infected Host. *Journal of Virology* **79**, 9492–9502 (2005).
44. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2005).

45. Aguilera, E. R., Erickson, A. K., Jesudhasan, P. R., Robinson, C. M. & Pfeiffer, J. K. Plaques Formed by Mutagenized Viral Populations Have Elevated Coinfection Frequencies. *mBio* **8**, e02020–16–12 (2017).
46. Botstein, D. A theory of modular evolution for bacteriophages. *Ann. N. Y. Acad. Sci.* **354**, 484–490 (1980).
47. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci USA* **96**, 2192–2197 (1999).
48. Hatfull, G. F. & Hendrix, R. W. Bacteriophages and their genomes. *Curr Opin Virol* **1**, 298–303 (2011).
49. Diemer, G. S. & Stedman, K. M. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol Direct* **7**, 13 (2012).
50. Roux, S. *et al.* Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nature Communications* **4**, 1–10 (2013).
51. Krupovic, M. *et al.* Multiple Layers of Chimerism in a Single-Stranded DNA Virus Discovered by Deep Sequencing. *Genome Biol Evol* **7**, 993–1001 (2015).
52. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**, 1047–1050 (2008).
53. Paez-Espino, D. *et al.* CRISPR Immunity Drives Rapid Phage Genome Evolution in *Streptococcus thermophilus*. *mBio* **6**, e00262–15–9 (2015).
54. Kourilsky, P. & Knapp, A. Lysogenization by bacteriophage lambda. III. Multiplicity dependent phenomena occurring upon infection by lambda. *Biochimie* **56**, 1517–1523 (1974).
55. Zeng, L. *et al.* Decision Making at a Subcellular Level Determines the Outcome of Bacteriophage Infection. *Cell* **141**, 682–691 (2010).

56. Trinh, J. T., Székely, T., Shao, Q., Balázsi, G. & Zeng, L. Cell fate decisions emerge as phages cooperate or compete inside their host. *Nature Communications* **8**, 14341–13 (2017).
57. Erez, Z. *et al.* Communication between viruses guides lysis–lysogeny decisions. *Nature* **541**, 488–493 (2017).
58. Carte, J. *et al.* The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Molecular Microbiology* **93**, 98–112 (2014).
59. Patterson, A. G. *et al.* Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. *Mol Cell* **64**, 1102–1108 (2016).
60. van Belkum, A. *et al.* Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant *Pseudomonas aeruginosa*. *mBio* **6**, e01796–15 (2015).
61. Hynes, A. P., Villion, M. & Moineau, S. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nature Communications* **5**, 4399 (2014).
62. Choi, K.-H. & Schweizer, H. P. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat Protoc* **1**, 153–161 (2006).

MATERIALS AND METHODS

Microbes

Pseudomonas aeruginosa strains (UCBPP-PA14 and PAO1) and *Escherichia coli* strains (DH5 α , for plasmid maintenance) were cultured on lysogeny broth (LB) agar or liquid media at 37 °C. LB was supplemented with gentamicin (50 μ g/mL for *P. aeruginosa*, 30 μ g/mL for *E. coli*) to maintain the pHERD30T plasmid or carbenicillin (250 μ g/mL for *P. aeruginosa*, 100 μ g/mL for *E. coli*) to maintain pHERD20T or pMMB67HE. To maintain pHERD30T and pMMB67HE in the same strain of *P. aeruginosa*, double selection of 30 μ g/mL gentamicin and 100 μ g/mL carbenicillin was employed. In all *P. aeruginosa* experiments, expression from pHERD20/30T was induced with 0.1% arabinose and expression from pMMB67HE was induced with 1mM Isopropyl β -D-1-thiogalactopyranoside (IPTG). *Escherichia coli* strains BL21 (DE3) were grown in LB broth supplemented with ampicillin (100 μ g/mL) to maintain pAcrIF4, or with ampicillin (100 μ g/mL) and kanamycin (50 μ g/mL) to maintain pCsy and pCRISPR together.

Phages

Pseudomonas aeruginosa DMS3m-like phages (JBD30, MP29, JBD88a, JBD24, LPB1, DMS3m and DMS3m derivatives) were amplified on PA14 Δ CRISPR or PAO1 and stored in SM buffer at 4 °C.

Construction of PA14 crRNA overexpression strains

PA14 CRISPR2 spacer-17 or CRISPR2 spacer-20 sequences flanked by PA14 Type I-F direct repeats were ordered as complementary ssDNA oligos (IDT), annealed, and ligated into the NcoI/HindIII site in pHERD30T to make pAB02 and pAB03, respectively. These constructs were transformed into PA14 WT, and expression induced with 0.1% arabinose.

Construction of PAO1::SpyCas9 expression strain

SpyCas9 expressed from the P_{LAC} promoter of pUC18T-mini-Tn7T-Gm (pBAO95) was integrated into the *P. aeruginosa* strain PAO1 chromosome by electroporation and Flp-mediated marker excision as previously described⁶². To generate the heterologous Type II-A PAO1 strain the PAO1-attTn7::pUC18T-miniTn7T- P_{LAC} -SpyCas9 strain was transformed with pMMB67HE- P_{LAC} -sgRNA (pBAO72) by electroporation. In all experiments with this strain, SpyCas9 and the sgRNA were induced with 1mM IPTG.

Construction of recombinant DMS3m_{acr} phages

DMS3m_{acrIF1} was generated previously¹⁵ by infecting cells containing a recombination plasmid bearing JBD30 genes 34-38 (the anti-CRISPR locus with large flanking regions). JBD30 naturally carries *acrIF1* and has high genetic similarity to DMS3m_{acrIE3}, permitting for the selection of recombinant DMS3m phages that acquired *acrIF1*. To generate the extended panel of DMS3m_{acr} phages in this work, recombination cassettes were generated with regions from up and downstream the anti-CRISPR gene from JBD30 and these fragments were assembled to flank the *acr* gene of interest on pHERD20T or pHERD30T (see **Table 2.1** for *acr* gene sources, **Table 2.2** for recombination plasmids) using Gibson assembly methods. In the case of AcrIF5, AcrIF6, and AcrIIA4 recombination cassettes, a ribosomal binding site was introduced between the *acr* and the downstream gene *aca1* to ensure proper expression of the *aca1* gene. Recombinant phages were generated by infecting cells bearing these recombination substrates. DMS3m_{acr} phages were screened for their ability to resist CRISPR targeting, and the insertion of the anti-CRISPR gene was confirmed by PCR. Virulent derivatives of DMS3m_{acr} phages were constructed by deleting *gp1* (C repressor) using materials and methods previously generated⁶.

Construction of DMS3m_{acr gp52::gent} phages

A recombination substrate (pAB45) with a gentamicin resistance cassette flanked by homology arms matching the DMS3m genome up and downstream of *gp52* (450 bp and 260 bp, respectively) was assembled into pHERD20T using Gibson assembly. This recombination cassette was transformed into PA14 Δ CRISPR lysogenized with either DMS3m_{acrE3} or DMS3m_{acrIF4}. These transformed lysogens were grown under gentamicin selection for 16 hours, then sub-cultured 1:100 into LB with gentamicin and 0.2 μ g/mL mitomycin C to induce the DMS3m_{acr} prophage. Supernatants were harvested after 24 hours of induction, and used to infect PA14 Δ CRISPR in liquid culture for 24 hours. These cells were then plated on gentamicin plates to select for cells that had acquired a prophage bearing the gentamicin resistance cassette, and gentamicin resistant lysogens were then re-induced with 0.2 μ g/mL mitomycin C to recover the recombinant phage.

Construction of DMS3m_{acr gp1-JBD30} Hybrid_{acr} phages

DMS3m_{acrE3} and JBD30_{acrE3} were used to co-infect PA14 Δ CRISPR and the infected cells were mixed with molten top agar and poured onto solid plates. After 24 hours of growth at 30 °C, the phages were harvested by flooding the plate with SM buffer and collecting and clarifying the supernatant. Phages were then used to infect PA14 Δ CRISPR expressing the DMS3m C repressor from pHERD30T (pAB80), and the infections were mixed with molten top agar and poured onto solid plates. After 24 hours of growth at 30 °C, individual plaques with DMS3m morphology were picked, purified 3x by passage in PA14 Δ CRISPR and screened as shown in **Fig. 2.5B**. The *acrIF1* gene was then knocked in to this hybrid phage using methods described above to generate DMS3m_{acrIF1 gp1-JBD30}.

Plaque forming unit quantification

Phage plaque forming units (PFU) were quantified by mixing 10 µl of phage with 150 µl of an overnight culture of host bacteria. The infection mixture was incubated at 37 °C for 10 minutes to promote phage adsorption, then mixed with 3 mL molten top agar and spread on an LB agar plate supplemented with 10 mM MgSO₄. After 16 hours of growth 30 °C, PFUs were quantified.

Phage titering

A bacterial lawn was generated by spreading 3 mL of top agar seeded with 150 µl of host bacteria on a LB agar plate supplemented with 10 mM MgSO₄. 3 µl of phage serially diluted in SM buffer was then spotted onto the lawn, and incubated at 30 °C for 16 hours.

Liquid culture phage infections

A *P. aeruginosa* overnight culture was diluted 1:100 in LB supplemented with 10 mM MgSO₄, required antibiotics and inducer. 140 µl of diluted bacteria were then infected with 10 µl of phage diluted in SM buffer in a 96 well Costar plate. These infections proceeded for 24 hours in a Synergy H1 microplate reader (BioTek, using Gen5 software) at 37 °C with continuous shaking. After 24 hours, phage was extracted by treating each sample with chloroform followed by centrifugation at 21,000 x g for 2 minutes.

Prophage acquisition and lysogen analysis

Overnight cultures of PA14 were subcultured at 1:100 for ~3 hours (OD_{600nm} = 0.3) in LB supplemented with 10 mM MgSO₄. 1 mL of cells was infected with 10 µl DMS3m_{acr gp52::gent} and incubated for 50 minutes at 37 °C, shaking at 100 rpm. The sample was then treated with a 10% volume of 10X gentamicin, spun down at 8,000xg, and resuspended in 200 µl of LB with 50 µg/mL gentamicin. 100 µl of sample was then plated (after further dilution, if required) on gentamicin selection plates and incubated at 37 °C. To analyze the lysogens, the resulting

colonies were grown for 16 hours in LB + 10 mM MgSO₄ (no selection), the supernatants harvested, and serial dilutions spotted onto lawns of non-targeting PA14 (0sp) or PA14 with 5 targeting spacers (5sp). Crude genomic DNA for PCR analysis was harvested from the lysogens by boiling 10 µl of culture in 0.02% SDS for 10 minutes.

Lysogen PCR

PCR amplification of 2 µl of crude genomic DNA harvested from lysogens was used to screen for the presence of *DMS3m-gp52* and the *gent* cassette using MyTaq (Bioline) polymerase with MyTaq GC buffer under standard conditions.

Csy complex purification

Csy genes and a synthetic CRISPR array were co-expressed on separate vectors (pCsy, pCRISPR) in *E. coli* BL21 (DE3) cells as previously described²⁸. Expression was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD₆₀₀ ~ 0.5. Cells were incubated overnight at 16°C, then pelleted by centrifugation (5,000 × g for 15 min at 4 °C), and resuspended in lysis buffer [50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 300 mM potassium chloride, 5% glycerol, 1 mM Tris(2- carboxyethyl) phosphine hydrochloride (TCEP), 1× protease inhibitor mixture (Thermo Scientific)]. Pellets were sonicated on ice for 3 × 2.5 min (1 s on, 3 s off), and then the lysate was clarified by centrifugation at 22,000 × g for 30 min at 4 °C. The Csy complex self-assembles *in vivo*, and the intact complex was affinity-purified over NiNTA Superflow resin (Qiagen) using 6xhis tags on Cas7f. Protein was eluted with lysis buffer supplemented with 300 mM imidazole and then concentrated (Corning Spin-X concentrators) at 4 °C before further purification over a Superdex 200 size-exclusion column (GE Healthcare) in 20 mM Hepes pH 7.5, 100 mM KCl, 5% glycerol, and 1 mM TCEP.

AcrIF4 purification

Gene 37 from phage JBD26 (AcrIF4) was cloned into a p15TV-L vector with N-terminal His6 tags (pAcrIF4) and expressed in *E. coli* BL21 (DE3) cells. Expression was induced with 0.5 mM IPTG at OD₆₀₀ ~ 0.5. Cells were incubated overnight at 16°C, then pelleted by centrifugation (5,000 × g for 15 min at 4 °C), and resuspended in lysis buffer containing 50 mM Tris, pH 7.5, 300 mM NaCl, 5% glycerol, 0.5x protease inhibitor cocktail (Thermo Scientific) and 1 mM TCEP. Cells were lysed by sonication and lysate was clarified by centrifugation as described above. AcrIF4 protein was affinity-purified over NiNTA Superflow resin (Qiagen) and eluted in lysis buffer supplemented with 300 mM imidazole, then concentrated (Corning Spin-X concentrators) at 4 °C before further purification over a Superdex 75 size-exclusion column (GE Healthcare) in 20 mM Tris, pH 7.5, 250 mM NaCl, 5% glycerol, 1 mM TCEP.

Surface plasmon resonance

Purified Csy complex was covalently immobilized by amine coupling to the surface of a carboxymethyl-dextran-modified (CM5) sensor chip (GE Healthcare). Purified 6his-tagged AcrIF4 was injected into the buffer flow in increasing concentrations (1.85 nM, 55.6 nM, 167 nM, 500 nM, 1.5 μM), and Csy complex-AcrIF4 binding events were recorded in real time. Experiments were conducted at 37°C, in 20 mM HEPES pH 7.5, 100 mM KCl, 1mM TCEP, 0.005% Tween.

QUANTIFICATION AND STATISTICAL ANALYSIS

All numerical data, with the exception of the surface plasmon resonance (SPR) data, were analyzed and plotted using GraphPad Prism 6.0 software. The SPR data were analyzed and plotted using Biocore evaluation software (GE). Below, we provide the details of the number of biological replicates as well as data quantification and presentation for the experimental methods utilized in this manuscript.

Efficiency of plaquing (Fig. 2.1A-B, Fig. 2.8B, Fig. 2.4A-L)

Efficiency of plaquing (EOP) was calculated as the ratio of the number of plaque forming units (PFUs) that formed on a targeting (+CRISPR, +sgRNA) strain of bacteria divided by the number of PFUs that formed on a non-targeting (Δ CRISPR, vector) strain. Each PFU measurement was performed in biological triplicate. The EOP data in Fig. 2.1A-B and Fig. 2.8B are displayed as the mean EOP \pm standard deviation (error bars) whereas the EOP data in Fig. 2.4A-L are displayed as individual replicate values overlaid with the mean EOP value \pm standard deviation.

Bacterial growth curves (Fig. 2.3A-F, Fig. 2.8C-F)

OD_{600nm} values were measured in biological triplicate for each experimental condition over a period of 12 hours, and the data displayed as the mean OD_{600nm} as a function of time (hours) \pm standard deviation (error bars).

Quantification of phage lytic replication (Fig. 2.3G-I, Fig. 2.8G-H, Fig. 2.5C)

Phage infections were performed in biological triplicate, and the phages harvested from each infection were quantified as plaque forming units (PFUs) on a non-targeting (Δ CRISPR, vector) strain. Values are displayed as the mean number of PFUs from the 3 experimental replicates, \pm standard deviation (error bars).

Quantification of phage lysogeny (Fig. 2.6A-B)

Phage lysogeny was measured as the number of lysogen forming units (LFUs) that formed under a given experimental condition. In our experimental setup, each sample was diluted at least 2-fold before quantification, meaning that the limit of detection (LoD) of this assay is 2 LFUs. Phage lysogeny experiments were performed in biological triplicate, and each replicate value is displayed.

Efficiency of lysogeny (Fig. 2.6C-F)

Efficiency of lysogeny (EOL) was calculated as ratio of the number of lysogen forming units (LFUs) that form under the targeting condition (5sp) divided by the number of LFUs that form under the non-targeting condition (0sp). Phage lysogeny experiments were performed in biological triplicate, and EOL is displayed as mean EOL +/- standard deviation (error bars).

Analysis of AcrIF4 binding kinetics (Fig. 2.1C, Fig. 2.2)

Data were fit with a model describing Langmuir binding (i.e. 1:1 binding between free analyte and immobilized ligand). Plotted residual data points scattered around zero and were <10% of R_{\max} , indicating good model fit. Kinetic rate constants were extracted from this curve fit using Biacore evaluation software (GE). Parameter significance was evaluated by assessing standard error (SE)/T-value (T-value = parameter value/standard error). This value provides a measure of how sensitive the model fit is to changes in the parameter value; high SE/low T-value indicates poor significance. SEs for k_a and k_d were both >21-fold lower than T-values, indicating good significance.

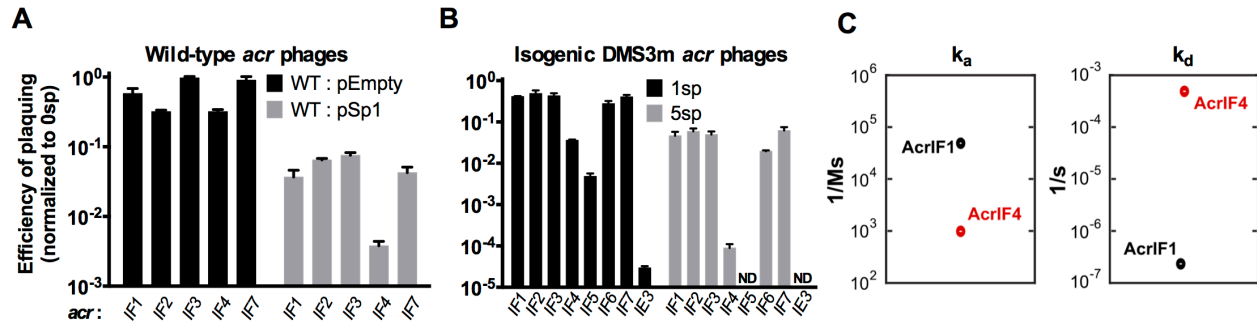


Figure 2.1. Anti-CRISPRs are imperfect CRISPR-Cas inhibitors

(A) Efficiency of plaquing (EOP) of 5 related phages bearing distinct *acrIF* genes (JBD30_{*acrIF1*}, MP29_{*acrIF2*}, JBD88a_{*acrIF3*}, JBD24_{*acrIF4*}, LPB1_{*acrIF7*}) on *Pseudomonas aeruginosa* strain PA14. Plaque forming units (PFUs) were quantified on wild-type PA14 with 1-2 natural targeting spacers (WT + pEmpty) or on PA14 overexpressing 1 targeting spacer (WT + pSp1), then normalized to the number of PFUs measured on a non-targeting PA14 derivative (0sp). Data are represented as the mean of 3 biological replicates +/- SD.

(B) EOP of isogenic DMS3m_{*acr*} phages with *acrIF1-7* or *acrIE3* in the DMS3m *acr* locus. EOP was calculated as PFU counts measured on WT PA14 with 1 targeting spacer (1sp) or a laboratory evolved PA14 derivative with 5 targeting spacers (5sp) normalized to PFU counts measured on non-targeting PA14 (0sp). Data are represented as the mean of 3 biological replicates +/- SD. ND, not detectable.

(C) Plot of association (k_a) and dissociation (k_d) rates for AcrIF1 (data adapted from Chowdhury et al. 2017) and AcrIF4 binding the PA14 Csy complex. AcrIF1 rate constants: $k_a = 5 \times 10^4$ (1/Ms), $k_d = 2 \times 10^{-7}$ (1/s), $K_D = 3 \times 10^{-11}$ M. AcrIF4 rate constants: $k_a = 1 \times 10^3$ (1/Ms), $k_d = 5 \times 10^{-4}$ (1/s), $K_D = 4 \times 10^{-7}$ (M). See Fig. 2.2 for AcrIF4 SPR sensogram.

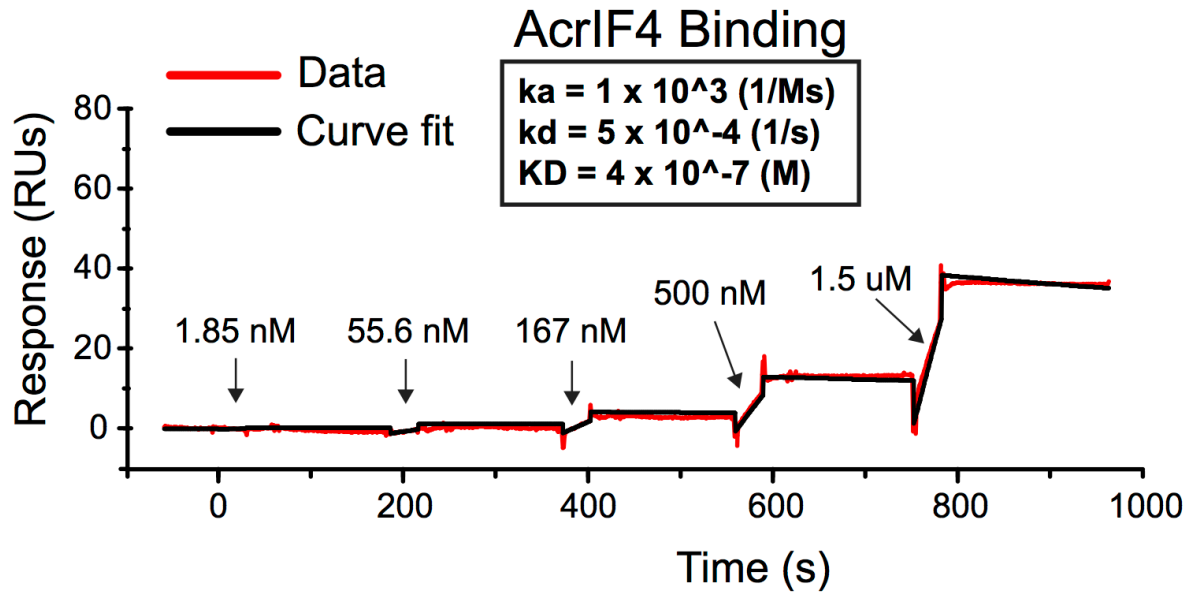
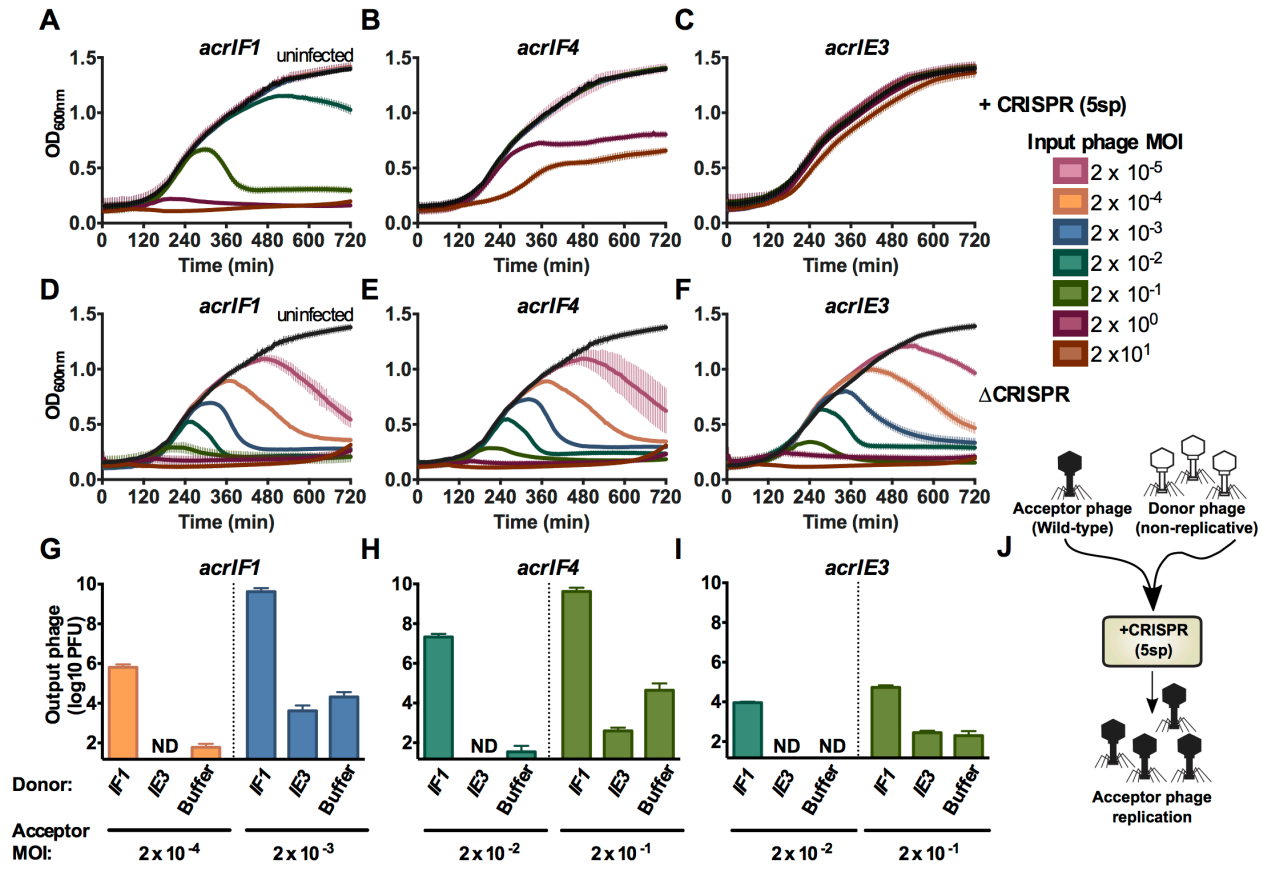


Figure 2.2. Sensogram of AcrIF4 binding the Csy complex, Related to Fig. 2.1C

Sensogram showing real-time binding of increasing concentrations of free AcrIF4 (1.85 nM, 55.6 nM, 167 nM, 500 nM, 1.5 μ M) to immobilized Csy complex. A model describing Langmuir binding (black line) was fit to the data to calculate binding constants (k_a , k_d , and K_D ; boxed inset).



