Review

CellPress

To acquire or resist: the complex biological effects of CRISPR–Cas systems

Joseph Bondy-Denomy and Alan R. Davidson

Departments of Molecular Genetics and Biochemistry, University of Toronto, Toronto, ON, Canada

Prokaryotic CRISPR–Cas (clustered regularly interspaced short palindromic repeat–CRISPR associated) systems provide a sophisticated adaptive immune system that offers protection against foreign DNA. These systems are widely distributed in prokaryotes and exert an important influence on bacterial behavior and evolution. However, interpreting the biological effects of a CRISPR–Cas system within a given species can be complicated because the outcome of rejecting foreign DNA does not always provide a fitness advantage, as foreign DNA uptake is sometimes beneficial. To address these issues, here we review data pertaining to the potential *in vivo* costs and benefits of CRISPR–Cas systems, novel functions for these systems, and how they may be inactivated.

The prevalence and importance of CRISPR–Cas systems in prokaryotes

CRISPR–Cas systems, which are composed of clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR associated (*cas*) genes, are found in a large number of diverse prokaryotic species and serve to specifically recognize and destroy foreign DNA, such as phage genomes. First observed in 1987 in *Escherichia coli* K12 [1] as an array of alternating repeats with interspersed spacers, this region was later given the CRISPR acronym [2]. The possibility that CRISPRs might form an immune system against phages and plasmids was raised in 2005 when three groups independently reported that spacer sequences possessed homology with foreign DNA elements [3–5]. The first experimental evidence of CRISPR–Cas-mediated adaptive immunity emerged in 2007 with the isolation of phage resistant *Streptococcus thermophilus* cells possessing new CRISPR spacers after a phage challenge. This novel acquisition from the phage genome into the bacterial CRISPR locus was shown to be causative of the phage resistance phenotype [6]. A notable aspect of CRISPR–Cas systems is that the spacers are incorporated from previously encountered foreign DNA elements, so that resistance to these elements is an acquired trait, similar to adaptive immunity seen in higher eukaryotes. The increased fitness provided by CRISPR–Cas systems is illustrated by the occurrence of CRISPR loci in 46% of bacteria and 84% of archaea (CRISPRdb) [7].

Since the first demonstration of CRISPR–Cas-mediated phage resistance, a number of CRISPR–Cas systems have been found in diverse bacteria with differing repeat sequences, Cas proteins, and modes of action. These systems have been grouped into three types (Types I–III), along with subtypes (e.g., Type I-E), on the basis of the *cas* genes they possess and their mode of action [8]. Despite the diversity of the *cas* genes and the organisms possessing CRISPRs, the loci are generally composed of multiple repeated sequences ranging from 21 to 48 bp, separated by 26 to 72 bp variable spacer sequences [9], with *cas* genes located adjacent to the CRISPR locus.

CRISPR–Cas systems have been the subject of intense investigation owing to their intriguing RNA-based mechanism of action. Parallels exist between CRISPR–Cas systems and the RNA interference (RNAi) systems in eukaryotes, and this similarity has, in part, fueled the investigation of these processes; however no homologous proteins have been identified between the CRISPR–Cas and RNAi machinery. For CRISPR–Cas system function, the CRISPR locus is transcribed, yielding a single precursor RNA that is processed within the repeat regions by a host and/or Cas protein into individual units of CRISPR RNAs (crRNAs) [10–12]. The mature crRNA subsequently nucleates the formation of a complex with Cas proteins that will survey the cell for invading DNA [10,13]. The crRNA–Cas complex recognizes and cleaves foreign DNA or RNA molecules at sites with complementarity to the crRNA, called protospacers [14–16] (Figure 1). In addition to requiring sequence identity between the spacer and protospacer, Type I and II CRISPR–Cas systems require a 2–5 nucleotide motif next to the protospacer, called the protospacer adjacent motif (PAM) [17].

Phage predation and horizontal transfer of DNA between bacterial species have massive effects on bacterial evolution, virulence, and physiology [18,19]. Owing to the widespread occurrence of CRISPR–Cas systems in prokaryotes and their proven role in inhibition of phage replication and foreign DNA uptake, there is no doubt that these

Corresponding author: Davidson, A.R. (alan.davidson@utoronto.ca).
Keywords: CRISPR–Cas; anti-CRISPR; mobile DNA; horizontal gene transfer; prophage; phage resistance.

