Alternative functions of CRISPR–Cas systems in the evolutionary arms race

Nature Reviews Microbiology

Mohanraju, Prarthana; Saha, Chinmoy; Baarlen, Peter; Louwen, Rogier; Staals, Raymond H.J. et al

https://doi.org/10.1038/s41579-021-00663-z

Alternative functions of CRISPR–Cas systems in the evolutionary arms race

Prarthana Mohanraju[1,2], *Chinmoy Saha*[3], *Peter van Baarlen*[4], *Rogier Louwen*[3], *Raymond H. J. Staals*[1]✉ *and John van der Oost*[1]✉

Abstract | CRISPR–Cas systems of bacteria and archaea comprise chromosomal loci with typical repetitive clusters and associated genes encoding a range of Cas proteins. Adaptation of CRISPR arrays occurs when virus-derived and plasmid-derived sequences are integrated as new CRISPR spacers. Cas proteins use CRISPR-derived RNA guides to specifically recognize and cleave nucleic acids of invading mobile genetic elements. Apart from this role as an adaptive immune system, some CRISPR-associated nucleases are hijacked by mobile genetic elements: viruses use them to attack their prokaryotic hosts, and transposons have adopted CRISPR systems for guided transposition. In addition, some CRISPR–Cas systems control the expression of genes involved in bacterial physiology and virulence. Moreover, pathogenic bacteria may use their Cas nuclease activity indirectly to evade the human immune system or directly to invade the nucleus and damage the chromosomal DNA of infected human cells. Thus, the evolutionary arms race has led to the expansion of exciting variations in CRISPR mechanisms and functionalities. In this Review, we explore the latest insights into the diverse functions of CRISPR–Cas systems beyond adaptive immunity and discuss the implications for the development of CRISPR-based applications.

Horizontal gene transfer
Exchange of genetic material between organisms that may be phylogenetically unrelated.

[1]*Laboratory of Microbiology, Wageningen University & Research, Wageningen, Netherlands.*

[2]*Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, Netherlands.*

[3]*Department of Medical Microbiology and Infectious Diseases, Erasmus University Medical Center, Rotterdam, Netherlands.*

[4]*Host–Microbe Interactomics Group, Wageningen University & Research, Wageningen, Netherlands.*

✉*e-mail: raymond.staals@wur.nl; john.vanderoost@wur.nl*

https://doi.org/10.1038/s41579-021-00663-z

Horizontal gene transfer is a widespread biological phenomenon that plays a major role in natural evolution, especially in prokaryotes (bacteria and archaea)[1,2]. Foreign DNA can enter a cell via dedicated uptake systems of the host (DNA remnants of a dead cell, including plasmids and transposons), via the conjugation machinery of a donor cell (conjugative plasmids) or via host transporters that are recognized and exploited by viruses for injection of their genome. Uptake and chromosomal integration of foreign DNA can be neutral, lead to a fitness gain (extra functionality) or result in fitness loss (disruption of functionality, or even cell death). Hence, at least on a population level, biological cells will benefit if there is a balance between active uptake of DNA on the one hand and appropriate defence against invading mobile genetic elements (MGEs) on the other hand[2].

The never-ending evolutionary arms race between prokaryotes and MGEs such as viruses, plasmids and transposons has resulted in a broad range of prokaryotic defence systems. Analogously to animal immune systems, prokaryotic immune systems should be able to distinguish between self and non-self. For example, restriction–modification systems recognize specific nucleotide modifications to discriminate the host DNA from the DNA of an invading MGE[3]. A range of prokaryotic innate immune systems have been described in detail[4], and many more have been discovered recently[5]. In addition, the CRISPR–Cas system was the first reported prokaryotic version of an adaptive immune system that is capable of acquiring genetic 'fingerprints' of invaders, storing that information in a genetic 'file' (that is, the CRISPR array) and, in case of a new invasion, using that information to specifically neutralize the invading MGE[6,7]. These systems are generally composed of multiple CRISPR-associated genes (*cas* genes) on the bacterial chromosome as well as a CRISPR array consisting of unique DNA sequences (spacers), which are acquired from invading MGEs through integration between identical repeat sequences. The spacers act as a 'memory bank' of prior infections and enable recognition of the invaders upon repeated infection.

CRISPR–Cas systems generally mediate adaptive immunity through three distinct phases: adaptation, expression (precursor CRISPR RNA (pre-crRNA) processing) and interference. First, during the adaptation phase, a subset of Cas proteins select and process DNA fragments of an invading MGE and insert them into the CRISPR array of the host genome. Selection of MGE DNA sequences, called 'protospacers', relies on the presence of a short protospacer adjacent motif (PAM), or in the case of RNA targeting type III and type VI CRISPR–Cas systems, an RNA PAM or a protospacer flanking sequence, respectively[8,9]. These protospacers are processed to a defined length by removal of the flanks,

including the motif. Next, they are added at the leader end of the CRISPR array as a new spacer, thereby expanding the memory of the immune system. Second, during the expression phase, the CRISPR array is transcribed as a long pre-crRNA and processed by specific ribonucleases into individual small crRNAs (the guides). Finally, during the interference phase, the effector complex detects an appropriate motif (PAM, RNA PAM or protospacer flanking sequence) on the foreign DNA, after which the crRNA guide screens for a complementary target sequence. When a match is found, activation of a nuclease leads to cleavage and eventually to neutralization of the invading MGE. The Cas proteins are remarkably diverse, resulting in CRISPR–Cas systems currently being grouped into two classes, six types and more than 30 subtypes[10] (BOX 1).

The role of CRISPR–Cas systems in defending prokaryotes from viruses and other foreign genetic elements is well documented[11–14]. However, a steadily growing list of examples has revealed that CRISPR–Cas systems are involved in different stages of the evolutionary arms race between prokaryotes and viruses.

Transposons
Segments of DNA that can move within and between genomes by integrating into target sites using one or more transposon-encoded enzymes (transposase, recombinase or integrase).

Mobile genetic elements
(MGEs). Clusters of selfish genes, either naked or packaged in capsid-like structures, that need to invade cellular organisms for their replication and proliferation.

Evolutionary arms race
The continuous arms race of developing infection and anti-infection strategies resulting in a rapid co-evolution of the parasite's offence systems and the host's defence systems.

CRISPR array
Genomic locus located adjacent to the CRISPR-associated genes (*cas* genes), consisting of multiple variable spacer sequences separated by tandem invariable repeats.

Spacers
Unique segments of DNA that are frequently derived from viral genomes and plasmids, and that are inserted between repeats in a CRISPR array.

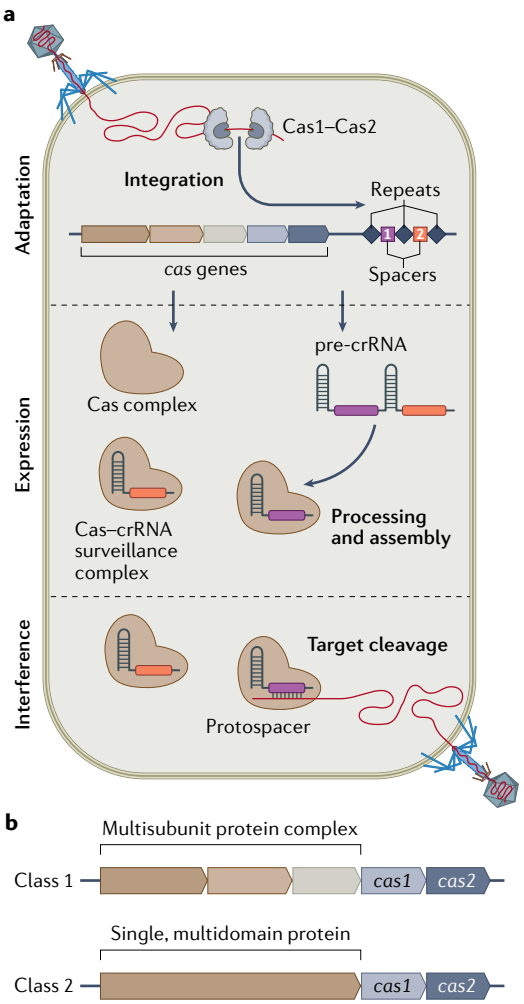
CRISPR RNA
Short RNA molecules, produced by primary processing of the long precursor transcript of a CRISPR array, consisting of a spacer flanked on one or both sides by repeat-derived handles. CRISPR RNAs guide the Cas protein(s) to target cognate foreign DNA or RNA.

Protospacer adjacent motif
(PAM). A short signature sequence flanking the protospacer that enables self–non-self discrimination. In most CRISPR–Cas systems, the PAM sequence is essential for both adaptation and target recognition.

Box 1 | **CRISPR–Cas basics: mechanism and classification**

Adaptation, expression and interference

CRISPR–Cas systems are encoded by genomic loci consisting of a CRISPR array and clusters of associated *cas* genes[6,7]. The mechanistic steps of adaptive CRISPR–Cas immunity include CRISPR adaptation, CRISPR expression and CRISPR interference (see the figure, part **a**). The CRISPR adaptation stage generally involves the acquisition of new spacers from foreign mobile genetic element (virus, plasmid or transposon) DNA and their addition to the CRISPR array in the bacterial or archaeal chromosome. The stepwise mechanism of spacer acquisition is as follows: (1) protospacer adjacent motif-containing double-stranded DNA fragments are bound and processed by the Cas1–Cas2 complex[154,155]; (2) the resulting prespacer fragment catalyses a nucleophilic attack of the leader-proximal repeat of the CRISPR array, generating nicks in either strand at both sides of the repeat[156]; (3) repeat strands are melted and the prespacer is integrated; (4) a second strand is synthesized by a non-Cas DNA polymerase; and (5) gaps are covalently linked by a non-Cas ligase[157]. During CRISPR expression, the entire CRISPR array is transcribed as a precursor CRISPR RNA (pre-crRNA) and is processed into mature crRNAs such that each crRNA contains a single spacer flanked on one side or both sides by part of a repeat from the CRISPR array[7]. In class 1 CRISPR–Cas systems, the crRNA guides are bound by a Cas protein complex (for example, Cascade)[7], whereas in class 2 systems, crRNAs are bound by a single, multidomain protein (for example, Cas9 or Cas12)[149,158] to form ribonucleoprotein complexes. In the case of DNA-targeting systems, the CRISPR interference stage is characterized by protospacer adjacent motif scanning followed by base-pairing of the transcribed spacer part of a crRNA–Cas ribonucleoprotein complex to a matching target DNA strand, while the non-target DNA strand is displaced[159,160]. When this results in complete guide–target base-pairing (R-loop configuration), then the nuclease protein (class 1, Cas3 or Cas10) or nuclease domains (class 2, RuvC or HNH) become activated, resulting in target cleavage, and hence neutralization of the invading mobile genetic elements[159,160].