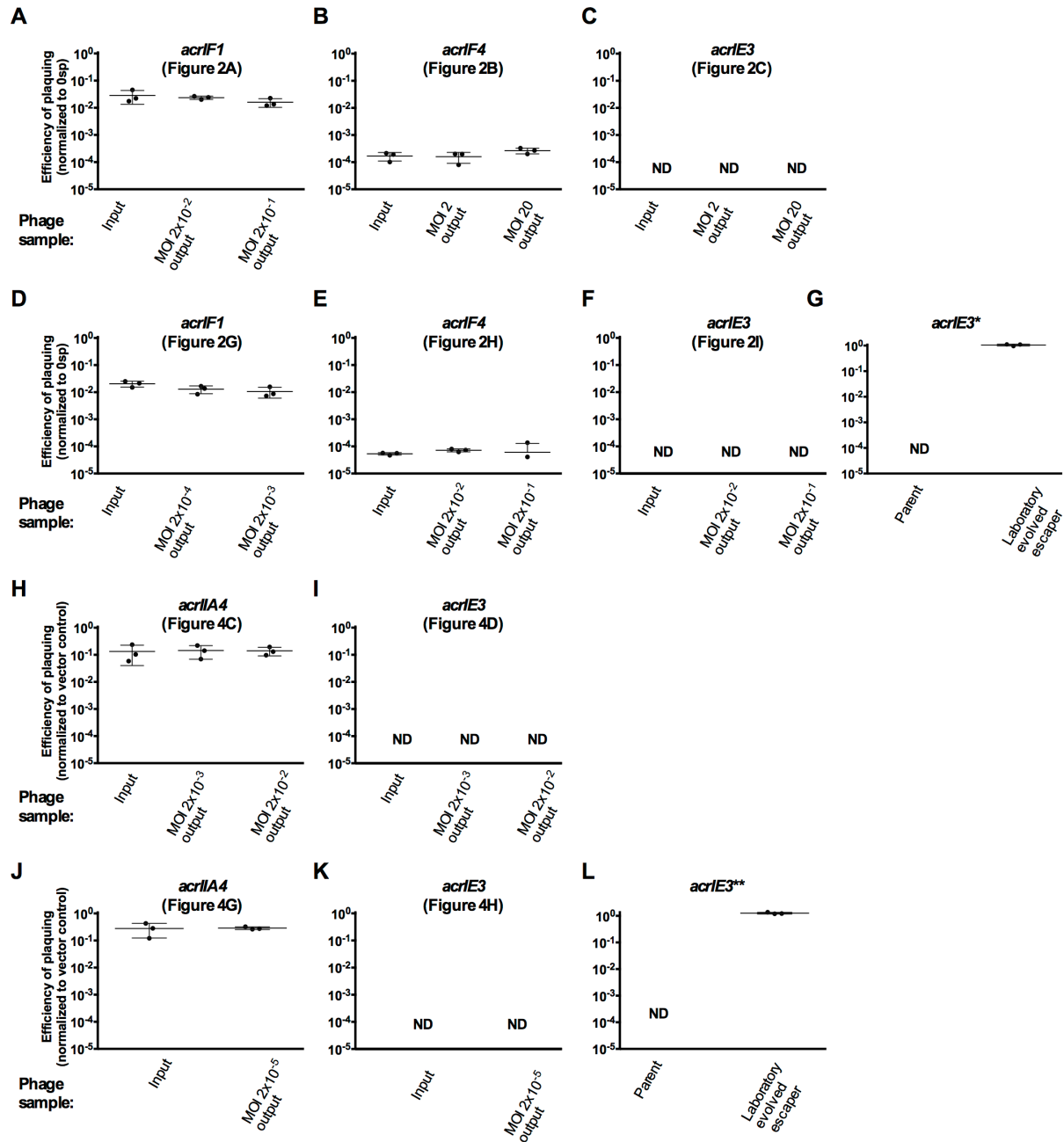


Figure 2.4. Output phages from liquid growth experiments remain CRISPR sensitive, Related to Fig. 2.3 and Fig. 2.8

(A-C) Efficiency of plaquing (EOP) of the original stocks of virulent DMS3m_{acr} (input) on PA14 5sp compared to EOP of DMS3m_{acr} harvested from high MOI infections in Figure 2A-C (MOI 2×10^{-2} , MOI 2×10^{-1} output for DMS3m_{acrIF1} and MOI 2 and MOI 20 output for DMS3m_{acrIF4} and DMS3m_{acrIE3}).

(D-F) EOP of acceptor output phages that amplified in the presence of AcrIF1 donor phages from Figure 2D-F on PA14 5sp compared to the original stock of virulent DMS3m_{acr} (input).

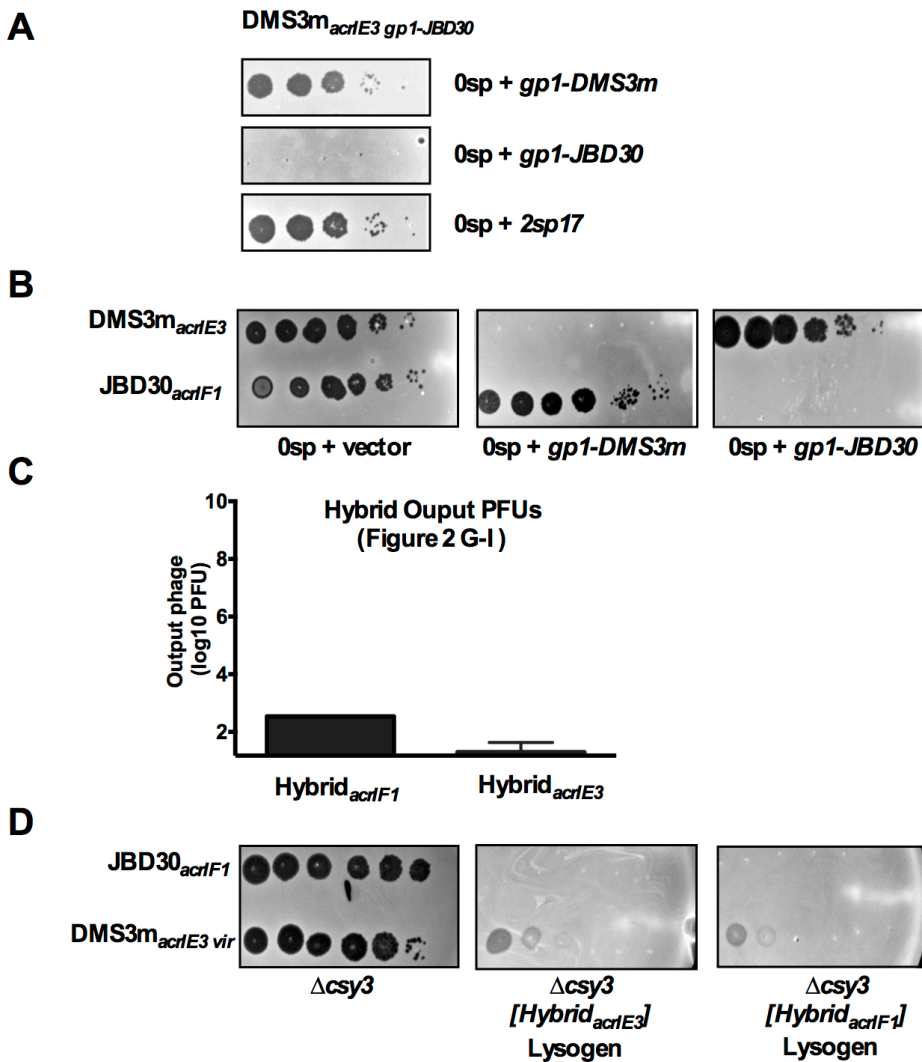


Figure 2.5. Generating and validating Hybrid_{acr} phages, Related to Fig 2.3 and Fig. 2.8

(A) 10-fold serial dilutions of hybrid DMS3m_{ΔacrE3 gp1-JBD30} plated on lawns of non-targeting (0sp) *Pseudomonas aeruginosa* PA14 expressing the DMS3m C repressor (*gp1-DMS3m*), the JBD30 C repressor (*gp1-JBD30*), or a crRNA which uniquely targets JBD30 (*2sp17*) outside of the immunity region.

(B) 10-fold serial dilutions of DMS3m_{ΔacrE3} and JBD30_{ΔacrF1} spotted on lawns of non-targeting (0sp) *Pseudomonas aeruginosa* PA14 expressing the DMS3m C repressor (*gp1-DMS3m*), the JBD30 C repressor (*gp1-JBD30*), or a vector control.

(C) Hybrid phage (Hybrid_{ΔacrF1} or Hybrid_{ΔacrE3}) harvested from infections of PA14 5sp expressing the JBD30 C repressor from experiments shown in Figure 2G-I. Hybrid PFUs were quantified on the 0sp PA14 strain. Data are represented as the mean of 3 biological replicates +/- SD.

(D) 10-fold serial dilutions of JBD30_{ΔacrF1} or virulent DMS3m_{ΔacrE3} spotted on lawns of PA14 Δcsy3 or PA14 Δcsy3 lysogenized with Hybrid_{ΔacrE3} or Hybrid_{ΔacrF1}. Despite being heteroimmune with respect to JBD30, the DMS3m phage is unable to replicate well on this lysogens due to other superinfection exclusion properties of DMS3m.

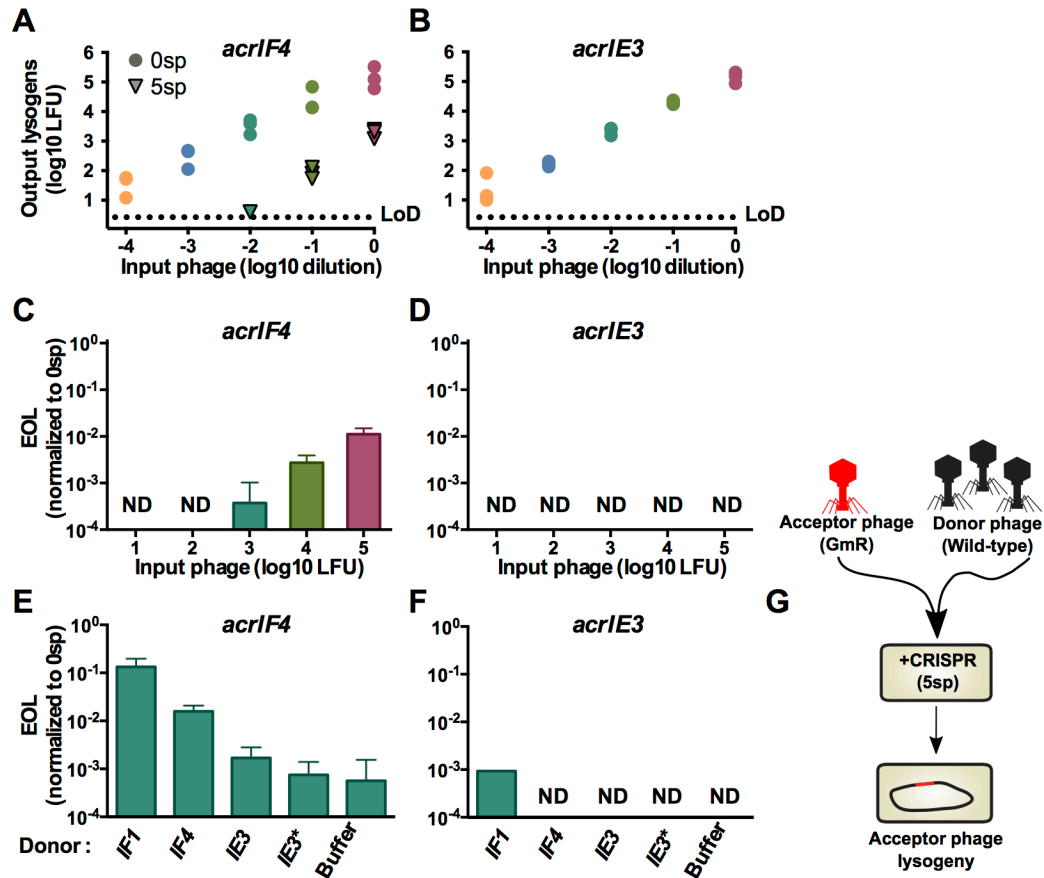


Figure 2.6. Immunosuppression facilitates acquisition of a marked prophage

(A,B) Acquisition of a marked DMS3m_{acrIF4 gp52::gent} (A) or DMS3m_{acrIE3 gp52::gent} (B) prophage by PA14 with 0 spacers (0sp, circles) or 5 targeting spacers (5sp, triangles). This experiment was performed in biological triplicate, and individual replicate values are displayed. LoD, limit of detection.

(C,D) Efficiency of lysogeny (EOL) of DMS3m_{acrIF4 gp52::gent} (A) and DMS3m_{acrIE3 gp52::gent} (B) in the presence of CRISPR targeting. EOL was calculated by dividing the output lysogens forming units (LFUs) from the strain with 5 targeting spacers (5sp) to the number of LFUs in PA14 with 0 targeting spacers (0sp). Data are represented as the mean of 3 biological replicates ± SD. ND, not detectable.

(E,F) EOL of 10³ LFUs of DMS3m_{acrIF4 gp52::gent} (E) and DMS3m_{acrIE3 gp52::gent} (F) in the presence of 10⁷ PFU of the indicated DMS3m_{acr} phage. Data are represented as the mean of 3 biological replicates ± SD. ND, not detectable. See Figure S4 for analysis of lysogen prophage content.

(G) Schematic of the experimental design in E-F, where a high MOI of wild-type “donor” phages is used to rescue a low MOI infection of marked “acceptor” phages.

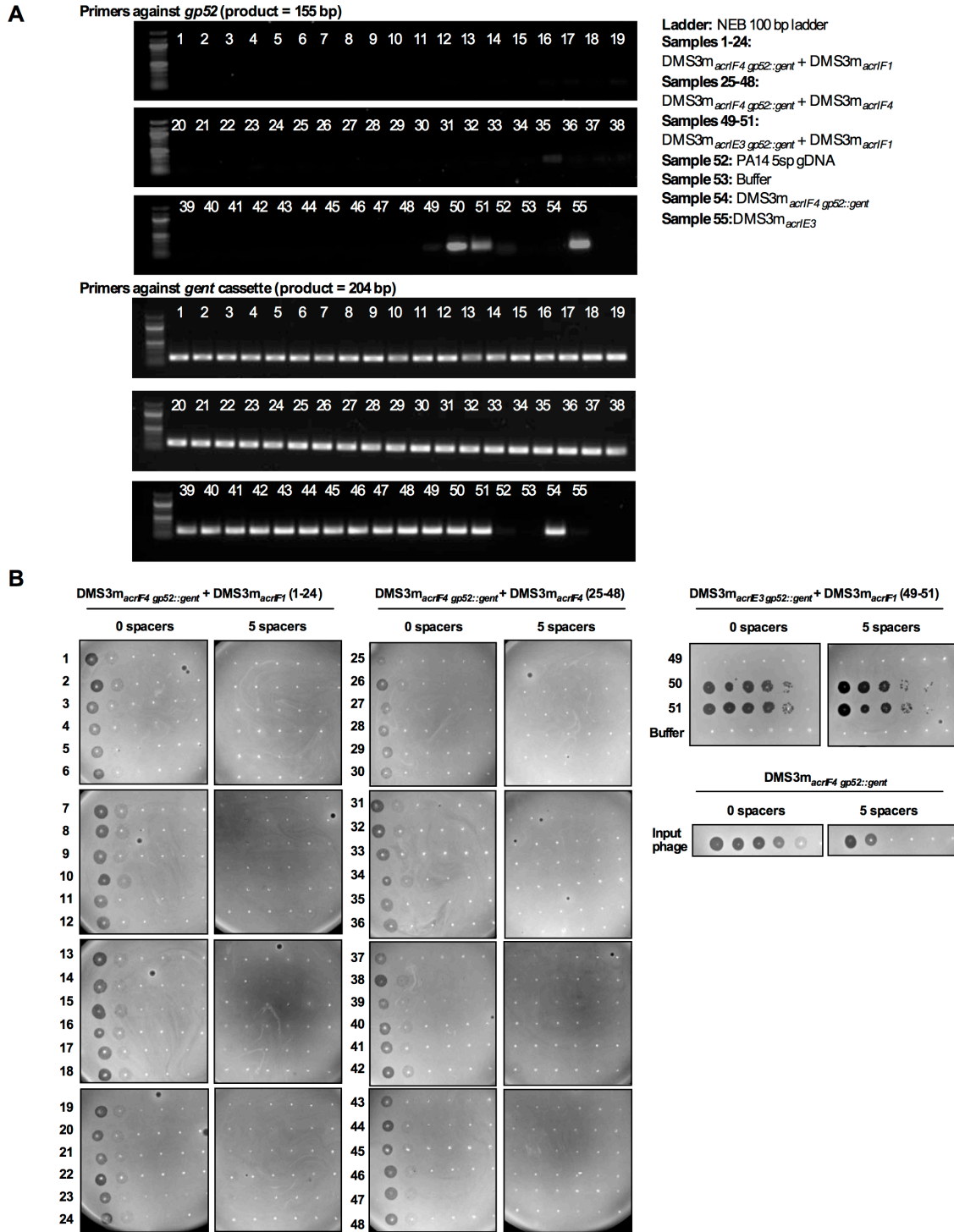


Figure 2.7. Prophage content of immunosuppressed lysogens, Related to Figure 2.6E-F
 (A) PCR of genomic DNA harvested from overnight cultures of lysogens from Figure 3E (1-48) and 3F (49-51) amplified with primers targeting *gp52-DMS3m* (top) or the gentamicin resistance cassette used to replace *gp52* in DMS3m_{acr gp52::gent} derivatives (bottom).

(B) 10-fold serial dilutions of supernatant harvested from overnight cultures of lysogens from Figure 3E (1-48) and 3F (49-51), spotted on a non-targeting (0sp) strain of PA14 and the 5 spacer (5sp) targeting strain of PA14. A faint clearing corresponds to induction of the gentamicin

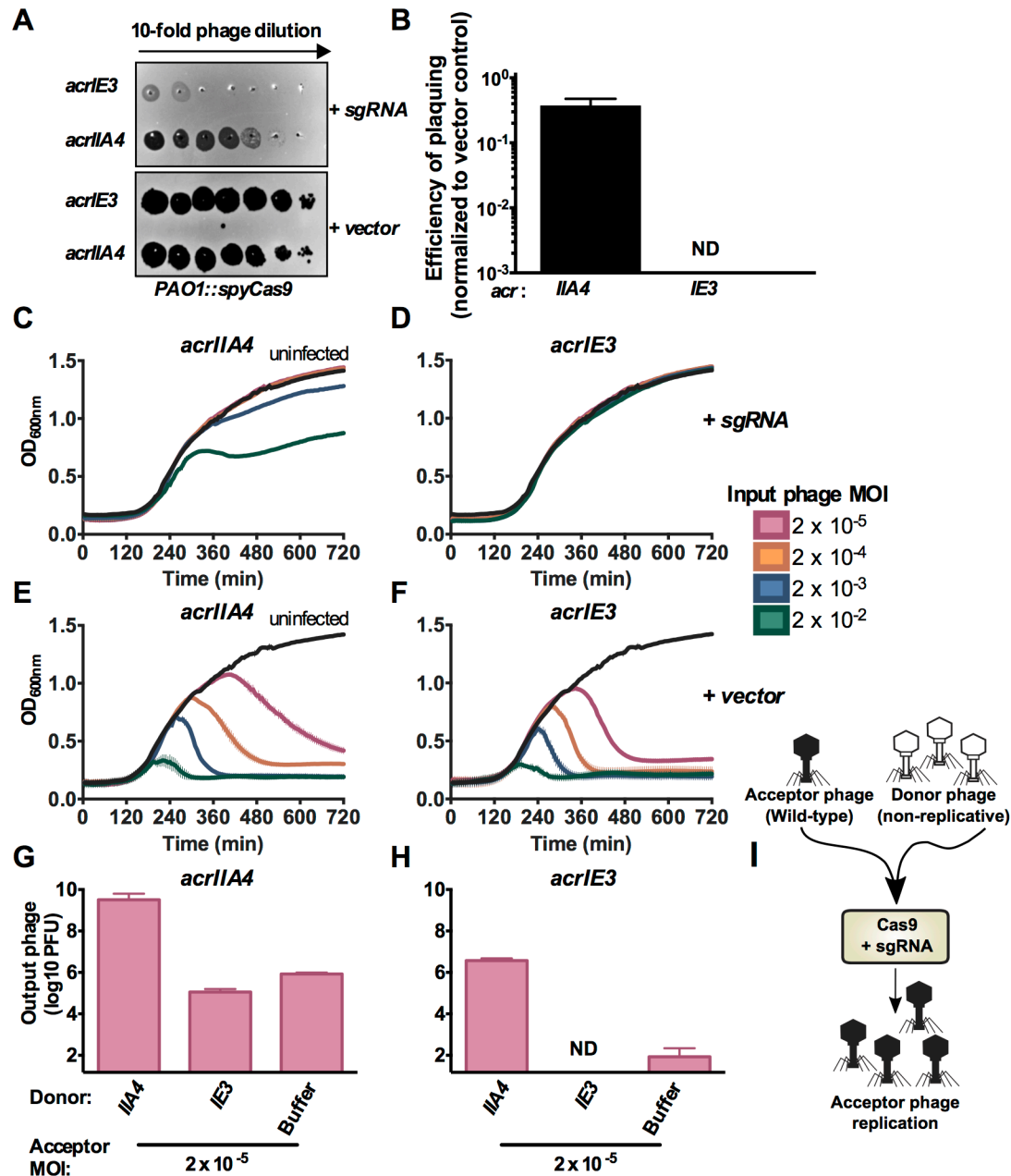


Figure 2.8. Cas9 anti-CRISPR AcrIIA4 requires cooperative infection to neutralize Type II-A CRISPR immunity

(A) 10-fold serial dilutions of DMS3m_{acrIE3} or DMS3m_{acrIIA4} plated on a lawn of *Pseudomonas aeruginosa* strain PAO1 expressing *Streptococcus pyogenes* Type II-A Cas9 (PAO1::SpyCas9) and single guide RNA (+ sgRNA) or non-targeting control (+ vector).