0966-842X/$ – see front matter
© 2014 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tim.2014.01.007

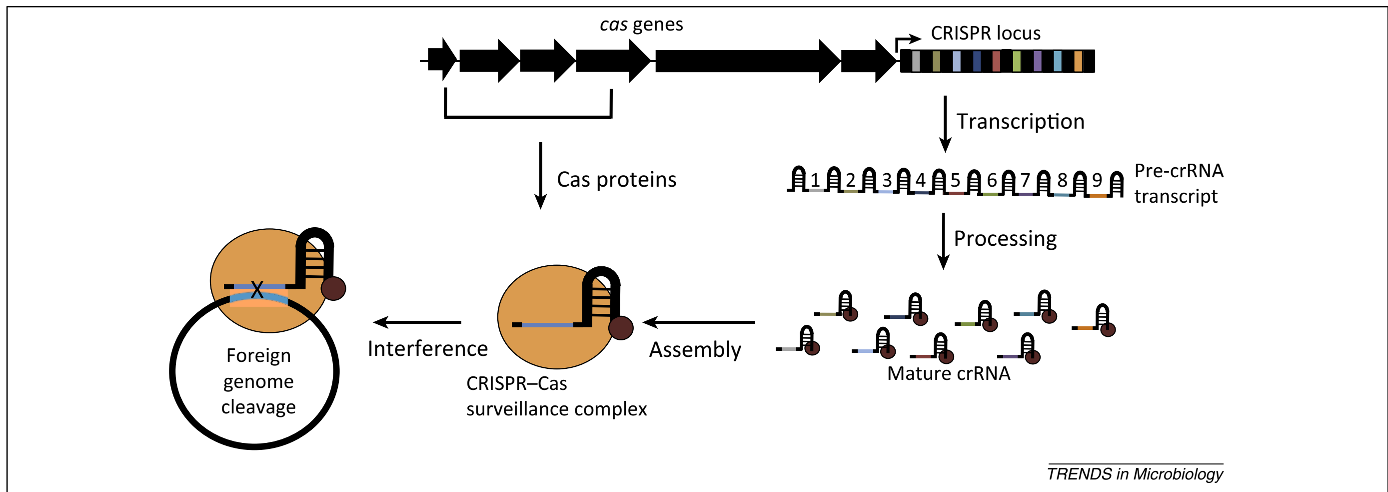

Figure 1. A generic schematic of the Type I CRISPR–Cas (clustered regularly interspaced short palindromic repeat–CRISPR associated) system is shown to demonstrate function. Thick black arrows represent the *cas* genes, which encode the members of the CRISPR–Cas surveillance complex as well as the genes for degradation of the target and spacer acquisition. The CRISPR locus is shown on the right with repeats as black boxes and spacers as colored boxes. The CRISPR locus is transcribed to produce the precursor CRISPR RNA (pre-crRNA) molecule, which is processed by a *cas* gene (brown circle) to produce single mature crRNAs. These are each assembled with other Cas proteins into a complex that will scan the cell for invading DNA. The spacer-derived crRNA sequence can recognize the protospacer sequence through complementary binding (blue line) and guide a cleavage event. The protospacer adjacent motif (PAM, not shown) is also required in Type I and II systems. CRISPR–Cas type-specific differences are not shown here. For specific mechanistic and structural insight, see [9,22,23].

systems play a crucial part in shaping phage genomes and bacterial populations both in the environment in general and in the human microbiome [20,21]. Although structural, biochemical, and *in vivo* studies have provided extensive insight into the molecular mechanisms of CRISPR–Cas function (reviewed in [22,23]), considerably fewer studies have focused on the endogenous operation of these systems and on the physiological ramifications of CRISPR–Cas activity. In particular, there are only a handful of cases in which a CRISPR–Cas system has been directly demonstrated to resist phage infection or uptake of foreign DNA (Table 1). The mere presence of an intact CRISPR–Cas system in a genome does not necessarily mean that it is functioning as a defense system in all conditions; some systems are inactive, repressed, suppressed, or performing alternative functions. In this review, we describe work that addresses the biological consequences of the presence or absence of a CRISPR–Cas system within a given prokaryotic species. We also outline the difficulties associated with identifying active CRISPR–Cas systems simply through genomic analysis, and we present examples of noncanonical CRISPR–Cas functions. This review emphasizes the complexities inherent to interpreting the effects that a CRISPR–Cas system might have on the fitness and physiology of a bacterial species.