Classification

Although the gene and repeat sequences and their arrangements that make up CRISPR–Cas loci are very diverse, it is possible to distinguish two main classes of CRISPR–Cas systems, which together contain six types and more than 30 subtypes[10]. The CRISPR adaptation module is very well conserved in class 1 and class 2 systems, consisting of a Cas1–Cas2 core, sometimes associated with Cas4 (REF.[10]). Processing of the pre-crRNA by the CRISPR expression module proceeds either by a Cas ribonuclease subunit (class 1) or domain (class 2; for example, Cas12a and Cas13) or by a non-Cas ribonuclease in the case of a crRNA–*trans*-activating crRNA (tracrRNA) pair (class 2; for example, Cas9 and Cas12b). The interference module is completely different in the two classes. In class 1 systems, this module consists of Cascade-like multisubunit complexes that bind the crRNA guide, complemented by one or more nuclease proteins, either alone or as part of the complex[10] (see the figure, part **b**). In class 2 systems, crRNAs (sometimes with a tracrRNA) are bound by a single, multidomain protein that includes one or more nuclease domains: HNH (Cas9), RuvC (Cas9 and Cas12) or HEPN (Cas13)[10] (see the figure, part **b**). The spacer part of the crRNA guide targets a complementary nucleic acid sequence, either DNA (type I, type II and type V systems) or RNA (type III and type VI systems).

Collateral cleavage
Nuclease activity exhibited by some Cas proteins leading to indiscriminate degradation of any nearby non-target single-stranded DNA or RNA, respectively, upon target recognition.

HD domain
A nuclease domain with a conserved catalytic site that includes a metal-binding histidine–aspartate (HD) pair. The HD domain of Cas3 and Cas10 in type I and type III CRISPR–Cas systems, respectively, is responsible for endonucleolytic degradation of DNA targets.

Palm domain
A domain typically found in nucleotide cyclases and polymerases (as part of their fingers, palm and thumb-like architecture). The palm domain in the type III Cas10 proteins is characterized by a conserved GGDD motif, which catalyses the cyclase reaction to form cyclic oligoadenylate messenger molecules from ATP molecules.

CRISPR-associated Rossman fold (CARF) domain
A domain often found fused to an effector domain with (ribo)nuclease activity or other catalytic activities. The CARF domain acts as a sensory domain that binds ligands (for example, cyclic oligoadenylate messenger molecules produced by Cas10 in type III systems) that allosterically activate the fused effector domain.

Burst size
The number of newly synthesized phage particles released from a bacterium infected by a single phage.

Cases have been described that range from regulation of gene expression to defence during invasion, not only by bacteria and archaea but also by MGEs. For instance, a wide range of anti-CRISPR systems have been described in phages and archaeal viruses, and examples of a further stage in the arms race have been discovered: anti-anti-CRISPR[15]. In addition, for most CRISPR–Cas types, collateral cleavage of host nucleic acids occurs, leading to cell dormancy and sometimes even cell death. Moreover, CRISPR–Cas systems have been adopted by different MGEs to assist either in invading their host or in targeting competing invaders. Finally, Cas9 of an invasive pathogenic bacterium has been demonstrated to enter the nucleus of human cells and cause major damage to the DNA, resulting in cell death[16]. In this Review, we describe the various non-defence phenomena of CRISPR–Cas systems, revealing exciting variations in CRISPR evolution, with relevant consequences for safe CRISPR-based genome editing applications.

Collateral damage in CRISPR–Cas systems

All CRISPR–Cas systems share the concept of using RNA-guided protein complexes to confer sequence-specific cleavage activity generally aimed at clearing invading MGEs. In addition, however, many CRISPR–Cas types appear to play a role in protecting the population rather than the individual. The latter feature is accomplished through 'collateral' RNase or DNase activity with no apparent sequence specificity, either by the effector Cas nuclease itself or by activation of auxiliary Cas nucleases. As the collateral activity does not discriminate between self and non-self, the outcome for the host can be either bacteriostatic (for example, growth arrest or dormancy) or even lethal, as discussed later. These findings shifted the paradigm of CRISPR–Cas being an immune system that protects the host from viral predation to a system that potentially sacrifices the infected host. This strategy prevents the production of new phages, which, in turn, protects neighbouring cells, akin to 'abortive infection'[17].

Collateral nuclease activity in class 1 CRISPR–Cas systems. The first reports of non-specific nuclease activity associated with CRISPR–Cas immunity came from studies of type III CRISPR–Cas systems, which provide sophisticated and multilayered immunity. Type III systems use crRNA-guided, multisubunit Cas complexes that bind complementary RNA targets arising from transcription of the invading DNA[18,19]. Binding of the target RNA triggers three different catalytic activities within the effector complex: cleavage of the target RNA by Cas7 (REFS[20–24]); non-specific DNase activity by the HD domain of Cas10 (REFS[24,25]); and secondary messenger production by the palm domain of Cas10 (REFS[26–28]). The type III effector complex is thought to operate in close proximity to actively transcribed regions on the phage DNA, potentially limiting unnecessary damage to the host DNA by the non-specific DNase activity of Cas10. However, the production of the secondary messenger molecules might have indirect detrimental consequences for the producing host. These molecules (typically cyclic oligoadenylates (cOAs)) bind and activate proteins containing a CRISPR-associated Rossman fold (CARF) domain. Many of the CARF proteins appear to confer sequence-non-specific RNase or DNase activities and have been reported to promote cell death and dormancy[27,29,30] (FIG. 1). For example, upon binding cOAs, the type III-A CARF protein Csm6 degrades RNAs in a sequence-non-specific fashion. This collateral activity was essential when late phage gene transcripts were targeted to prevent completion of the lytic phase of replication[29]. For plasmid targeting, the collateral activity of Csm6 induced a growth arrest to prevent the invading plasmid from accumulating in the host[30]. From the findings taken together, it appears that Csm6 operates as a fallback mechanism in case targeting by the type III complex fails to clear the invader. As collateral activity can be detrimental for the host, cellular cOA levels are tightly regulated. This is achieved by temporal activation of the palm domain of Cas10 (ceasing cOA production when the cleaved target RNA dissociates from the type III effector complex), dedicated ring nucleases (that cleave cOAs) and/or intrinsic ring-nuclease activity associated with particular proteins of the CARF family[31–34].

The repertoire of predicted catalytic activities associated with CARF proteins includes nucleases, proteases, transcriptional regulators and adenine deaminases[32]. Whether all these different catalytic activities contribute to type III CRISPR–Cas immunity is currently unknown. In *Thermus thermophilus*, for example, the CARF protein Can1 nicks supercoiled DNA, thereby interfering with the replication kinetics of invading nucleic acids[35]. By contrast, the CARF protein Card1 from *Treponema succinifaciens* displays single-stranded, sequence-non-specific DNase and RNase activity, inducing cellular dormancy when targeting invading phages or plasmids[36].

Also type I systems may affect host fitness, as exemplified by the type I-E system of *Escherichia coli* and the type I-F system of *Pectobacterium atrosepticum*[37,38]. In these examples, phage infection of cells carrying a spacer targeting the phage resulted in cell death and reduced the burst size of infected individuals. Although the exact mechanism remains to be determined, the observed lethality could arise from the processive nuclease activity of Cas3 and/or irreversible phage-induced damage to the host[37] (FIG. 1).

Collateral nuclease activity in class 2 CRISPR–Cas systems. Collateral damage is also widespread among the single-subunit class 2 systems. The type VI systems specifically target RNA, via the crRNA-guided effector protein Cas13. Upon binding of a cognate RNA target, the two HEPN domains of Cas13 are rearranged to form a single active site that can cleave both target and non-target RNAs[39]. In *E. coli*, this promiscuous RNase activity of Cas13 has been reported to cause growth arrest in cells carrying a spacer targeting the foreign transcript[40,41] (FIG. 1). Similarly to the aforementioned class 1 examples, this growth arrest prevents phages from completing their replication cycle, thereby reducing the release of new phages[42].

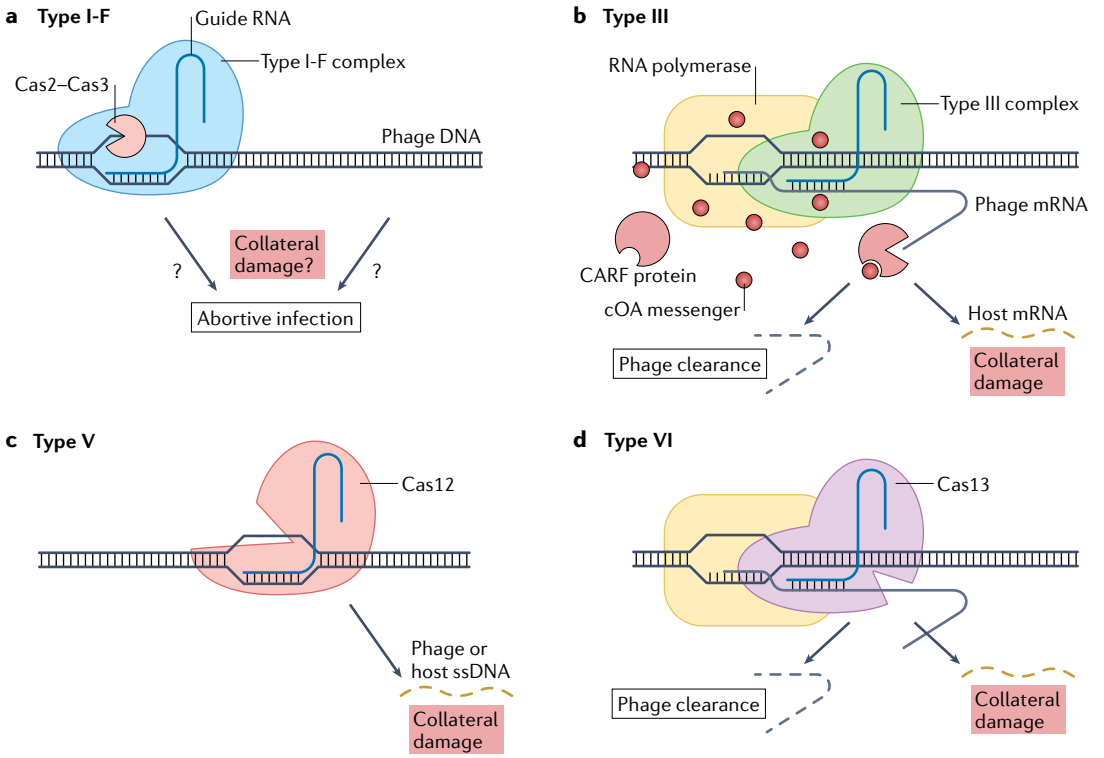


Fig. 1 | **Collateral damage associated with CRISPR–Cas systems. a** | Type I-F complex (blue) binding to its target DNA recruits the nuclease Cas2–Cas3 (red) to degrade viral DNA and triggers abortive infection through an unknown mechanism. **b** | Type III complex (green) binding to its target RNA arising from transcription of the viral DNA. This results in the production of cyclic oligoadenylate (cOA) messenger molecules (red circles) that allosterically activate sequence-non-specific CRISPR-associated Rossman fold (CARF) nucleases (light purple), cleaving both phage and host RNAs. **c** | Cas12 (pink) binding to its target DNA induces a conformational change that traps Cas12 in an active state with an accessible active site that can engage in non-specific cleavage of single-stranded DNAs (ssDNAs). **d** | Cas13 (purple) binds the transcribed viral target RNA, inducing a conformational change forming an accessible active site on the periphery of Cas13, cleaving both phage and host RNAs.

