REVIEW

doi:10.1038/nature15386

CRISPR–Cas immunity in prokaryotes

Luciano A. Marraffini[1]

Prokaryotic organisms are threatened by a large array of viruses and have developed numerous defence strategies. Among these, only clustered, regularly interspaced short palindromic repeat (CRISPR)–Cas systems provide adaptive immunity against foreign elements. Upon viral injection, a small sequence of the viral genome, known as a spacer, is integrated into the CRISPR locus to immunize the host cell. Spacers are transcribed into small RNA guides that direct the cleavage of the viral DNA by Cas nucleases. Immunization through spacer acquisition enables a unique form of evolution whereby a population not only rapidly acquires resistance to its predators but also passes this resistance mechanism vertically to its progeny.

Prokaryotic viruses (bacteriophages or phages) are the most abundant life form, outnumbering their hosts by a factor of 10 (refs 1, 2). Since the beginning of the study of phages in the laboratory[3,4], investigators isolated and characterized bacteria that were resistant to phage infection[5–7]. This led to the discovery of many different antiviral defence mechanisms (Box 1). CRISPR loci and CRISPR-associated (*cas*) genes encode a unique defence mechanism that provides rapid and robust adaptation to the rapidly evolving viruses (primarily double-stranded (ds)DNA viruses and other foreign DNA) of archaea and bacteria. CRISPR loci consist of an array of short (approximately 30–40 base pairs (bp)) and partially palindromic, repetitive sequences interspaced by equally short 'spacer' sequences (Fig. 1) from viral or plasmid origin. Spacers are at the centre of CRISPR defence as they specify immunity against phages or plasmids that contain a complementary sequence[8,9]. The acquisition and utilization of spacer sequences constitute the two main stages of CRISPR immunity (Fig. 1). In the first stage, also known as 'adaptation' or 'spacer acquisition' (Fig. 1a), sequences from the viral genome are integrated into the CRISPR array (that is, the host is immunized). The second stage, where immunity is executed, can be further divided into two phases (Fig. 1b). First, in the guide RNA biogenesis phase, the CRISPR array is transcribed and processed to generate short RNAs containing one spacer sequence. Second, in the targeting phase, the spacer sequences in these RNAs are used as guides to direct the cleavage of the viral genome by the Cas endonucleases. As is the case with other major antiviral defence systems (Box 1), phages can escape CRISPR immunity through mutations in the target region that prevent its recognition and/or cleavage[10,11]. However, in contrast to the other defence systems, the rapid acquisition of new spacers endows CRISPR-Cas loci with an in-built mechanism to counterattack these phage 'escapers'[10,12–14]. This distinctive property makes CRISPR-Cas immunity a unique form of heritable and adaptive immunity.

Discovery of CRISPR–Cas immune systems

CRISPR-*cas* loci and their function in antiviral and antiplasmid defence were first analysed *in silico*. The first description of CRISPR loci appeared in 1987, after the sequencing of the *iap* gene of *Escherichia coli*[15]. An 'unusual' repeat cluster was found downstream of this gene but due to the lack of homology to other known sequences the authors concluded that "the biological significance of these sequences is not known". The first comprehensive report of CRISPR sequences was only possible after the accumulation of prokaryote genomes in GenBank[16], and established CRISPR loci as common clusters of repetitive sequences in bacteria and archaea. In 2002, the isolation and sequencing of small non-coding RNA species of *Archaeoglobus fulgidus* revealed that CRISPR loci are transcribed into small RNAs[17]. Also in 2002, the *cas* genes were identified as a gene family associated with CRISPR loci[18]. Sequence homology indicated that many of the Cas proteins participate in chemical reactions involving nucleic acids. In 2005 it was discovered that spacers match sequences present in phages and plasmids[19–21], and that the more spacers present in *Streptococcus thermophilus* strains, the fewer phages that were able to infect them[19]. These findings suggested a role for CRISPR-Cas systems in the prevention of phage infection and plasmid conjugation. All these data were incorporated into a model for CRISPR-Cas immunity reminiscent of RNA interference in eukaryotes, in which small CRISPR RNAs (crRNAs) had an antisense function against phage or plasmid transcripts, thereby preventing the propagation of these genetic elements[20,22], with the Cas proteins functioning as the effectors of the immunity mechanism[22].

The initial experimental work on CRISPR-Cas systems concentrated on testing this model. Three foundational studies uncovered the basic features of CRISPR-Cas immunity: adaptation, the role of crRNA guides and the targeting of the invading nucleic acid (Fig. 1). The first study demonstrated the suspected function of CRISPR-Cas in the prevention of phage infection in *S. thermophilus*[8]. The work demonstrated a fundamental aspect of CRISPR-Cas immunity that was not foreseen in the models: that immunity is adaptive. To test for an involvement of CRISPR-*cas* loci in antiphage defence, Barrangou and colleagues infected *S. thermophilus* with phage and examined the CRISPR locus of the phage-resistant bacteria[8]. They observed that the resistant mutants harboured one or more new spacer sequences that perfectly matched a region of the genome of the infecting phage, which showed that new spacers (and immunity) are acquired during infection. This study also implicated the *cas* genes in CRISPR-mediated defence. A second study (Brouns and colleagues[23]) demonstrated the requirement of crRNAs for immunity that was proposed in the early models. Brouns and colleagues discovered a Cas ribonucleoprotein complex harbouring the RNase responsible for the generation of small (mature) crRNAs in *E. coli*[23]. Disruption of the complex or mutation of the catalytic residues of the RNase subunit prevented the accumulation of crRNAs and at the same time abrogated CRISPR immunity against phages. A third study demonstrated the role of CRISPR-Cas systems in the prevention of plasmid conjugation[9]. Marraffini and Sontheimer showed that a spacer present in the CRISPR locus of *Staphylococcus epidermidis*, which matched a region of the *nickase* gene of staphylococcal plasmids, prevented the conjugative transfer of such plasmids. The study also revealed

[1]Laboratory of Bacteriology, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA.