(B) Efficiency of plaquing of DMS3m_{acrIIA4} and DMS3m_{acrIE3} was calculated by normalizing PFU counts on a targeting strain of PAO1::SpyCas9 (+sgRNA) to PFU counts on a non-targeting strain of PAO1::SpyCas9 (+vector). Data are represented as the mean of 3 biological replicates +/- SD. ND, not detectable.

(C-F) 12-hr growth curves of *P. aeruginosa* strain PAO1::SpyCas9 expressing a targeting sgRNA (+sgRNA) infected with virulent DMS3m_{acrIIA4} (C) or DMS3m_{acrIE3} (D) at multiplicities of infection (MOI, rainbow colors) from 2×10^{-5} to 2×10^{-2} . Growth curves of uninfected cells are

shown in black. As a control, a non-targeting strain of PA01::SpyCas9 (+vector) was infected with DMS3m_{acrIIA4} (E) or DMS3m_{acrIE3} (F) under the same conditions. OD_{600nm} values are represented as the mean of 3 biological replicates ± SD (vertical lines).

(G-H) Replication of virulent DMS3m_{acrIIA4} (G) or DMS3m_{acrIE3} (H) (acceptor phage) in the presence of 10⁷ PFU (MOI 2) hybrid phage (donor) in PA01::SpyCas9 + sgRNA expressing the JBD30 C repressor. Phages were harvested after 24 hr and DMS3m_{acr} phage PFUs quantified on PA01::SpyCas9 + vector expressing the JBD30 C repressor. Phage output is represented as the mean of 3 biological replicates ± SD. ND, not detectable.

(I) Schematic of the experimental design in G-H, where a high MOI of non-replicative “donor” phages is used to rescue a low MOI infection of wild-type “acceptor” phages.

Table 2.1. DMS3m_{acr} phage genotypes

Phage	acr gene source	Acr Accession
DMS3m _{acrIE3} (parent phage, no manipulation)	DMS3m (gp30)	WP_003723290.1
DMS3m _{acrIF1}	JBD30 (gp35)	YP_007392342.1
DMS3m _{acrIF2}	MP29 (gp29)	YP_002332454.1
DMS3m _{acrIF3}	JBD88a (gp33)	YP_007392440.1
DMS3m _{acrIF4}	JBD26 (gp37)	WP_016068584.1
DMS3m _{acrIF5}	JBD5 (gp36)	YP_007392740.1
DMS3m _{acrIF6}	<i>Pseudomonas aeruginosa</i> strain PSE05 (prophage)	WP_043884810.1
DMS3m _{acrIF7}	LPB1 (gp29)	YP_009146150.1
DMS3m _{acrIIA4}	<i>Listeria monocytogenes</i> J0161 (prophage)	WP_003723290.1

Table 2.2. Plasmids used in this study

Plasmid	Backbone	Purpose
pAB02	pHERD30T	crRNA overexpression: CRISPR2 spacer 17
pAB03	pHERD30T	crRNA overexpression: CRISPR2 spacer 20
pAB58	pHERD30T	generating DMS3m _{acrIF2}
pAB59	pHERD30T	generating DMS3m _{acrIF3}
pAB21	pHERD20T	generating DMS3m _{acrIF4}
pJZ69	pHERD30T	generating DMS3m _{acrIF5}
pJZ70	pHERD30T	generating DMS3m _{acrIF6}
pAB24	pHERD20T	generating DMS3m _{acrIF7}
pJZ69	pHERD30T	generating DMS3m _{acrIIA4}
pAB45	pHERD20T	generating DMS3m _{acr gp52::gent}
pAB77,78	pHERD20T and 30T	C repressor (JBD30) overexpression
pAB79,80	pHERD20T and 30T	C repressor (DMS3) overexpression
pBAO72	pMMB67HE	sgRNA targeting DMS3m
pBAO95	pUC18T-mini-Tn7T-Gm	Insertion of SpyCas9 into the PAO1 tn7 site

Chapter 3: Discovery of widespread Type I and Type V CRISPR-Cas inhibitors

ABSTRACT

Bacterial CRISPR-Cas systems protect their host from bacteriophages and other mobile genetic elements. Mobile elements, in turn, encode various anti-CRISPR (Acr) proteins to inhibit the immune function of CRISPR-Cas. To date, Acr proteins have been discovered for type I (subtypes I-D, I-E, and I-F) and type II (II-A and II-C) but not other CRISPR systems. Here we report the discovery of 12 *acr* genes, including inhibitors of type V-A and I-C CRISPR-Cas systems. The *acr* genes reported here provide useful biotechnological tools and mark the discovery of *acr* loci in many bacteria and phages.

MAIN

The discovery of bacterial CRISPR-Cas systems that prevent infection by bacterial viruses (phages) has opened a paradigm for bacterial immunity while yielding exciting tools for targeted genome editing. CRISPR systems destroy phage genomes, and in turn, phages express anti-CRISPR (Acr) proteins that directly inhibit Cas effectors^{1,2}. Six distinct types (I-VI) of CRISPR systems are spread widely across the bacterial world³, but Acr proteins have only been discovered for type I and II CRISPR systems^{1,3-6}. Given the prevalence and diversity of CRISPR systems, we predict that Acr proteins against other types await discovery.

Anti-CRISPR proteins do not have conserved sequences or structures and only share their relatively small size, making *de novo* prediction of *acr* function challenging⁶. However, *acr* genes often cluster together with other *acr* genes or are adjacent to highly conserved anti-CRISPR associated genes (*aca* genes, **Table 3.1**) in “*acr* loci”^{7,8}. In this work, we sought to identify *acr* genes in bacteria and phages that are not homologous to previously identified *acr* or *aca* genes.

Acr proteins were first discovered in *Pseudomonas aeruginosa*, inhibiting type I-F and I-E CRISPR systems^{1,9}. *P. aeruginosa* strains also encode a third CRISPR subtype (type I-C), which lacks known inhibitors¹⁰. We engineered *P. aeruginosa* to target phage JBD30 with type I-C CRISPR-Cas and used it in parallel with existing type I-E (strain SMC4386) and I-F (strain PA14) CRISPR strains to screen for additional *acr* candidates.

Homologs of *aca1* were searched for in *Pseudomonas* genomes, and 7 gene families not previously tested for anti-CRISPR function were identified upstream of *aca1* (**Fig. 3.1A**). Three genes inhibited the type I-E CRISPR-Cas system (*acrIE5-7*), one inhibited type I-F (*acrIF11*), restoring the plaquing of a targeted phage, and two genes had no inhibitory activity (*orf1*, *orf2*) (**Fig. 3.1B**, **Table 3.2**). Another gene exhibited dual I-E and I-F inhibition, and domain analysis revealed a chimera of previously identified *acrIE4* and *acrIF7* (*acrIE4-F7*). No type I-C inhibitors were identified. The type I-F inhibitor *acrIF11* was commonly represented in both the *P. aeruginosa* mobilome and in over 50 species of diverse Proteobacteria (**Fig. 3.2**, **Table 3.3**). *acrIF11* is often associated with genes encoding DNA-binding motifs, which we have designated *aca4-7* (**Table 3.1**, **Table 3.4**). To confirm that these *aca* genes can be used to facilitate *acr* discovery, we used *aca4* to discover an additional *Pseudomonas* anti-CRISPR, *acrIF12* (**Fig. 3.1A-B**).

Given the widespread nature of *acrIF11*, we next used it to discover Acr proteins against CRISPR systems where they have not yet been found: type I-C, a minimal Class 1 system and type V-A CRISPR-Cas12a (Cpf1), a Class 2 single effector system that has high efficiency in genome editing¹¹⁻¹³. To find AcrIC and AcrVA proteins, we first searched for genomes encoding CRISPR spacers that match a target protospacer elsewhere in the same genome (**Fig. 3.3A**). The tolerance of this “self-targeting” in viable bacteria indicates potential inhibition of the CRISPR system⁴, since genome cleavage would result in bacterial death.

The Gram negative bovine pathogen *Moraxella bovoculi*^{14,15} is a Cas12a-containing organism¹¹ where four of the seven genomes feature Type V-A self-targeting, and one strain

(58069) also features self-targeting by type I-C. Although no previously described *acr* or *aca* genes were present in this strain, an *acrIF11* homolog was found in phages infecting the human pathogen *M. catarrhalis*¹⁶, a close relative of *M. bovoculi*. Genes adjacent to *acrIF11* in *M. catarrhalis* had homologs in the self-targeting *M. bovoculi* strains (**Fig. 3.3B**), and together these genes were selected as candidate *acr* genes. Genes were screened against type I-C and I-F systems introduced above, as well a heterologous *M. bovoculi* Type V-A system that was transplanted into a strain of *P. aeruginosa*. Using this panel of strains, we successfully identified Type I-C (AcrIC1), Type I-F (AcrIF13-14), and Type V-A inhibitors (AcrVA1-3) (**Fig. 3.3C-E**). This discovery of Type V-A inhibitors in *M. bovoculi* was in good agreement with the independent discovery of AcrVA1 reported in a companion paper¹⁷

AcrVA1 inhibits Cas12a in bacteria, and it also potently inhibits Cas12a in human cells¹⁸. This strong inhibitory activity may stem from the unique mechanism of action of AcrVA1. AcrVA1 is a multiturnover enzyme which cleaves the guide RNA, leaving Cas12a unable to find target DNA^{19,20}. This makes AcrVA1 the first enzymatic anti-CRISPR to be described, though we hypothesize many mobile genetic elements may need to employ multi-turnover Acrs to protect themselves from CRISPR-Cas immunity. Previous work has shown that stoichiometric inhibitors require phage cooperation to fully neutralize CRISPR-Cas immunity²¹, a strategy that is highly dependent on local phage concentration and multiplicity of infection. This concentration dependent strategy is likely unsuitable for classes of mobile genetic elements have low reproductive rates, such as phages with small burst sizes. MGEs that cannot multiply infect cells, such as plasmids, would also be at a disadvantage. A prediction then is that plasmids and phages with small burst sizes may be especially likely to require hyper-potent mechanisms of immune neutralization. Future work should explore the autonomy that hyper-potent Acrs could provide to MGEs, and diverse MGEs should continue to be mined for new Acrs as they will likely yield new and exciting mechanisms of CRISPR-Cas neutralization.

Here we report the discovery of a broadly distributed type I-F Acr protein (AcrIF11) that served as a marker for *acr* loci and led to the identification of type I-C and V-A CRISPR inhibitors. One of these *acrVA* genes (*acrVA1*) is a potent RNase^{19,20} that inhibits Cas12a in bacteria and human cells¹⁸. The strategy described herein enabled the identification of many widespread anti-CRISPR proteins, which may prove useful in future anti-CRISPR discovery.

REFERENCES

1. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
2. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
3. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67–78 (2017).
4. Rauch, B. J. *et al.* Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150–158.e10 (2017).
5. Pawluk, A. *et al.* Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* **167**, 1829–1838.e9 (2016).
6. Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs. *Annual Review of Virology* **4**, 37–59 (2017).
7. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: discovery, mechanism and function. *Nat Rev Micro* **16**, 12–17 (2018).
8. Pawluk, A. *et al.* Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology* **1**, 1–6 (2016).
9. Pawluk, A., Bondy-Denomy, J., Cheung, V. H. W., Maxwell, K. L. & Davidson, A. R. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* **5**, e00896–e00896–14 (2014).
10. van Belkum, A. *et al.* Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant *Pseudomonas aeruginosa*. *mBio* **6**, e01796–15 (2015).
11. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771 (2015).

12. Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A. & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517–521 (2016).
13. Kleinstiver, B. P. *et al.* Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nature Biotechnology* **34**, 869–874 (2016).
14. Angelos, J. A., Spinks, P. Q., Ball, L. M. & George, L. W. *Moraxella bovoculi* sp nov, isolated from calves with infectious bovine keratoconjunctivitis. *Int. J. Syst. Evol. Microbiol.* **57**, 789–795 (2007).
15. Dickey, A. M. *et al.* Large genomic differences between *Moraxella bovoculi* isolates acquired from the eyes of cattle with infectious bovine keratoconjunctivitis versus the deep nasopharynx of asymptomatic cattle. *Vet. Res.* **47**, 31 (2016).
16. Ariff, A. *et al.* Novel *Moraxella catarrhalis* prophages display hyperconserved non-structural genes despite their genomic diversity. *BMC Genomics* **16**, 860 (2015).
17. Watters, K.E. *et al.* Systematic discovery of natural CRISPR-Cas12a inhibitors. *Science* **362**, 236–239 (2018).
18. Marino, N.D. *et al.* Discovery of widespread type I and type V CRISPR-Cas inhibitors. *Science* **362**, 240–242 (2018).
19. Zhang, H. *et al.* Structural basis for the Inhibition of CRISPR-Cas12a by Anti-CRISPR Proteins. *Cell Host & Microbe* **25**, 815–826 (2019).
20. Knott, G.J. *et al.* Broad-spectrum enzymatic inhibition of CRISPR-Cas12a. *Nat. Struct. Mol. Bio.* **4**, 315–321 (2019).
21. Borges, A.L. *et al.* Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. *Cell* **174**, 917–925 (2018)

Materials and Methods

Bacterial strains and growth conditions

Pseudomonas aeruginosa strains UCBPP-PA14 (PA14), SMC4386, and PAO1 were used in this study. The strains were grown at 37 °C in lysogeny broth (LB) agar or liquid medium, which was supplemented with 50 µg ml⁻¹ gentamicin, 30 µg ml⁻¹ tetracycline, or 250 µg ml⁻¹ carbenicillin as needed to retain plasmids or other selectable markers.

Phage isolation

Phage lysates were generated by mixing 10 µl phage lysate with 150 µl overnight culture of *P. aeruginosa* and pre-adsorbing for 15 min at 37 °C. The resulting mixture was then added to molten 0.7% top agar and plated on 1% LB agar overnight at 30 °C or 37 °C. The phage plaques were harvested in SM buffer, centrifuged to pellet bacteria, treated with chloroform, and stored at 4 °C.

Bacterial transformations

Transformations of *P. aeruginosa* strains were performed using standard electroporation protocols. Briefly, one mL of overnight culture was washed twice in 300 mM sucrose and concentrated tenfold. The resulting competent cells were transformed with 20 – 200 ng plasmid, incubated in antibiotic-free LB for 1 hr at 37 °C, plated on LB agar with selective media, and grown overnight at 37 °C. Bacterial transformations for cloning were performed using *E. coli* DH5α (NEB) and *E. coli* Stellar competent cells (Takara) according to the manufacturer's instructions.

Discovery of *acr* genes using *aca1*

All bacterial genome sequences used in this study were downloaded from NCBI. BLASTp was used to search the nonredundant protein database for Aca1 homologs (seed Aca1 accession: YP_007392343.1, e value < 0.005) in *Pseudomonas sp.* (taxid: 286) Individual genomes encoding an Aca1 homolog were then manually surveyed for *aca1* associated genes. This approach was extended to discover the Aca4 (WP_034011523.1) associated anti-CRISPR AcrIF12. tBLASTn searches to identify orthologs of VA2 in self-targeting *Moraxella bovoculi* strains were performed using the protein sequence in *Moraxella catarrhalis* BC8 strain (EGE18855.1) as the query and *Moraxella bovoculi* genome accessions as the subject (accessions: 58069 genome, CP011374.1; 58069 plasmid, CP011375.1; 22581, CP011376.1; 33362, CP011379.1; 28389, CP011378.1). Other searches for orthologs in *Moraxella sp.* were performed using BLASTp.

Discovery of anti-CRISPR associated (*aca*) gene families

Genomes with homologs of AcrIF11 were manually examined for novel anti-CRISPR associated (*aca*) genes. A gene was designated as an *aca* if it fit the following criteria: I) directly downstream of an AcrIF11 homolog in the same orientation, II) a non-identical homolog of this gene exists in the same orientation relative to a non-identical homolog of AcrIF11, and III) predicted in high confidence to contain a DNA-binding domain based on structural prediction using HHPred (probability >90%, E < 0.0005). Genes that fit these three criteria were then grouped into sequence families, requiring that a given gene have >40% sequence identity to at least one member of the family for family membership.

Type I-C CRISPR-Cas expression in *Pseudomonas aeruginosa*

Reconstitution of the Type I-C system from a *P. aeruginosa* isolate in the Bondy-Denomy lab into PAO1 was achieved by amplifying the four effector cas genes (*cas3-5-8-7*) from genomic

DNA by PCR and cloning the resulting fragment into the integrative, IPTG-inducible pUC18T-mini-Tn7T-LAC plasmid to generate the pJW31 vector. This plasmid was then electroporated into PAO1 and chromosomal integration was selected for using 50 $\mu\text{g ml}^{-1}$ gentamicin. After chromosomal integration of the insert was confirmed, the gentamicin selectable marker was removed using flippase-mediated excision at the flippase recognition target (FRT) sites of the construct; the resulting strain was named LL76. CRISPR RNAs (crRNAs) consisting of a spacer that targets JBD30 phage (see **Table 3.5** for the sequence) and two flanking repeats were cloned into the mini-CTX2 (AF140577) vector, and the resulting vector was electroporated into LL76. Stable integration of the vector at the *attB* site was selected for using 30 $\mu\text{g ml}^{-1}$ tetracycline. Targeting was confirmed in the resulting strain (LL77) using phage challenge assays, as described in the “bacteriophage plaque assays” section.

Type V-A CRISPR-Cas expression in *Pseudomonas aeruginosa*

Human codon-optimized MbCas12a (*Moraxella bovoculi* 237) was amplified from the pTE4495 plasmid (Addgene #80338) by PCR and cloned into pTN7C130, a mini-Tn7 vector that integrates into the attTn7 site of *P. aeruginosa*. The pTN7C130 vector expresses MbCas12a off the araBAD promoter upon arabinose induction and contains a gentamicin selectable marker. The resulting construct, pTN7C130-MbCas12a, was used to transform the PAO1 strain of *P. aeruginosa*, and stable integration of the vector was selected for using 50 $\mu\text{g ml}^{-1}$ gentamicin and confirmed by PCR. After integration, flippase was used to excise the gentamicin selectable marker from the flippase recognition target (FRT) sites of the construct.

CRISPR RNAs (crRNAs) for MbCas12a were generated by designing oligonucleotides with spacers that target gp23 and gp24 (see **Table 3.5** for sequences) in JBD30 phage flanked by two direct repeats of the MbCas12a crRNA. The flanking repeats consist only of the sequence retained after crRNA maturation. The oligos were annealed and phosphorylated using T4

polynucleotide kinase (PNK) and ligated into NcoI and HindIII sites of pHERD30T. A fragment of the resulting plasmid that includes the *araC* gene, pBAD promoter, and crRNA sequence was then amplified by PCR and cloned into the mini-CTX2 plasmid. The resulting constructs were then used to transform the PAO1 tn7::MbCas12a strain, and stable integration was selected for using 30 $\mu\text{g ml}^{-1}$ tetracycline. The parental strain encoding MbCas12a but no crRNA was used as the “no crRNA” control.

Cloning of candidate anti-CRISPR genes

All candidate genes were cloned into the pHERD30T shuttle vector, which replicates in both *E. coli* and *P. aeruginosa*. Novel genes found upstream of *aca1* in *Pseudomonas sp.* were synthesized as gBlocks (IDT) and cloned into the SacI/PstI site of pHERD30T, which has an arabinose-inducible promoter and gentamicin selectable marker. Candidate genes derived from *Moraxella bovoculi* strains were amplified from the genomic DNA of 58069 and 22581 by PCR, whereas genes derived from *Moraxella catarrhalis* were synthesized as gBlocks (IDT). These inserts were cloned using Gibson assembly into the NcoI and HindIII sites of pHERD30T. All plasmids were sequenced using primers outside of the multiple cloning site. All constructs are listed in **Table 3.5**.

Bacteriophage plaque assays

Plaque assays were performed using 1.5% LB agar plates and 0.7% LB top agar, both of which were supplemented with 10 mM MgSO₄. 150 μl overnight culture was resuspended in 3-4 ml molten top agar and plated on LB agar to create a bacterial lawn. Ten-fold serial dilutions of phage were then spotted onto the plate and incubated overnight at 30 °C. Agar plates and/or top agar were supplemented with 0.5 – 1mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and 0.1-0.3% arabinose for assays performed with the LL77 (I-C) strain and with 0.1-0.3% arabinose for assays performed with the SMC4386 (I-E), PA14 (I-F), and PAO1 tn7::MbCas12a (V-A)

strains. The PA14 Δ CRISPR1 Δ CRISPR2 (SMC5454) strain, which lacks its endogenous CRISPR arrays 1 and 2, was used as the “no crRNA” control for type I-F assays. PAO1 strains encoding MbCas12a but no crRNA were used as the “no crRNA” control for type V-A assays. For type I-C assays, the “uninduced” control was plated on agar lacking IPTG. Agar plates were supplemented with 50 $\mu\text{g ml}^{-1}$ gentamicin for pHERD30T retention, as specified in the text. Anti-CRISPR activity was assessed by measuring replication of the CRISPR-sensitive phages JBD30 (V-A, I-C), JBD8 (I-E) and DMS3m (I-F) on bacterial lawns relative to the vector control. JBD30, JBD8, and DMS3m are closely related phages, differing slightly at protospacer sequences. Plate images were obtained using Gel Doc EZ Gel Documentation System (BioRad) and Image Lab (BioRad) software.