The costs and benefits of CRISPR–Cas systems

Because resisting lytic phage growth is expected to always benefit a bacterium, the evolutionary advantage of possessing an active CRISPR–Cas system to destroy the genomes of these phages is clear. Indeed, models of the relationship between CRISPR–Cas and lytic phage have confirmed this assertion [24]. Further, indiscriminate foreign DNA insertions within the genome (e.g., genomic islands and prophages) can often be detrimental owing to the misregulation of new genes or the interruption of essential genes. Bacteria have developed multiple mechanisms to prevent invasions by such detrimental DNA in addition to CRISPR–Cas, such as restriction endonucleases, abortive infection systems, and phage adsorption/entry inhibitors [18]. However, invasion by DNA encoded on temperate phages, plasmids, and conjugative elements could, in some cases, be beneficial owing to advantageous genes from these sources being incorporated into the bacterial/host genome. Supporting this idea, many bacterial species acquire foreign DNA through natural competence and/or have evolved strategies to safely incorporate foreign DNA, including histone-like nucleoid-structuring protein (H-NS) mediated silencing of foreign DNA [25]. Horizontally acquired traits, such as antibiotic resistance and virulence factors, increase the fitness of many bacterial species. This presents a potential cost of CRISPR–Cas (i.e., destroying beneficial foreign DNA) in addition to the energetic costs of maintaining and producing the CRISPR–Cas surveillance system. Thus, owing to the unpredictable effects of foreign DNA uptake, evaluating the evolutionary pressure for the maintenance or loss of a CRISPR–Cas system is complicated [26].

Experiments using temperate phages, which are able to integrate their genomes into the bacterial genome, have demonstrated that CRISPR–Cas systems can block the uptake of foreign DNA into bacterial genomes. CRISPR–Cas systems can inhibit temperate phages during infection [27] or after integration has taken place [28]. The experiments by Edgar and Qimron [28] first demonstrated that bacterial genomic DNA (i.e., a lambda prophage) is not intrinsically protected from the CRISPR–Cas system, and several other experimental approaches have subsequently confirmed that a CRISPR–Cas system can kill the cell when a spacer co-exists with a chromosomal protospacer [29–31]. The ability of the CRISPR–Cas systems to target the host genome explains why, in general, very few perfect matches between CRISPR spacers and the host genome are observed [32].

Because prophages can be beneficial to the host bacterium, CRISPR–Cas systems could reduce fitness when they eliminate or prevent the acquisition of beneficial

Table 1. CRISPR–Cas systems with characterized *in vivo* effects[a]

Bacterial species	CRISPR–Cas results	Type	Refs
Laboratory experiments			
Streptococcus thermophilus	Phage targeting, spacer acquisition	II-A	[6]
Staphylococcus epidermidis	Plasmid targeting	III-A	[15]
Pseudomonas aeruginosa	Phage targeting, spacer acquisition	I-F	[27]
Pectobacterium atrosepticum	crRNA targeting self not tolerated, spacers matching mobile genetic elements	I-F	[29]
Streptococcus agalactiae	CRISPR blocked conjugation, 40% of 949 spacers target mobile genetic elements	II-A	[68]
Sulfolobus spp.	Phage and plasmid targeting, spacer acquisition	I-A	[31,69,70]
Neisseria meningitides	Natural transformation blocked	II-C	[41]
Francisella tularensis subsp. *novicida*	Endogenous gene regulation	II-B	[60]
Vibrio cholerae	Phage-encoded CRISPR–Cas system	I-F	[50]
Lactococcus lactis	Plasmid-encoded CRISPR–Cas targets phage	III-A	[45]
Haloferax volcanii	Plasmid targeting, spacers matching viruses	I-B	[71]
Streptococcus pyogenes	Plasmid targeting	II-A	[10]
Thermococcus kodakarensis	Plasmid targeting	I-A, I-B	[72]
Escherichia coli	Plasmid targeting	I-F	[73]
Inference/natural spacer matches			
Streptococcus pyogenes	Many spacers matching phages, inverse correlation with prophage	II-A, I-C	[33]
Enterococcus sp.	Spacers matching horizontal elements, inverse correlation with plasmid	II-A	[34]
Campylobacter jejuni	crRNA expression and processing	II-C	[35]
Mycoplasma gallisepticum	Related strains which infect different hosts possess different and rapidly evolving spacers	II	[74]
Xanthomonas oryzae	139 out of 203 (68%) unique spacers match phage	I-C	[75]
Leptospirillum group II	Lateral transfer of CRISPR loci and subsequent locus expansion over time	III	[76,77]
Yersinia pestis	High CRISPR locus diversity among isolates, spacers with matches to prophages	I-F	[78]
Francisella tularensis subsp. *novicida*	CRISPR spacers with matches to phage and prophage	II	[59]
Clostridium difficile	crRNA expression and processing	I-B	[47]
Sulfolobus islandicus	Extensive spacer re-assortment and diversity among related strains	I-A	[79]
Erwinia amylovora	Spacers matching a plasmid correlate with the absence of that plasmid	I-E, I-F	[80]
Porphyromonas gingivalis	1,187 diverse spacers in 60 strains, matches to insertion sequences	I-C, III-B	[81]
Repressed			
Escherichia coli	Repressed by H-NS, does not block plasmid acquisition	I-E	[52,55]
Salmonella enterica serovar Typhimurium	Repressed	I-E	[56]
Suppressed			
Pseudomonas aeruginosa	Phage-encoded anti-CRISPRs inhibit CRISPR–Cas	I-F	[57]