The type V-A CRISPR–Cas systems include the single-subunit protein Cas12a, which functions as a crRNA-guided nuclease. Cas12a has a bilobed clamp-like architecture consisting of a recognition lobe and a nuclease lobe. In Cas12a, the nuclease lobe bears a single RuvC domain that contains the active site for making staggered cuts in both the target strand and the non-target strand of a double-stranded (dsDNA) substrate. In-depth structural analyses have revealed how a single Cas12a active site (RuvC) can cleave both strands[43,44]. In short, target DNA binding to the crRNA induces a conformational change in Cas12a that initially results in RuvC-based cleavage of the displaced non-target strand, after which the guide-bound target strand is repositioned such that it can also be cleaved by the same RuvC domain. Thereafter, the Cas12a complex remains associated with the DNA in an active state, allowing sequence-non-specific cleavage of single-stranded DNA (FIG. 1). It is unclear how long the DNA–Cas12a complex remains in its active state and how relevant any non-specific DNA cleavage activity is for type V immunity. The activation of collateral activity upon target binding has also been demonstrated in other Cas12 variants[45,46].

The sequence-non-specific DNA cleavage activity of Cas12 variants appears not to hamper its application as a genome editing tool, probably because single-stranded DNA rarely occurs in cells. In the case of Cas13, collateral targeting of RNA may be an issue, although reports in the literature are somewhat conflicting[47,48]. For other applications, collateral activity can be very useful, such as the sensitive CRISPR–Cas nucleic acid detection platforms DETECTR and SHERLOCK, which are based on Cas12 and Cas13, respectively[45,49]. More recently, the first type III CRISPR–Cas-based detection platforms were developed[50,51].

From the findings taken together, next to the well-established role of CRISPR–Cas in providing sequence-specific protection against invading MGEs, many CRISPR–Cas systems seem to be equipped with a mechanism that can potentially kill the host by degrading essential biomolecules, such as DNA, RNA and likely also proteins[10,52]. Although the molecular details might differ from system to system, they all appear to be geared towards providing population protection in case rapid MGE clearance within the infected individual fails. This 'safety net' comes at the cost of sacrificing individuals, but indirectly benefits clone mates by limiting the phage pandemic. In turn, these dynamics influence how functional spacers are selected for within (sub)populations in a complex ecological setting.

Helix–turn–helix (HTH) domain
A widespread domain found in many proteins that bind DNA. The domain is characterized by two α-helices that bind the major groove of double-stranded DNA.

Counteracting and hijacking of CRISPR–Cas by MGEs

The CRISPR–Cas system can in theory be used by any organism (or MGE) that contains the encoding DNA and is able to either produce Cas proteins by itself or use hosts for the production of the encoded Cas proteins. In other cases, MGEs can encode factors that counteract the host-encoded CRISPR–Cas systems they encounter during invasion.

Anti-CRISPRs and anti-anti-CRISPRs. There are various ways by which phages and MGEs can circumvent or even interfere with CRISPR-based immunity. Phages can counteract CRISPR–Cas defence either passively through point mutations (in the PAM or seed, the sequence on the crRNA where base-pairing with the target is initiated) or deletions of CRISPR-targeted regions[53,54] or actively through using anti-CRISPR inhibitors[55–57] (FIG. 2). The first anti-CRISPR proteins (Acr proteins) were discovered in prophages in *Pseudomonas aeruginosa* that encode distinct Acr protein families, inhibiting either the type I-E CRISPR–Cas system or the type I-F CRISPR–Cas system[55,58]. Acr proteins have a wide range of mechanisms. They can bind directly to Cas proteins, preventing them from binding or cleaving phage DNA[56,57], prevent guide RNA loading[59], induce non-specific DNA binding[60,61], trigger Cas protein degradation[62] or degrade second messengers[63].

Since the identification of Acr proteins, 92 distinct families of Acr proteins[64], inhibiting type I, type II, type III, type V and type VI CRISPR–Cas systems, have been identified. The inhibition mechanisms are manifold, including blocking target binding by either DNA mimicry or steric occlusion, and preventing DNA cleavage by either obstructing the nuclease domain or enzymatic activity[57,65] (FIG. 2).

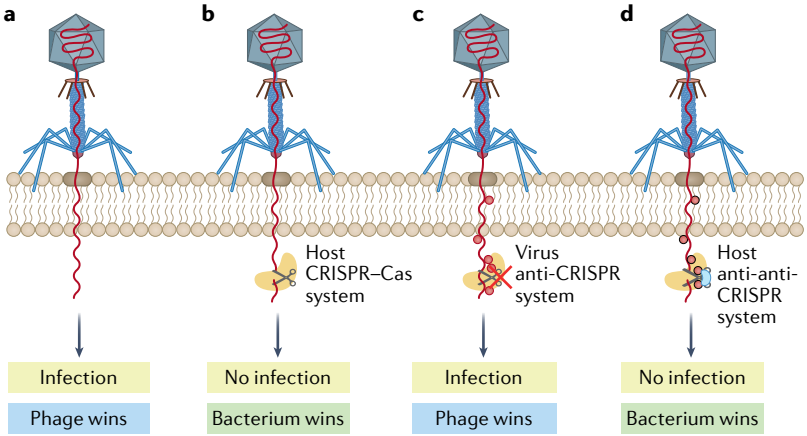


Fig. 2 | **CRISPR–Cas in the evolutionary arms race between phages and bacteria.** **a** | Lytic phage infecting a phage-sensitive host. Phage DNA enters the bacterial cell cytoplasm via an inner-membrane protein (brown) and the lytic cycle of the phage starts. Most bacterial cells will undergo lysis on completion of the phage lytic cycle. **b** | Phage infecting a phage-resistant cell harbouring an active CRISPR–Cas system (yellow). The phage genome is cleaved at specific sites by the Cas nuclease. **c** | The genome of a phage encoding anti-CRISPR proteins (Acr proteins) (red) that interact with different CRISPR–Cas components, preventing Cas proteins from binding or cleaving phage DNA. **d** | Some bacteria encode repressors (blue) that silence phage Acr expression, thus preventing infection by Acr-encoding phages.

An anti-CRISPR protein targeting Cas9 (AcrIIA1) in *Listeria monocytogenes* has a unique bifunctional mechanism[66]. The carboxy-terminal domain of AcrIIA1 binds and inactivates Cas9, whereas its amino-terminal helix–turn–helix (HTH) domain is a transcriptional repressor, silencing the strong Acr promoter. The full-length AcrIIA1 uses its two-domain architecture to sense Cas9 levels, tuning Acr expression accordingly. As expected in the course of an arms race, orthologues of the HTH domain also reside on bacterial genomes, often next to the *cas9* gene. The bacterial genes encode repressors of phage Acr expression, implying yet another layer in the arms race: an anti-anti-CRISPR system (FIG. 2). Another intriguing mechanism to prevent inactivation of a CRISPR–Cas system by MGEs was discovered in a bifunctional type I-B system of some archaea[15]. This system not only provides adaptive immunity but also acts as a transcriptional repressor of a toxin–antitoxin system. Apart from using crRNAs for defence, the Cascade complex uses the antitoxin RNA as a partially complementary guide to repress expression of the toxin gene. In case of inactivation of the Cascade-encoding operon by integration of transposons, the toxin RNA results in dormancy, and eventually cell death[15]. Again, counteracting the 'anti-CRISPR' transposon activity could be considered an anti-anti-CRISPR mechanism.

Phage-encoded CRISPR–Cas systems. Phages have been estimated to outnumber bacteria by a factor of 10 (REF.[67]), leading to a strong selection pressure for phages to compete not only with their hosts but also with other phages. Hence, it is not surprising that phages typically encode an extensive set of functions[68]. An unexpected feature of phages, however, was the discovery of CRISPR arrays that were encoded by prophages of *Clostridioides difficile*[69]. In a subsequent study of viromes from humans, CRISPR arrays were found in the genomes of enriched phage particles[70]. Analysis of the spacers of these CRISPR arrays led to the suggestion that the CRISPR–Cas system might participate in phage–phage competition[70]. Subsequently, a metagenomic study of a single human gut viral community revealed that the detected CRISPR spacers were highly dynamic over time[71]. However, the biological relevance of phage-encoded CRISPR–Cas systems remained elusive.

Soon thereafter, it was demonstrated that *Vibrio* phage ICP1, which infects *Vibrio cholera*, has 'hijacked' a complete CRISPR–Cas system for its own defence and persistence[72]. The type I-F CRISPR–Cas system of this phage contains a spacer that targets a phage-inducible chromosomal island-like element (PLE) on the genome of its bacterial host *V. cholerae* (FIG. 3a). This PLE resembles the *Staphylococcus aureus* pathogenicity island, which upon phage infection is excised from the genome, replicates and hijacks phage capsids from the invading phage for its own transduction[73]. Similarly to the *S. aureus* pathogenicity island, the *V. cholerae* PLE is activated, excised and packaged during phage infection. PLE activity is characterized by accelerated cell lysis and, instead of promoting release of new phage progeny, it contributes to the propagation of the PLE[74] (FIG. 3a). However, during infection by phages that encode a

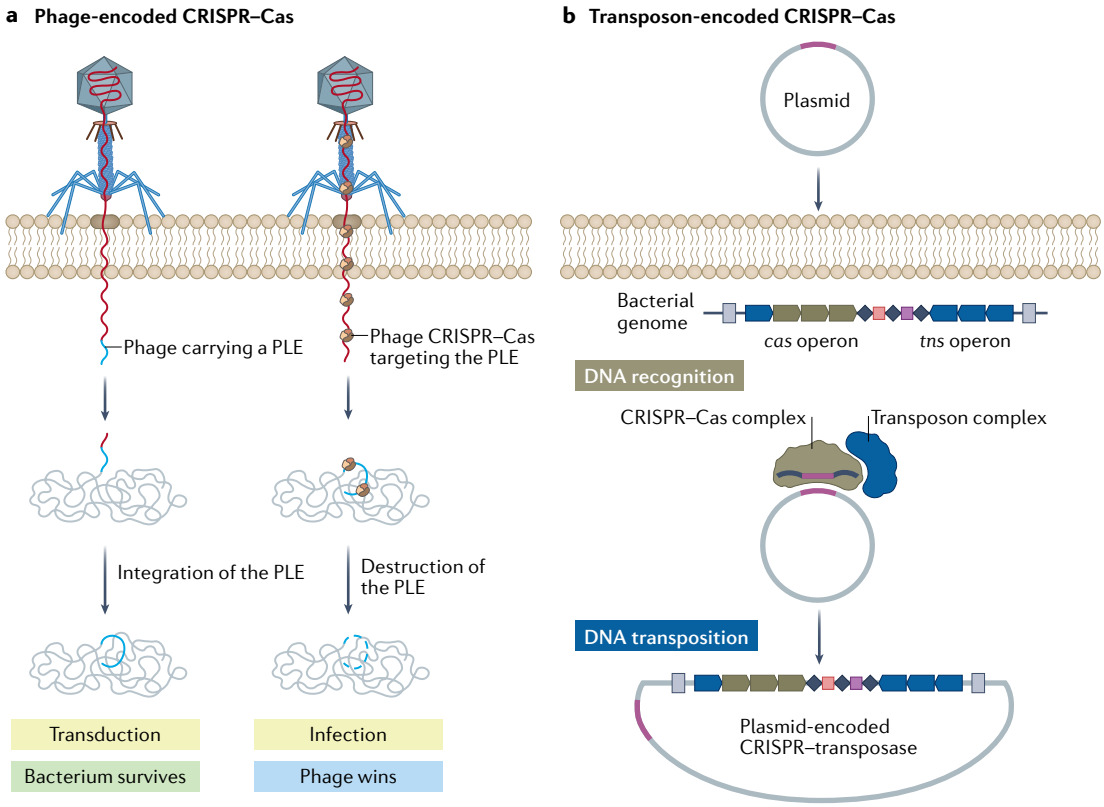


Fig. 3 | **Hijacking of CRISPR–Cas by mobile genetic elements. a** | Lytic *Vibrio* phage ICP1 infects *Vibrio cholerae*, leading to excision, replication and integration of the phage-inducible chromosomal island-like element (PLE). However, when ICP1-encoded CRISPR–Cas and CRISPR RNA (crRNA) complexes are expressed, the CRISPR–Cas machinery targets and inactivates the PLE, leading to phage propagation. **b** | CRISPR–Cas-mediated transposition. Upon entry of a foreign plasmid, the transposon-encoded CRISPR–Cas–crRNA complex in the host bacterial genome recruits the Tn7 transposon complex to a target site in the plasmid that is complementary to the crRNA sequence. This leads to site-specific transposition within the plasmid, close to the target site.