Phylogenetic reconstructions

Homologs of AcrIF11 (accession: WP_038819808.1) were acquired through 3 iterations of psiBLASTp search the non-redundant protein database. Only hits with > 70% coverage and an E value < 0.0005 were included in the generation of the position specific scoring matrix (PSSM). A non-redundant set of high confidence homologs (> 70% coverage, E value < 0.0005) represented in unique species of bacteria were then aligned using NCBI COBALT using default settings and a phylogeny was generated in Cobalt using the fastest minimum evolution method employing a maximum sequence difference of 0.85 and Grishin distance to calculate the tree. The resulting phylogeny was then displayed as a phylogenetic tree using iTOL: Interactive Tree of Life.

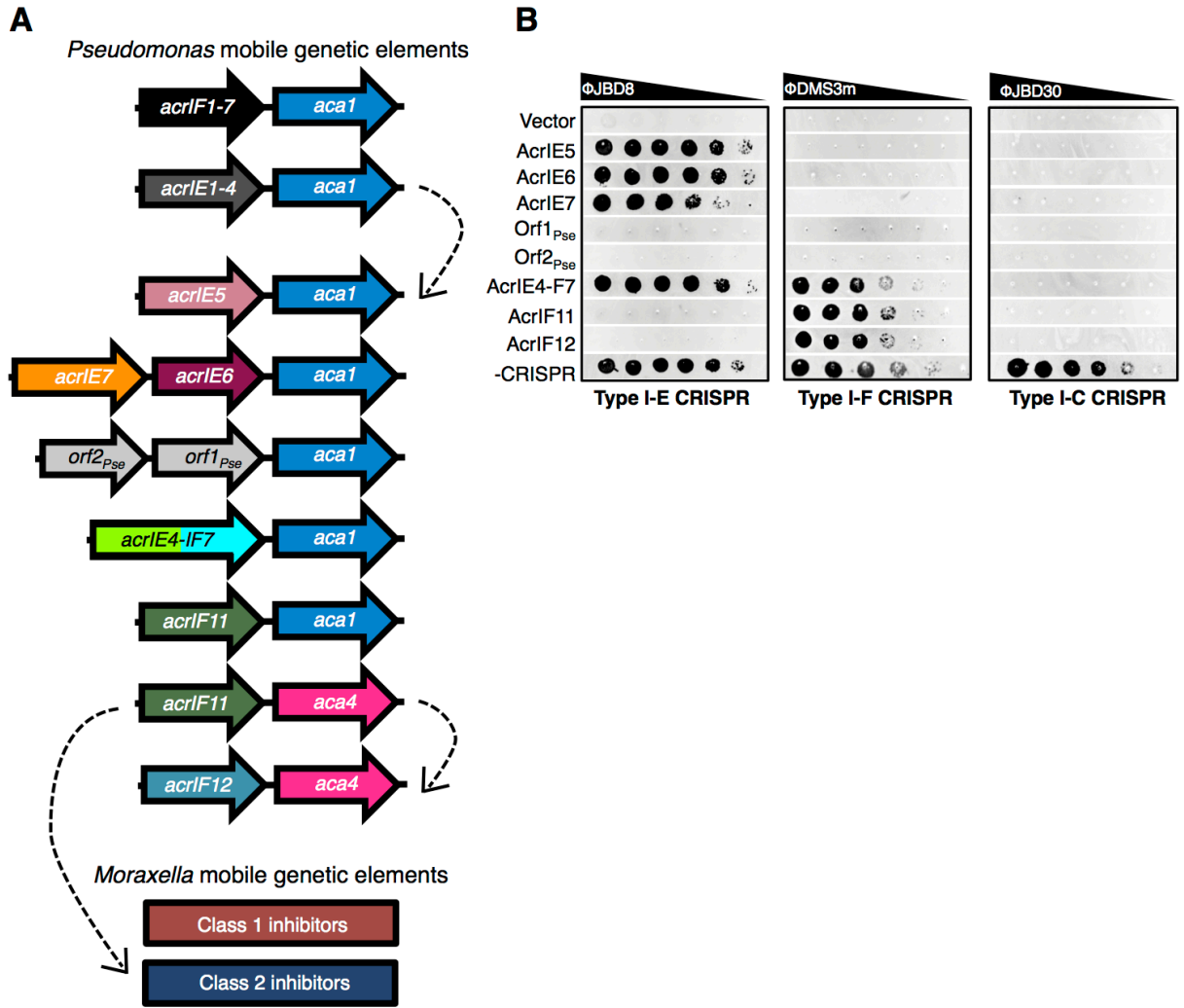


Figure 3.1. The discovery of a widespread Type I inhibitor

(A) The associations of novel Type I-E (IE5-7) and Type I-F (IF11-12) anti-CRISPRs with anti-CRISPR associated (*Aca1*, *Aca4*) genes in *Pseudomonas* sp. *acrIE4-7* is a chimera of two previously characterized Type I anti-CRISPRs (IE4 and IF7), and ORF1 and ORF2 did not manifest anti-CRISPR activity.

(B) 10-fold serial dilutions of a Type I-E, Type I-F, or Type I-C CRISPR-targeted phage (φ JBD8 and φ DMS3m, φ JBD30 respectively) plated on lawns of *Pseudomonas aeruginosa* expressing the indicated CRISPR-Cas systems. A restoration of phage plaquing relative to the vector control indicates inhibition of CRISPR-Cas immunity by the expression of the specified plasmid-borne anti-CRISPR. Phages were spotted on (-) CRISPR-Cas strains to measure phage replication in the complete absence of CRISPR-Cas immunity (bottom row).

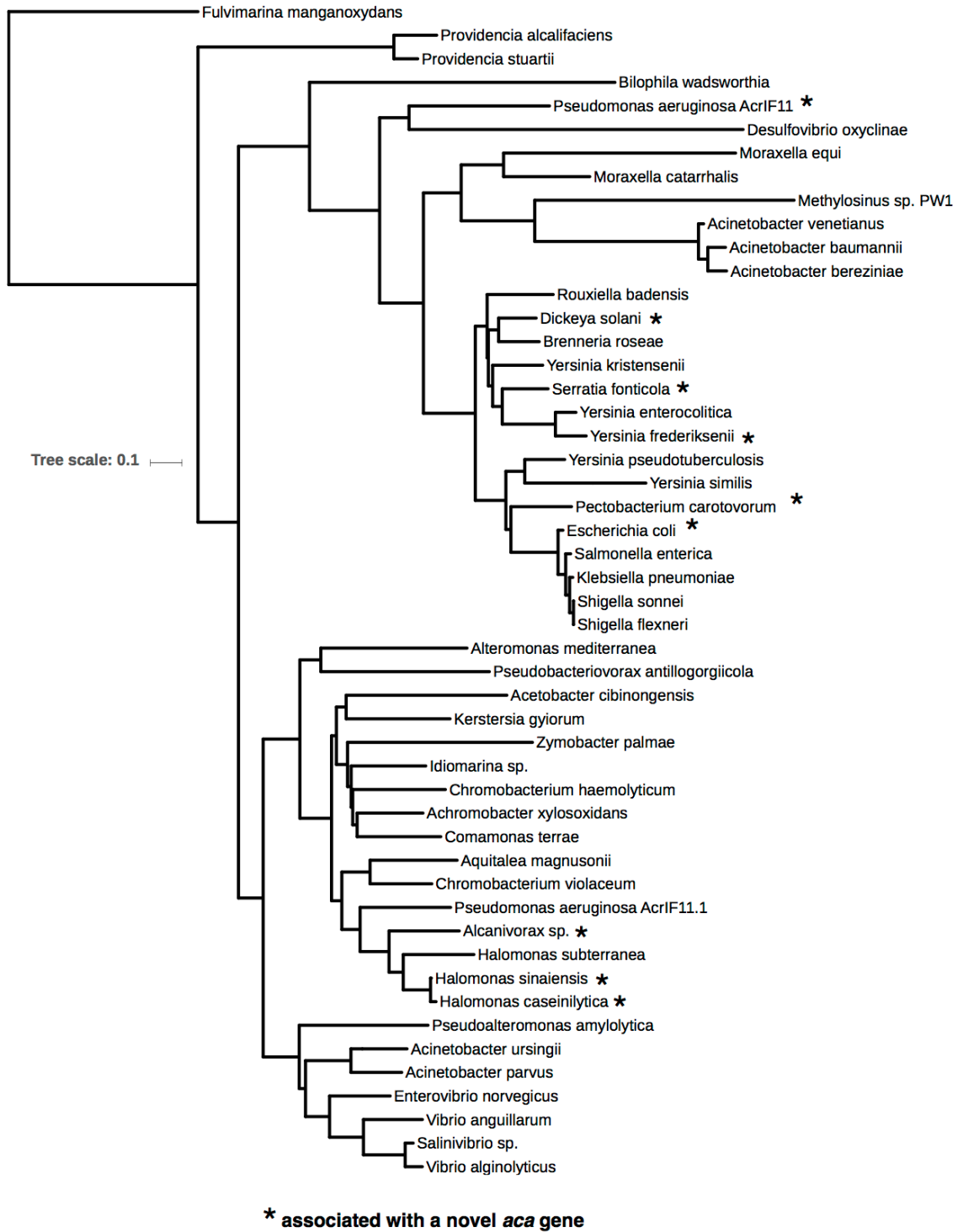


Figure 3.2. AcrIF11 phylogenetic tree

An unrooted phylogenetic tree of full-length homologs of AcrIF11 with all branches labeled with species names. Species in which AcrIF11 is associated with a novel *aca* gene (*aca4-7*) are marked with asterisks.

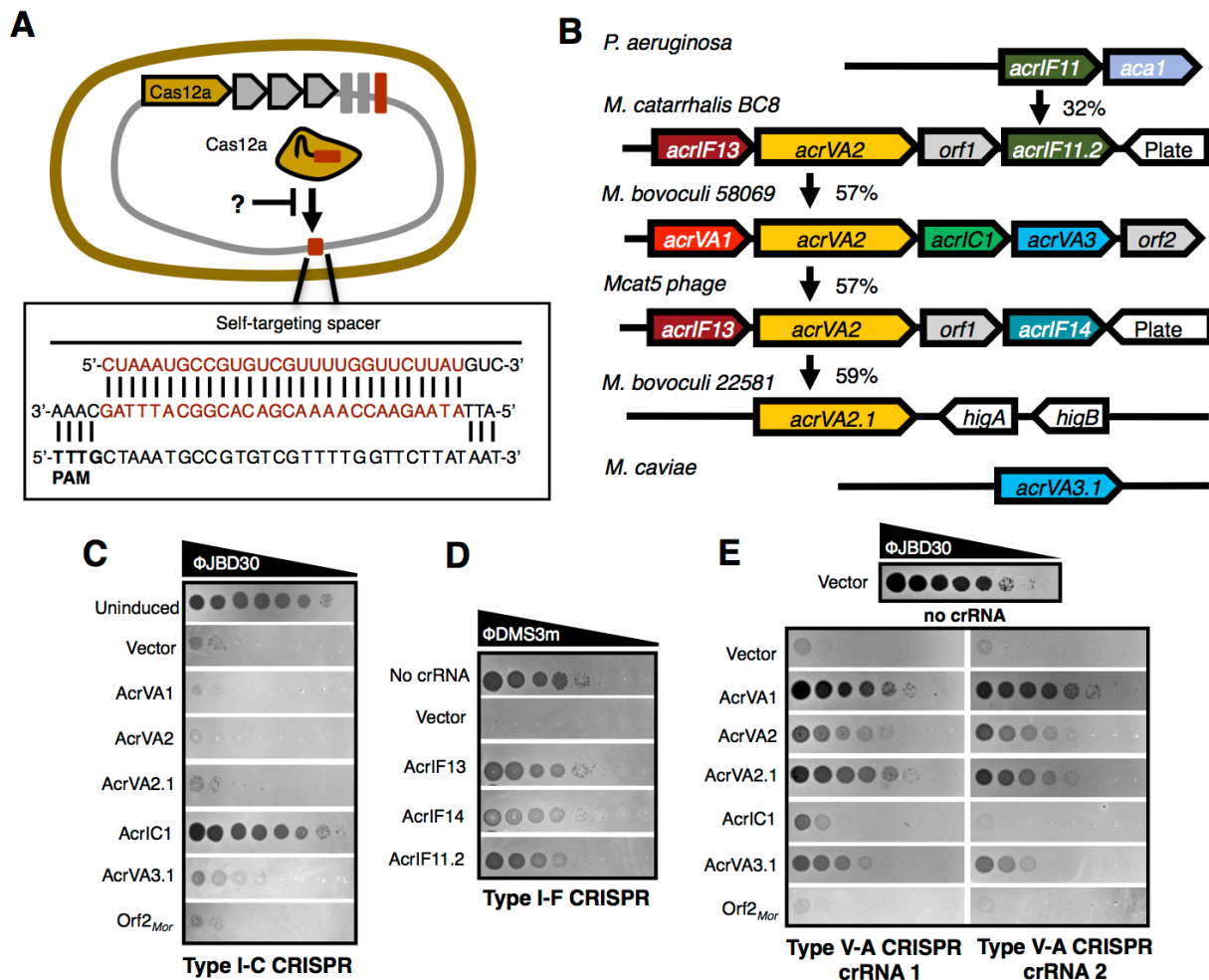


Figure 3.3. Guilt-by-association in self-targeting strains leads to Type V-A and Type I-C anti-CRISPR proteins

(A) *Moraxella bovoculi* exhibits intragenomic self-targeting, which is the co-occurrence of a spacer encoded by an apparently functional CRISPR-Cas12 system and its target protospacer within the same genome.

(B) Schematic showing the presence of AcrIF11 orthologs in anti-CRISPR loci within *Moraxella catarrhalis* and the use of guilt-by-association to unveil novel Type V-A and Type I-C inhibitors in *Moraxella bovoculi*.

(C) Tenfold serial dilutions of JBD30 phage were applied to bacterial lawns of *P. aeruginosa* PAO1 strain engineered to express MbCas12a, phage-targeting crRNA, and a candidate gene or vector control. “No crRNA” control was included to demonstrate loss of Cas12a targeting.

(D) Tenfold serial dilutions of JBD30 phage were applied to bacterial lawns of *P. aeruginosa* PAO1 strain engineered to express the Type I-C system upon induction with IPTG as well as phage-targeting crRNA and a candidate gene or vector control. Uninduced control lacks IPTG.

(E) Tenfold serial dilutions of DMS3m phage were applied to bacterial lawns of *P. aeruginosa* strain UCBPP-PA14 transformed with candidate gene or vector control. PA14ΔCRISPR-Cas strain was included as a control for loss of targeting.

Table 3.1. A table of previously discovered *aca* genes (*aca1-3*) and novel *aca* genes found in this study (*aca4-7*)

All *aca* proteins are predicted with high confidence to contain helix-turn-helix motifs as predicted by HHPred.

Name	HHPred : Protein motifs	HHPred Probability, e value	Discovery	Citation
Aca1	Helix-turn-helix, DNA binding	Probability = 98%, e value = 1.6E-6	Associated with Type I-F and Type I-E inhibitors	Bondy-Denomy et al, Nature 2013, Pawluk et al. mBio 2014, Pawluk et al. Nature Micro 2016
Aca2	Helix-turn-helix, DNA binding	Probability = 98%, e value = 5E-8	Associated with Type I-F and Type II-C inhibitors	Pawluk et al. Nature Micro 2016, Pawluk et al. Cell 2016
Aca3	Helix-turn-helix, DNA binding	Probability = 98%, e value = 4.2E-8	Associated with Type II-C inhibitors	Pawluk et al. Cell 2016
Aca4	Helix-turn-helix, DNA binding	Probability = 99%, e value = 3.1E-9	Associated with AcrIF11 and AcrIF12 in <i>Pseudomonas sp.</i>	This study
Aca5	Helix-turn-helix, DNA binding	Probability = 97%, e value = 5.6E-5	Associated with AcrIF11 in <i>Pectobacterium carotovorum</i> , <i>Yersinia frederiksenii</i> , <i>Escherichia coli</i> , <i>Serratia fonticola</i> , <i>Dickeya solani</i> , and <i>Enterobacter cloacae</i> complex members	This study
Aca6	Helix-turn-helix, DNA binding	Probability = 98%, e value = 7.8E-7	Associated with AcrIF11 in <i>Alcanivorax sp.</i>	This study
Aca7	Helix-turn-helix, DNA binding	Probability = 99%, e value = 7.2E-9	Associated with AcrIF11 in <i>Halomonas sp.</i>	This study

Table 3.2. Protein sequences and accession numbers of Acr and Aca proteins found in this study

Name	Accession	Protein Sequences
AcrIE4-F7	WP_064584002.1	MSTQYTYQQIAEDFRLWSEYVDTAGEMSKDEFNS LSTEDKVRLQVEAFGEEKSPKFSTKVTTKPDFDGF QFYIEAGRDFDGDAYTEAYGVAVPTNIAARIQAQA AELNAGEWLLVEHEA
AcrIE5	WP_074973300.1	MSNDRNGIINQIIDYTGDRDHAERIYEELRADDRI YFDDSVGLDRQGLLIREVDVLMVAVAEIE
AcrIE6	WP_087937214.1	MNNDTEVLEQQIKAFELLADELKDRLPTLEILSPMY TAVMVTYDLIGKQLASRRRAELIEILEEQYPGHAADL SIKNLCP
AcrIE7	WP_087937215.1	MIGSEKQVNWAKSIIKEVEAWEAIGVDVREVAAF LRSISDARVIIDNRNLIHFQSSGISYSLESSPLNSPIF LRRFSACSVGFEEIPTALQRIRSVYTAKLLEDE
AcrIF11	WP_038819808.1	MSMELFHGSYEEISEIRD SGVFGGLFGAHEKETAL SHGETLHRIISPLPLTDYALNYEIESAWEVALDVAG GDENVAEAIMAKACESDSNDGWELQRLRGVLAV RLGYTSVEMEDEHGTTWLCPLPGCTVEKI
AcrIF11.1	WP_033936089.1	MEIFHTSPVEITTINTQGRFGFLCFAADEYVMTA GDHVTYRIKVDSDIIMAGSIFYHERAADLSGLVE RVMQLTGCEDETAELISQRIDVFNLDDIDASDAA ELSWEIQAITAKAAKTLGFRGVSMQDEQGT CYMID MLGHDAELVRVK
AcrIF11.2	EGE18857.1	MTTLYHGSHENTAPVIKIGFAAFLPADNVFDGIFAN GDKNVARSHGDFIYAYEVDSIATNDDLDCDEAIQII AKELYIDEETAAPIAEAVAYEESLAEFEEHIMPRSC GDCADFGWEMQRLRGVIARKLGFDAVECVDEHG VSHLIVNANIRGSIA
AcrIF12	ABR13388.1	MAYEKTWHRDYAAESLKRAETSRWTQDANLEWT QLALECAQVVHLARQVGEELGNEKIIGIADTVLSTI EAHSQATYRRPCYKRITTAQTHLLAVTLLERFGSA RRVANAVWQLTDDEIDQAKA
AcrIF13	EGE18854.1	MKLLNIKINEFAVTANTEAGDELYLQLPHTPDSQH SINHEPLDDDDFVKEVQEICDEYFGKGDRTLARLS YAGGQAYDSYTEEDGVYTTNTGDQFVEHSYADY YNVEVYCKADLV
AcrIF14	AKI27193.1	MKKIEMIEISQNRQNLTAFLHISEIKAINAKLADGVD VDKKSFDEICSI VLEQYQAKQISNKQASEIFETLAK ANKSFKIEKFRCSHGYN EIKYSPDHEAYLFYCKG GQGQLNKLIAENGRFM

Name	Accession	Protein Sequences
Orf1(Pse)	SDJ61947.1	MGVVVLIIRLKRWSLHLERKLGEAGKAGIWEFH RSESSYTTDGRRTTFRNAALRPAEPKEGQTVEVFC SDSREPEEQWRAVGEGVARYE
Orf2(Pse)	WP_084336955.1	MLSVLFFWLYFYALFFIRFASSNKRARGRMQR ALVSIALEWGMRELMSRSFTTRIDHLQEVSRLGR GVARLRLGHSGRNLMPLILERRDGTGLTLKLDPKA DPDEALRQLARGGIHVRVYSKYGERMRVVVDAP QAISILRDELVDRE
Aca1	YP_007392343	MRFPGVKTPDASNHDPDPRLRGLLKKAGISQRR AAELLGLSDRVMRYLSEDIKEGYRPAPYTVQFAL ECLANDPPSA
Aca2	WP_019933869.1	MTHYELQALRKLMLLEVSEAAREIGDVSPRSWQY WESGRSPVPDDVANQIRNLDMRYQLLELRTEQI EKAGKPIQLNFYRTLDDYEAVTGKRDVVSWRLTQ AVAATLFAEGDVTLVEQGGLTLE
Aca3	WP_049360086.1	MKKFEAPEIGYTPANLKALRKQFGLTQAQVAEITG TKTGYSVRRWEAAIDAKNRADMPLVKWQKLLDSL K
Aca4	ABR13385.1	MTEEQFSALAEMLRLRGGPGEDAARLVLVNGLKP TDAARKTGITPQAVNKTLSRCRRGIELAKRVFT
Aca5	WP_039494319.1	MSLTEYIDKNFAGNKAARHMGVDAQAVNKWIK SEWVSTTDDNKIYLSSVRREIPPVA
Aca6	WP_035450933.1	MTAMKEWRARMGWSQRRAAQELGVTLPTYQSW EKGIRLSGDGSPIDPPLTALLAAAAREKGLPPIS
Aca7	WP_064702654.1	MIDARKHYDPNLAPELVRRALAVTGTQKELAERLD VSRTYLQLLGKQKSMYSYAVQVMLEQVIQDGET
AcrlC1	AKG19229.1	MNNLKKTAITHDGVFAYKNTETVIGSVGRNDIVMAI DATHGEFNDKNFIIYADTNGNPIYLGAYLDDNND AHIDLAVGACNEDDDFDEKEIHEMIAEQMELAKRY QELGDTVHGTTRLAFDDDDGYMTVRLDQQAYPDY RPENDDKHIMWRALALTATGKELEVFWLVEDYED EEVNSWDFDIADDWREL
Orf1(Mor)	EGE18856.1	MSKNKTPDYVLRANANYRKKHTTNKSLQLHNEKD ADIIQALQNETKSFNALMKDILRNHYNLNQNG

Name	Accession	Protein Sequences
Orf2(Mor)	AKG19231.1	MNNPKTPEYTRKAIRAYEKNLVRKSVTFDVRKDD DMELLMKIEQDGRFTAQIARTALLEHLQK
AcrVA1	AKG19227.1	MYEAKERYAKKKMQENTKIDTLTDEQHDALAQLC AFRHKFHNSKDSLFLSESAFSGEFSFEMQSDENS KLREVLPTIEWSFYDNSHIPDDSFREWFNFANYS ELSETIQEQGLELDLDDDETYELVYDELYTEAMGE YEELNQDIEKYLRRIDEEHGTQYCPTGFARLR
AcrVA2	AKG19228.1	MHHTIARMNAFNKAFANAKDCYKKMQAWHLLNK PKHAFFPMQNTPALDNGLAALYELRGGKEDAHILS ILSRLYLYGAWRNTLGIYQLDEEIIKDCKELPDDTP TSIFLNLPDWCVYVDISSAQIATFDDGVAKHIKGFW AIYDIVEMNGINHVDLDFVVDTDTDNVIYVPQPFIL SSGQSVAEVLDYGASLFDDDTSNLIKLLPYLLW LCVAEPDITYKGLPVSREELTRPKHSINKKTGAFVT PSEPFYQIGERLGSEVRRYQSIIDGEQKRNRPH KRPHIRRGHWGHWYWGQTGQAKEFRVRW QPAVFNNSGRVSS
AcrVA2.1	AKG12143.1	MHHTIARMNAFNKAFGNAKDCYKKMQAWHLNKN PKHIFSPQLNTLSLNEGLAALYELHGGKEDEHILSI LCCLYLYGTWRNTLGIYQLDEEIIKDCKELPDDTPT SIFLNLPDWCVYVDISSAKIATIDGGVAKHIKGFVAI YDNIEMHGVNHDVLFNFIIDTDTDNVIYVPQSLISS EMSVAESLDYGLTLFGYDESNELVKGMLPYLLWL CVAEPDITHKGLPVSREELTKPKHGINKKTGAFVT PSEPFYQIGERLGGEVRRYQSLIDDEKNQNRH HTKRPHIRRGHWGHWYWGQTGQAKEFKVRWQPA VFNNSGV
AcrVA3	AKG19230.1	MVGKSKIDWQSIDWTKTNAQIAQECGRAYNTVCK MRGKLGKSHQGAKSPPKDKGISRPQPHLNRLEY QALATAKAKASPKAGRFETNTKAKTWTLSKSPDNK TYFTNLMHFVRTNPHLFDPPDDVWVRTKSNQVE WCRASSGLALLAKRKKAPLSWKGWRLISLTKDNK
AcrVA3.1	OOR90252.1	MIAHQKNRRADWESVDWTKHNDEIAQLLSRHPDS VAKMRTKFGAQGMARPKPRRKYKVTRKAVPPPH TQELATAAAKISPKSGRYETNVNAKRWLIISPSGQ RFEFNLQHFVRNHPELFAKADTVWKRQGGKRG TGGEYCNASNGLAQAARLNIGWKGWQAKIIG

Table 3.3 List of all accessions of AcrIF11 homologs, related to Figure 3.2

List of the accession numbers for AcrIF11 homologs represented in the AcrIF11 phylogenetic tree in Fig. 3.2. The species in which each listed accession is found is listed on the right.