[a]Abbreviations: Cas, CRISPR associated; CRISPR, clustered regularly interspaced short palindromic repeat; crRNA, CRISPR RNA.

prophages. An organism with a large prophage population, *Streptococcus pyogenes*, has many virulence factors that are prophage-encoded. It has been observed that *S. pyogenes* strains with large numbers of prophages (up to eight in one genome) generally possess fewer spacers in their Type II and Type I-C CRISPR loci [33]. Although 27 out of the 41 different CRISPR spacers that were present in these strains matched streptococcal phage genomes, no single strain possessed a spacer that matched a resident prophage within the same strain. These data imply that CRISPR–Cas systems prevent the integration of phage genomes even though the phage DNA may be beneficial. Thus, the net outcome on fitness is a balance between the selective pressures that favor the presence of a specific prophage and those that favor the presence of a given CRISPR spacer, given the inability of the CRISPR–Cas system to distinguish a beneficial phage from a detrimental one.

The role of CRISPR–Cas is not limited to resisting phage infection; plasmids can also be targeted. For example, conjugation efficiency into *Staphylococcus epidermidis* was reduced by greater than 10^4-fold when the conjugated plasmid possessed a protospacer matching a spacer in the Type III-A CRISPR locus that is present in this species [15]. To assess the role of naturally occurring CRISPR–Cas systems in preventing the acquisition of plasmids, retrospective analyses have been conducted in the human pathogens *Enterococcus faecalis* and *Enterococcus faecium*. The presence of a Type II CRISPR–Cas system was found to have a significant inverse correlation with the presence of horizontally acquired antibiotic resistance genes [34], suggesting that CRISPR–Cas systems might be functioning in what would seem to be a non-beneficial manner by preventing the acquisition of useful genes. Further, RNA sequencing studies on four *Campylobacter jejuni* strains revealed that two had nonfunctional Type II-C systems (i.e., obvious *cas* gene mutations or deletions), whereas the other two strains possessed intact CRISPR–Cas systems that produced mature crRNAs. The strains with defective CRISPR–Cas systems each possessed a prophage or virulence-conferring plasmid that was not present in the other two strains, in which there was a CRISPR spacer matching the virulence plasmid [35]. These data indicate that this system is likely to be excluding potentially beneficial genes

from *C. jejuni*. Recently, an experimental approach designed to make the cell 'choose' between an antibiotic resistance-bearing plasmid and its Type III-A CRISPR–Cas system showed that *S. epidermidis* lost CRISPR–Cas function through a number of different mechanisms to acquire the plasmid under times of selection [36]. No mutations of the plasmid protospacer were seen, which is in contrast to lytic phage experiments that find evasion mutations in the phage protospacer and PAM [37]. Taken together, these cases illustrate the difficulty in rationalizing the effects and consequences of possessing an active CRISPR–Cas system, and they again show that selection does not always favor the maintenance of these systems.