BOX 1

Mechanisms of antiviral defence in prokaryotes

A common mechanism of defence against phage is the prevention of phage adsorption and/or genome injection. Bacteria can switch on or off the expression of phage receptors[105,106] or secrete polysaccharides that limit the access to the receptor[107]. Bacterial hosts can also express membrane proteins that prevent injection of the viral genetic material into the cell[108]. Abortive infection is another common antiviral strategy in which the infected host cell sacrifices itself to prevent phage propagation[109]. In most of the cases studied to date, phage infection is detected by a sensor protein that activates a cell death mechanism such as membrane depolarization[110] or inhibition of bacterial translation[111]. Other abortive infection mechanisms work through a phage-triggered activation of toxin–antitoxin systems that kills the host[112,113]. Restriction-modification is the best studied defence system of prokaryotes and confers protection by cleaving the invading viral genome[114]. This system utilizes nucleases that recognize and cleave short DNA motifs. The target sequences of restriction nucleases are present in both the viral and host genome, but the bacterial chromosome is protected through methylation of the target DNA sequence. Therefore, restriction-modification systems display both endonuclease and methylase activities with the same sequence specificity. Other defence systems recently discovered, and less well characterized, are the bacteriophage exclusion (BREX) system, which blocks phage replication[115], and the prokaryotic Argonaute pathway, which attacks foreign nucleic acid elements[116,117].

As expected from the highly dynamic virus–host interactions that occur in prokaryotic ecosystems, phages have evolved several counter-defence strategies to overcome host defences[118]. Phages can adapt to use different receptors[119] or produce degrading enzymes to eliminate the extracellular polysaccharides that occlude receptors[120]. Mutations in the viral proteins that trigger abortive infection allow phages to bypass this resistance mechanism[109], and when methylation of the invading phage DNA occurs faster than its cleavage by restriction endonucleases, the infecting phage can overcome this barrier and rapidly propagate throughout the whole host population[114]. To adapt to phage evasion strategies the host population needs to restore the lost mechanism of defence, in most cases through the selection of infrequent (and sometimes costly) mutations. By contrast, the acquisition of spacer sequences into CRISPR-Cas loci constitutes an in-built mechanism that provides a rapid and efficient response against phages that escape immunity through the introduction of mutations in the target site[10,12–14].

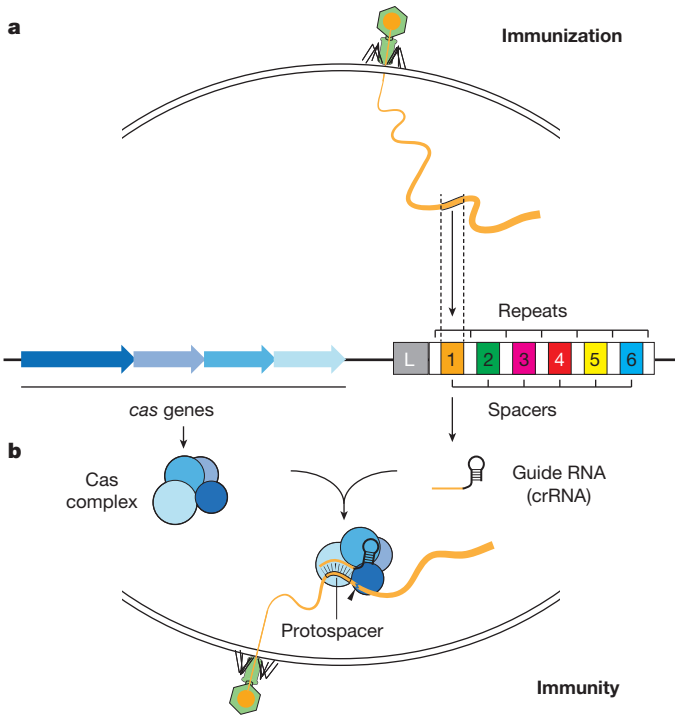

Figure 1 | Stages of CRISPR-Cas immunity. CRISPR loci are a cluster of short DNA repeats (white boxes) separated by equally short spacer sequences of phage and plasmid origin (coloured, numbered boxes). This repeat/spacer array is flanked by an operon of CRISPR-associated (*cas*) genes (blue-tone arrows) that encode the machinery for the immunization and immunity stages of the system. The CRISPR array is preceded by a leader sequence (grey box) containing the promoter for its expression. **a**, In the immunization stage, spacer sequences are captured upon entry of the foreign DNA into the cell and integrated into the first position of the CRISPR array. **b**, In the immunity stage the spacer is used to target invading DNA that carries a cognate sequence for destruction. Spacers are transcribed and processed into small CRISPR RNAs (crRNAs) in the 'crRNA biogenesis' phase. These small RNAs act as antisense guides for Cas RNA-guided nucleases (which usually form a complex) that locate and cleave the target sequence (black arrowhead) in the invader's genome during the 'targeting' phase.

that as opposed to the proposed RNA-interference-like mechanism, CRISPR-Cas systems provide immunity by targeting DNA, rather than RNA. To show this, a self-splicing intron was placed in the *nickase* target sequence, creating a scenario in which the DNA target was interrupted but the RNA target was intact after splicing. Immunity to conjugation was lost, demonstrating the requirement of an intact DNA, but not RNA, target for CRISPR-Cas function. These results also suggested the existence of crRNA-guided, programmable Cas DNA nucleases, opening up the possibility for the development of CRISPR-based DNA manipulation tools[24].