CRISPR–Cas system targeting the PLE, viral crRNAs and Cas proteins form CRISPR–Cas interference complexes that target and inactivate the antiviral defence system encoded by the PLE[72] (FIG. 3a). This type of interference allows the phage to complete its lytic replication cycle. Moreover, the ICP1 phage CRISPR–Cas system could adapt and acquire new spacers that target the PLE[72]. Sequencing of environmental samples indicates that more phages possess CRISPR arrays and/or *cas* genes[75,76]. Although the functionality of these systems awaits further investigation, hijacking of CRISPR–Cas systems by phages is a common strategy to promote their fitness.

Metagenomics studies have the power to discover uncultured 'novel' organisms and MGEs in diverse habitats, and to catalogue their genomic features. Metagenomics has led to the discovery of huge phages that are ubiquitous in a wide range of ecosystems[77]. The largest phage genomes, with lengths of more than 200 kb, encode tRNAs, tRNA synthetases, tRNA-modification enzymes, ribosomal proteins, translation initiation and elongation factors, and class 1 and class 2 CRISPR–Cas systems[77]. The phage-encoded class 2 systems include variants of type II (Cas9) and type V (Cas12i, Cas12f and uncharacterized type V-U systems). Some of these CRISPR–Cas systems were predicted to play relevant roles in the regulation of transcription and translation of their bacterial host, most likely redirecting host protein synthesis towards phage proteins. Other phage CRISPR–Cas systems have been predicted to cleave bacterial chromosomes, which may be advantageous during the phage infection cycle[77]. In addition, spacers with sequence identity to core structural and regulatory genes of other phages or prophages may suggest roles in targeting competing phages[77].

Most phage-encoded Cas effectors may have nuclease activity. For example, a novel type V effector protein, Cas12j encoded in the 'Biggiephage' clade, has been experimentally demonstrated to be capable of dsDNA cleavage[78,79]. Nevertheless, some appear to have an incomplete catalytic site, suggesting a role in target silencing rather than cleavage[77]. Notably, many of the phage-encoded CRISPR–Cas systems appear to lack CRISPR spacer acquisition machinery (Cas1, Cas2 and Cas4) and generally harbour compact CRISPR arrays (2–22 spacers)[77].

Other examples of phages with partial CRISPR–Cas systems are the CP8 and CP30A phages of *Campylobacter jejuni*, which encode a Cas4-like protein to stimulate acquisition of host-derived spacers by the *C. jejuni* type II-C CRISPR–Cas system, which lacks *cas4* (REFS[80,81]). Newly acquired spacers include self-targeting

spacers that might indirectly provide a benefit for phages, as they may exert a selective pressure on the host to inactivate its own CRISPR–Cas system or to tolerate changes in gene regulation[80].

Plasmid-encoded CRISPR–Cas systems. Phylogenetic analyses of some of the core components of CRISPR–Cas systems (that is, Cas1 and repeat sequences of the CRISPR array) as well as experimental studies suggest that CRISPR–Cas systems are frequently horizontally transferred[10,82–84]. Although some CRISPR–Cas systems have established a stable relationship with their respective host by integrating into the chromosome, many type I and most type IV CRISPR–Cas systems are still encoded on plasmids. A subset of these plasmid-encoded systems are minimalistic variants without the *cas* genes involved in adaptation and/or interference, suggesting that they may have evolved a function other than adaptive immunity (akin to the transposon-encoded systems described in more detail later). Of special interest are the elusive type IV CRISPR–Cas systems, which are almost exclusively located on plasmids. The type IV-A and type IV-B subtypes encode multisubunit Cascade-like effector complexes[85,86] and typically do not have *cas* genes required for adaptation or interference, and only type IV-A systems have associated CRISPR arrays. These systems may play a role in plasmid maintenance and/or interplasmid competition in a nuclease-independent manner. For example, the type IV-A system from *P. aeruginosa* utilizes the activity of a DinG helicase (Csf4) to confer RNA-guided plasmid interference in vivo[87]. Furthermore, the spacers of the type IV-A CRISPR arrays have a strong bias for targeting plasmid genes involved in their mobility, suggesting a role in excluding other plasmids from invading the same host[88]. This agrees well with the observation that type IV-A3 variants are exclusively encoded by special classes of conjugative plasmids.

Adoption of CRISPR-Cas by transposons. Several Tn*7*-like transposons have been found to be associated with short CRISPR arrays along with adaptation-deficient and nuclease-deficient type I-B, type I-F and type V-K CRISPR–Cas systems[89]. CRISPR arrays present on integrated transposons may contain combinations of spacers homologous to plasmid and phage sequences and, in some cases, bacterial chromosomal sequences adjacent to the transposon. The presence of *cas* genes responsible for target recognition suggested that these 'minimal' CRISPR–Cas systems may have been repurposed to allow guided transposition of these MGEs. The CRISPR–Cas-associated Tn*7*-like transposons contain typical transposition genes, including *tnsA*, *tnsB*, *tnsC* and *tnsQ* (a *tnsD* homologue)[89], similarly to the canonical Tn*7* heterotrimeric TnsABC complex[90,91], but lack *tnsD* and *tnsE*, which are typically responsible for the attachment site[92,93]. Indeed, experimental studies of the systems from *V. cholerae* and cyanobacteria have demonstrated that crRNAs specifically guide the Tn*7*-like transposase machinery to complementary genomic target sites when expressed in *E. coli*[94–96] (FIG. 3b). Directed transposition is mediated using a specialized guide RNA with an atypical repeat structure and mismatches, using a differentially regulated guide RNA or via the *tniQ*/*tnsD* pathway of canonical Tn*7* transposons[96,97]. The potential of these CRISPR-associated transposons to control the efficient introduction of large DNA fragments (at least up to 10 kb in length) may provide starting points for novel genome engineering tools[98,99].

Autoregulation of CRISPR–Cas systems

CRISPR–Cas immunity may coincide with increased levels of autoimmune toxicity. In archaeal *Sulfolobus* species, both the type I-A interference complex and a transcriptional regulator, Csa3b, are required for the repression of CRISPR–Cas interference[100]. Csa3b binds to the promoter region of the operon encoding the type I-A CRISPR–Cas system, facilitating binding of Cascade to the promoter region. Upon viral infection, redistribution of the Cascade complex onto crRNA-matching viral target sequences alleviates the transcriptional repression[100], triggering a fast transcriptional activation response and potentially a strong immune response to viral infection[100,101].

Cas9 systems use dual guides: a crRNA and a ('short') *trans*-activating crRNA (tracrRNA). Interestingly, Cas9 utilizes a 'long tracrRNA' to act as an autoregulator. The 5′ extension of the long tracrRNA substitutes the crRNA and has the potential to partially base-pair with the promoter upstream of the *cas9* gene. The relatively short region of guide–target complementarity (11 bp rather than 20 bp) does not allow activation of the Cas9 nuclease domains but is sufficient to block transcription initiation (TABLE 1). The *cas9* promoter serves as a crucial, regulatory node that controls the expression of all CRISPR–Cas components[102]. Derepression by an unknown mechanism causes a major induction of the expression of *cas* genes, crRNAs and tracrRNAs, resulting in a substantial increase in CRISPR adaptation and interference. Apart from saving on production costs, tight control of CRISPR–Cas expression may also prevent the acquisition of deleterious, self-targeting guides.

Regulation of endogenous genes by CRISPR–Cas

Even before the role of CRISPR–Cas systems in defence was demonstrated experimentally, it was predicted that some CRISPR–Cas systems could regulate prokaryotic genes via self-targeting guide RNAs and a mechanism that is analogous to RNA interference in eukaryotes[103]. Although type III and type VI systems use their crRNA guides to specifically target RNA, most CRISPR–Cas systems (types I, II and V) target dsDNA (BOX 1). In the latter case, the coupling between self-targeting guides and active nucleases poses the obvious problem of autoimmunity[104]. Still, multiple cases of regulatory roles of DNA-targeting CRISPR–Cas systems have been reported, affecting bacterial physiology in different ways[105].

Regulation of group behaviour. The predatory bacterium *Myxococcus xanthus* responds to starvation by halting swarming of cells and the aggregation of cells into fruiting bodies, resembling group behaviour (social behaviour) of cells[106]. During nutrient limitation,

trans-activating crRNA (tracrRNA). RNA encoded by all known type II and some type V CRISPR–Cas systems that includes an antirepeat part that base-pairs with the repeat portion of CRISPR RNA (crRNA) to form a functional guide RNA. tracrRNA is essential for crRNA maturation and target interference in the respective CRISPR–Cas systems.

Table 1 | **Summary of the different RNAs used by CRISPR–Cas systems for target cleavage and gene regulation**

Classification	Cas nuclease	Guide RNA	Second RNA	Target	Full or partial complementarity (guide–target)	PAM	Effect	Refs
Class 1								
Type I	Cas3	crRNA	None	dsDNA	Full	5′ PAM	Cleavage	7
	Cas3	crRNA	None	dsDNA (regulation of biofilm)	Partial	5′ PAM	Silencing	118
	Cas3	crRNA	None	RNA (*lasR*)	Partial (9 nt)	5′ PAM	Silencing	130
	Cas3	creA	None	dsDNA (*creT* promoter)	Partial (11 nt)	5′-PAM	Silencing	15
Type III	Cas10	crRNA	None	RNA	Full	rPAM	Cleavage	8
Class 2								
Type II	Cas9	crRNA	tracrRNA	dsDNA	Full	3′ PAM	Cleavage	148
	Cas9	5′ extended-tracrRNA	tracrRNA	dsDNA (*cas9* promoter)	Partial (11 nt)	3′ PAM	Silencing	102
	Cas9	scaRNA	tracrRNA	dsDNA (*blp* promoter)	Partial (11 nt)	3′ PAM	Silencing	140
Type V	Cas12a, Cas12f2, Cas12l or Cas12h	crRNA	None	dsDNA	Full	5′ PAM	Cleavage	46,149
	Cas12b, Cas12e, Cas12f1 or Cas12g	crRNA	tracrRNA	dsDNA	Full	5′ PAM	Cleavage	45,46,150–152
	Cas12c or Cas12d	crRNA	scoutRNA	dsDNA	Full	5′ PAM	Cleavage	153
Type VI	Cas13	crRNA	None	RNA	Full	PFS	Cleavage	40

creA, CRISPR RNA-resembling antitoxin RNA; crRNA, CRISPR RNA; dsDNA, double-stranded DNA; nt, nucleotide; PAM, protospacer adjacent motif; PFS, protospacer flanking sequence; rPAM, RNA protospacer adjacent motif; scaRNA, small CRISPR-associated RNA; scoutRNA, short-complementarity untranslated RNA; tracrRNA, *trans*-activating CRISPR RNA.