In addition to phages and plasmids, CRISPR loci can interfere with competence and natural transformation. Here we present three examples of different scenarios with regards to the link between CRISPR–Cas systems and competency. In the first example, the CRISPR–Cas system directly interferes with competency and was probably selected against. Second, a scenario is discussed where the CRISPR–Cas system may fend off detrimental DNA while allowing a competent state to proceed. In the final example, a recent report is presented that demonstrates that the CRISPR–Cas system interferes with competency but is nevertheless maintained. Naturally competent bacteria such as *Streptococcus pneumoniae* can take up foreign DNA and be transformed, for example, to produce a capsule, thereby increasing the virulence of the strain in a mouse infection [38]. Interestingly, *S. pneumoniae* is naturally devoid of CRISPR–Cas systems, whereas related streptococci possess active systems of different types. When a Type II system from *S. thermophilus* was artificially introduced into *S. pneumoniae* and targeted towards capsule genes, the process of transformation no longer occurred, and mice were protected during infection [39]. At low frequencies, the introduced CRISPR–Cas system was lost, leading to DNA acquisition and a lethal mouse infection. Although an artificial set up, these results suggest that the absence of a CRISPR–Cas system has been adaptive for this organism and the cost of detrimental foreign DNA invasion (i.e., lytic phage) has perhaps been balanced by the benefit of natural competence. As a counter-example, only 30% of strains of the bacterium *Aggregatibacter actinomycetemcomitans* have maintained competency, and the others have lost this ability. The loss of competence in this species is often followed by the loss or degradation of the CRISPR–Cas system, suggesting that CRISPR–Cas may be useful in fending off detrimental foreign DNA during times of competence. Consistent with this, spacers in competent strains are enriched for phage and plasmid targets. Strains that have lost competence and much of their CRISPR–Cas system seem to acquire more foreign DNA through non-competency mechanisms (e.g., phage) than those that have maintained competence [40]. The few spacers found in non-competent strains are enriched for self-targeting sequences, suggesting a potential role in gene regulation or perhaps indiscriminate spacer acquisition in an inactive system. A final example relating competence with CRISPR–Cas function is in *Neisseria meningitidis*. The Type II-C CRISPR–Cas system of this naturally competent species seems to be highly active because isolates possesses a diverse collection of spacers, with ~97% of all database matches being to *N. meningitidis* and *N. gonorrhoeae* genomes. Despite this, most spacers do not have a match within their own genome, and those that do come with PAM mutations that are likely to eliminate self-targeting. Further, it has been shown experimentally that this CRISPR–Cas system can indeed block the transformation process. Thus, the CRISPR–Cas system seems to be able to limit inter-strain and interspecies genetic exchange within this genus, but is maintained [41]. Although the selective pressures that drive these variable outcomes relating to competency are not entirely clear, horizontal transfer and CRISPR–Cas systems certainly have a complex relationship that elicits different phenotypes over the course of evolution.

Mobile CRISPR–Cas systems

CRISPR–Cas systems are thought of as a way to resist foreign DNA invasion in the cell, but, paradoxically, these systems are found on plasmids and megaplasmids [42]. In the cyanobacterium *Synechocystis*, three CRISPR–Cas loci belonging to Type I-D and Type III were found on a single 103 kb plasmid that produced highly transcribed and processed crRNAs [43]. A Type I-C CRISPR–Cas system comprising 48 spacers was also found on a linear plasmid in *Streptomyces rochei* with no matches to any putative targets [44]. Despite the absence of chromosomally encoded CRISPRs in *Lactococcus lactis*, this organism seems to have an active plasmid-encoded Type III-A system that is self-transmissible and contains many spacers that match phage targets [45]. This sharing of plasmid-encoded CRISPR–Cas systems can result in the ability of many different related strains to exclude detrimental foreign DNA without the need to each independently acquire CRISPR spacers against common phage targets. These plasmids also enable horizontal transfer of CRISPR–Cas systems, which is likely to explain the distribution of these systems in a way that does not necessarily match the phylogeny of their host [8].

In addition to plasmid-encoded CRISPR–Cas systems, examples have emerged of phage-encoded systems. Before the roles of CRISPR–Cas systems were fully appreciated, several CRISPR–Cas loci were found in mobile elements in *Clostridium difficile*, including two in prophages [46]. Later studies revealed that *C. difficile* isolates have many CRISPR arrays (up to 34 in one isolate), along with *cas* genes, and expression was detected from all 12 of the Type I-B CRISPR loci in one strain, five of which were found in prophages [47]. In addition, metagenomic studies of the human gut have revealed examples of prophage-encoded CRISPR arrays, representing a large diversity of CRISPR types and spacers with matches to co-existing viral populations [48,49]. *Vibrio cholera* phage ICP1 encodes a functional Type I-F CRISPR–Cas system that it uses to neutralize a phage-inducible chromosomal island-like element that would otherwise mediate phage resistance. The phage CRISPR array has spacers matching the island, and protospacer mutations on the island evade CRISPR targeting, thus preventing phage infection [50]. The phage can acquire new spacers against this island to shift the balance

back to favor the phage. In these cases, either a mobile element uses its own functional CRISPR–Cas system as a means to invade the host or a newly acquired and established CRISPR–Cas system provides a fitness advantage to the recipient in times of detrimental DNA exposure. Overall, the occurrence of CRISPR–Cas systems encoded by mobile DNA again emphasizes the difficulty in distinguishing between 'good' and 'bad' foreign DNA.