AcrIF11 Accession	Species
WP_038819808.1	<i>Pseudomonas aeruginosa</i> (AcrIF11)
WP_102394900.1	<i>Enterovibrio norvegicus</i>
WP_033936089.1	<i>Pseudomonas aeruginosa</i> (AcrIF11.1)
WP_087698854.1	<i>Chromobacterium violaceum</i>
WP_049175110.1	<i>Acinetobacter ursingii</i>
WP_004681960.1	<i>Acinetobacter parvus</i>
WP_062681378.1	<i>Achromobacter xylosoxidans</i>
KTG25401.1	<i>Idiomarina</i> sp.
WP_059284897.1	<i>Aquitalea magnusonii</i>
WP_107732478.1	<i>Chromobacterium haemolyticum</i>
WP_071971444.1	<i>Alteromonas mediterranea</i>
WP_086652143.1	<i>Acetobacter cibirongensis</i>
OHU91773.1	<i>Pseudoalteromonas amylolytica</i>
WP_064700809.1	<i>Halomonas sinaiensis</i>
WP_064702655.1	<i>Halomonas caseinilytica</i>
WP_066478200.1	<i>Comamonas terrae</i>
WP_068370878.1	<i>Kerstersia gyiorum</i>
WP_057083778.1	<i>Dickeya solani</i>
WP_074032235.1	<i>Serratia fonticola</i>
WP_039494318.1	<i>Pectobacterium carotovorum</i>
WP_077457760.1	<i>Salinivibrio</i> sp. IB872
WP_064369479.1	<i>Vibrio alginolyticus</i>
WP_041946990.1	<i>Vibrio anguillarum</i>
WP_036292019.1	<i>Methylosinus</i> sp. PW1
WP_017725053.1	<i>Acinetobacter baumannii</i>
WP_061524032.1	<i>Acinetobacter venetianus</i>
WP_004824702.1	<i>Acinetobacter bereziniae</i>
WP_049556453.1	<i>Yersinia kristensenii</i>
WP_109055423.1	<i>Brenneria roseae</i>
WP_097468739.1	<i>Escherichia coli</i>
OZT63688.1	<i>Salmonella enterica</i>
PKT06451.1	<i>Klebsiella pneumoniae</i>
WP_084913096.1	<i>Rouxiiella badensis</i>
WP_050090803.1	<i>Yersinia pseudotuberculosis</i>
WP_050879812.1	<i>Yersinia enterocolitica</i>
WP_050296286.1	<i>Yersinia frederiksenii</i>
WP_079326564.1	<i>Moraxella equi</i>

AcrIF11 Accession	Species
WP_003671754.1	<i>Moraxella catarrhalis</i> (AcrIF11.2)
WP_026949101.1	<i>Alcanivorax</i> sp.
WP_092828131.1	<i>Halomonas subterranea</i>
WP_027705017.1	<i>Zymobacter palmae</i>
SMF80656.1	<i>Pseudobacteriovorax antillogorgiicola</i>
WP_016360505.1	<i>Bilophila wadsworthia</i>
SMC32303.1	<i>Fulvimarina manganoxydans</i>
WP_051420249.1	<i>Providencia alcalifaciens</i>
WP_060561196.1	<i>Providencia stuartii</i>
WP_004247747.1	<i>Proteus mirabilis</i>
WP_086368795.1	<i>Photobacterium damsela</i>
WP_078005047.1	<i>Izhakiella australiensis</i>
WP_018125160.1	<i>Desulfovibrio oxycliniae</i>
OYL21963.1	<i>Shigella sonnei</i>
PAY74230.1	<i>Shigella flexneri</i>
CFQ72446.1	<i>Yersinia similis</i>

Table 3.4. List of Aca accessions

Representative homologs of each Aca protein (Aca1-7) and its associated AcrIF11 homolog listed by accession number as well as the species of origin.

Species	Aca	AcrIF11 accession	Aca accession
<i>Pseudomonas aeruginosa</i>	Aca1	WP_038819808.1	WP_033971918.1
<i>Pseudomonas aeruginosa</i>	Aca4	WP_034011523.1	WP_079381596.1
<i>Pectobacterium carotovorum</i>	Aca5	WP_039494318.1	WP_039494319.1
<i>Yersinia frederiksenii</i>	Aca5	WP_050101208.1	WP_050101207.1
<i>Escherichia coli</i>	Aca5	WP_000765122.1	WP_012565004.1
<i>Serratia fonticola</i>	Aca5	WP_074032235.1	WP_074032234.1
<i>Dickeya solani</i>	Aca5	WP_057083778.1	WP_057083779.1
<i>Pectobacterium carotovorum</i>	Aca5	WP_039558031.1	WP_039558032.1
<i>Enterobacter cloacae complex</i>	Aca5	WP_045331704.1	WP_072050017.1
<i>Alcanivorax sp.</i>	Aca6	WP_026949101.1	WP_035450933.1
<i>Alcanivorax sp.</i>	Aca6	WP_063139756.1	WP_063139755.1
<i>Halomonas caseinilytica</i>	Aca7	WP_064702655.1	WP_064702654.1
<i>Halomonas sinaiensis</i>	Aca7	WP_064700809.1	WP_064700810.1

Table 3.5. Plasmids used for expression in *Pseudomonas aeruginosa*.

Plasmid ID	Plasmid use	Backbone	Description
NM100	Express acrVA1	pHERD30T	p30T-acrVA1
NM101	Express acrVA2	pHERD30T	p30T-acrVA2
NM102	Express acrIC1	pHERD30T	p30T-acrIC1
NM103	Express acrVA3	pHERD30T	p30T-acrVA3
NM104	Express ORF2 (Mor)	pHERD30T	p30T-ORF2 (Mor)
NM105	Express acrIF13	pHERD30T	p30T-acrIF13
NM106	Express acrIF14	pHERD30T	p30T-acrIF14
NM107	Express VA2.1	pHERD30T	p30T-acrVA2.1
NM108	Express VA3	pHERD30T	p30T-acrVA3
NM109	Express VA3.1	pHERD30T	p30T-acrVA3.1
NM110	Type V-A crRNA -JBD30 gp23	mini-CTX2	mini-CTX2-crRNA g23
NM111	Type V-A crRNA-JBD30 gp24	mini-CTX2	mini-CTX2-crRNA g24
JDB500	Express MbCas12a (237)	pTN7C130	pTN7-MbCas12a (VA)
JDB501	Type V-A crRNA - JBD30 gp23	pHERD30T	p30T-crRNA g23
JDB502	Type V-A crRNA-JBD30 gp24	pHERD30T	p30T-crRNA g24
JZ_83	Express acrIE5	pHERD30T	p30T-acrIE5
JZ_99	Express acrIE6	pHERD30T	p30T-acrIE6
JZ_100	Express acrIE7	pHERD30T	p30T-acrIE7
JZ_127	Express ORF1 (Pse)	pHERD30T	p30T- ORF1 (Pse)
JZ_297	Express ORF2 (Pse)	pHERD30T	p30T- ORF2 (Pse)
JZ_298	Express acrIE4-IF7	pHERD30T	p30T-acrIE4-IF7
JZ_299	Express acrIF11	pHERD30T	p30T-acrIF11
JZ_300	Express acrIF12	pHERD30T	p30T- acrIF12
JZ_303	Express F11.1	pHERD30T	p30T-acrIF11.1
JZ_309	Express F11.2	pHERD30T	p30T-acrIF11.2
pJW31	Express type I-C cas3-5-8-7 genes	pUC18T-mini-Tn7T-LAC	pUC18T-cas3-5-8-7
LL7724	Express type I-C crRNA against JBD30 gp24	mini-CTX2	mini-CTX2-crRNA g24 (IC)

Chapter 4: Bacterial alginate regulators and phage homologs repress CRISPR-Cas immunity

ABSTRACT

CRISPR-Cas systems are adaptive immune systems that protect bacteria from bacteriophage (phage) infection¹. To provide immunity, RNA-guided protein surveillance complexes recognize foreign nucleic acids, triggering their destruction by Cas nucleases². While the essential requirements for immune activity are well understood, the physiological cues that regulate CRISPR-Cas expression are not. Here, a forward genetic screen identifies a two-component system (KinB/AlgB), previously characterized in regulating *Pseudomonas aeruginosa* alginate biosynthesis^{3,4}, as a regulator of the expression and activity of the *P. aeruginosa* Type I-F CRISPR-Cas system. Downstream of KinB/AlgB, activators of alginate production AlgU (a σ^E orthologue) and AlgR, repress CRISPR-Cas activity during planktonic and surface-associated growth⁵. AmrZ, another alginate regulator⁶, is triggered to repress CRISPR-Cas immunity during surface-association. *Pseudomonas* phages and plasmids have taken advantage of this regulatory scheme, and carry hijacked homologs of AmrZ that repress CRISPR-Cas expression and activity. This suggests that while CRISPR-Cas regulation may be important to limit self-toxicity, endogenous repressive pathways represent a vulnerability for parasite manipulation.

MAIN

Type I CRISPR-Cas systems are comprised of a multi-subunit RNA-guided surveillance complex, a trans-acting nuclease (Cas3)^{2,7,8}, and proteins dedicated to spacer acquisition, Cas1 and Cas2⁹. *Pseudomonas aeruginosa* has become a powerful model organism for studying Type I CRISPR-Cas mechanisms¹⁰⁻¹⁵, functions¹⁶⁻¹⁹, evolution²⁰⁻²², and interactions with phages utilizing anti-CRISPR proteins²³⁻²⁶. The *P. aeruginosa* strain PA14 possesses a naturally active Type I-F CRISPR-Cas immune system, comprised of two CRISPR arrays, an operon encoding

surveillance complex subunits Csy1-4¹², and a separate operon encoding Cas1 and a Cas2-3 fusion protein. Quorum sensing has been shown to activate CRISPR-Cas expression in *Pseudomonas aeruginosa*²⁷, as well as other species of bacteria²⁸. However, little is known regarding the factors that temper CRISPR-Cas activity and mitigate the risk of acquiring and expressing a nucleolytic immune system.

To discover new CRISPR-Cas regulators in *P. aeruginosa*, we utilized a *P. aeruginosa* strain PA14 engineered to express *lacZ* in place of the *csy3* gene (*csy3::lacZ*)¹⁷. This strain was subjected to C9 *mariner* transposon mutagenesis and ~40,000 colonies screened on X-gal plates. Multiple independent insertions were identified within *lacZ* and upstream genes (*csy1* and *csy2*), and thirty mutants with transposon insertions outside of this region were isolated and mapped (**Table 4.1**). Four independent insertions were identified in a single gene, *kinB*, which resulted in decreased β -galactosidase production on solid plates (**Fig. 4.1a**) and ~30% less *csy3::lacZ* activity in liquid culture compared to the unmutagenized parent (**Fig. 4.1b**). We selected *kinB* (a sensor kinase/phosphatase) for follow-up study as it was the only gene with >1 independent transposon insertion and displayed the largest β -galactosidase activity change.

We measured the ability of *kinB::Tn* insertions to limit the survival and replication of phage when introduced into the wild-type (CRISPR active) strain. Phages used to assay activity are: DMS3_{*acrIE3*} which is an untargeted control phage, DMS3m_{*acrIE3*}¹⁸, which is fully targeted by the PA14 Type I-F CRISPR-Cas system, and phage DMS3m_{*acrIF4*}, which is partially targeted, by virtue of encoding a “weak” anti-CRISPR, *acrIF4*, that binds to the surveillance complex to inhibit CRISPR-Cas function^{23,25,29}. The *kinB::Tn* strains remained resistant to DMS3m_{*acrIE3*} infection, but DMS3m_{*acrIF4*} formed 10-fold more plaques relative to WT, demonstrating attenuated CRISPR-Cas activity (**Fig. 4.1c, Fig. 4.2a**). This defect was complemented by expression of *kinB* *in trans* (Extended Data Fig. 2c). Growth of control phage DMS3_{*acrIE3*} was not

impacted in the absence of *kinB* (**Fig. 4.1c**, **Fig. 4.2a**). Furthermore, two other phages that are partially targeted, JBD26 (naturally possessing *acrIF4*) and JBD25 (a phage with no Acr that is targeted by a weak spacer that provides incomplete immunity) also showed increased survival in the *kinB:Tn* strain (**Fig. 4.1d**) relative to WT PA14. Survival of a phage with a weak anti-CRISPR or one that is targeted by a less active spacer is therefore a sensitive barometer for perturbations in CRISPR-Cas levels. Together, these data confirm that in the absence of *kinB*, *csy* gene expression and phage targeting are decreased.

KinB is a sensor kinase/phosphatase in a two-component system with response regulator AlgB. The KinB/AlgB system has a large regulon within *P. aeruginosa*, and controls the biosynthesis of the extracellular polysaccharide alginate⁴. This pathway is well-studied due to the recurrent isolation of alginate-overproducing (mucoid) *P. aeruginosa* from the lungs of cystic fibrosis patients, where alginate plays an important role in the formation of antibiotic resistant biofilms during chronic infection. The absence of KinB function results in the accumulation of the phosphorylated form of the response regulator AlgB (P-AlgB), while the phosphorylation of AlgB has been attributed to unknown kinases^{30,31} (**Fig. 4.2b**). P-AlgB activates the periplasmic protease AlgW (a DegS homolog), which degrades MucA, liberating sigma factor AlgU^{3,32,33} (**Fig. 4.2b**). AlgU positively regulates many factors involved in alginate production, including AlgR, AlgD, AlgB, and AmrZ^{5,6,34}.

WT *kinB* or kinase inactive H385A *kinB* complemented an in-frame $\Delta kinB$ deletion, restoring CRISPR targeting of DMS3m_{*acrIF4*} (**Fig. 4.2c**). However, a P390S *kinB* mutant incapable of dephosphorylating the response regulator AlgB did not complement, and in fact decreased CRISPR-Cas activity further (**Fig. 4.2c**). A $\Delta kinB \Delta algB$ double mutant restored CRISPR-Cas targeting to levels two-fold above WT (**Fig. 4.3b**), confirming the role of this signaling pathway. A strain lacking *algB* ($\Delta algB$) or possessing a D59N mutant that cannot be phosphorylated also

elevated CRISPR-Cas activity two-fold, supporting the repressive role of P-AlgB (**Fig. 4.2d**). These data show that accumulation of high levels of P-AlgB (achieved in *kinB::Tn*, $\Delta kinB$, or *kinB P390S*) leads to CRISPR-Cas repression.

We next assayed anti-phage immunity in $\Delta algU$ and $\Delta algR$ backgrounds, revealing increased targeting of DMS3m_{acrIF4} but not control phage DMS3_{acrIE3} in both knockouts (**Fig. 4.2e**). Complementation restored CRISPR-Cas levels (**Fig. 4.2e**), demonstrating that AlgU and AlgR repress CRISPR-Cas immunity. Double knockouts of each gene combined with $\Delta kinB$ also demonstrated increased CRISPR-Cas immunity, consistent with these factors acting as repressors downstream of KinB (**Fig. 4.3a**). All changes in DMS3m_{acrIF4} phage replication and survival are CRISPR-dependent, as double knockouts (*kinB*, *algB*, *algU*, *algR* mutants combined with *csy3::lacZ*, a loss of function mutation) revealed plaquing equivalent to *csy3::lacZ* alone (**Fig. 4.3b**). β -galactosidase activity was measured in these strains during growth in liquid culture, revealing a peak in *csy* expression around 8 h, with repression of this operon during entry into stationary phase (**Fig. 4.2f**). As suggested by the phage targeting experiments, a marked increase in expression of the *csy* operon was noted for both $\Delta algR$ and $\Delta algU$ strains, with a decrease in *csy* expression for $\Delta kinB$.

Next, we performed RT-qPCR of the *cas3* and the *csy3* transcripts in the mutant strains. We measured the relative abundance of Cas3 and Csy complex protein by fusing a sfCherry tag to the endogenous *cas3* or *csy1* gene in the mutant backgrounds, and used fluorescence as a proxy for protein abundance. We found that *kinB* loss decreased expression of both the *cas3* and *csy* operon, resulting in lower *cas3* and *csy3* transcripts and Cas3-sfCherry and Csy1-sfCherry levels relative to WT (**Fig. 4.4a,b**). Conversely, we observed increased levels of *cas3* and *csy3* transcripts and Cas3-sfCherry and Csy1-sfcherry in the $\Delta algR$ and $\Delta algU$ mutants relative to WT (**Fig. 4.4a,b**). These data demonstrate that this pathway controls the levels of

both Cas3 and the Csy complex in the bacterial cell by transcriptionally controlling the *cas3* and *csy* operons.

As Cas3-sfCherry was expressed at low levels relative to Csy1-sfCherry, and is also known to be subject to post-translational control by Cas1¹⁴, we sought to dissect the relative contribution of nuclease versus surveillance complex dysregulation in driving the immune phenotypes of the KinB/AlgB pathway mutants. To specifically measure the anti-phage activity of the Csy complex, we developed a Cas3-independent bioassay to read out the activity of the surveillance complex in the cell. Through the rational design of crRNAs to target an early phage promoter (P_{E1} , P_{E2}), we observed inhibition of phage survival in a *P. aeruginosa* strain with a nuclease dead Cas3 (dCas3), while an ORF-targeting crRNA (ORF1) was ineffective (**Fig. 4.4c**). This CRISPR-based transcriptional interference (CRISPRi) effect was remarkably strong enough to completely limit phage replication in the absence of Cas3 nuclease activity for crRNA P_{E2} . Phage inhibition via CRISPRi occurred when infecting with a phage that expressed the inhibitor of Cas3 recruitment, AcrIF3, but not an inhibitor that blocks Csy complex-phage DNA binding²⁵, AcrIF1 (**Fig. 4.4c**). We selected P_{E1} as a moderately-functional CRISPRi spacer and expressed it in KinB/AlgB pathway mutants. We observed decreased CRISPRi activity against phage DMS3m_{acrIF3} in the $\Delta kinB$ background, but increased CRISPRi in $\Delta algR$ and $\Delta algU$, (**Fig. 4.4d**, compare F3 and F1 phage). This demonstrates modulation of *csy* gene expression is sufficient to impact phage targeting in a Cas3-independent manner. We conclude that the KinB/AlgB pathway regulates Cas3 and Csy complex levels, and repression of Csy complex levels has a large impact on anti-phage immunity.

To identify downstream CRISPR-Cas regulators in the AlgU regulon³⁵, we focused on another factor involved in alginate production, the alginate and motility and regulator Z, *amrZ*^{6,36}. We generated a knockout of *amrZ* and observed a CRISPR-dependent increase in efficiency of

immunity (EOI) against phage DMS3m_{acrIF4} (**Fig. 4.5a, Fig. 4.3b**). This was complemented when *amrZ* was expressed *in trans* (**Fig. 4.5a**). A $\Delta kinB\Delta amrZ$ double knockout also showed increased CRISPR-Cas activity, consistent with its role as a repressor downstream of KinB (**Fig. 4.3a**). However, when we measured *cas3* and *csy3* transcript levels and Cas3-sfCherry and Csy1-sfCherry levels in $\Delta amrZ$, neither transcript or protein levels differed from WT (**Fig. 4.6a,b**). In considering these discrepant results, we realized that the anti-phage plaque assay is performed on solid plates whereas RNA quantification and sfCherry fluorescence measurements were conducted on liquid culture samples. To measure anti-phage activity of $\Delta amrZ$ in planktonic growth, we challenged WT and $\Delta amrZ$ with 10^6 PFU (MOI = 0.2) of virulent DMS3m_{acrIF4} in liquid culture. Both strains succumbed to phage infection with similar kinetics (**Fig. 4.5b**), and phage replication did not differ significantly between the two strains (**Fig. 4.5c**). Phage replication in the absence of CRISPR-Cas immunity also did not differ between the two strains (**Fig. 4.5c**). This demonstrates that under our conditions, AmrZ does not control CRISPR-Cas during planktonic growth.

To test the hypothesis that AmrZ is a surface-activated repressor of CRISPR-Cas, we measured the levels of Csy complex during surface association and planktonic growth in WT and $\Delta amrZ$ cells using an endogenous Csy1-sfCherry reporter over a period of 30 h. In WT cells, the levels of Csy complex were attenuated during surface-association relative to planktonic growth (~50% reduction of peak Csy1-cherry levels, **Fig. 4.5d**), but in the absence of AmrZ, Csy complex levels during surface association increased to levels comparable to those in planktonic growth (**Fig. 4.5d**). Deletion of *amrZ* did not impact Csy complex levels in liquid culture at any timepoint (**Fig. 4.5e**). To increase the levels of AmrZ during planktonic growth, we ectopically expressed AmrZ from a high copy plasmid, and measured the impact on the our transcriptional reporter *csy3::lacZ* and our translational reporters Csy1-sfCherry and Cas3-sfCherry. Here, high levels of AmrZ in liquid growth reduced β -galactosidase activity of the *csy3::lacZ* reporter (**Fig. 4.6c**)

and strongly limited expression of Csy1-sfCherry and Cas3-sfCherry (**Fig. 4.5f**). These results suggest that low AmrZ activity in planktonic growth underlies its surface-activated control of CRISPR-Cas. In contrast to AmrZ, overexpression of AlgU only moderately impacted Csy complex and Cas3 levels and AlgR did not impact the levels of either reporter when overexpressed (**Fig. 4.5f**).