Molecular mechanisms of CRISPR–Cas immunity

All CRISPR-Cas systems harbour the *cas1* and *cas2* genes, which are central to the immunization stage of the pathway (see below). However, based on the accessory *cas* gene content, each system can be classified into three different types[25]. While all types use the same basic molecular mechanism to achieve immunity, that is, through crRNA-guided nucleases, they differ in the biogenesis of crRNAs and the targeting requirements. Type I CRISPR-Cas immunity is mediated by the Cascade complex and the Cas3 nuclease (Fig. 2a)[23]. One of the subunits of Cascade, Cas6e, is a repeat-specific endoribonuclease that cleaves the precursor crRNA that is generated by transcription of the full CRISPR array[23,26]. This cleavage produces short crRNAs that remain associated with Cascade and that are used by the complex to locate a complementary sequence in the target DNA[23,27,28], known as the protospacer. Another subunit, Cas8 (also known as CasA or Cse1), recognizes a short sequence motif located immediately upstream of the target sequence recognized by the crRNA[29]. Sequence motifs adjacent to the targets specified by CRISPR spacers were first identified in type II systems[11,19,30] and subsequently named as 'protospacer adjacent motif', or PAM[31]. PAM recognition is required for type I CRISPR-Cas immunity[32], and the absence of a PAM in the repeat sequences prevents the targeting of the spacers within the CRISPR array by their complementary crRNAs; that is, it prevents an autoimmune reaction. The presence of a PAM promotes Cascade binding to its target[29,33,34] and the formation of an R-loop between the crRNA spacer sequence and the dsDNA[35–37]. The first 8 bp at the 5′ end of the crRNA–DNA duplex are critical for immunity and define a 'seed' sequence within the target[28,32]. Mutant viruses containing mutations in this region can escape type I CRISPR immunity in *E. coli*[32]. An exception is mutations in the sixth nucleotide of the seed, which do not affect CRISPR immunity. The recent crystal structure of the interaction between a guide crRNA and its cognate

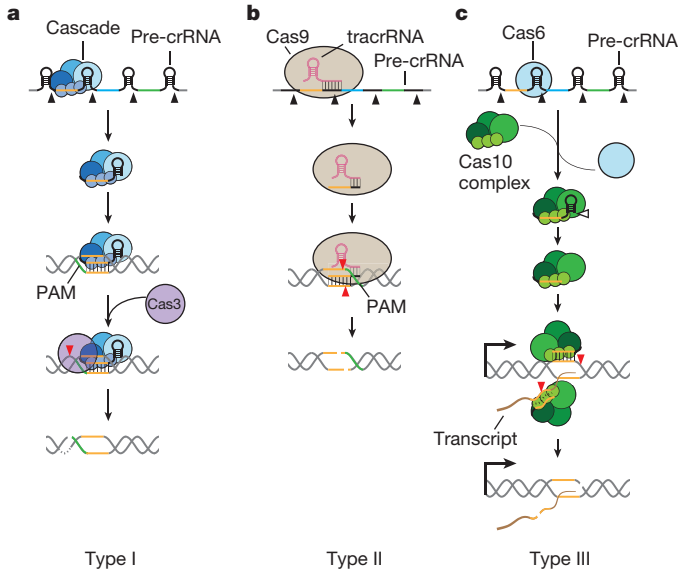

Figure 2 | Immunity mechanisms of the different CRISPR-Cas types. **a**, Type I systems. A Cas protein complex known as Cascade cleaves at the base of the stem–loop structure of each repeat in the long precursor crRNA (pre-crRNA, black arrowheads), which generates short crRNA guides. The Cascade–crRNA complex scans the target DNA for a matching sequence (known as protospacer), which is flanked by a protospacer-adjacent motif (PAM, in green). Annealing of the crRNA to the target strand forms an R-loop; the Cas3 nuclease is recruited and cleaves the target downstream of the PAM (red arrowhead) and also degrades the opposite strand. **b**, Type II systems. These systems encode another small RNA known as trans-encoded crRNA (tracrRNA) which is bound by Cas9 and has regions of complementarity to the repeat sequences in the pre-crRNA. The repeat/tracrRNA dsRNA is cleaved by RNase III to generate crRNA guides for the Cas9 nuclease (black arrowheads). This nuclease cleaves both strands of the protospacer/crRNA R-loop (red arrowhead). A PAM (in green) is located downstream of the target sequence. **c**, Type III systems. Cas6 is a repeat-specific endoribonuclease that cleaves the pre-crRNA at the base of the stem–loop structure of each repeat (black arrowhead). The crRNA is loaded into the Cas10 complex where it is further trimmed at the 3′ end to generate a mature crRNA (white arrowhead). The Cas10 complex requires target transcription to cleave the non-template strand of the protospacer DNA and it is also capable of crRNA-guided transcript cleavage (red arrowheads).

single-stranded (ss)DNA target within the Cascade complex showed that the crRNA–ssDNA interaction forms a non-canonical ribbon structure in which every sixth nucleotide is rotated out of the double helix and therefore not engaged in the formation of a base pair[38–40]. Finally, target recognition by Cascade triggers the recruitment and activity of Cas3, a nuclease that introduces ssDNA breaks into the target virus or plasmid[33,41–44] to initiate their degradation.

Type II CRISPR-Cas systems require only one *cas* gene, *cas9*, to execute immunity in the presence of an existing targeting spacer sequence (Fig. 2b)[45]. However, as opposed to the other CRISPR types, two small RNAs are needed for immunity: the crRNA and the trans-encoded crRNA (tracrRNA)[46]. The tracrRNA forms a secondary structure that mediates its association with Cas9 (refs 47–49) but also has a region that is complementary to the repeat sequences of the CRISPR array[46]. The dsRNA formed between the tracrRNA and the precursor crRNA is cleaved by RNase III, resulting in the cleavage of each repeat and the processing of the long CRISPR transcript into small crRNA guides[46]. Type II immunity also requires a PAM, and mutations in this motif are the most common mechanism of escape from CRISPR immunity by the targeted viruses[11]. In contrast to type I, the PAM is located immediately downstream of the target sequence[11,19] and is recognized by a PAM-binding domain present in Cas9 (refs 47–49). Type II CRISPR-Cas immunity results in the introduction of crRNA-specific dsDNA breaks in the invading DNA[50] that require two nuclease domains: HNH and RuvC[45]. Each of these domains cleaves one DNA strand of the protospacer sequence[51,52] and the tracrRNA is absolutely required for cleavage[52]. The first step in target recognition is the transient binding of Cas9 to PAM sequences within the target DNA, which promotes the melting of the two DNA strands immediately upstream of the PAM[53]. A productive interaction in this region of the target, between 6–8 bases of the spacer sequence of the crRNA guide and the melted DNA (the 'seed' sequence of type II systems[11,54]), triggers the formation of an R-loop and target cleavage[37,53].