Inactivation of CRISPR–Cas systems

To gauge the impact of the CRISPR–Cas system in any particular species, it is essential to determine whether the system is active under natural and/or laboratory conditions. This is a relevant concern because few species with CRISPR–Cas systems have been shown to be active through experimental challenge with phage or plasmids (Table 1). Most notably, the Type I-E system of *E. coli* is one of the most thoroughly studied systems, yet it is repressed under laboratory conditions [51]. It may also be repressed in some natural conditions because among the many spacers in the CRISPR loci of *E. coli* strains, there is little interstrain diversity observed when comparing 290 different strains [52], and few matches to sequenced phages or plasmids. These observations are consistent with a CRISPR–Cas system that has been inactive for more than 200,000 years [53,54]. Further, no CRISPR–Cas mediated exclusion of antibiotic resistance-encoding plasmids was observed among 263 *E. coli* isolates [54]. This situation contrasts with the case of *E. faecalis*, in which many matches between plasmids and CRISPR spacers were observed [34]. To elicit anti-phage activity from this system, it can be activated either by overexpression of transcriptional activator LeuO or elimination of H-NS [55], which suggests that this system has maintained functionality and is perhaps performing an alternative function. A Type I-E CRISPR–Cas system of *Salmonella enterica* serovar Typhimurium was shown to be repressed in a similar manner to the *E. coli* Type I-E system [56].

Even when a bacterial strain possesses a CRISPR–Cas system with many spacers matching existing phages, this system may still be inactive towards a given target. The most commonly observed evasion mechanism for phages has been via mutation in the protospacer or PAM region [37]. Recently, an alternative approach to evading CRISPR–Cas activity that does not require mutation (and is thus harder to predict bioinformatically) was described in *Pseudomonas aeruginosa*. Prophages in various strains of this species express genes that specifically inactivate the resident Type I-F CRISPR–Cas system, leading to co-existing spacer–protospacer matches in the same genome [57]. Prophage-borne anti-CRISPR genes may mediate a more transient inactivation of the CRISPR–Cas system than is seen in *E. coli* because prophages can be acquired readily. In addition, prophages can potentially be lost more often than a typical chromosomal region through prophage excision followed by a failed lytic cycle. It is notable that anti-CRISPR genes have also been observed on other mobile DNA elements in *P. aeruginosa* [57]; thus, the influence of this group of genes in this species may be profound. Although no homologues to the *P. aeruginosa* genes encoding anti-CRISPRs have been detected outside this genus, distinct classes of anti-CRISPR genes might exist in phages or mobile DNA elements of other species. In fact, given the potential evolutionary advantage provided by such genes, it would be surprising if they did not exist elsewhere. Of course, extending the theme of the complexity of these systems, phage-borne anti-CRISPR genes are a boon during lytic growth, but a bane for lysogens because they render the cell more susceptible to further phage infection [57]. A clue for the presence of an anti-CRISPR mechanism in a given strain would be co-existing spacer–protospacer matches with the correct PAM in the same genome or widespread mobile elements that seem to be recalcitrant to CRISPR–Cas targeting in organisms with intact systems.