the *M. xanthus* CRISPR–Cas system can modulate expression of endogenous genes to regulate group behaviour, fruiting body formation and spore development[107]. Addition of nutrients to developing *M. xanthus* cells results in the degradation of the transcriptional activators MrpC and FruA belonging to the development (*dev*) operon, hence halting development and sporulation[108,109]. The *dev* operon also encodes a type I-C CRISPR–Cas system[103,110,111] composed of seven *cas* genes (also known as *dev* genes in *M. xanthus*) and a downstream CRISPR array with 22 spacers. Disruption of *cas7* (*devR*), *cas5* (*devS*) and *cas8c* (*devT*) of the *M. xanthus* CRISPR–Cas system markedly impairs sporulation, possibly due to decreased *fruA* transcript and FruA protein levels[111–113]. The mechanism by which the *dev* operon promotes sporulation involves downregulating the *devI* gene, which is an inhibitor of sporulation. Downregulation is achieved by a complex comprising three proteins: DevRST, which corresponds to a Cascade subcomplex consisting of Cas7, Cas5 and Cas8, respectively[114,115]. Most likely, the DevRST–crRNA complex regulates *M. xanthus* development through binding the target DNA without cleaving it, because of a suboptimal PAM, because of partial guide–target base-pairing (TABLE 1) or because of a disrupted nuclease activity. *M. xanthus* fruiting body formation is further influenced (directly or indirectly) by a type III-B CRISPR–Cas system, which also regulates exopolysaccharide production and type IV pilus-mediated chemotaxis (swarming behaviour). Both crRNA processing and the associated *cas* genes are required for this regulatory activity[116]. The study of gene regulatory functions of CRISPR–Cas systems is still in its infancy, and systematic analyses are needed to obtain a full appreciation of CRISPR–Cas regulatory functions.

Regulation of biofilm formation. CRISPR–Cas-mediated gene regulation by partial guide RNA–target sequence complementarity (TABLE 1) is supported by studies on the regulation of biofilm formation and swarming ability in *P. aeruginosa* UCBPP-PA14 (REFS[117–119]). Infection of a *P. aeruginosa* isolate with the Mu-like lysogenic phage DMS3 blocked biofilm formation and swarming motility[119]. Surprisingly, these phenotypical changes correlated with mutations in the *cas* genes of the type I-F CRISPR–Cas system[119] and required nuclease and helicase activities of Cas3 (REF.[117]). The proposed mechanism postulates that a partial complementarity between a crRNA and a gene of the DSM3 prophage might allow sufficiently strong binding of Cascade to induce moderate levels of DNA nicking by Cas3. These events and the concomitant DNA damage may trigger a DNA repair SOS response that blocks biofilm formation[118]. Likewise, Cas3 belonging to type I CRISPR–Cas systems of the pathogenic bacteria *Streptococcus mutans* and *Porphyromonas gingivalis* has been reported to play an important role in virulence and biofilm formation[120,121].

mRNA cleavage. Cas9 nucleases are traditionally known as DNA-targeting enzymes. However, RNA targeting by the *Streptococcus pyogenes* Cas9 was accomplished when combined with synthetic PAM-presenting oligonucleotides[122]. Later, it was found that some Cas9 variants[123,124] can target nucleic acids in a PAM-independent manner. For instance, *Neisseria meningitidis*

SOS response
A coordinated cellular response to genotoxic stress comprising an error-prone DNA repair system that allows restarting of stalled replication forks past lesions or errors.

Cas9 (type II- C) cleaves single-stranded RNA in vitro, whereas *S. aureus* Cas9 (type II-A) and *C. jejuni* Cas9 (CjeCas9, type II- C) cleave single-stranded RNA both in vitro and in vivo[123–125], potentially acting as an additional line of defence. Interestingly, CjeCas9 could use native sets of crRNA and tracrRNA to bind and, in some cases, to cleave endogenous RNAs. These crRNAs exhibit partial complementarity to their targets, and the associated DNA sequences are not flanked by appropriate PAMs, thereby preventing genomic cleavage[125,126]. Besides, the tracrRNA could hybridize to partially complementary cellular transcripts, leading to the biogenesis of 'non-canonical' crRNAs (ncrRNAs) with variable sizes capable of guiding CjeCas9 to target DNA. The ability of a reprogrammable tracrRNA to convert any cellular RNA into a functional Cas9–crRNA complex led to the development of LEOPARD (leveraging engineered tracrRNAs and on-target DNAs for parallel RNA detection), a platform for multiplexed RNA detection. Upon sensing of a target RNA, reprogrammable tracrRNA and ncrRNA enable the cleavage of fluorescent DNA sensors by Cas9. Notably, a physiological role for ncrRNAs in *C. jejuni* has not yet been reported[126].

Additionally, *S. aureus* Cas9 could inhibit gene expression through programmable RNA targeting in *E. coli*[124]. Further work is needed to outline the prevalence of RNA targeting by natural Cas9 variants from different bacteria, and to elucidate the pathways (if any) for how RNA targeting drives cell gene expression and metabolism.

Roles of CRISPR–Cas in bacterial virulence

So far, we have discussed the versatile role of CRISPR–Cas systems during the evolutionary arms race between MGEs and prokaryotes. However, CRISPR–Cas activity is also linked to the arms race between pathogenic bacteria and their eukaryotic hosts, through controlling virulence either directly or indirectly.

CRISPR–Cas-mediated gene regulation to promote virulence. Several notorious human, animal and plant pathogens possess CRISPR–Cas systems that are involved in controlling the pathogenicity process through endogenous gene regulation[127]. Deletion of the *cas3* gene from the type I-E system of *Salmonella enterica* subsp. *enterica* serovar Enteritidis decreases its capacity to infect murine, porcine and human cell lines, as well as chickens upon oral intubation[128]. The *S. enterica cas3* gene appears to (directly or indirectly) modulate the expression of genes involved in quorum sensing, the type III secretion system and the flagellum, thereby modulating virulence and biofilm formation of *S. enterica*[128]. An additional observation concerns the induction of *cas7* expression upon infection of human macrophages by *S. enterica* subsp. *enterica* serovar Typhi[129]. These findings may reflect a role of type I CRISPR–Cas systems in bacterial pathogenicity. A molecular explanation for an active role of type I systems in bacterial pathogenicity has been provided by an analysis of the type I-F CRISPR–Cas system in *P. aeruginosa* UCBPP-PA14. This system includes a spacer that bears partial

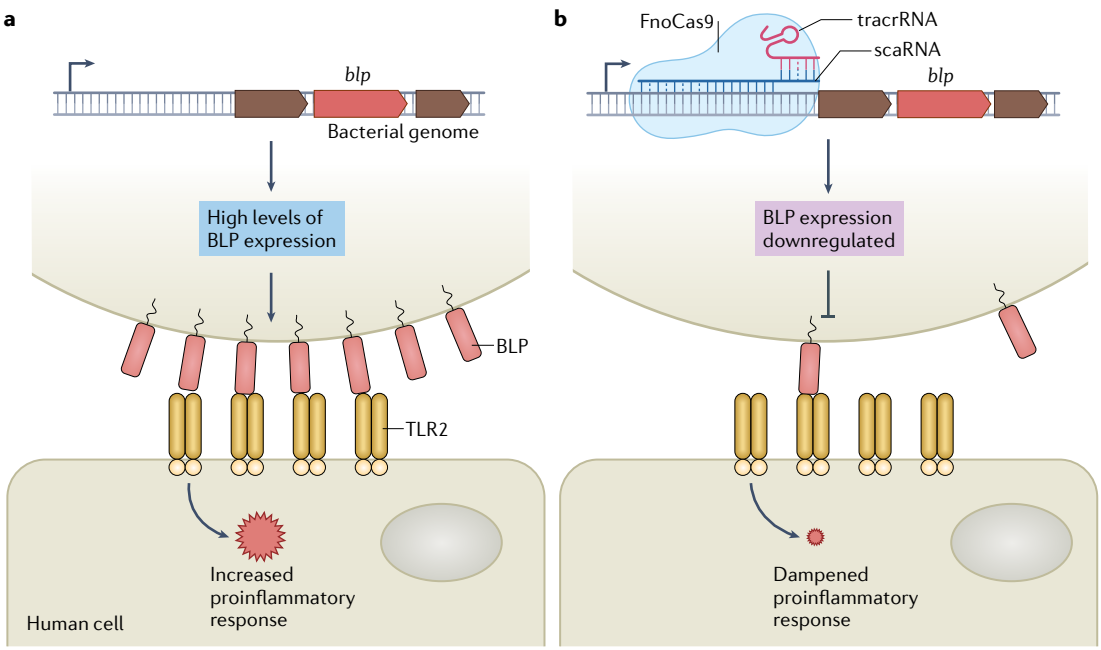


Fig. 4 | **Function of CRISPR–Cas in the regulation of bacterial virulence. a** | Bacterial membrane-exposed lipoprotein (BLP) is a virulence-associated cell surface protein. During in vivo infection by *Francisella novicida*, the presence of high levels of BLP on its cell surface is recognized by Toll-like receptor 2 (TLR2) on host human immune cells. This triggers a proinflammatory response aimed at clearing the pathogen. **b** | The *F. novicida* type II-B CRISPR–Cas system comprises four *cas* genes, a downstream CRISPR locus and a *trans*-activating CRISPR RNA (tracrRNA). Another small RNA, known as small CRISPR-associated (scaRNA), is transcribed from an independent promoter upstream of the CRISPR locus. Binding of the scaRNA–tracrRNA–*F. novicida* Cas9 (FnoCas9) complex to the bacterial lipoprotein gene (*blp*) leads to its transcriptional repression. Low levels of BLPs lead to a dampened TLR2-dependent proinflammatory response.

complementarity (8–12 nucleotides) to the chromosomally encoded *lasR* gene. LasR is a master regulator of bacterial quorum sensing and is responsible for modulating the expression of virulence-associated factors that may be recognized by Toll-like receptor 4 (TLR4) in mammals[130]. In this *P. aeruginosa* strain, virulence-associated CRISPR–Cas activity appears to function not through cleavage of chromosomal DNA but rather through specific crRNA-dependent binding and subsequent cleavage of the *lasR* mRNA[130]. Indeed, downregulation of this regulator led to a diminished proinflammatory host immune response in human cells and in mouse models. The target within the *lasR* mRNA spans nine continuous nucleotides (or 12 including 1 mismatch) that base-pair with the 3′ end of the crRNA guide. A three-nucleotide recognition sequence (5′-GGN) immediately upstream of the 8–12-nucleotide *lasR* target, as well as a functional Cas3 nuclease, is required for successful target mRNA degradation[130]. It would be of interest to extend this study to mRNA degradation by type I CRISPR–Cas systems in clinically relevant bacteria[131].