In type III CRISPR-Cas systems, the precursor crRNA is cleaved by a repeat-specific endoribonuclease, Cas6, which is not part of a complex[55]. As a result of this processing, 8 nucleotides of the repeat sequence remain at the 5′ end of the spacer sequence in the crRNA[55,56], a sequence known as the crRNA tag. By an as yet unknown mechanism, the small crRNAs generated after Cas6 cleavage are transferred to a larger complex[57], the Cas10–Csm or Cas10–Cmr complex for type III-A or III-B systems, respectively. Within these complexes the crRNAs undergo a process of maturation whereby the 3′ end is trimmed at 6-nucleotide intervals[58–60]. As opposed to type I and II systems, in which targeting relies strictly on the recognition of DNA sequences, type III CRISPR-Cas immunity also requires target transcription[61,62] and a crRNA complementary to the non-template strand of the DNA target and to the transcript[62]. Both DNA[9,61–63] and RNA[63–68] are targeted by type III CRISPR-Cas systems, resulting in the co-transcriptional crRNA-guided cleavage of the target DNA and its transcripts[63,69]. The Cas10 complex contains both nucleolytic activities: the palm domain of Cas10 is required for cleavage of the non-template DNA strand[63], and backbone subunits Csm3 (refs 63, 65) or Cmr4 (ref. 66), for type III-A or III-B systems, respectively, are responsible for cleavage of the RNA transcripts. To date, no PAM requirements have been observed for type III CRISPR-Cas targeting. To avoid targeting of the CRISPR locus, type III systems rely on the differential base pairing between the crRNA tag and the sequences flanking the protospacer[56]. Whereas the absence of complete complementarity between these sequences licenses DNA targeting, full complementarity between the crRNA tag and the repeat sequence in the CRISPR locus prevents DNA targeting[56,63,67] and thus autoimmunity. The biological significance of this elaborate targeting mechanism is beginning to be elucidated. The transcription requirement for targeting offers the possibility of immunological tolerance of mobile genetic elements with non-transcribed regions, such as prophages with the potential to provide a fitness advantage to the host[62]. Moreover, when the prophage re-activates and enters the lytic cycle, its genome is transcribed and type III targeting resumes, resulting in clearance of the infection and preventing the death of the host cell. Studies have shown that RNA cleavage can protect against RNA viruses[66]; however, the role of RNA targeting in immunity against DNA elements and possibly regulation of gene expression remains to be determined.

CRISPR immunization

The mechanisms by which new spacer sequences are added to the CRISPR locus to immunize the host are beginning to be understood. This process can be divided into two stages: the selection of protospacer sequences from the invader DNA and their integration into the CRISPR array (Fig. 3). With a few exceptions, the spacer acquisition mechanism has been studied in detail in the *E. coli* type I CRISPR-Cas system. This is because an early report showed that overexpression of type I Cas1 and Cas2 is sufficient for the expansion of the *E. coli* CRISPR array of this organism and thus established a very simple and elegant experimental system to study this phenomenon[70].

Central to the mechanism of selection of new spacer sequences is the prevention of autoimmunity; that is, the ability of the acquisition machinery to distinguish self (chromosomal) from non-self (invading) DNA. Failure to do so leads to the death of the host, a scenario that has been tested and exploited for the use of CRISPR-Cas systems as genome editing tools of bacteria[54,71–73] and as antimicrobials[74–76]. Applying the Cas1–Cas2 overexpression system—in which autoimmunity is avoided

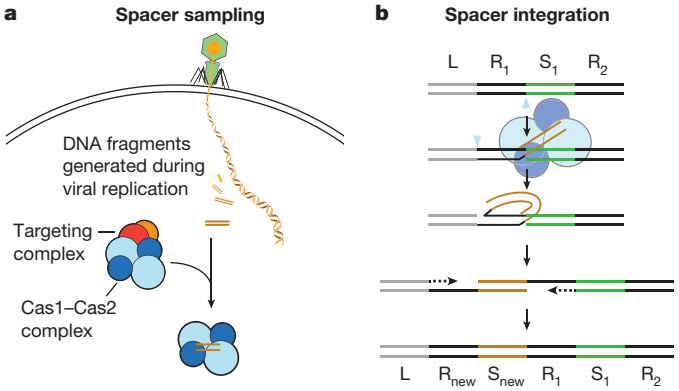

Figure 3 | Mechanism of CRISPR immunization. **a**, The first step of CRISPR immunization is the sampling of the spacer sequences. These are believed to be generated from non-specific DNA breaks that occur during replication of the virus or plasmid. The fragments generated are captured by the Cas1–Cas2 complex, with the participation of the targeting machinery for the recognition of DNA sequences carrying a functional PAM. **b**, The Cas1–Cas2 complex catalyses the integration of the spacer into the first position of the CRISPR array. Cas1 performs two concerted cleavage-ligation reactions whereby the 5′ end of each repeat strand is cleaved (blue arrowhead) and ligated to the 3′ ends of the spacer. This mechanism generates two ssDNA gaps on the repeat sequences that flank the inserted spacer, which presumably are filled by DNA polymerase (dotted arrow). L, leader; R_1, first repeat; R_2, second repeat; R_{new}, new repeat; S_1, first spacer; S_{new}, new spacer.