Novel CRISPR–Cas functions

Upon the discovery of CRISPR loci through bioinformatic means, it was hypothesized that they might be involved in gene regulation, analogous to functions of RNAi in eukaryotes [58]. This hypothesis was strengthened in 2009 when it was shown that the Type III-B CRISPR–Cas system of *Pyrococcus furiosus* does indeed target RNA *in vitro* [16], shortly after DNA targeting had been demonstrated in the Type III-A system of *Staphylococcus epidermidis* [15]. Although the major role of most CRISPR–Cas systems certainly seems to be resisting invasion by foreign DNA, novel roles of CRISPR–Cas systems in gene regulation have emerged in recent years. For example, *Francisella tularensis* subsp. *novicida* seems to have an active Type II CRISPR–Cas system with a full suite of *cas* genes and arrays with multiple spacers matching phages [59]. In addition, this system mediates repression of an endogenous lipoprotein-encoding gene through imperfect base pairing between a *trans*-activating crRNA (tracrRNA) and the target transcript, in a region spanning the start codon [60]. This endogenous gene regulation is necessary for full virulence of *F. novicida* in mice and represents the first characterized demonstration of CRISPR–Cas-mediated gene regulation along with a role in virulence. An unprocessed orphan precursor crRNA (i.e., not associated with any *cas* gene) in *Listeria monocytogenes* was found to be expressed as a five-repeat unit containing spacers that matched a host mRNA in two locations. This crRNA formed a duplex with the target mRNA *in vitro*, and overexpression of the crRNA seemed to stabilize the target, causing an increase of mRNA levels *in vivo* [61]. In *P. aeruginosa*, when phage DMS3 is present as a prophage, it mediates the inhibition of biofilm formation [62]. This inhibition is dependent on full activity of the *P. aeruginosa* CRISPR–Cas system and also required a CRISPR spacer matching a region of DMS3 with five mismatches, demonstrating that CRISPR–Cas recognition and function might be more plastic than originally thought [63]. Although DNA cleavage in this case is unlikely, because it would kill the cell, there might be an effect on transcript production from this region due to CRISPR–Cas complex recruitment. On mutation of the DMS3-targeted region to four or zero mismatches with the crRNA, the CRISPR–Cas system resisted phage infection, demonstrating a gradient of activity in the presence of mismatches [27]. These data show that not all mismatches will necessarily abolish cleavage or

Table 2. Genomic hallmarks of active versus inactive CRISPR–Cas systems[a]

Hallmark	Description
Intact CRISPR locus and *cas* genes	An intact system is defined as one that is similar to an active system (i.e., a system of the same type with conserved repeat and Cas protein sequences), that has an intact CRISPR locus, and that has a complete set of *cas* genes with no obviously debilitating mutations (e.g., frameshifts or nonsense mutations)
Spacer diversity	Many different CRISPR spacers are found when comparing strains of a given species. This suggests that each strain is actively acquiring new and diverse spacers. *Leptospirillum* isolates in microbial biofilms are closely related with distinct CRISPRs, and no two strains had same CRISPR spacers [76]. *Pseudomonas aeruginosa* has spacer diversity in related isolates [82], and *Streptococcus agalactiae* has 109 unique spacer arrangements in 124 strains [68], whereas many *Escherichia coli* and *Salmonella enterica* isolates have identical arrays [52,56]. The ancestral end (leader distal) of the array can be used to compare loci in related strains, and leader end diversity demonstrates recent acquisition. Caveat: horizontal transfer of 'successful' CRISPR arrays could permeate through a population, thus appearing homogenous
Protospacers match known phages or plasmids	In species with known active systems, matches are observed to phages (e.g., *P. aeruginosa*) [27], plasmids (e.g., *Staphylococcus epidermidis*) [15], and other genomes (e.g., *Neisseria meningitidis*) [41]. These matches may be 100% complementary, indicating the potential for interference, or may contain mismatches that could indicate mutations to evade the CRISPR–Cas system. Species with no matches to sequenced elements may indicate function has been lost. Caveat: some species may have very few sequenced phages and plasmids for comparison
Self-targeting	Strains with active CRISPR–Cas systems generally do not display protospacers with perfect complementarity to self-spacers. Perfect self-matches mapping to mobile regions indicate active suppression [57] or repression [55] mechanisms, possibly encoded on the mobile element. Mismatches to mobile elements may indicate a functioning system being evaded, and mismatches mapping to the 'core' genome might indicate the potential for alternative/regulatory functions. Caveat: a perfect match might not be relevant if coupled with a PAM mutation [29]
Lower frequency of mobile elements	Strains with active CRISPR–Cas systems generally are less likely to possess mobile elements. This can be established through intraspecies comparisons. More CRISPRs correlate with fewer prophages and plasmids in *Streptococcus pyogenes* [33] and *Enterococcus* spp. [34]. Caveat: widespread suppression mechanisms of the CRISPR–Cas system could allow acquisition of foreign elements despite matches

[a]Abbreviations: Cas, CRISPR associated; CRISPR, clustered regularly interspaced short palindromic repeat; PAM, protospacer adjacent motif.

recognition by the CRISPR–Cas system, consistent with work in the Type I-E CRISPR–Cas system of *E. coli* showing that certain mismatches do not eliminate target DNA binding [64]. The knowledge that protospacers with up to five mismatches might still mediate *in vivo* function presents a challenge when attempting to identify CRISPR–Cas targets bioinformatically, although excellent tools exist to approach this problem, such as CRISPRTarget [65].