We next considered if phages and other mobile genetic elements had evolved mechanisms to manipulate this CRISPR-Cas repressive pathway. Inspired by the discovery of a *Paraburkholderia* phage that carried a distant homolog of AmrZ³⁷, we searched the NCBI database for AmrZ homologs on *Pseudomonas* mobile genetic elements (MGE). Excitingly, we identified 14 diverse *Pseudomonas* mobile elements carrying AmrZ homologs (**Table 4.2**). These MGEs included obligately lytic and temperate Myophages, temperate Siphophages, and plasmids. AmrZ has been structurally characterized in complex with operator DNA³⁸, and these mobile AmrZ homologs showed perfect conservation of critical DNA-interacting residues in the ribbon-helix-helix domain, suggesting conserved binding specificity (**Fig. 4.7a, b**, red residues/arrowheads). To test if these mobilized AmrZ variants were capable of regulating CRISPR-Cas activity in *Pseudomonas aeruginosa*, we assayed the ability of 6 MGE-encoded AmrZ homologs to complement the $\Delta amrZ$ strain. Five out of six homologs complemented the $\Delta amrZ$ mutant to various degrees, indicating they were active in the PA14 transcriptional network and were *bona fide* CRISPR-Cas regulators (**Fig. 4.7c**). Next, each gene was expressed in WT cells, revealing 3 *P. aeruginosa* phage AmrZ homologs (AmrZ_{PaBG}, AmrZ_{phi3}, AmrZ_{JBD68}) inhibited Csy complex biogenesis (**Fig. 4.7d**).

We next studied the anti-CRISPR function of these mobilized AmrZ homologs in the context of the phage life cycle. By inserting the two most potent phage AmrZ homologs, *amrZ*_{phi3} and *amrZ*_{PaBG} into the anti-CRISPR locus of phage DMS3m, we compared the anti-CRISPR capacity

of these repressors relative to *bona-fide* Type I-F inhibitor AcrIF4 and the negative control inhibitor AcrIE3. While the AmrZ homologs provided no protection during lytic growth (likely because they cannot act on previously synthesized CRISPR-Cas complexes) (**Fig. 4.7e**), they were able to significantly reduce the expression and activity of the CRISPR-Cas complex during lysogeny (**Fig. 4.7f,g**). By lysogenizing a strain of PA14 with a catalytically dead Cas3 and an endogenously tagged copy of Csy1-sfCherry, we demonstrated that the presence of AmrZ_{Phi3} or AmrZ_{PaBG} reduced Csy complex levels more than 50% of an unlysogenized control, while AcrIF3 and AcrIF4 did not reduce Csy complex levels (**Fig. 4.7f**). To measure the activity of the Csy complex in these lysogens, we programmed the Csy complex to transcriptionally repress the *phzM* gene, which is responsible for the generation of the green pigment pyocyanin. De-repression of *phzM* expression can be quantified by measuring accumulation of the pyocyanin pigment in an overnight culture. We found that AmrZ_{Phi3} or AmrZ_{PaBG} de-repressed *phzM* to a similar extent as AcrIF4 (**Fig. 4.7g**), demonstrating anti-CRISPR activity for these hijacked CRISPR-Cas repressors.

Regulation of bacterial processes is highly variable across species, reflecting niche-specific adaptations. Here, a genetic screen reveals that the KinB/AlgB two-component system regulates CRISPR-Cas in *P. aeruginosa*. Removal of KinB or inactivation of its phosphatase activity leads to the accumulation of P-AlgB, activating CRISPR-Cas repressors AlgU, AlgR, and AmrZ. This pathway also drives alginate production, which is responsible for the formation of the characteristic mucoid biofilms of cystic fibrosis *P. aeruginosa* isolates^{3,39,40}. We show that P-AlgB (via *kinB* deletion), AlgU, and AlgR repress CRISPR-Cas activity during surface-association and planktonic growth, and AmrZ is triggered to further repress CRISPR-Cas during surface-association. Some *Pseudomonas* genetic parasites encode hijacked AmrZ homologs, which retain their ability to repress CRISPR-Cas expression and inhibit CRISPR-Cas biogenesis during lysogeny. Strikingly, we have identified multiply-lysogenized strains of *Pseudomonas*

aeruginosa with as many as 4 independent copies of AmrZ on mobile elements in addition to host AmrZ (**Table 4.3**). The evolutionary success of AmrZ in the *Pseudomonas* mobilome and core genome suggests a “guns for hire⁴¹” role for this gene in the arms race between bacteria and their parasites.

We and others observe CRISPR-Cas activation^{27,28} during exponential growth, where phage infection risk is high (i.e. metabolically active, well-mixed planktonic culture²¹). Surface-association lessens infection risk, as spatial structure limits phage dispersal and prevents a phage bloom from overtaking the entire bacterial population⁴². Though not measured here, spatial stratification and polysaccharide secretion in a mucoid biofilm likely also provide high levels of intrinsic phage resistance.

The observation that CRISPR-Cas expression and surface-association/biofilm formation are inversely regulated is supported by our analysis of a previously published PA14 RNAseq data set⁴³ and proteomic data set⁴⁴, which show activation of CRISPR-Cas expression in exponential phase, and repression during stationary phase and biofilm growth at 24 and 48 h (**Fig. 4.8a**). Cas proteins are still detected in stationary phase and biofilm growth, suggesting the cells retain some immunity after transcriptional shutdown (**Fig. 4.8b**). Furthermore, previous studies show that the *P. aeruginosa* genome is hyper-sensitive to CRISPR-induced DNA damage during surface-association and biofilm formation, leading to cell death when a mismatched prophage sequence target is present in the chromosome^{16,17}. This suggests that CRISPR auto-immunity costs are also dependent on the growth state and physical environment of the cell.

Here, we identify a CRISPR-Cas repressive pathway in *P. aeruginosa*. We speculate that the ability to control CRISPR-Cas activity during lifestyle transitions may be essential for *P. aeruginosa* to safely maintain a CRISPR-Cas system by limiting self-toxicity. In our discovery of MGE-encoded CRISPR-Cas repressors we reveal an unexpected cost to CRISPR-Cas

regulation: the evolution of CRISPR-Cas repression has created an Achilles Heel that is exploited by genetic parasites.

REFERENCES

1. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
2. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
3. Damron, F. H., Qiu, D. & Yu, H. D. The *Pseudomonas aeruginosa* sensor kinase KinB negatively controls alginate production through AlgW-dependent MucA proteolysis. *J. Bacteriol.* **191**, 2285–2295 (2009).
4. Damron, F. H. *et al.* Analysis of the *Pseudomonas aeruginosa* Regulon Controlled by the Sensor Kinase KinB and Sigma Factor RpoN. *J. Bacteriol.* **194**, 1317–1330 (2012).
5. Wozniak, D. J. & Ohman, D. E. Transcriptional analysis of the *Pseudomonas aeruginosa* genes *algR*, *algB*, and *algD* reveals a hierarchy of alginate gene expression which is modulated by *algT*. *J. Bacteriol.* **176**, 6007–6014 (1994).
6. Baynham, P. J. & Wozniak, D. J. Identification and characterization of AlgZ, an AlgT-dependent DNA-binding protein required for *Pseudomonas aeruginosa* *algD* transcription. *Molecular Microbiology* **22**, 97–108 (1996).
7. Hochstrasser, M. L. *et al.* CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proceedings of the National Academy of Sciences* **111**, 6618–6623 (2014).
8. Westra, E. R. *et al.* CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. *Mol Cell* **46**, 595–605 (2012).
9. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520**, 505–510 (2015).

10. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science* **329**, 1355–1358 (2010).
11. Wiedenheft, B. *et al.* Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* **17**, 904–912 (2009).
12. Wiedenheft, B. *et al.* RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences* **108**, 10092–10097 (2011).
13. Rollins, M. F., Schuman, J. T., Paulus, K., Bukhari, H. S. T. & Wiedenheft, B. Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Research* (2015). doi:10.1093/nar/gkv094
14. Rollins, M. F. *et al.* Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. *Proceedings of the National Academy of Sciences* **23**, 201616395–E5121 (2017).
15. Chowdhury, S. *et al.* Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. *Cell* **169**, 47–57.e11 (2017).
16. Zegans, M. E. *et al.* Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* **191**, 210–219 (2009).
17. 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).
18. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **194**, 5728–5738 (2012).
19. Vorontsova, D. *et al.* Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. *Nucleic Acids Research* **43**, 10848–10860 (2015).

20. van Belkum, A. *et al.* Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant *Pseudomonas aeruginosa*. *mBio* **6**, e01796–15 (2015).
21. Westra, E. R. *et al.* Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. *Curr. Biol.* **25**, 1043–1049 (2015).
22. van Houte, S. *et al.* The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* **532**, 385–388 (2016).
23. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
24. Pawluk, A., Bondy-Denomy, J., Cheung, V. H. W., Maxwell, K. L. & Davidson, A. R. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* **5**, e00896–e00896–14 (2014).
25. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
26. Pawluk, A. *et al.* Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology* **1**, 1–6 (2016).
27. Høyland-Kroghsbo, N. M. *et al.* Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proceedings of the National Academy of Sciences* **114**, 201617415–135 (2016).
28. Patterson, A. G. *et al.* Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. *Mol Cell* **64**, 1102–1108 (2016).
29. Borges, A. L. *et al.* Bacteriophage Cooperation Suppresses CRISPR-Cas3 and Cas9 Immunity. *Cell* **174**, 917–925.e10 (2018).
30. Chand, N. S. *et al.* The Sensor Kinase KinB Regulates Virulence in Acute *Pseudomonas aeruginosa* Infection. *J. Bacteriol.* **193**, 2989–2999 (2011).

31. Chand, N. S., Clatworthy, A. E. & Hung, D. T. The Two-Component Sensor KinB Acts as a Phosphatase To Regulate *Pseudomonas aeruginosa* Virulence. *J. Bacteriol.* **194**, 6537–6547 (2012).
32. Cezairliyan, B. O. & Sauer, R. T. Control of *Pseudomonas aeruginosa* AlgW protease cleavage of MucA by peptide signals and MucB. *Molecular Microbiology* **72**, 368–379 (2009).
33. Schurr, M. J., Yu, H., Martinez-Salazar, J. M., Boucher, J. C. & Deretic, V. Control of AlgU, a member of the sigma E-like family of stress sigma factors, by the negative regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis. *J. Bacteriol.* **178**, 4997–5004 (1996).
34. Tart, A. H., Blanks, M. J. & Wozniak, D. J. The AlgT-dependent transcriptional regulator AmrZ (AlgZ) inhibits flagellum biosynthesis in mucoid, nonmotile *Pseudomonas aeruginosa* cystic fibrosis isolates. *J. Bacteriol.* **188**, 6483–6489 (2006).
35. Schulz, S. *et al.* Elucidation of sigma factor-associated networks in *Pseudomonas aeruginosa* reveals a modular architecture with limited and function-specific crosstalk. *PLoS Pathog* **11**, e1004744 (2015).
36. Wozniak, D. J. *et al.* Alginate is not a significant component of the extracellular polysaccharide matrix of PA14 and PAO1 *Pseudomonas aeruginosa* biofilms. *Proc Natl Acad Sci USA* **100**, 7907–7912 (2003).
37. Pratama, A. A. & van Elsas, J. D. A novel inducible prophage from the mycosphere inhabitant *Paraburkholderia terrae* BS437. *Sci Rep* **7**, 9156 (2017).
38. Pryor, E. E. *et al.* The transcription factor AmrZ utilizes multiple DNA binding modes to recognize activator and repressor sequences of *Pseudomonas aeruginosa* virulence genes. *PLoS Pathog* **8**, e1002648 (2012).
39. Martin, D. W. *et al.* Mechanism of conversion to mucoidy in *Pseudomonas aeruginosa* infecting cystic fibrosis patients. *Proc Natl Acad Sci USA* **90**, 8377–8381 (1993).

40. Jones, A. K. *et al.* Activation of the *Pseudomonas aeruginosa* AlgU regulon through mucA mutation inhibits cyclic AMP/Vfr signaling. *J. Bacteriol.* **192**, 5709–5717 (2010).
41. Koonin, E. V. & Makarova, K. S. Mobile Genetic Elements and Evolution of CRISPR-Cas Systems: All the Way There and Back. *Genome Biol Evol* **9**, 2812–2825 (2017).
42. Heilmann, S., Sneppen, K. & Krishna, S. Sustainability of virulence in a phage-bacterial ecosystem. *Journal of Virology* **84**, 3016–3022 (2010).
43. Dötsch, A. *et al.* The *Pseudomonas aeruginosa* Transcriptional Landscape Is Shaped by Environmental Heterogeneity and Genetic Variation. *mBio* **6**, e00749–15–10 (2015).
44. Erdmann, J., Preusse, M., Khaledi, A., Pich, A. & Häussler, S. Environment-driven changes of mRNA and protein levels in *Pseudomonas aeruginosa*. *Environmental Microbiology* **20**, 3952–3963 (2018).
45. Smale, S. T. β -Galactosidase Assay. *Cold Spring Harb Protoc* **2010**, pdb.prot5423 (2010).
46. Arndt, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research* **44**, W16–21 (2016).

MATERIALS AND METHODS

Bacterial strains and bacteriophages. *P. aeruginosa* UCBPP-PA14 (PA14) strains and *E. Coli* strains (**Table 4.4**) were grown on lysogeny broth (LB) agar or liquid at 37 °C. Media was supplemented with gentamicin (50 µg ml⁻¹ for *P. aeruginosa* and 30 µg ml⁻¹ for *E. Coli*) to maintain the pHERD30T plasmid or carbenicillin (250 µg ml⁻¹) for *P. aeruginosa* or ampicillin (100 µg ml⁻¹) for *E. coli* containing the pHERD20T plasmids. pHERD plasmids were induced with 0.1% arabinose. Bacteriophage stocks (**Table 4.4**) were prepared as described previously¹⁸. In brief, 3 ml of SM buffer was added to plate lysates of the desired purified bacteriophage and incubated at room temperature for 15 minutes. SM buffer containing phages was collected and 100 µl of chloroform was added. This was centrifuged at 10,000 x g for 5 minutes and supernatant containing phages was transferred to a screw cap storage tube and incubated at 4 °C.

Transposon mutagenesis screen. The *csy::lacZ* reporter strain was subjected to transposon mutagenesis and colonies were isolated on plates containing X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). ~50,000 colonies were visually examined for increased or decreased levels of β-galactosidase and insertions mapped by semi-random PCR. To conduct transposon mutagenesis, overnights of PA14 *csy3::lacZ* and *E. coli* containing the pBTK30 Tn suicide vector were mixed in a 1:2 ratio (donor : recipient) for conjugation. Mixed cells were centrifuged at 4,000 x g for 10 minutes to pellet cells. 100 µl of resuspended conjugation pellet was then spotted on LB agar plates and incubated at 37 °C for 6h. Conjugation spots were collected and resuspended in LB liquid media. Conjugation was then plated on an LB agar plates supplemented with nalidixic acid (30 µg ml⁻¹) and gentamicin (50 µg ml⁻¹). Surviving colonies containing Tn insertions were collected into 1ml of LB liquid media. Serial dilution of were prepared and plated on LB agar plates supplemented with x-gal (200 µg ml⁻¹) and gentamicin (50 µg ml⁻¹) and nalidixic acid (30 µg ml⁻¹). Plates were incubated at 37 °C for 24 h to allow for

colonies to change color. Colonies displaying changed expression levels as compared to the unmutagenized parental strain (PA14 *csy3::lacZ* no pBTK30) were then isolated onto secondary LB agar plates with X-gal, gentamicin, and nalidixic acid at the stated concentrations. Genomic DNA (gDNA) was collected from isolated single colonies by resuspending bacterial colonies in 0.02% SDS and boiling the sample for 15 minutes. Samples were then centrifuged at 10,000 x g and supernatants containing gDNA were collected. Transposon insertion junctions were mapped using semi-random PCR (Supplementary 2). PCR samples were sequenced and reads were then mapped to the *P. aeruginosa* UCBPP-PA14 genome using BLAST. Expression changes were then verified via modified β -galactosidase assay in liquid culture.

Plaque assays. Plaque assays were performed on LB agar plates (1.5% agar) with LB top agar (0.7% agar), supplemented MgSO₄ (10 mM final concentration) and gentamicin (50 $\mu\text{g ml}^{-1}$) and arabinose (0.1%) as needed for plasmid maintenance and induction. Spot titrations were done by mixing 150 μl of a *P. aeruginosa* overnight culture with 3 ml of top agar, which was dispersed evenly on a LB MgSO₄ plate. 3 μl of 10-fold phage dilutions were then spotted on the surface. Plates were incubated overnight at 30 °C. To count plaques, full plate assays were used, except when CRISPR-targeting was so strong that discrete plaques could not be accurately measured. In this case, spot titrations are shown. For full plate assays, 10 μl of the phage dilution giving single plaques was incubated with 150 μl of *P. aeruginosa* overnight for 10 minutes at 37 °C. 3 ml of top agar was then added and the mixture was dispersed evenly on a LB MgSO₄ plate. Individual plaques were then counted to assess differences in efficiency of bacterial immunity and phage. Efficiency of immunity (EOI) of a bacterial mutant relative to WT was calculated by dividing the number of plaque-forming units (PFUs) formed on WT by the number of PFUs formed on the mutant strain. EOI>1 means less plaques formed on the mutant than on WT, so the mutant was more immune to phage infection than WT. EOI<1 means more plaques formed on the mutant, so the mutant was less immune to phage infection than WT. Efficiency of

plaquing (EOP) of a phage (**Fig. 4.7e**) was calculated by dividing the number of PFUs formed on WT (CRISPR+) by the number of PFUs formed on Δ CRISPR. An EOP of 1 means that CRISPR does not impact phage replication, and EOP of 0 means that the phage cannot replicate in the presence of CRISPR.

β -galactosidase assay. A β -galactosidase assay described previously⁴⁵ was used to measure *lacZ* activity in transcriptional fusions. Bacterial cultures were grown overnight at 37 °C. Cultures were then diluted 1:100 into LB liquid medium supplemented with the desired antibiotic, and incubated at 37 °C with shaking until the desired time point was reached. Culture density was measured with a spectrophotometer (OD_{600}) and 200 μ l of the sample was added to 800 μ l of permeabilization solution. Cells were mixed via inversion and vortexed for 1 minute to permeabilize the cells. 200 μ l of ONPG (4 mg ml⁻¹) was added and samples were incubated at 30 °C until sample turned yellow. Enzymatic reaction was stopped by addition of 300 μ l of 1M Na₂CO₃. Samples were centrifuged at 13,000 x g for 5 minutes to remove debris and 200 μ l of supernatant was moved to a 96-well plate to read absorbance at 420 nm and at 550 nm. Miller units were calculated using the Miller equation: $(1000 \cdot OD_{420 \text{ nm}} - 1.75 \cdot OD_{550 \text{ nm}}) / (T_{\text{min}} \cdot V_{\text{mLs}} \cdot OD_{600 \text{ nm}})$.

Phage transduction of *kinB::Tn* alleles. Transposon insertions in *kinB* from a *csy3::lacZ* background were transduced into WT PA14 to enable testing of CRISPR-Cas function with the same transposon insertion. Phage phiPA3 was used to infect the donor strain (*kinB::Tn*), on plates with top agar overlays, using $\sim 10^4$ PFU to generate near confluent lysis. Plates were soaked in 3-4 mL of phage SM buffer and 2 mL collected over chloroform, vortexed, and pelleted to isolate transducing phage in the supernatant. Lysates were used to infect recipient strains (WT PA14). $\sim 10^8$ PFU were used to infect a culture at an MOI of 1. After 30 minutes of static incubation on the bench, cultures were gently shaken at 37 °C for 20 min and then

pelleted at 5000 x g. Cells were washed twice with LB, and subsequently incubated at 37 °C for 1 hour to allow recombination and gentamicin resistance outgrowth. Cultures were pelleted and resuspended in 200 µL of LB, and plated on LB plates containing gentamicin. Controls included uninfected cells and cells infected with phages not propagated on a gentamicin resistant donor strain. Additionally, phage lysate was directly plated under selection to confirm no residual donor strain in the phage preparation. Plates were incubated overnight at 37 °C and their identity (i.e. CRISPR-Cas intact) confirmed with a plaque assay using DMS3m_{acrIE3} as the target phage and PCR of the *kinB* locus.

Introduction of *csy3::lacZ* *P. aeruginosa* UCBPP-PA14 strains. The *lacZ* gene was introduced into PA14 strains of interest via allelic replacement. Recombination vector pMQ30 containing *lacZ* flanked by homology arms matching *csy2* and *csy4* was introduced via conjugation. PA14 strains and *E. coli* containing vector were mixed at a ratio of 1:2 (recipient:donor). Mixture was heat shocked at 42 °C for 10 min. Mating spot was then plated on a LB agar plate and incubated overnight for 30 °C. Mating spot was then collected, resuspended in 1 ml of LB liquid media and plated on VBMM plates supplemented with 50 µg/mL gentamicin to select for colonies with the integrated homology plasmid. Colonies were cultured overnight in LB in the absence of selection at 37 °C, and were then diluted and counterselected on no salt LB (NSLB) agar plates supplemented with 15% sucrose. Surviving colonies were then grown on LB agar plates supplemented with gentamicin and X-gal to check for *lacZ* insertion via color change and *lacZ* insertion was further verified via PCR.

Reverse transcriptase quantitative PCR (RT-PCR). Total RNA extracts were harvested using an acid-phenol chloroform extraction from liquid cultures subcultured 1:100 and grown for 8 h in LB media. RNA treated with DNase (Ambion) to remove DNA and 1ng of total RNA was used in a series of RT-qPCR reactions. Reactions were conducted in a BioRad CFX connect qPCR

cycler, using clear BioRad plates with the Luna Universal One-Step Reaction Mix (NEB). A standard curve for each primer set was generated using pooled RNA samples. The housekeeping gene, *rpsL*, was used for normalization, and gene specific primers against *cas3* and *csy3* (Table 4.5) were used to quantify expression from the *cas* and *csy* operons. For RT-qPCR reactions, 1 ng of total RNA was used in each reaction, performed in triplicate. Reverse transcription was conducted using to generate cDNA using Luna WarmStart® RT Enzyme Mix (NEB). Standard curves were used to calculate the relative abundance of target transcripts, and *cas3* and *csy3* transcript levels were then normalized to *rpsL* levels.

Generation of endogenous Csy1-sfCherry and Cas3-sfCherry reporters

Endogenous Csy1-sfCherry and Cas3-sfCherry reporters were constructed similar to the construction of *csy3::lacZ*. We initially verified that tagging of sfCherry at the N-terminus of Csy1 and Cas3 are functional, when expressed from a plasmid. pMQ30-sfCherry-Csy1, which contains sfCherry sequence flanked by 657 bp upstream of *csy1* and 701 bp downstream of *csy1* start codon, was cloned in pMQ30 plasmid between HindIII and BamHI sites using Gibson assembly. pMQ30-sfCherry-Cas3, which contains sfCherry sequence flanked by 353 bp upstream of *cas3* and 350 bp downstream of the *cas3* start codon, was cloned in pMQ30 plasmid between HindIII and BamHI sites using Gibson assembly. The 4 bp that overlap between the end of *cas1* and the beginning of *cas3* were duplicated in the final construct. Both pMQ30-sfCherry-Csy1 and pMQ30-sfCherry-Cas3 contains ggaggcgggtggagcc sequence (encoding GGGGA) as linker between sfCherry and the respective tagged proteins. The Csy1-sfCherry and Cas3-sfCherry construct were introduced into PA14 strains of interest via allelic replacement. Strains containing appropriate insertion were verified via PCR.

sfCherry reporter profiling. Liquid: Cells were diluted 1:100 from an overnight culture into fresh LB (with 0.1% arabinose and 50 ug/mL gentimicin if required for plasmid induction and

maintenance), and grown for the indicated number of hours in biological triplicate. 500 μ l of each sample was then spun down at 8,000xg for 2 minutes, and resuspended in 500 μ l of M9 media. Samples were loaded in to a 96 well plate (150 μ l /well) in technical triplicate and red fluorescence and OD600 were measured using a Biotek H4 Synergy 96 well plate reader. M9 media alone was measured to obtain a background fluorescence and absorbance reading.