by using an *E. coli* strain that lacks the Cascade and Cas3 immunity machinery—several studies determined that there is a strong preference for the integration of plasmid over chromosomal spacer sequences[70,77,78]. Recently, it has been shown that Chi sites, 8-nucleotide motifs present approximately once every 5 kb in the *E. coli* genome, limit the acquisition of chromosomal sequences[79]. Chi sites are over-represented in the *E. coli* chromosome[80]; therefore, this mechanism favours the acquisition of spacer sequences from foreign elements. The origin of the substrates for CRISPR adaptation is not yet known. In *E. coli*, the overexpression of Cas1–Cas2 in the absence of immunity leads to the acquisition of spacer sequences derived from regions limited by Chi sites on one side and by *Ter* sites on the other[79]. There are two *Ter* sites in the *E. coli* genome, *TerA* and *TerC*, which are diametrically opposed from the origin of replication to stall a faster replication fork and ensure proper chromosomal decatenation at the end of a round of DNA replication[81]. Replisome stalling at either *Ter* site leads to chromosome breaks that are processed by the RecBCD exonuclease complex, generating 3′OH overhang ends for RecA-mediated homologous recombination repair of the DNA lesion[82]. To avoid excessive DNA degradation, RecBCD activity is inhibited by Chi sequences[83]. Therefore, the chromosomal acquisition data obtained in *E. coli* (ref. 79) suggest that the nucleolytic processing of dsDNA breaks generated during replication is the source of new spacer sequences for the CRISPR-Cas acquisition machinery. This model is supported by a reduction in spacer acquisition from the host chromosome when DNA replication is chemically inhibited and in mutants that lack the RecBCD machinery[79]. Although not directly tested, this work predicts that the DNA degradation that follows the formation of spontaneous breaks during viral or plasmid replication produces DNA fragments that are captured by the Cas adaptation machinery for their incorporation as non-self spacers into the CRISPR array (Fig. 3a).

A second aspect of spacer selection relates to the PAM requirement for targeting: how does the acquisition machinery ensure that a new spacer will match a protospacer flanked by the correct PAM? Recently it was revealed that Cas9, which has the ability to recognize the PAM, is used to select new spacers with the correct PAM sequence[84] (Fig. 3a). Genetic studies showed that Cas9 is absolutely required for type II CRISPR-Cas adaptation[84,85]. While the nuclease activity does not contribute to the role of Cas9 in spacer acquisition[84,85], mutation of the PAM-binding domain results in the incorporation of new spacers that match DNA targets but without conservation of their flanking sequences; that is, without a PAM[84]. Supporting a role for Cas9 in the acquisition of functional spacer sequences it was shown that this nuclease forms a complex with Cas1, Cas2 and Csn2, proteins exclusively involved in CRISPR adaptation, and that swapping *cas9* alleles that recognize different PAM sequences results in the incorporation of spacers with PAMs corresponding to the *cas9* allele used[84].

Cas1 and Cas2, without Cascade, are sufficient for *E.coli* type I spacer acquisition. However, it is not clear whether the new spacers acquired in the absence of Cascade display an absolute conservation of the PAM sequence—AAG for this CRISPR-Cas system[70,86,87]. Interestingly, the first (5′ end) nucleotide of *E. coli* spacer sequences is invariably G; that is, the last nucleotide of the PAM[88–90]. Therefore it is possible that, similarly to Cas9 in type II systems, the PAM recognition feature of the Cascade complex is involved in the selection of functional spacers. Although this remains to be tested, studies have found another role for Cascade in type I adaptation, in a phenomenon known as 'priming'[86]. It has been shown that the presence of pre-existing (priming) spacers with partial homology (and therefore unable to provide full immunity) to the ssDNA phage M13 increases the rate of spacer acquisition by several orders of magnitude[86]. In addition, new spacers acquired in the presence of a priming spacer have a strong strand bias, producing a crRNA guide that matches the same strand matched by the priming spacer[86,87]. Primed acquisition facilitates adaptation against invaders that are related in sequence to previous invaders (that have partial matches to pre-existing spacers) and also against escapers (that is, phage containing target mutations that prevent CRISPR-Cas immunity). The mechanism of primed acquisition is yet to be resolved. Primed adaptation requires the Cascade complex[86], and the partial match between the crRNA-target sequences that trigger this process results in a distinct mode of Cascade binding to the target DNA[35]. It is believed that this binding results in a low incidence of target cleavage that, similarly to the DNA breaks that occur at replication stall sites[79], would generate the substrates for spacer acquisition. Interestingly, primed acquisition reveals a coupling between the immunity and immunization stages of CRISPR defence.

In the second stage of CRISPR immunization, after the selection of the spacer by the acquisition machinery, spacers are integrated into the CRISPR array in a reaction that resembles retroviral integration (Fig. 3b). Recently, it was shown that *E. coli* Cas1 and Cas2 perform the spacer integration reaction. Studies using the Cas1–Cas2 overexpression system demonstrated the presence of integration intermediates *in vivo*[91]. Using specific probes for Southern blot assays it was demonstrated that each strand of the first repeat sequence is separated and ligated to the 3′ end of the new spacer sequence. This reaction is catalysed by the Cas1–Cas2 complex, which is composed of two Cas1–Cas2 units, and mutations that prevent the interaction between these proteins[78] or eliminate Cas1 nuclease activity[70,78,91] abrogate spacer acquisition. The complex binds to *E. coli* CRISPR sequences *in vitro*[78] and, when incubated with a 33-nucleotide dsDNA substrate (which mimics a captured spacer sequence), it mediates the covalent addition of this dsDNA into plasmids harbouring the CRISPR array[92]. This reaction requires free 3′ OH ends on the 33-nucleotide spacer substrate, which presumably perform a direct nucleophilic attack on the phosphodiester bond between the first repeat and spacer sequences[92] (Fig. 3b). Importantly, deep sequencing of the reaction products revealed a strong bias for the integration of spacers with a C nucleotide at the 3′ end (that is, the complementary base to the first (5′ end) G of the spacer sequence[92]). Collectively, these results indicate that the Cas1–Cas2 complex provides the orientation specificity of the integration reaction, which probably occurs by two consecutive 3′ OH attacks on opposite sides of the repeat sequence, with the first nucleophilic attack by the 3′-end C of the spacer sequence on the 5′ end of the bottom strand of the repeat. After the repair of the repeat gaps that result from this reaction, a new repeat-spacer unit is added into the CRISPR array. This mechanism

is consistent with early work that demonstrated that the additional repeat is derived from the first repeat of the CRISPR array[70]. This study showed that a 'labelling' mutation introduced in the first, but not the second, repeat sequence is incorporated into the new repeat after spacer acquisition.