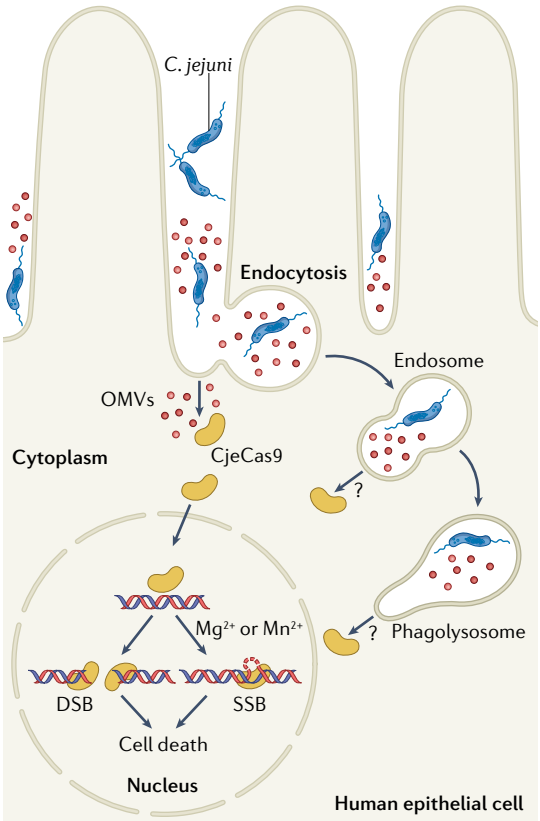


Fig. 5 | **Direct role of CRISPR–Cas in bacterial virulence.** The proposed model of Cas9-mediated DNA breaks in human cells leading to subsequent cell death. During *Campylobacter jejuni* infection of intestinal epithelial cells, *C. jejuni* Cas9 (CjeCas9) is likely released via the outer membrane vesicles (OMVs) that are taken up by host cells. During the invasion of intestinal epithelial cells, *C. jejuni* bacteria generate *C. jejuni*-containing vesicles (CCVs). During endosomal maturation of early and late CCVs, *C. jejuni* may also secrete CjeCas9. Upon release from the OMVs or CCVs into the cytoplasm, CjeCas9 can autonomously translocate into the nucleus. In the presence of Mg^{2+} and/or Mn^{2+} ions, the nuclease activity of CjeCas9 is activated, leading to CjeCas9-induced single-strand breaks (SSBs) and double-strand breaks (DSBs), leading to cell stress and death.

Type II CRISPR–Cas systems bearing the hallmark Cas9 protein are over-represented in the genomes of bacteria that interact with eukaryotic hosts[132–134]. In *S. pyogenes* GAS-M1T1-5448 (type II-A), deletion of the *cas9* gene leads to decrease in the master regulator protein Mga, which in turn downregulates ScpA and streptococcal inhibitor of complement (SIC) proteins, which are essential to inactivate the host complement immune defence[135]. Cas9 also downregulates the CovR–CovS two-component system in *S. pyogenes*, which regulates capsule genes that confer antiphagocytic properties[135,136]. In *Francisella novicida*, Cas9 (FnoCas9) transcriptionally modulates pathogenesis by repressing the expression of a bacterial membrane-exposed lipoprotein, thereby preventing its recognition by human TLR2 (FIG. 4). Indeed, reduced production of the bacterial membrane-exposed lipoprotein led to dampening of the TLR2-dependent immune response, thus promoting bacterial virulence[137]. The *F. novicida* type II-B CRISPR–Cas system regulates *blp* transcription through small CRISPR-associated RNA (scaRNA), which substitutes crRNA (TABLE 1) and tracrRNA[138–140] (FIG. 4). The scaRNA is a small RNA transcribed from an independent promoter near the CRISPR locus that acts similarly to crRNA: the 3′ half of the scaRNA hybridizes with the 5′ part of the tracrRNA, whereas the 5′ half of the scaRNA directs FnoCas9 to a regulon (FTN_1104-1101) that includes the *blp* gene. Binding of the scaRNA–tracrRNA–FnoCas9 complex to the corresponding transcripts leads to transcriptional repression of these genes, thus promoting virulence (FIG. 4). Robust and specific gene repression requires a PAM-dependent interaction of FnoCas9, and a perfect complementarity of 11 consecutive bases between scaRNA and the native 5′ untranslated regions (FIG. 4). As described earlier herein (autoregulation by long tracrRNA)[102], genomic DNA cleavage by FnoCas9 is likely prevented through partial complementarity of the scaRNA to the genomic target site[140]. Interestingly, upregulation of *blp* genes in *cas9*-deletion mutants has also been observed in *Streptococcus agalactiae* GD201008-001 (type II-A) and *Riemerella anatipestifer* (type II-C)[141,142].

In *N. meningitidis* and *C. jejuni*, similar potential roles have been reported for Cas9 in virulence, including adherence onto, invasion into and replication within human lung tissue, as well as translocation across human intestinal epithelial cells through endogenous gene regulation[138,143]. Moreover, it has been suggested that the amount of sialylated lipooligosaccharides on the outer surface of *C. jejuni* is controlled by its type II CRISPR–Cas system as well[143], but the molecular mechanism remains elusive.

Our understanding of the (indirect) contribution of CRISPR–Cas systems to bacterial virulence is starting to take shape, with new mechanisms still to be discovered. For instance, collateral DNA damage by the type III CRISPR–Cas system in *S. aureus* and *Staphylococcus*

Adeno-associated virus delivery
Transduction of genes to cells and organisms using adeno-associated viruses, which is generally considered safer than use of adenoviral and retroviral vectors. It can be used to transduce genes into both proliferating and non-proliferating cells, and can impart long-term expression in non-dividing cells.

Base editing
Genome editing technology that consists of a catalytically inactive CRISPR–Cas nuclease fused to a single-stranded DNA deaminase and, in some cases, to proteins that manipulate DNA repair machinery; cytosine base editors catalyse the conversion of C•G base pairs to T•A base pairs; and adenine base editors catalyse the conversion of A•T base pairs to G•C base pairs .

Prime editing
Genome editing technology based on the fusion proteins formed between a Cas9 nickase (inactivated HNH nuclease domain) and an engineered reverse transcriptase domain, including a synthetic single guide (prime editing guide RNA) consisting of CRISPR RNA, *trans*-activating CRISPR RNA and a prime editing extension.

epidermidis contributes to the high mutation rates (important for the rapid development of antibiotic resistance), which is characteristic of these bacterial pathogens[144].

Direct role in virulence. Apart from the aforementioned examples in which Cas9 indirectly controls virulence through modulating the expression of key genes to evade host immune systems, a direct role of Cas9 has also been described in the damage of host nuclear DNA[16]. Upon infection of human cells, *C. jejuni* secretes its guide-free Cas9 (CjeCas9) nuclease, either via bacterial outer membrane vesicles or upon lysis of the bacterial cell, into the cytoplasm of infected host cells (FIG. 5). Next, a native nuclear localization signal of CjeCas9 (probably within the well-conserved bridging helix domain) enables its transport to the nucleus, where it catalyses metal-dependent, non-specific DNA cleavage leading to cell death[16] (FIG. 5). This guide-free, non-specific cleavage activity of CjeCas9 agrees well with in vitro analyses[16,145]. Mn^{2+}-dependent or Mg^{2+}-dependent nicking of dsDNA has been demonstrated by guide-free FnoCas9, but not by *S. pyogenes* Cas9 (REFS[16,145]). Overall, it was concluded that CjeCas9 appears to play an active role in attacking the nuclear DNA of human cells, whereas many other Cas9 variants, including *S.* pyogenes Cas9, are likely mainly involved in protecting its bacterial host against invading MGEs.

Conclusions and outlook

The evolutionary arms race has been a major driver of evolution in general, and for the development of attack and defence systems in particular. In line with that, discoveries of new CRISPR–Cas systems are still being made, with major variations in terms of composition, mechanism and/or function[10]. The astonishing diversity is a beautiful example of evolution by tinkering[146], strongly suggesting that CRISPR–Cas systems are derived from counterparts of MGEs (transposons) and toxin–antitoxin systems, as well as host repair, recombination and sensing systems (nucleases, helicases and polymerases)[10,14,147]. Vice versa, CRISPR–Cas systems (or parts thereof) have been hijacked by different types of MGEs to increase their fitness in the course of the evolutionary arms race.

The well-known role in defence of bacteria and archaea against their parasites is generally based on cleavage of genomes of invading MGEs. On top of that, in this Review we explored many recent examples of CRISPR–Cas playing another important role, namely in the control of gene expression. A general trend, in both class 1 and class 2 systems, appears to be that partial guide–target base-pairing results in the binding of a target sequence, but not in its cleavage. Therefore, CRISPR–Cas can be used to control group behaviour of bacteria, for instance during stress and/or pathogenicity. In pathogenic bacteria, CRISPR–Cas systems have been found to coordinate evasion of the host (human) immune response to infection, or even to enter the nucleus of an infected (human) host cell and directly cause severe DNA damage. Apart from these host–parasite conflicts, CRISPR–Cas systems also may play a role in the competition between MGEs in their selfish attempts to prevent other phages and plasmids sharing the same host. In addition, a wide range of virus-encoded anti-CRISPR systems have been discovered that disrupt all types of CRISPR–Cas systems. Finally, the first examples of anti-anti-CRISPR systems were reported recently.

The reoccurring strategy of partial base-pairing by CRISPR–Cas systems to control gene expression should also be taken into account when one is analysing off-target effects, especially in the case of human gene therapy strategies. Moreover, the recently described damage caused by a guide-free Cas9 variant[16] emphasizes that guide saturation of Cas nucleases may be key for therapeutic applications. Another recent discovery that might have major implications for in vivo genome editing in humans is the wide variety of small variants of Cas12. Some of these systems may turn out to be well suited for adeno-associated virus delivery, either of the natural enzymes or of synthetic base editing and prime editing variants. It is anticipated that the ongoing multidisciplinary exploration of CRISPR–Cas systems will result in a continuous gain of fundamental insights, some of which may serve as a basis for innovative applications soon.

Published online: 06 January 2022

1. Koonin, E. V. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Res* **5**, 1805 (2016).
2. Vogan, A. A. & Higgs, P. G. The advantages and disadvantages of horizontal gene transfer and the emergence of the first species. *Biol. Direct* **6**, 1 (2011).
3. Tock, M. R. & Dryden, D. T. The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* **8**, 466–472 (2005).
4. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).
5. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
6. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
7. Brouns, S. J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
This article provides the first insights into the molecular mechanism of antiviral defence by CRISPR–Cas, including the use of designed CRISPRs for dedicated DNA interference.
8. Elmore, J. R. et al. Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR-Cas system. *Genes Dev.* **30**, 447–459 (2016).
9. Smargon, A. A. et al. Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630.e7 (2017).
10. Makarova, K. S. et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
This article is the latest published overview of CRISPR–Cas functional diversity, with emphasis on novel class 2 CRISPR–Cas systems, derived CRISPR–Cas variants and ancillary CRISPR-linked genes.
11. Koonin, E. V. & Makarova, K. S. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol.* **10**, 679–686 (2013).
12. van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **12**, 479–492 (2014).
13. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
14. Mohanraju, P. et al. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
15. Li, M. et al. Toxin-antitoxin RNA pairs safeguard CRISPR-Cas systems. *Science* **372**, eabe5601 (2021).
16. Saha, C. et al. Guide-free Cas9 from pathogenic Campylobacter jejuni bacteria causes severe damage to DNA. *Sci. Adv.* **6**, eaaz4849 (2020).
This study demonstrates that guide-free Cas9 from *C. jejuni* causes guide-independent, non-specific host-cell DNA damage.
17. Lopatina, A., Tal, N. & Sorek, R. Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).
18. 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).
19. Samai, P. et al. Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. *Cell* **161**, 1164–1174 (2015).
20. Benda, C. et al. Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. *Mol. Cell* **56**, 43–54 (2014).