Solid: Cells were diluted 1:100 from an overnight culture into fresh LB and 20 μ l plated onto individual wells in biological triplicate in a 24 well plate with each well containing solidified 1.5% LB Agar (with 0.1% arabinose and 50 μ g/mL gentimicin if required for plasmid induction and maintenance). The 24 well plate was then covered with a breathable Aeraseal, and incubated at 37 °C with no shaking. At the indicated timepoint, cells were harvested by flooding each well with 500 μ l of M9 buffer, and were spun down at 8,000xg for 2 minutes, and resuspended in 500 μ l of M9 media. Samples were loaded in to a 96 well plate (150 μ l /well) in technical triplicate and red fluorescence (excitation 580 nm, emission 610 nm) and OD600 were measured using a Biotek H4 Synergy 96 well plate reader. M9 media alone was measured to obtain a background fluorescence and absorbance reading. To calculate the relative fluorescence units for each sample, the background fluorescence and background OD600 values obtained were subtracted from the sample values, and the sample fluorescence was then normalized to the sample OD600.

Generation of PA14 Δ *amrZ* using the endogenous I-F CRISPR-Cas system.

Complementary oligonucleotides encoding a crRNA targeting the *amrZ* gene of PA14 were annealed and ligated into the multiple cloning site of the pHERD30T vector. A fragment possessing homology arms flanking the desired mutation (500 bp upstream and 500 bp downstream) around *amrZ* was cloned into a distinct location (NheI site) of the same vector via Gibson assembly. The new plasmid containing both a crRNA and homology region was introduced into WT PA14 via electroporation. Transformation efficiency dropped dramatically in

the presence of the crRNA due to the toxicity caused by self-targeting. All surviving colonies had the desired clean deletion of the *amrZ* gene. Deletions were confirmed by PCR of the region of interest and subsequent Sanger sequencing of the amplicon. A 2000 bp region flanking *amrZ* was PCR amplified and sequencing primers were designed to sequence both the deletion junction and outside of the original 500 bp flanking regions to confirm the removal of the *amrZ* gene.

Liquid phage infection assay. Liquid phage infections were performed as described in²⁹. In brief, an overnight culture of cells was diluted 1:100 into fresh media, and infected with 10^6 PFU virulent phage DMS3m_{acrIF4} in biological triplicate in a 96 well Costar plate. Cells were incubated at 37 °C with constant rotation and OD600 measured every 5 minutes in a Biotek H4 Synergy plate reader. Phage were harvested from each well and quantified by plaque assay after 24 h. In these experiments, all strains used in the assay carried 2 spacers against the DMS3m_{acrIF4} phage to prevent phage escape: one endogenous spacer (CRISPR2_sp1), and the other spacer was provided on a pHERD30T plasmid.

AmrZ homolog discovery and characterization. BLASTp was used to search the nonredundant protein database for AmrZ homologs (accession: ABJ12639.1) in *Pseudomonas* sp. (taxid: 286) in May 2019. This homolog list (e value > 0.001) was then examined for homologs found on phage or plasmid genomes. Representative homologs were aligned using Clustal and the alignment visualized in Jalview, and key conserved residues were mapped onto the structure in Pymol (PDB ID: 3QOQ). Select homologs were synthesized (TWIST Biosciences) and cloned into the SacI/PstI site of the arabinose-inducible plasmid pHERD30T using Gibson assembly. Vectors were electroporated into *Pseudomonas aeruginosa* strains for functional testing, where they were induced with 0.1% arabinose and maintained with 50 ug/mL gentimicin.

Construction of recombinant DMS3m AmrZ phages. Phages were generated as previously described²⁹. Briefly, Gibson assembly was used to generate a recombination plasmid on pHERD 30T with *amrZ_{phi3}* or *amrZ_{PaBG}* flanked by homology arms up and downstream of the DMS3m Acr locus. This plasmid was transformed into PA14 Δ CRISPR, and infected with phage DMS3m_{acr-gent} – a phage that is sick as the result of the insertion of a large gentamicin resistance cassette into its anti-CRISPR locus. Healthy plaques resulting from the recombination were screened for their incorporation of *amrZ_{phi3}* or *amrZ_{PaBG}* into the anti-CRISPR locus with PCR.

Construction of PA14 lysogens. Lysogens were obtained by first spotting phage onto a bacterial lawn, then streaking out surviving colonies from phage spots. These colonies were screened for phage resistance using a cross streak method, and lysogeny verified by prophage induction.

Pyocyanin Repression Assay. The pyocyanin repression assay was performed as previously described²⁵. Lysogens were transformed with a plasmid encoding a Type I-F crRNA targeting the promoter region of the gene *phzM*, which is required for the synthesis of the green pigment pyocyanin. As a control, each lysogen was also transformed with the empty vector plasmid. These strains were grown overnight (~16 h) in 5 mLs of LB media supplemented with 50 $\mu\text{g ml}^{-1}$ gentamicin and 0.01% arabinose, to induce expression of the crRNA. Pyocyanin was extracted with an equal volume of chloroform, and then mixed with a half-volume of 0.2 M HCl, which produces a pink color proportional to the amount of pyocyanin and can be quantified by measuring absorbance at 520 nm. The absorbance value of each crRNA-expressing lysogen was expressed as a percentage of the pyocyanin level measured in the empty vector control lysogen. Samples were measured in technical triplicate.

Statistical Testing. This study used a two-tailed unpaired Student's T-test for statistical testing. In all cases, sample size is $n = 3$, degrees of freedom is $n-1$, confidence interval is 95%. In all plots, bar or data point height is equivalent to the mean and error bars are shown as ± 1 *standard deviation (SD).

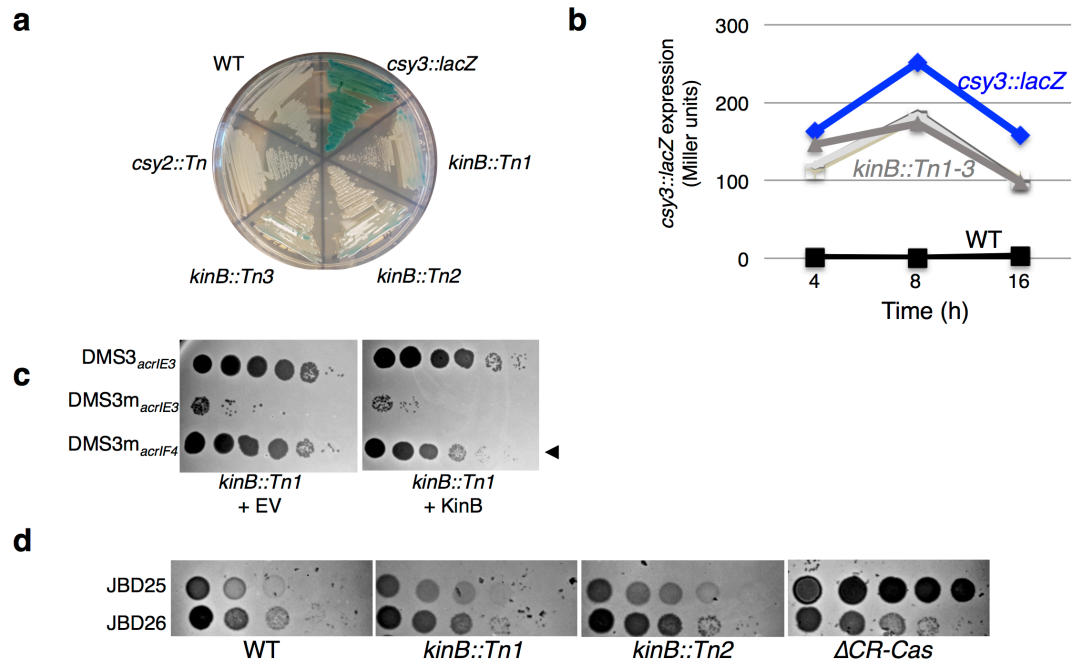


Figure 4.1. Characterization of *kinB::Tn* mutants

a. A streak plate on X-gal plates, showing strains involved in this study and isolated transposon (Tn) insertions. *csy3::lacZ* is a derivative of WT PA14, and is the unmutagenized parent of *kinB::Tn 1-3*.

b. β -galactosidase measurements of strains grown in liquid culture for the indicated time. Measurements for the unmutagenized (*csy3::lacZ*) parent strain and three isolated *kinB* transposon mutants (*kinB::Tn1-3*) are shown, as well as a control PA14 culture with no *lacZ* insertion.

c. Phage titration on lawns of the *kinB::Tn1* mutant transformed with empty vector or *kinB*.

d. Spot titration of phages JBD26 (CR2_sp17, sp20-targeted, possessing *acrIF4*), JBD25 (CR1_sp1 targeted) on *kinB::Tn* mutants and Δ CRISPR-Cas. These experiments have been replicated at least 2 times with consistent results.

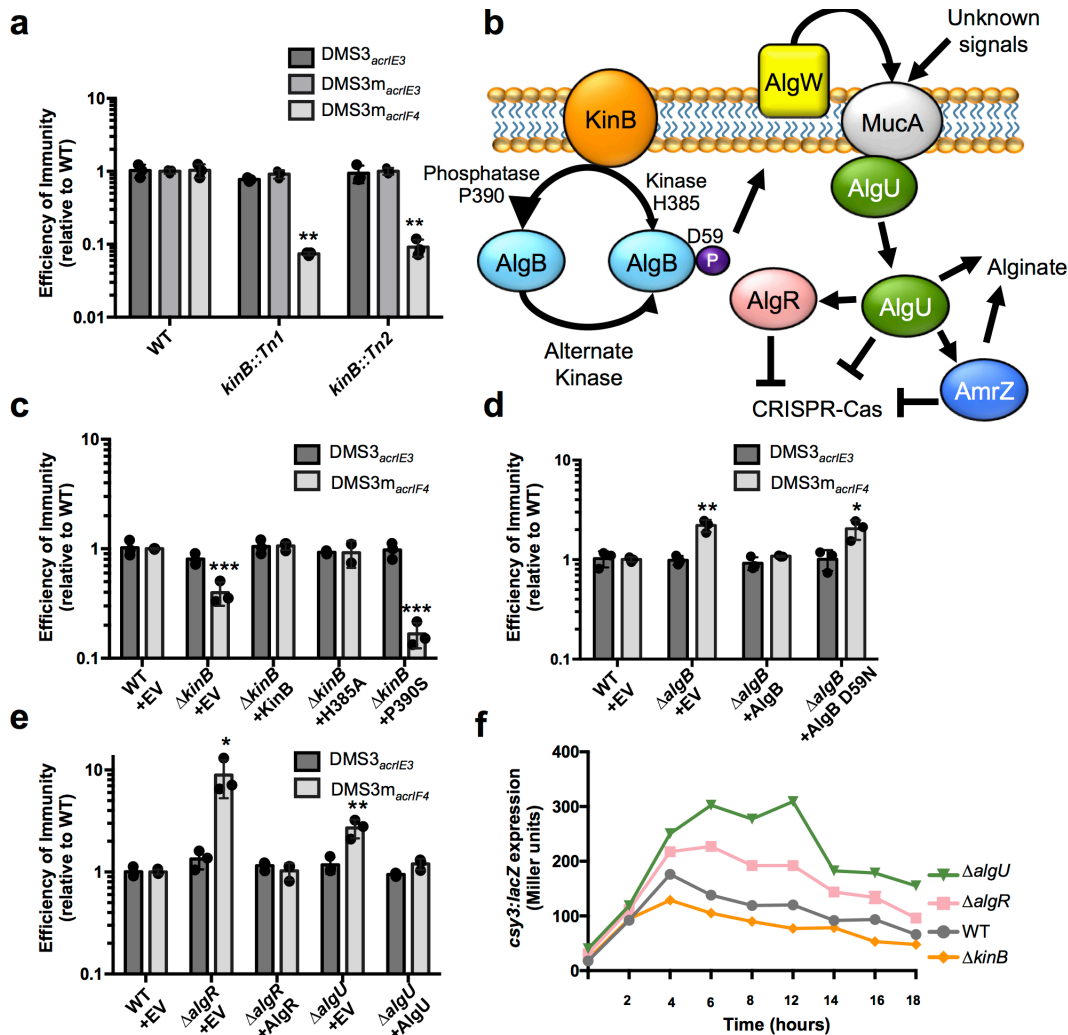


Figure 4.2: A forward genetic screen identifies a role for an alginate-activating pathway in repressing CRISPR-Cas immunity

a. Efficiency of immunity (EOI) against isogenic phages DMS3_{acrIE3} (non-targeted), DMS3m_{acrIE3} (no I-F anti-CRISPR, CRISPR- targeted), and DMS3m_{acrIF4} (weak I-F anti-CRISPR, CRISPR-targeted). Plaque forming units (PFUs) are presented as a ratio relative to the number of PFUs measured on WT PA14, quantified on two independent *kinB* transposon mutants (*kinB::Tn1* and *kinB::Tn2*). Tn mutants show altered EOI against DMS3m_{acrIF4} relative to WT (*Tn1*, $P = 2.9 \times 10^{-3}$, *Tn2*, $P = 3.2 \times 10^{-3}$)

b. A cartoon summarizing the KinB/AlgB two component system and downstream effects, based on prior work⁴⁶ (see text) with CRISPR-Cas regulation added.

c,d,e. EOI measurements for indicated $\Delta algB$, $\Delta kinB$, $\Delta algR$, and $\Delta algU$ strains with complementation. Mutants show altered EOI against DMS3m_{acrIF4} relative to WT ($\Delta kinB + EV$, $P = 4.30 \times 10^{-4}$, $\Delta kinB + P390S$, $P = 5.6 \times 10^{-6}$, $\Delta algB + EV$, $P = 2.8 \times 10^{-3}$, $\Delta algB + D59N$, $P = 1.8 \times 10^{-2}$, $\Delta algR + EV$, $P = 1.9 \times 10^{-2}$, $\Delta algU + EV$, $P = 6.6 \times 10^{-3}$).

f. *csy3::lacZ* β -galactosidase activity over time in the indicated strain backgrounds. Experiment was replicated twice with fewer timepoints and consistent results seen. All EOI data are represented as the mean of 3 biological replicates +/- SD and β -galactosidase reporter activity is represented as the mean of 3 technical replicates. Two-tailed unpaired Student's T-test was used to calculate P values, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

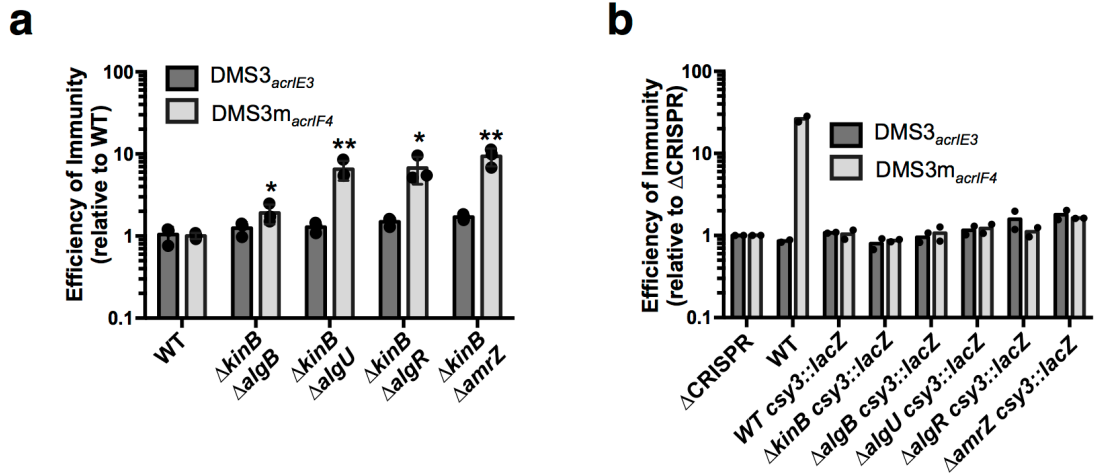


Figure 4.3. Double knockouts of pathway members

a-b. Efficiency of immunity measurements for indicated mutants relative to WT.
a. Double knockouts show $\Delta kinB$ combined with *algB*, *algU*, *algR*, or *amrZ*. EOI measurements are shown as the mean of 3 biological replicates, +/- S.D. Mutants show increased EOI against DMS3_{macrIF4} relative to WT ($\Delta kinB \Delta algB$, $P = 3.8 \times 10^{-2}$, $\Delta kinB \Delta algU$, $P = 5.9 \times 10^{-3}$, $\Delta kinB \Delta algR$, $P = 1.5 \times 10^{-2}$, $\Delta kinB \Delta amrZ$, $P = 3.2 \times 10^{-3}$) Two-tailed unpaired Student's T-test was used to calculate P value, * $p < 0.05$, ** $p < 0.01$.
b. Indicated knockouts were combined with *csy3::lacZ*, EOI shown as the mean of two biological replicates.. These experiments have been replicated at least 2 times with consistent results.

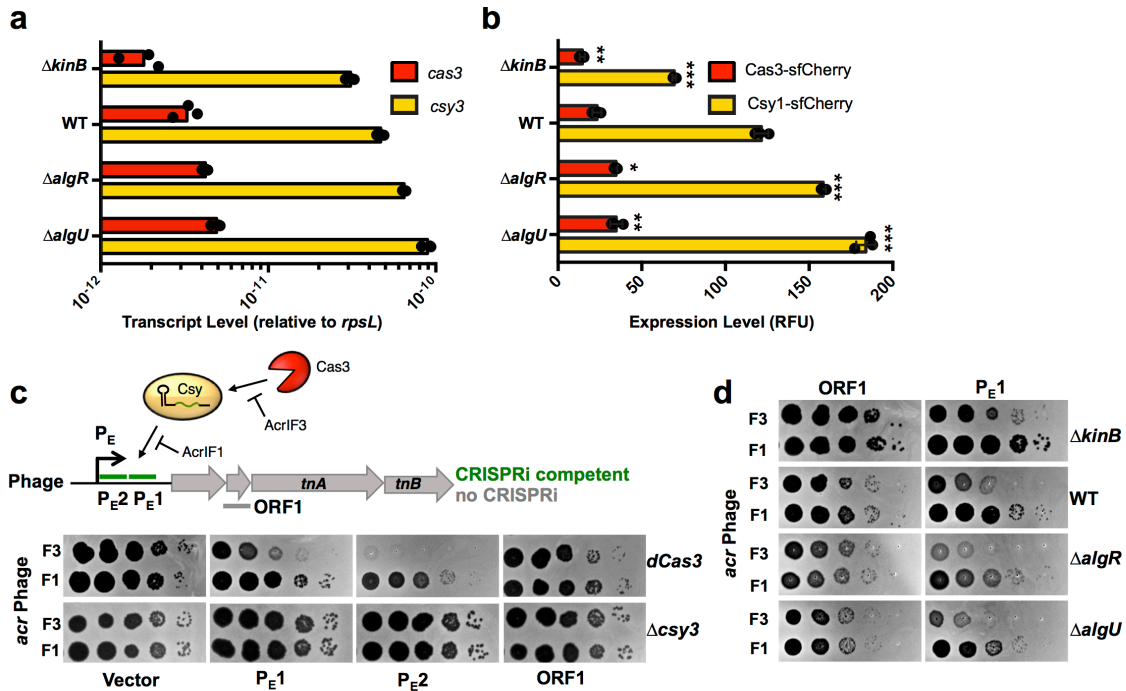


Figure 4.4. The KinB/AlgB pathway modulates Cas3 and Csy protein and RNA levels

a. qRT-PCR measurements of transcript levels of *cas3* (red) and *csy3* (yellow) normalized to the housekeeping gene *rpsL* after 8 h of growth in liquid culture. Measurements are represented as the mean of 3 technical replicates.

b. Measurement of the fluorescence levels of Cas3-sfCherry (red) or Csy1-sfCherry (yellow) reporter strains after 10 h of growth in liquid culture. Fluorescence measurements are represented as the mean of 3 biological replicates +/- SD. Mutants show altered Cas3-sfCherry levels ($\Delta kinB$, $P = 7.8 \times 10^{-3}$ $\Delta algR$, $P = 1.5 \times 10^{-4}$ $\Delta algU$, $P = 1.1 \times 10^{-4}$) and Csy1-sfCherry levels relative to WT ($\Delta kinB$, $P = 3.3 \times 10^{-5}$, $\Delta algR$, $P = 1.5 \times 10^{-4}$, $\Delta algU$, $P = 1.1 \times 10^{-4}$). Two-tailed unpaired Student's T-test was used to calculate P values, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

c. Spot titration of F3 (DMS3m_{acrIF3}) or F1 (DMS3m_{acrIF1}) on *dCas3* (dead Cas3) or $\Delta csy3$ (active Cas3, no Csy complex) strains. Phages are targeted by natural spacer CR2_sp1, as well as crRNAs designed to target DMS3m genome in positions designated on ORF map.

d. Spot titration of DMS3m_{acrIF3} and DMS3m_{acrIF1} phages on WT PA14 or deletion mutants expressing the indicated crRNA. Plaquing experiments were replicated 3 times and consistent results seen.