Perspectives

An extensive body of work has established the function of CRISPR-Cas systems as the adaptive immune system of prokaryotes, which function in protection against viral infection and plasmid invasion. One intriguing aspect is the role of these systems in the regulation of horizontal gene transfer (HGT) between microorganisms. HGT is fundamental for the generation of the genetic diversity required for prokaryotic evolution and is achieved through three major routes: phage transduction, plasmid conjugation and DNA transformation[93]. In addition to the activity of CRISPR-Cas against phages, it has been experimentally demonstrated that CRISPR-Cas systems can prevent plasmid conjugation[9] and the transformation of naturally competent bacteria[94,95]. Therefore, in principle, CRISPR immunity can prevent the major routes of HGT and therefore have an important role in limiting the evolution of bacteria and archaea. This drawback of CRISPR-Cas systems has often been considered a cause for the inconsistent distribution of these loci[96,97], which are present in only ~50% of bacterial and ~90% of archaeal genomes[98]. A recent study investigated the relationship between the number of spacer sequences (an indicator of CRISPR activity) and the estimated number of HGT transfer events for all the genomes present in GenBank[99]. If CRISPR immunity limits HGT, a negative correlation between these two values is expected. Since this correlation was not evident, the study suggests that there is no impact of CRISPR-Cas on gene transfer among prokaryotes over evolutionary timescales. Future work should determine the effects of CRISPR-Cas immunity on HGT at the population level as well as other potential disadvantages of the system, such as a high incidence of autoimmunity and/or off-target effects, which could impact on its distribution among prokaryotic organisms.

Another puzzling aspect of CRISPR-Cas systems is their high diversity[25,100]. For each of the three CRISPR types there are many subtypes encoding divergent sets of Cas proteins. The biological significance of this diversity is not known. A driving force for the diversification of CRISPR-Cas systems could be imposed by phage-encoded anti-CRISPR mechanisms. It has been shown that phages encode small proteins that specifically inhibit type I-F CRISPR immunity[101], which could conceivably lead to the evolution of another type or subtype to overcome the inhibition of the existing CRISPR-Cas system. In this hypothetical scenario CRISPR-Cas systems would diversify as a result of the arms race with the viruses they target. It is also possible that the mechanistic differences between types and subtypes provide advantages for different CRISPR-Cas systems in different conditions imposed by the lifestyle or the environment of the host. An example of this is the type III systems, which tolerate untranscribed or inert mobile genetic elements[62]. Different conditions where this tolerance is beneficial or detrimental will determine the selection of type III systems in different hosts. Similarly, mechanistic differences between other types or subtypes that have an impact on the fitness of the host could have a role in the distribution of the CRISPR systems. Furthermore, additional functions of particular CRISPR subtypes could explain their diversification. For example, the type II-C CRISPR-Cas system of the bacterial pathogen *Francisella novicida* harbours a tracrRNA with sequence homology to an endogenous lipoprotein gene[102,103]. When complexed with Cas9, the tracrRNA associates with another small RNA, the small CRISPR-associated RNA (scaRNA), to mediate the silencing of the lipoprotein gene during *F. novicida* infection. Another example is the RNA cleavage feature of type III (Cas10) complexes, which offers the possibility of regulation of gene expression by these systems[67,104] that could lead to the preferential selection of these systems over other types that cannot perform this function.

Similarly to restriction-modification systems, the possibilities to manipulate DNA sequences in a predictable manner by CRISPR-Cas immunity have led to numerous applications in molecular biology, most notably the genetic engineering of a range of different cell types[24]. This highlights the importance of the study of basic biological problems, particularly the interactions between prokaryotes and the mobile genetic elements that invade them, for the development of innovative biotechnological applications. In this context it is exciting to consider the different mechanisms of immunity present in the understudied (or even presently undiscovered) CRISPR-Cas systems, as well as the many uncharacterized prokaryotic defence systems carried by the archaeal and bacterial domains of life that inhabit every corner of our planet.

Received 29 May 2015; accepted 7 August 2015.