21. Hale, C. R., Cocozaki, A., Li, H., Terns, R. M. & Terns, M. P. Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. *Genes Dev.* **28**, 2432–2443 (2014).
22. Tamulaitis, G. et al. Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. *Mol. Cell* **56**, 506–517 (2014).
23. Zhu, X. & Ye, K. Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. *Nucleic Acids Res.* **43**, 1257–1267 (2015).
24. Estrella, M. A., Kuo, F. T. & Bailey, S. RNA-activated DNA cleavage by the type III-B CRISPR-Cas effector complex. *Genes Dev.* **30**, 460–470 (2016).
25. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C. & Siksnys, V. Spatiotemporal control of type III-A CRISPR-cas immunity: coupling DNA degradation with the target RNA recognition. *Mol. Cell* **62**, 295–306 (2016).
26. Zhu, Y., Klompe, S. E., Vlot, M., van der Oost, J. & Staals, R. H. J. Shooting the messenger: RNA-targetting CRISPR-Cas systems. *Biosci. Rep.* **38**, BSR20170788 (2018).
27. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605–609 (2017).
28. Niewoehner, O. et al. Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543–548 (2017).
Together with Kazlauskiene et al. (2017), this study demonstrates the involvement of secondary messenger molecules (cOAs) in type III CRISPR–Cas immunity.
29. 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).
30. Rostol, J. T. & Marraffini, L. A. Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat. Microbiol.* **4**, 656–662 (2019).
31. Athukoralage, J. S., Rouillon, C., Graham, S., Gruschow, S. & White, M. F. Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277–280 (2018).
32. Makarova, K. S. et al. Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antivirus defense. *Nucleic Acids Res.* **48**, 8828–8847 (2020).
33. Jia, N., Jones, R., Yang, G., Ouerfelli, O. & Patel, D. J. CRISPR-Cas III-A Csm6 CARF domain is a ring nuclease triggering stepwise cA4 Cleavage with ApA>p formation terminating RNase activity. *Mol. Cell* **75**, 944–956.e6 (2019).
34. Athukoralage, J. S., Graham, S., Gruschow, S., Rouillon, C. & White, M. F. A type III CRISPR ancillary ribonuclease degrades its cyclic oligoadenylate activator. *J. Mol. Biol.* **431**, 2894–2899 (2019).
35. McMahon, S. A. et al. Structure and mechanism of a type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. *Nat. Commun.* **11**, 500 (2020).
36. Rostol, J. T. et al. The Card1 nuclease provides defence during type III CRISPR immunity. *Nature* **590**, 624–629 (2021).
37. Watson, B. N. J. et al. Type I-F CRISPR-Cas resistance against virulent phages results in abortive infection and provides population-level immunity. *Nat. Commun.* **10**, 5526 (2019).
38. Strotskaya, A. et al. The action of Escherichia coli CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Res.* **45**, 1946–1957 (2017).
39. Liu, L. et al. Two Distant catalytic sites are responsible for C2c2 RNase activities. *Cell* **168**, 121–134.e12 (2017).
40. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
41. Meeske, A. J. & Marraffini, L. A. RNA guide complementarity prevents self-targeting in type VI CRISPR systems. *Mol. Cell* **71**, 791–801.e3 (2018).
42. Meeske, A. J., Nakandakari-Higa, S. & Marraffini, L. A. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* **570**, 241–245 (2019).
43. Stella, S. et al. Conformational activation promotes CRISPR-Cas12a catalysis and resetting of the endonuclease activity. *Cell* **175**, 1856–1871.e21 (2018).
44. Swarts, D. C. & Jinek, M. Mechanistic insights into the cis- and trans-acting DNase activities of Cas12a. *Mol. Cell* **73**, 589–600.e4 (2019).
45. Harrington, L. B. et al. Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* **362**, 839–842 (2018).
46. Yan, W. X. et al. Functionally diverse type V CRISPR-Cas systems. *Science* **363**, 88–91 (2019).
47. Abudayyeh, O. O. et al. RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284 (2017).
48. Wang, Q. et al. The CRISPR-Cas13a gene-editing system induces collateral cleavage of RNA in glioma cells. *Adv. Sci.* **6**, 1901299 (2019).
49. Gootenberg, J. S. et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **356**, 438–442 (2017).
50. Santiago-Frangos, A. et al. Intrinsic signal amplification by type III CRISPR-Cas systems provides a sequence-specific SARS-CoV-2 diagnostic. *Cell Rep. Med.* **2**, 100319 (2021).
51. Steens, J. A. et al. SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation. *Nat. Commun.* **12**, 5033 (2021).
52. van Beljouw, S. P. B. et al. The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. *Science* **373**, 1349–1353 (2021).
53. Deveau, H., Garneau, J. E. & Moineau, S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* **64**, 475–493 (2010).
54. Semenova, E. et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. USA* **108**, 10098–10103 (2011).
55. 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).
This article presents the discovery of phage-encoded protein inhibitors of CRISPR–Cas systems, called 'Acr proteins'.
56. Bondy-Denomy, J. et al. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
57. Trasanidou, D. et al. Keeping CRISPR in check: diverse mechanisms of phage-encoded anti-CRISPRS. *Fems Microbiol. Lett.* **366**, fnz098 (2019).
58. Pawluk, A., Bondy-Denomy, J., Cheung, V. H., 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 (2014).
59. Thavalingam, A. et al. Inhibition of CRISPR-Cas9 ribonucleoprotein complex assembly by anti-CRISPR AcrIIC2. *Nat. Commun.* **10**, 2806 (2019).
60. Hirschi, M. et al. AcrIF9 tethers non-sequence specific dsDNA to the CRISPR RNA-guided surveillance complex. *Nat. Commun.* **11**, 2730 (2020).
61. Lu, W. T., Trost, C. N., Muller-Esparza, H., Randau, L. & Davidson, A. R. Anti-CRISPR AcrIF9 functions by inducing the CRISPR-Cas complex to bind DNA non-specifically. *Nucleic Acids Res.* **49**, 3381–3393 (2021).
62. Osuna, B. A. et al. Listeria phages induce Cas9 degradation to protect lysogenic genomes. *Cell Host Microbe* **28**, 31–40.e9 (2020).
63. Athukoralage, J. S. et al. An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. *Nature* **577**, 572–575 (2020).
64. Bondy-Denomy, J. et al. A unified resource for tracking anti-CRISPR names. *Crispr J.* **1**, 304–305 (2018).
65. Wiegand, T., Karambelkar, S., Bondy-Denomy, J. & Wiedenheft, B. Structures and strategies of anti-CRISPR-mediated immune suppression. *Annu. Rev. Microbiol.* **74**, 21–37 (2020).
66. Osuna, B. A. et al. Critical anti-CRISPR locus repression by a bi-functional Cas9 inhibitor. *Cell Host Microbe* **28**, 23–30.e5 (2020).
67. Stern, A. & Sorek, R. The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43–51 (2011).
68. Refardt, D. Within-host competition determines reproductive success of temperate bacteriophages. *ISME J.* **5**, 1451–1460 (2011).
69. Sebaihia, M. et al. The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nat. Genet.* **38**, 779–786 (2006).
70. Minot, S. et al. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* **21**, 1616–1625 (2011).
71. Minot, S. et al. Rapid evolution of the human gut virome. *Proc. Natl Acad. Sci. USA* **110**, 12450–12455 (2013).
72. 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. *Nature* **494**, 489–491 (2013).
This study, for the first time, identifies a CRISPR–Cas system encoded in a phage that is used to disarm a phage-inhibitory chromosomal island in its bacterial host.
73. Ram, G. et al. Staphylococcal pathogenicity island interference with helper phage reproduction is a paradigm of molecular parasitism. *Proc. Natl Acad. Sci. USA* **109**, 16300–16305 (2012).
74. O'Hara, B. J., Barth, Z. K., McKitterick, A. C. & Seed, K. D. A highly specific phage defense system is a conserved feature of the Vibrio cholerae mobilome. *PLoS Genet.* **13**, e1006838 (2017).
75. Bellas, C. M., Anesio, A. M. & Barker, G. Analysis of virus genomes from glacial environments reveals novel virus groups with unusual host interactions. *Front. Microbiol.* **6**, 656 (2015).
76. Chenard, C., Wirth, J. F. & Suttle, C. A. Viruses infecting a freshwater filamentous cyanobacterium (Nostoc sp.) encode a functional CRISPR array and a proteobacterial DNA polymerase B. *mBio* **7**, e00667-16 (2016).
77. Al-Shayeb, B. et al. Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).
78. Pausch, P. et al. CRISPR-CasPhi from huge phages is a hypercompact genome editor. *Science* **369**, 333–337 (2020).
79. Carabias, A. et al. Structure of the mini-RNA-guided endonuclease CRISPR-Cas12j3. *Nat. Commun.* **12**, 4476 (2021).
80. Hooton, S. P. T. & Connerton, I. F. Campylobacter jejuni acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. *Front. Microbiol.* **5**, 744 (2015).
81. Hooton, S. P., Brathwaite, K. J. & Connerton, I. F. The bacteriophage carrier state of campylobacter jejuni features changes in host non-coding RNAs and the acquisition of new host-derived CRISPR spacer sequences. *Front. Microbiol.* **7**, 355 (2016).
82. Godde, J. S. & Bickerton, A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* **62**, 718–729 (2006).
83. Varble, A., Meaden, S., Barrangou, R., Westra, E. R. & Marraffini, L. A. Recombination between phages and CRISPR-cas loci facilitates horizontal gene transfer in staphylococci. *Nat. Microbiol.* **4**, 956–963 (2019).
84. Watson, B. N. J., Staals, R. H. J. & Fineran, P. C. CRISPR-Cas-mediated phage resistance enhances horizontal gene transfer by transduction. *mBio* 9, e02406-17 (2018).
85. Ozcan, A. et al. Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat. Microbiol.* **4**, 89–96 (2019).
86. Zhou, Y. et al. Structure of a type IV CRISPR-Cas ribonucleoprotein complex. *iScience* **24**, 102201 (2021).
87. Crowley, V. M. et al. A type IV-A CRISPR-cas system in pseudomonas aeruginosa mediates RNA-guided plasmid interference. *Vivo. Crispr J.* **2**, 434–440 (2019).
88. Pinilla-Redondo, R. et al. Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids. *Nucleic Acids Res.* **48**, 2000–2012 (2020).
89. Peters, J. E., Makarova, K. S., Shmakov, S. & Koonin, E. V. Recruitment of CRISPR-Cas systems by Tn7-like transposons. *Proc. Natl Acad. Sci. USA* **114**, E7358–E7366 (2017).
This article provides the first description of derived CRISPR–Cas systems encoded in Tn*7*-like transposons and predicts their function in RNA-guided transposition.
90. Sarnovsky, R. J., May, E. W. & Craig, N. L. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. *EMBO J.* **15**, 6348–6361 (1996).
91. Peters, J. E. & Craig, N. L. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. *Genes Dev.* **15**, 737–747 (2001).
92. Waddell, C. S. & Craig, N. L. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. *Genes Dev.* **2**, 137–149 (1988).