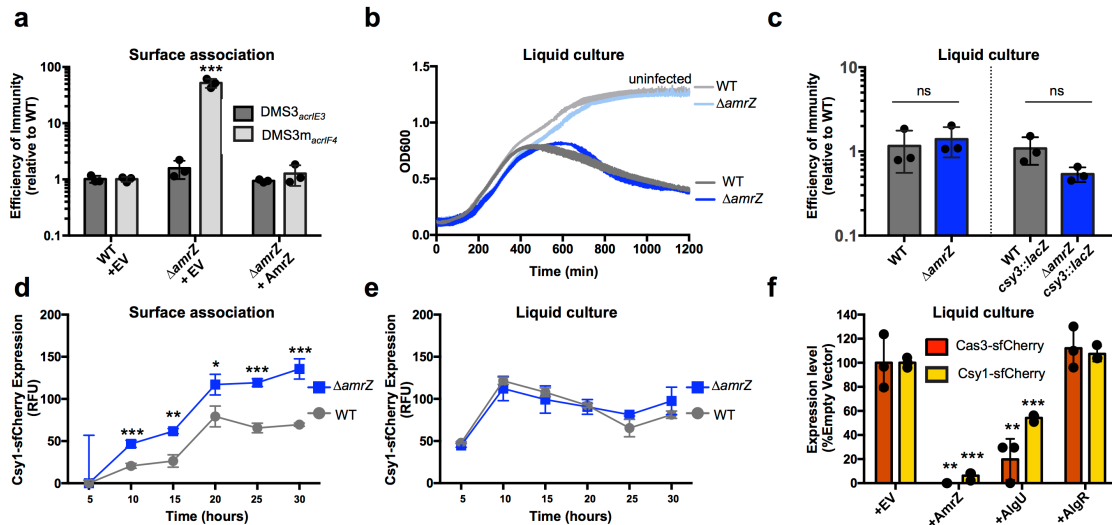


Figure 4.5. AmrZ is a surface-activated repressor of CRISPR-Cas immunity

a. Efficiency of immunity (EOI) against phages DMS3_{acrIE3} (non-targeted) and DMS3m_{acrIF4} (CRISPR-targeted). Plaque forming units (PFUs) were quantified on $\Delta amrZ$ or the complemented strain, then represented as a ratio of the number of PFUs measured on WT PA14. $\Delta amrZ + EV$ shows increased EOI against DMS3m_{acrIF4} relative to WT ($P = 7.3 \times 10^{-4}$). EOI measurements are represented as the mean of 3 biological replicates +/- SD.

b. Growth curves of PA14 WT and $\Delta amrZ$ infected with 10^6 PFU of virulent DMS3m_{acrIF4} alongside uninfected controls.

c. EOI against virulent DMS3m_{acrIF4} in liquid culture of WT and $\Delta amrZ$ strains (CRISPR active) or WT *csy3::lacZ* and $\Delta amrZ$ *csy3::lacZ* (CRISPR inactive). PFUs were quantified after 24 h from $\Delta amrZ$ or *amrZ* *csy3::lacZ*, then represented as a ratio of PFUs from WT or WT *csy3::lacZ*, respectively. OD600 and EOI measurements are represented as the mean of 3 biological replicates +/- SD. $\Delta amrZ$ and $\Delta amrZ$ *csy3::lacZ* show no significant difference of EOI relative to WT and WT *csy3::lacZ*, respectively ($\Delta amrZ$, $P = 0.6$ $\Delta amrZ$ *csy3::lacZ*, $P = 0.08$).

d, e. Timecourse of the fluorescence levels of Csy1-sfCherry reporter strains during surface-association (d) or planktonic growth (e). $\Delta amrZ$ has increased Csy1-sfCherry levels during surface association relative to WT (10 h, $P = 8.9 \times 10^{-4}$, 15 h, $P = 1.5 \times 10^{-3}$, 20 h, $P = 2.0 \times 10^{-2}$, 25 h, $P = 2.2 \times 10^{-4}$, 30 h, $P = 7.0 \times 10^{-4}$) **f.** Normalized fluorescence measurements of WT Cas3-sfCherry (red) or Csy1-sfCherry (yellow) overexpressing the indicated transcription factor after 10 h growth in liquid culture. AmrZ and AlgU overexpression reduced Cas3-sfCherry (AmrZ, $P = 1.5 \times 10^{-3}$, AlgU, $P = 7.8 \times 10^{-3}$) and Csy1-sfCherry (AmrZ, $P = 7.5 \times 10^{-6}$, AlgU, $P = 9.9 \times 10^{-5}$) levels relative to WT. Fluorescence measurements are represented as the mean of 3 biological replicates +/- SD. Two-tailed unpaired Student's T-test was used to calculate P values, ns = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

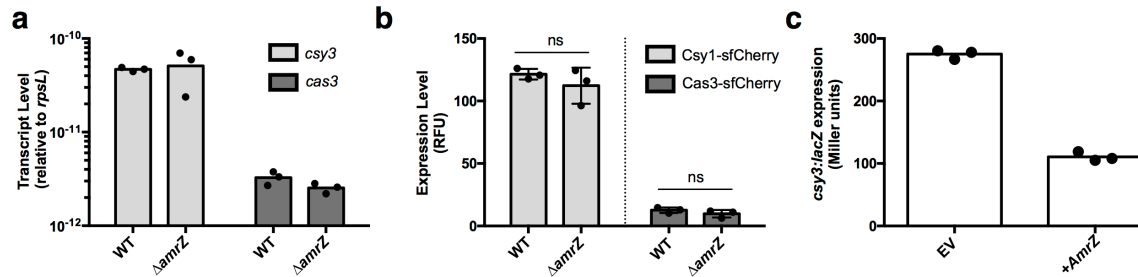


Figure 4.6. AmrZ activity in liquid growth

a. qRT-PCR measurements of transcript levels of *csy3* (light grey) and *cas3* (dark grey) normalized to the housekeeping gene *rpsL* after 8 h of growth in liquid culture. Measurements are represented as the mean of 3 technical replicates.

b. Measurement of the fluorescence levels of Csy1-sfCherry (light grey) or Cas3-sfCherry (dark grey) reporter strains after 10 h of growth in liquid culture. Fluorescence measurements are represented as the mean of 3 biological replicates +/- SD. Cas3-sfCherry ($P = 0.26$) and Csy1-sfCherry levels ($P = 0.35$) in $\Delta amrZ$ did not differ significantly from WT. Two-tailed unpaired Student's T-test was used to calculate P value, ns = not significant

c. *csy3::lacZ* β -galactosidase activity from PA14 WT *csy3::lacZ* transformed with either empty vector (EV) or a plasmid overexpressing AmrZ (+AmrZ). β -galactosidase reporter activity was measured after 8 h in liquid growth and is represented as the mean of 3 technical replicates. Experiment was replicated two times with consistent results.

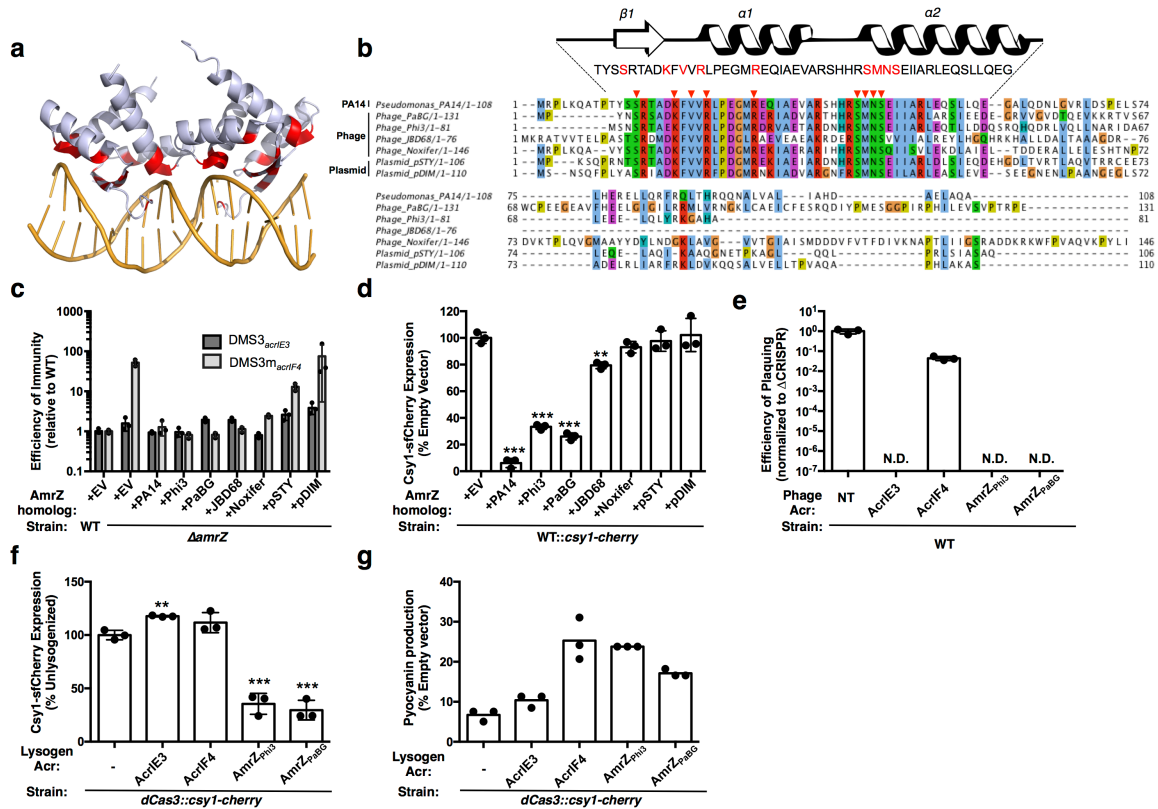


Figure 4.7. Phage-derived AmrZ homologs control CRISPR-Cas immunity

a. Structure of an AmrZ tetramer bound to 18bp of operator DNA³⁸ with DNA-contacting residues highlighted in red.

b. Alignment of six mobile AmrZ homologs and the native PA14 AmrZ homolog, with the ribbon-helix-helix DNA binding domain schematized and DNA-contacting residues indicated with red arrows and text.

c. Efficiency of immunity (EOI) against DMS3_{acrIE3} (non-targeted) and DMS3m_{acrIF4} (CRISPR-targeted). Plaque forming units (PFUs) were quantified on $\Delta amrZ$ or the strains complemented with AmrZ homologs, and represented as a ratio to the number of PFUs measured on WT PA14. Measurements are represented as the mean of 3 biological replicates +/- SD.

d. Normalized fluorescence levels of Csy1-sfCherry reporter strains expressing AmrZ homologs after 10 h of growth in liquid culture, shown as mean of 3 biological replicates, +/- SD. AmrZ homologs from PA14, Phi3, PaBG, and JBD68 repressed Csy1-sfCherry relative to WT (PA14, $P = 7.5 \times 10^{-6}$, Phi3, $P = 1.5 \times 10^{-5}$, PaBG, $P = 1.3 \times 10^{-5}$, JBD68, $P = 1.9 \times 10^{-3}$).

e. Efficiency of plaquing (EOP) of non-targeted DMS3_{acrIE3} phage (NT) or targeted DMS3m_{acr} phages. EOP is the ratio of PFUs on PA14 WT over PFUs formed on PA14 Δ CRISPR, represented as the mean of 3 biological replicates, +/- S.D. N.D. = not detectable.

f. Fluorescence levels of *dCas3::csy1-sfCherry* after 16 h liquid growth lysogenized with the indicated DMS3m_{acr} phage, normalized to the unlysogenized control (-), and represented as the mean of 3 biological replicates +/- SD. Expression of AmrZ_{Phi3} ($P = 4.9 \times 10^{-4}$) and AmrZ_{PaBG} ($P = 2.8 \times 10^{-4}$) from a prophage repressed Csy1-sfCherry expression relative to an unlysogenized control.

g. Pyocyanin production from *dCas3::csy1-sfCherry* reporter strains lysogenized with the indicated DMS3m_{acr} phage or the unlysogenized control (-) after 16 h of growth in liquid culture. Pyocyanin levels during *phzM*-targeting are shown as a percentage of pyocyanin levels in an empty vector control, and represented as the mean of 3 technical replicates. Experiment was replicated three times and consistent results seen. Two-tailed unpaired Student's T-test was used to calculate *P* values, ns = not significant, **p* < 0.05, ***p* < 0.01, ****p* < 0.001.

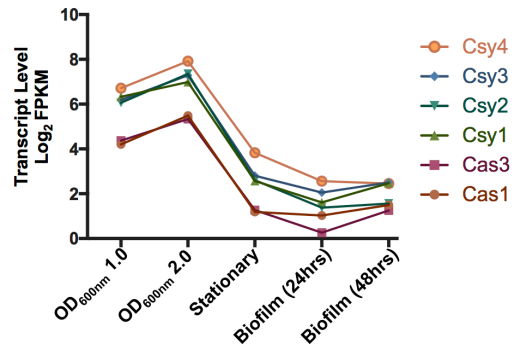
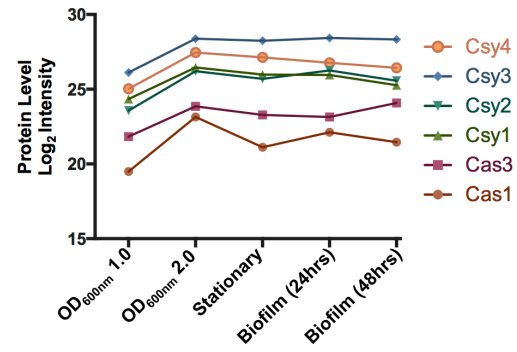
a**b**

Figure 4.8. Cas and Csy RNA and protein levels across growth conditions

a. Log₂ of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) shown for each I-F cas gene in PA14 in the indicated growth condition⁴³.

b. Log₂ of protein levels for each of the I-F Cas proteins in PA14 in the indicated growth condition⁴⁴.

Table 4.1. Mapped insertions from transposon mutagenesis screen

All independent transposon insertions identified and mapped by visual screening with increased or decreased *csy3::lacZ* β -galactosidase activity. β -galactosidase activity is expressed as a percentage of the unmutagenized parent strain, and measurements were taken at a single timepoint after 8 h of growth in liquid culture. The insertion location in the PA14 genome is shown, along with the measured level of β -galactosidase enzyme at the 8 hour timepoint. These measurements were not determined (N/D) for strains with a growth defect.

Gene with insertion	Transposon Location	β -gal activity (% unmutagenized)
pchH	797705	127%
pchH	798159	131%
S-type pyocin	1196879	145%
putative membrane protein	1400718	118%
minD	1915545	N/D (growth defect)
deaD	2373899	N/D (growth defect)
putative Zn-dependent oxidoreductase	2820446	115% (growth defect)
gnyL	3434168	87%
bacA	3490006	75%
Intergenic; zipA and smc +	3979746	N/D (growth defect)
lasR	4085810	135%
oxidoreductase FMN binding	4188602	N/D (growth defect)
pyoS3A	4404303	145%
tolA	4595505	158%
purM	4618060	134%
Intergenic; fstA and fstZ	5104077	111% (growth defect)
cytochrome c1 precursor	5126449	94%
putative plasmid stabilization protein	5347104	N/D (growth defect)
paraquat inducible protein	5532785	108%
glycosyl transferase	5889967	81%
gltB	5943637	108%
yhiH/yhil	6162134	108%
crc	6275250	91%
kinB (3)	6447811	53%
kinB (2)	6447945	50%
kinB (1)	6448519	59%
kinB (5)	6449345	98%
kinB (4)	6449373	54%
polA	6457284	N/D (growth defect)
gidB	6530337	88%
gidB	6530337	88%

Table 4.2. Mobile AmrZ homologs

AmrZ homologs listed by the genome that encodes them, the accession number, and the mobile genetic element type.

Name	Accession	MGE type
PA14 AmrZ	ABJ12639.1	
Pseudomonas phage Noxifer	ARV77275.1	Lytic Myovirus
Pseudomonas phage phi3	YP_009276432.1	Integrated prophage
Pseudomonas phage PaBG	YP_008433620.1	Lytic Myovirus
Pseudomonas phage SM1	ALT58107.1	Siphoviridae (temperate)
Pseudomonas phage F10	YP_001293379.1	Siphoviridae (temperate)
Pseudomonas phage JBD68	ARM70500.1	Siphoviridae (temperate)
Pseudomonas sp. VLB120 plasmid pSTY	AGZ38169.1	Plasmid
Pseudomonas putida plasmid pKF715B	BAW27310.1	Plasmid
Pseudomonas veronii plasmid PVE plasmid	SBW85251.1	Plasmid
Pseudomonas koreensis plasmid p3	AVX93364.1	Plasmid
Pseudomonas sp. Leaf58 plasmid pBASL58	AYG48213.1	Plasmid
Pseudomonas sp. XWY-1 plasmid	AUZ62175.1	Plasmid
Pseudomonas putida KF715C pA870	BAW26592.1	Plasmid
Pseudomonas putida S12 plasmid pTTS12	AJA17154.1	Plasmid
Pseudomonas putida p12969-DIM	ALZ46341.1	Plasmid

Table 4.3. AmrZ copy number analysis of two *Pseudomonas aeruginosa* strains

AmrZ copy number analysis of two different strains of *Pseudomonas aeruginosa*. AmrZ homologs listed by accession number and their genomic coordinates. Phaster⁴⁶ was used to identify the prophages encoding mobile AmrZ copies.

Pseudomonas aeruginosa strain FDAARGOS_570, CP033835.1	Accession	Genomic coordinates
AmrZ-1	AYZ87048.1	6274191 to 6274418, JBD68-like prophage
AmrZ-2	AYZ83165.1	1992281 to 1992508, JBD68-like prophage
AmrZ-3	AYZ81620.1	292461 to 292688, JBD68-like prophage
AmrZ-4	AYZ82758.1	1562237 to 1562392, phi3-like prophage
AmrZ-5	AYZ86193.1	5356309 to 5356461, core genome (endogenous AmrZ)

Pseudomonas aeruginosa strain PA11803, CP015003.1	Accession	Genomic coordinates
AmrZ-1	AOX38089.1	1366154 to 1366381, unknown prophage type
AmrZ-2	AOX38026.1	1323851 to 1324078, JBD68-like prophage
AmrZ-3	AOX37649.1	937493 to 937720, JBD68-like prophage
AmrZ-4	AOX37592.1	896299 to 896526, JBD68-like prophage
AmrZ-5	AOX38565.1	1866517 to 1866654, core genome (endogenous AmrZ)

Table 4.4. Strains and phages used in this study

Bacteria: Pseudomonas aeruginosa strain PA14 derivatives	Source
<i>PA14 WT csy3::lacZ</i>	George O'Toole Lab
<i>ΔkinB csy3::lacZ</i>	This study
<i>ΔalgB csy3::lacZ</i>	This study
<i>ΔalgU csy3::lacZ</i>	This study
<i>ΔalgR csy3::lacZ</i>	This study
<i>ΔamrZ csy3::lacZ</i>	This study
<i>ΔkinB</i>	Deborah Hung lab
<i>ΔalgB</i>	Deborah Hung lab
<i>ΔalgU</i>	Deborah Hung lab
<i>ΔalgR</i>	Deborah Hung lab
<i>ΔamrZ</i>	This study
<i>kinB::Tn1</i>	This study
<i>kinB::Tn2</i>	This study
<i>dCas3</i>	George O'Toole Lab
<i>Δcsy3</i>	George O'Toole Lab
<i>WT PA14 cas3-sfCherry</i>	This study
<i>ΔkinB cas3-sfCherry</i>	This study
<i>ΔalgU cas3-sfCherry</i>	This study
<i>ΔalgR cas3 sf-Cherry</i>	This study
<i>ΔamrZ cas3-sfCherry</i>	This study
<i>WT PA14 csy1-sfCherry</i>	This study
<i>ΔkinB csy1-sfCherry</i>	This study
<i>ΔalgU csy1-sfCherry</i>	This study
<i>ΔalgR csy1 sf-Cherry</i>	This study
<i>ΔamrZ csy1-sfCherry</i>	This study
<i>WT::dCas3 csy1-sfCherry</i>	This study
Phage: DMS3 and DMS3m derivatives	Source
<i>DMS3acrIE3</i>	George O'Toole lab
<i>DMS3macrIE3</i>	George O'Toole lab
<i>DMS3macrIE3 vir</i>	George O'Toole lab
<i>DMS3macrIF1</i>	Alan Davidson Lab
<i>DMS3macrIF3</i>	Borges et. al. Cell, 2018
<i>DMS3macrIF4</i>	Borges et. al. Cell, 2018
<i>DMS3macrIF4 vir</i>	Borges et. al. Cell, 2018
<i>DMS3mamrZPhi3</i>	This study
<i>DMS3mamrZPaBG</i>	This study

Supplementary Table 4.5. Plasmids and primers used in this study

Plasmids	Source
pBTK30 <i>mariner</i> C9 mini-transposon suicide vector	Stephen Lory lab
pHERD30T (gentamicin resistance)	Alan Davidson lab
pHERD20T (carbenicillin resistance)	Alan Davidson lab
pKinB pHERD20T	Deborah Hung lab
pKinB pHERD30T	This study
pKinB P390S pHERD20T	Deborah Hung lab
pKinB P390S pHERD30T	This study
pKinB H385A pHERD20T	Deborah Hung lab
pKinB H385A pHERD30T	This study
pAlgB pHERD20T	Deborah Hung lab
pAlgB pHERD30T	This study
pAlgB D59N pHERD20T	Deborah Hung lab
pAlgB D59N pHERD30T	This study
pAlgU pHERD20T	Deborah Hung lab
pAlgU pHERD30T	This study
pAlgR pHERD30T	This study
pAmrZ pHERD30T	This study
pAmrZ KO pHERD30T	This study
pAmrZ-Phi3 pHERD30T	This study
pAmrZ-PaBG pHERD30T	This study
pAmrZ-JBD68 pHERD30T	This study
pAmrZ-Noxifer pHERD30T	This study
pAmrZ-pSTY pHERD30T	This study
pAmrZ-pDIM pHERD30T	This study

crRNA Plasmids: construct DNA top strand, repeats in lower case, spacer in upper case	Source
IF crRNA- phzM pHERD30T gttcactgcccgtgtaggcagctaagaaaGAAAGAATAAAATTACAACCTTGGCTACA ACCTgttcactgcccgtgtaggcagctaagaaa	Alan Davidson lab, Bondy-Denomy et. al. Nature, 2015
IF crRNA - ORF1 pHERD30T gttcactgcccgtgtaggcagctaagaaaGAGGGCCAAGGCCGCCGATTTCGAGCC AGTCGAGttcactgcccgtgtaggcagctaagaaa	This study
IF crRNA - PE1 pHERD30T gttcactgcccgtgtaggcagctaagaaaGCTATAGAGCTTGTTTTCAACTTCTATA GCGgttcactgcccgtgtaggcagctaagaaa	This study
IF crRNA - PE2 pHERD30T gttcactgcccgtgtaggcagctaagaaaTATAGAAGCTTAAAACCGCTATAGAGCT TGTTgttcactgcccgtgtaggcagctaagaaa	This study

Semi-Random PCR primers	Source
Round1-Pa(1) GGCCACGCGTCTGACTAGTACNNNNNNNNNGATAT	Goodman et. al, Dev. Cell, 2004. Stephen Lory lab
Round1-BTK30 CACCGCTGCGTTCGGTCAAGGTTTC	Goodman et. al, Dev. Cell, 2004. Stephen Lory lab
Round2-Pa GGCCACGCGTCTGACTAGTAC	Goodman et. al, Dev. Cell, 2004. Stephen Lory lab
Round2-BTK30 CGAACCGAACAGGCTTATGTCAATTC	Goodman et. al, Dev. Cell, 2004. Stephen Lory lab
BTK30-Seq TGGTGCTGACCCCGGATGAAG	Goodman et. al, Dev. Cell, 2004. Stephen Lory lab

RT-qPCR primers	Source
csy3-intF aagaccaaggaccgtgacc	Bondy-Denomy et. al, Nature, 2013.
csy3-intR agccctgatcgttcacgtag	Bondy-Denomy et. al, Nature, 2013.
cas3-intF ggttgatcgtcagccatcat	Bondy-Denomy et. al, Nature, 2013.
cas3-intR ggcctttctttgctct	Bondy-Denomy et. al, Nature, 2013.
rpsL-intF3 gtcgacaagagcgacgtg	Bondy-Denomy et. al, Nature, 2013.
rpsL-intR3 aacggttggtcagacgtaca	Bondy-Denomy et. al, Nature, 2013.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

6D5EE33C58934F6... Author Signature

4/27/2020
Date