1. Bergh, O., Borsheim, K. Y., Bratbak, G. & Heldal, M. High abundance of viruses found in aquatic environments. *Nature* **340,** 467–468 (1989).
2. Chibani-Chennoufi, S., Bruttin, A., Dillmann, M. L. & Brussow, H. Phage-host interaction: an ecological perspective. *J. Bacteriol.* **186,** 3677–3686 (2004).
3. d'Herelle, F. Sur un microbe invisible antagoniste des bacilles dysentériques. *C.R. Acad. Sci. Paris* **165,** 373–375 (1917).
4. Twort, F. W. An investigation on the nature of ultra-microscopic viruses. *Lancet* **186,** 1241–1243 (1915).
5. Burnet, F. M. Further observations on the nature of bacterial resistance to bacteriophage. *J. Pathol. Bacteriol.* **32,** 349–354 (1929).
6. Gratia, A. Studies on the D'herelle phenomenon. *J. Exp. Med.* **34,** 115–126 (1921).
7. Luria, S. E. & Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28,** 491–511 (1943).
8. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
   A study that demonstrated that CRISPR-Cas loci provide acquired immunity against bacteriophages.
9. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322,** 1843–1845 (2008).
   A paper showing that CRISPR-Cas loci target DNA molecules in a sequence-specific manner, highlighting for the first time the potential for the technological applications of these systems.
10. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320,** 1047–1050 (2008).
    This study revealed the arms race between CRISPR-Cas systems and viruses in their natural habitat.
11. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1390–1400 (2008).
12. 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).
13. Paez-Espino, D. *et al.* CRISPR immunity drives rapid phage genome evolution in Streptococcus thermophilus. *MBio* **6,** e00262–15 (2015).
14. Weinberger, A. D. *et al.* Persisting viral sequences shape microbial CRISPR-based immunity. *PLOS Comput. Biol.* **8,** e1002475 (2012).
15. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169,** 5429–5433 (1987).
16. Mojica, F. J., Diez-Villasenor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36,** 244–246 (2000).
    First description of CRISPR loci as a new family of repetitive sequences in prokaryotes.
17. Tang, T. H. *et al.* Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA* **99,** 7536–7541 (2002).
18. Jansen, R., Embden, J. D., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565–1575 (2002).
    First description of *cas* sequences as a family of genes associated with CRISPR repeats.
19. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151,** 2551–2561 (2005).
    This paper, along with references 20 and 21, made the discovery that spacer sequences match viruses and plasmids, and suggested a defence function for CRISPR-Cas systems.
20. Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60,** 174–182 (2005).
21. 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* **151,** 653–663 (2005).
22. 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).
This work provided the first comprehensive model for the mechanism of CRISPR-Cas immunity.
23. Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
Study demonstrating the central role of crRNA guides and Cas ribonucleoproteins in CRISPR immunity.
24. Pennisi, E. The CRISPR craze. *Science* **341,** 833–836 (2013).
25. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nature Rev. Microbiol.* **9,** 467–477 (2011).
26. 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).
27. Jore, M. M. *et al.* Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nature Struct. Mol. Biol.* **18,** 529–536 (2011).
28. 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).
29. Sashital, D. G., Wiedenheft, B. & Doudna, J. A. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* **46,** 606–615 (2012).
30. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1401–1412 (2008).
31. Mojica, F. J., 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).
32. 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).
33. Sinkunas, T. *et al. In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* **32,** 385–394 (2013).
34. 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).
35. Blosser, T. R. *et al.* Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. *Mol. Cell* **58,** 60–70 (2015).
36. Rutkauskas, M. *et al.* Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell Rep.* http://dx.doi.org/10.1016/j.celrep.2015.01.067 (2015).
37. Szczelkun, M. D. *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl Acad. Sci. USA* **111,** 9798–9803 (2014).
38. Jackson, R. N. *et al.* Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*. *Science* **345,** 1473–1479 (2014).
39. Mulepati, S., Heroux, A. & Bailey, S. Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* **345,** 1479–1484 (2014).
40. Zhao, H. *et al.* Crystal structure of the RNA-guided immune surveillance Cascade complex in *Escherichia coli*. *Nature* **515,** 147–150 (2014).
41. Hochstrasser, M. L. *et al.* CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proc. Natl Acad. Sci. USA* **111,** 6618–6623 (2014).
42. Huo, Y. *et al.* Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nature Struct. Mol. Biol.* **21,** 771–777 (2014).
43. 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.* **288,** 22184–22192 (2013).
44. Sinkunas, T. *et al.* Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **30,** 1335–1342 (2011).
45. Sapranauskas, R. *et al.* The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39,** 9275–9282 (2011).
46. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).
47. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513,** 569–573 (2014).
48. Jinek, M. *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343,** 1247997 (2014).
49. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156,** 935–949 (2014).
50. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71 (2010).
This work showed that the crRNA-guided DNA targeting by CRISPR-Cas systems results in sequence-specific DNA cleavage.
51. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA* **109,** E2579–E2586 (2012).
52. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).
53. 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).
54. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnol.* **31,** 233–239 (2013).
55. Carte, J., Wang, R., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22,** 3489–3496 (2008).
56. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463,** 568–571 (2010).
57. Sokolowski, R. D., Graham, S. & White, M. F. Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.* **42,** 6532–6541 (2014).
58. Hale, C., Kleppe, K., Terns, R. M. & Terns, M. P. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **14,** 2572–2579 (2008).
59. Hatoum-Aslan, A., Maniv, I. & Marraffini, L. A. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl Acad. Sci. USA* **108,** 21218–21222 (2011).
60. Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W. & Marraffini, L. A. A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J. Biol. Chem.* **288,** 27888–27897 (2013).
61. Deng, L., Garrett, R. A., Shah, S. A., Peng, X. & She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* **87,** 1088–1099 (2013).
62. 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).
63. Samai, P. *et al.* Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. *Cell* **161,** 1164–1174 (2015).
64. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139,** 945–956 (2009).
First demonstration that some CRISPR-Cas systems can cleave RNA molecules.
65. Staals, R. H. *et al.* RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol. Cell* **56,** 518–530 (2014).
66. Tamulaitis, G. *et al.* Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol. Cell* **56,** 506–517 (2014).
67. Zebec, Z., Manica, A., Zhang, J., White, M. F. & Schleper, C. CRISPR-mediated targeted mRNA degradation in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.* **42,** 5280–5288 (2014).
68. Zhang, J. *et al.* Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* **45,** 303–313 (2012).
69. Peng, W., Feng, M., Feng, X., Liang, Y. X. & She, Q. An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res.* **43,** 406–417 (2014).
70. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **40,** 5569–5576 (2012).
71. Oh, J. H. & van Pijkeren, J. P. CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* **42,** e131 (2014).
72. Jiang, Y. *et al.* Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* **81,** 2506–2514 (2015).
73. Huang, H., Zheng, G., Jiang, W., Hu, H. & Lu, Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim. Biophys. Sin.* **47,** 231–243 (2015).
74. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nature Biotechnol.* **32,** 1146–1150 (2014).
75. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nature Biotechnol.* **32,** 1141–1145 (2014).
76. Gomaa, A. A. *et al.* Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *MBio* **5,** e00928–e00913 (2014).
77. Díez-Villaseñor, C., Guzman, N. M., Almendros, C., Garcia-Martinez, J. & Mojica, F. J. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. *RNA Biol.* **10,** 792–802 (2013).
78. Nuñez, J. K. *et al.* Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nature Struct. Mol. Biol.* **21,** 528–534 (2014).
79. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520,** 505–510 (2015).
This paper showed that dsDNA breaks generated during replication trigger spacer acquisition.
80. El Karoui, M., Biaudet, V., Schbath, S. & Gruss, A. Characteristics of Chi distribution on different bacterial genomes. *Res. Microbiol.* **150,** 579–587 (1999).
81. Neylon, C., Kralicek, A. V., Hill, T. M. & Dixon, N. E. Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-Ter complex. *Microbiol. Mol. Biol. Rev.* **69,** 501–526 (2005).
82. Dillingham, M. S. & Kowalczykowski, S. C. RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* **72,** 642–671 (2008).
83. Smith, G. R. How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view. *Microbiol. Mol. Biol. Rev.* **76,** 217–228 (2012).
84. Heler, R. *et al.* Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* **519,** 199–202 (2015).
85. Wei, Y., Terns, R. M. & Terns, M. P. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev.* **29,** 356–361 (2015).
86. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* **3,** 945 (2012).