93. Waddell, C. S. & Craig, N. L. Tn7 transposition: recognition of the attTn7 target sequence. *Proc. Natl Acad. Sci. USA* **86**, 3958–3962 (1989).
94. Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S. & Sternberg, S. H. Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219–225 (2019).
Complementary to Strecker et al. (2019), this work reveals that interference-deficient subtype I-F CRISPR–Cas systems encoded in Tn7-like transposons enable crRNA-guided transposition.
95. Strecker, J. et al. RNA-guided DNA insertion with CRISPR-associated transposases. *Science* **365**, 48–53 (2019).
Complementary to Klompe et al. (2019), this work reveals that inactivated CRISPR–Cas subtype V-U5 effector proteins encoded in Tn7-like transposons form a complex with the transposase subunit and enable crRNA-guided transposition.
96. Saito, M. et al. Dual modes of CRISPR-associated transposon homing. *Cell* **184**, 2441–2453.e18 (2021).
97. Petassi, M. T., Hsieh, S. C. & Peters, J. E. Guide RNA categorization enables target site choice in Tn7-CRISPR-Cas transposons. *Cell* **183**, 1757–1771. e18 (2020).
98. Vo, P. L. H. et al. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat. Biotechnol.* **39**, 480–489 (2020).
99. Rubin, B. E. et al. Targeted genome editing of bacteria within microbial communities. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.17.209189 (2020).
100. He, F., Vestergaard, G., Peng, W., She, Q. & Peng, X. CRISPR-Cas type I-A cascade complex couples viral infection surveillance to host transcriptional regulation in the dependence of Csa3b. *Nucleic Acids Res.* **45**, 1902–1913 (2017).
101. Quax, T. E. et al. Massive activation of archaeal defense genes during viral infection. *J. Virol.* **87**, 8419–8428 (2013).
102. Workman, R. E. et al. A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression. *Cell* **184**, 675–688.e19 (2021).
103. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct.* **1**, 7 (2006).
104. Stern, A., Keren, L., Wurtzel, O., Amitai, G. & Sorek, R. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* **26**, 335–340 (2010).
105. Sampson, T. R. & Weiss, D. S. CRISPR-Cas systems: new players in gene regulation and bacterial physiology. *Front. Cell Infect. Microbiol.* **4**, 37 (2014).
106. Jelsbak, L. & Sogaard-Andersen, L. Pattern formation: fruiting body morphogenesis in Myxococcus xanthus. *Curr. Opin. Microbiol.* **3**, 637–642 (2000).
107. Kaiser, D., Robinson, M. & Kroos, L. Myxobacteria, polarity, and multicellular morphogenesis. *Cold Spring Harb. Perspect. Biol.* **2**, a000380 (2010).
108. Rajagopalan, R. & Kroos, L. Nutrient-regulated proteolysis of MrpC halts expression of genes important for commitment to sporulation during Myxococcus xanthus development. *J. Bacteriol.* **196**, 2736–2747 (2014).
109. Campbell, A. et al. Combinatorial regulation of the dev operon by MrpC2 and FruA during Myxococcus xanthus development. *J. Bacteriol.* **197**, 240–251 (2015).
110. SogaardAndersen, L., Slack, F. J., Kimsey, H. & Kaiser, D. Intercellular C-signaling in Myxococcus xanthus involves a branched signal transduction pathway. *Genes Dev.* **10**, 740–754 (1996).
111. Viswanathan, P., Murphy, K., Julien, B., Garza, A. G. & Kroos, L. Regulation of dev, an operon that includes genes essential for Myxococcus xanthus development and CRISPR-associated genes and repeats. *J. Bacteriol.* **189**, 3738–3750 (2007).
112. Boysen, A., Ellehauge, E., Julien, B. & Sogaard-Andersen, L. The DevT protein stimulates synthesis of FruA, a signal transduction protein required for fruiting body morphogenesis in Myxococcus xanthus. *J. Bacteriol.* **184**, 1540–1546 (2002).
113. Thonymeyer, L. & Kaiser, D. Devrs, an autoregulated and essential genetic-locus for fruiting body development in myxococcus-xanthus. *J. Bacteriol.* **175**, 7450–7462 (1993).
114. Rajagopalan, R., Wielgoss, S., Lippert, G., Velicer, G. J. & Kroos, L. devI Is an evolutionarily young negative regulator of Myxococcus xanthus development. *J. Bacteriol.* **197**, 1249–1262 (2015).
115. Rajagopalan, R. & Kroos, L. The dev operon regulates the timing of sporulation during Myxococcus xanthus development. *J. Bacteriol.* **199**, e00788-16 (2017).
116. Wallace, R. A., Black, W. P., Yang, X. S. & Yang, Z. M. A CRISPR with roles in Myxococcus xanthus development and exopolysaccharide production. *J. Bacteriol.* **196**, 4036–4043 (2014).
117. 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).
118. 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).
119. 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).
120. Tang, B. et al. Deletion of cas3 gene in Streptococcus mutans affects biofilm formation and increases fluoride sensitivity. *Arch. Oral. Biol.* **99**, 190–197 (2019).
121. Solbiati, J., Duran-Pinedo, A., Godoy Rocha, F., Gibson, F. C. 3rd & Frias-Lopez, J. Virulence of the pathogen Porphyromonas gingivalis is controlled by the CRISPR-Cas protein Cas3. *mSystems* **5**, e00852-20 (2020).
122. O'Connell, M. R. et al. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263–266 (2014).
123. Rousseau, B. A., Hou, Z., Gramelspacher, M. J. & Zhang, Y. Programmable RNA cleavage and recognition by a natural CRISPR-Cas9 system from Neisseria meningitidis. *Mol. Cell* **69**, 906–914.e4 (2018).
124. Strutt, S. C., Torrez, R. M., Kaya, E., Negrete, O. A. & Doudna, J. A. RNA-dependent RNA targeting by CRISPR-Cas9. *eLife* **7**, e32724 (2018).
125. Dugar, G. et al. CRISPR RNA-dependent binding and cleavage of endogenous RNAs by the Campylobacter jejuni Cas9. *Mol. Cell* **69**, 893–905.e7 (2018).
126. Jiao, C. et al. Noncanonical crRNAs derived from host transcripts enable multiplexable RNA detection by Cas9. *Science* **372**, 941–948 (2021).
127. Medina-Aparicio, L., Davila, S., Rebollar-Flores, J. E., Calva, E. & Hernandez-Lucas, I. The CRISPR-Cas system in Enterobacteriaceae. *Pathog. Dis.* **76**, fty002 (2018).
128. Cui, L. et al. CRISPR-cas3 of Salmonella upregulates bacterial biofilm formation and virulence to host cells by targeting quorum-sensing systems. *Pathogens* **9**, 53 (2020).
129. Faucher, S. P., Curtiss, R. III & Daigle, F. Selective capture of Salmonella enterica serovar typhi genes expressed in macrophages that are absent from the Salmonella enterica serovar Typhimurium genome. *Infect. Immun.* **73**, 5217–5221 (2005).
130. 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).
131. Muller-Esparza, H. & Randau, L. Commentary: type I CRISPR-Cas targets endogenous genes and regulates virulence to evade mammalian host immunity. *Front. Microbiol.* **8**, 319 (2017).
132. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**, e60 (2005).
133. Makarova, K. S. et al. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477 (2011).
134. Chylinski, K., Le Rhun, A. & Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* **10**, 726–737 (2013).
135. Gao, N. J. et al. Functional and proteomic analysis of Streptococcus pyogenes virulence upon loss of its native Cas9 nuclease. *Front. Microbiol.* **10**, 1967 (2019).
136. Sarkar, P. & Sumby, P. Regulatory gene mutation: a driving force behind group a Streptococcus strain- and serotype-specific variation. *Mol. Microbiol.* **103**, 576–589 (2017).
137. Jones, C. L., Sampson, T. R., Nakaya, H. I., Pulendran, B. & Weiss, D. S. Repression of bacterial lipoprotein production by Francisella novicida facilitates evasion of innate immune recognition. *Cell Microbiol.* **14**, 1531–1543 (2012).
138. 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. *Nature* **497**, 254–257 (2013).
139. Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y. L. & Weiss, D. S. Author correction: a CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* **570**, E30–E31 (2019).
140. Ratner, H. K. et al. Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence. *Mol. Cell* **75**, 498–510.e5 (2019).
141. Ma, K. et al. cas9 enhances bacterial virulence by repressing the regR transcriptional regulator in streptococcus agalactiae. *Infect. Immun.* **86**, e00552-17 (2018).
142. Wang, Y. et al. Cas9 regulated gene expression and pathogenicity in Riemerella anatipestifer. *Microb. Pathog.* **136**, 103706 (2019).
143. Louwen, R. et al. A novel link between Campylobacter jejuni bacteriophage defence, virulence and Guillain-Barre syndrome. *Eur. J. Clin. Microbiol. Infect. Dis.* **32**, 207–226 (2013).
144. Mo, C. Y. et al. Type III-A CRISPR immunity promotes mutagenesis of staphylococci. *Nature* **592**, 611–615 (2021).
145. Sundaresan, R., Parameshwaran, H. P., Yogesha, S. D., Keilbarth, M. W. & Rajan, R. RNA-independent DNA cleavage activities of Cas9 and Cas12a. *Cell Rep.* **21**, 3728–3739 (2017).
146. Jacob, F. Evolution and tinkering. *Science* **196**, 1161–1166 (1977).
147. Koonin, E. V. & Makarova, K. S. Origins and evolution of CRISPR-Cas systems. *Phil. Trans. R. Soc. B* **374**, 20180087 (2019).
148. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
149. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
150. Liu, J. J. et al. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* **566**, 218–223 (2019).
151. Shmakov, S. et al. Discovery and functional characterization of diverse class 2 CRISPR-cas systems. *Mol. Cell* **60**, 385–397 (2015).
152. Karvelis, T. et al. PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Res.* **48**, 5016–5023 (2020).
153. Harrington, L. B. et al. A scoutRNA is required for some type V CRISPR-Cas systems. *Mol. Cell* **79**, 416–424.e5 (2020).
154. Levy, A. et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520**, 505–510 (2015).
155. Dupuis, M. E., Villion, M., Magadan, A. H. & Moineau, S. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat. Commun.* **4**, 2087 (2013).
156. Arslan, Z., Hermanns, V., Wurm, R., Wagner, R. & Pul, U. Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Res.* **42**, 7884–7893 (2014).
157. Nunez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193–198 (2015).
158. Garneau, J. E. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
159. Wiedenheft, B. et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl Acad. Sci. USA* **108**, 10092–10097 (2011).
160. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).

Acknowledgements
The authors acknowledge financial support from the Dutch Research Council (NWO) (NWO TOP-grant 714.015.001) and the European Research Council (ERC-AdG-834279) to J.v.d.O. R.H.J.S was supported by the NWO (VIDI grant VI.Vidi.203.074). C.S. is a graduate student at the Erasmus Postgraduate School of Molecular Medicine and is partially supported by the I&I Fund (Erasmus Vrienden Fonds) and LSH-TKI foundation grant LSHM18006. R.L. is supported by the Department of Medical Microbiology and Infectious Diseases and the Department of Bioinformatics, Erasmus University Medical Center.

Author contributions
P.M., R.H.J.S. and J.v.d.O. researched data for the article. All authors contributed substantially to the discussion of the content. P.M., R.H.J.S. and J.v.d.O. wrote the article. All authors reviewed/edited the manuscript before submission.

Competing interests
J.v.d.O. is scientific advisor of NTrans Technologies, and J.v.d.O. and R.H.J.S. are scientific advisors of Scope Biosciences. P.M., P.v.B., R.L., R.H.J.S. and J.v.d.O. are included as inventors on CRISPR–Cas-related patents. C.S. declares no competing interests.

Peer review information
Nature Reviews Microbiology thanks Luciano Marraffini and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.