This study showed that pre-existing spacers with partial homology to an invader sequence enhance the acquisition of new spacers.
87. Swarts, D. C., Mosterd, C., van Passel, M. W. & Brouns, S. J. CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* **7,** e35888 (2012).
88. Goren, M. G., Yosef, I., Auster, O. & Qimron, U. Experimental definition of a clustered regularly interspaced short palindromic duplicon in *Escherichia coli*. *J. Mol. Biol.* **423,** 14–16 (2012).
89. Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A. & Severinov, K. High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in *E. coli*. *RNA Biol.* **10,** 716–725 (2013).
90. Shmakov, S. *et al.* Pervasive generation of oppositely oriented spacers during CRISPR adaptation. *Nucleic Acids Res.* **42,** 5907–5916 (2014).
91. 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).
92. Nuñez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519,** 193–198 (2015). **This study showed the molecular mechanism of spacer integration.**
93. Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Rev. Microbiol.* **3,** 711–721 (2005).
94. 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).
95. Zhang, Y. *et al.* Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* **50,** 488–503 (2013).
96. Jiang, W. *et al.* Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.* **9,** e1003844 (2013).
97. Marraffini, L. A. CRISPR-Cas immunity against phages: its effects on the evolution and survival of bacterial pathogens. *PLoS Pathog.* **9,** e1003765 (2013).
98. Grissa, I., Vergnaud, G. & Pourcel, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8,** 172 (2007).
99. Gophna, U. *et al.* No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *ISME J.* **9,** 2021–2027 (2015).
100. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. The basic building blocks and evolution of CRISPR-cas systems. *Biochem. Soc. Trans.* **41,** 1392–1400 (2013).
101. 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).
102. 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).
103. Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y. L. & Weiss, D. S. Corrigendum: A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* **501,** 262 (2013).
104. Hale, C. R. *et al.* Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* **45,** 292–302 (2012).
105. Liu, M. *et al.* Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. *Science* **295,** 2091–2094 (2002).
106. Zaleski, P., Wojciechowski, M. & Piekarowicz, A. The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection. *Microbiology* **151,** 3361–3369 (2005).
107. Hanlon, G. W., Denyer, S. P., Olliff, C. J. & Ibrahim, L. J. Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through Pseudomonas aeruginosa biofilms. *Appl. Environ. Microbiol.* **67,** 2746–2753 (2001).
108. Lu, M. J. & Henning, U. Superinfection exclusion by T-even-type coliphages. *Trends Microbiol.* **2,** 137–139 (1994).
109. Molineux, I. J. Host-parasite interactions: recent developments in the genetics of abortive phage infections. *New Biol.* **3,** 230–236 (1991).
110. Parma, D. H. *et al.* The Rex system of bacteriophage lambda: tolerance and altruistic cell death. *Genes Dev.* **6,** 497–510 (1992).
111. Bingham, R., Ekunwe, S. I., Falk, S., Snyder, L. & Kleanthous, C. The major head protein of bacteriophage T4 binds specifically to elongation factor Tu. *J. Biol. Chem.* **275,** 23219–23226 (2000).
112. Aizenman, E., Engelberg-Kulka, H. & Glaser, G. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3′,5′-bispyrophosphate: a model for programmed bacterial cell death. *Proc. Natl Acad. Sci. USA* **93,** 6059–6063 (1996).
113. Fineran, P. C. *et al.* The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl Acad. Sci. USA* **106,** 894–899 (2009).
114. Bickle, T. A. & Kruger, D. H. Biology of DNA restriction. *Microbiol. Rev.* **57,** 434–450 (1993).
115. Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* **34,** 169–183 (2015).
116. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* **51,** 594–605 (2013).
117. Swarts, D. C. *et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507,** 258–261 (2014).
118. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol.* **8,** 317–327 (2010).
119. Doulatov, S. *et al.* Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. *Nature* **431,** 476–481 (2004).
120. Sutherland, I. W., Hughes, K. A., Skillman, L. C. & Tait, K. The interaction of phage and biofilms. *FEMS Microbiol. Lett.* **232,** 1–6 (2004).

Acknowledgements L.A.M. is supported by the Rita Allen Scholars Program, an Irma T. Hirschl Award, a Sinsheimer Foundation Award and a NIH Director's New Innovator Award (1DP2AI104556-01).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence should be addressed to L.A.M. (marraffini@rockefeller.edu).