Finally, additional roles for *cas* genes have been previously described, but they do not seem to be dependent on the CRISPR locus and thus represent isolated gene functions. For example, Cas1 and Cas2 proteins, which have been implicated in DNA acquisition, have individually been attributed to playing a role in DNA repair [66] and in *Legionella pneumophila* virulence [67], respectively. We expect that future studies will uncover many more instances of alternative roles for CRISPR–Cas systems that may involve varying degrees of mismatch to protospacers.

Concluding remarks

CRISPR–Cas systems provide a powerful means for bacteria to destroy potentially harmful foreign DNA, and the common occurrence of these systems within bacterial genomes emphasizes the positive influence that these systems must have on evolutionary fitness. However, as summarized here, the current literature provides many examples in which the acquisition of foreign DNA may be advantageous for an organism and the possession of an active CRISPR–Cas system could be non-adaptive. This is likely to explain the observation that <50% of sequenced bacteria possess these systems, despite their ability to be horizontally transferred. The principle conclusion of this review is that the net biological outcome of a CRISPR–Cas system within a given organism in a given environment is difficult to predict. This uncertainty arises because the balance between the beneficial and detrimental effects of foreign DNA depends on the nature of the DNA being acquired, and this property varies among species. In addition, some systems may be inactive (i.e., suppressed, repressed, or defective) or be performing alternative roles. For improvement of our understanding of the complex biological outcomes that can result from the presence of CRISPR–Cas systems, future studies must focus on *in vivo* characterization of more systems operating in diverse species. As shown in Table 1, only a small fraction of systems have been analyzed *in vivo*. Future *in vivo* work would address the important questions of what percentage of seemingly intact CRISPR–Cas systems are actually able to resist the invasion of foreign DNA, and also whether CRISPR–Cas systems are commonly performing alternative functions (Box 1). The recent discovery of anti-CRISPR genes in *P. aeruginosa* raises the possibility that many such systems may exist; thus, the identification and characterization of more CRISPR–Cas suppression mechanisms will be crucial for assessing the general impact of CRISPR–Cas systems. An intuitive extension of anti-CRISPR findings may be *cas* or host genes that inhibit these elements, acting as 'anti-anti-CRISPRs', a phenomenon that has been observed with restriction enzymes [18].

Common genomic hallmarks of active and inactive CRISPR–Cas systems are outlined in Table 2 to facilitate

Box 1. Outstanding questions

- What percentage of CRISPR–Cas (clustered regularly interspaced short palindromic repeat–CRISPR associated) systems are actively excluding foreign DNA?
- To what extent does non-canonical CRISPR–Cas function shape prokaryotic biology?
- How prevalent are CRISPR–Cas suppression and repression mechanisms?
- What are the evolutionary downsides to possessing a CRISPR–Cas system?

genomic analyses of a strain or species of interest. Finally, we encourage all investigators who are analyzing bacterial genomes to carefully study the spacer content of CRISPR loci and to clearly enumerate the percentage of spacers that match known invasive DNA (e.g., plasmids and phages), that are self-complementary, or that are unique in individual strains of a given species. This information is often difficult to find in publications, yet it provides crucial insight into whether a given CRISPR–Cas system is functional. Increased accumulation of data pertaining to the *in vivo* functioning of CRISPR–Cas systems will allow accurate interpretation of the roles that these systems are playing in various bacterial species. Because CRISPR–Cas systems can provide a unique 'fossil record' of encounters with foreign DNA within bacterial species, this knowledge will greatly improve our understanding of bacterial evolution and the impact of horizontal gene transfer on the environment and human health.

Acknowledgments
CRISPR–Cas research in the Davidson laboratory is currently funded by Canadian Institute of Health Research (CIHR) grant MOP-130482 and was previously funded by CIHR grant XNE86943. We thank April Pawluk for comments on the manuscript.

References
1 Ishino, Y. *et al.* (1987) 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
2 Jansen, R. *et al.* (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575
3 Mojica, F.J.M. *et al.* (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182
4 Pourcel, C. *et al.* (2005) 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
5 Bolotin, A. *et al.* (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151, 2551–2561
6 Barrangou, R. *et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712
7 Grissa, I. *et al.* (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57
8 Makarova, K.S. *et al.* (2011) Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* 9, 467–477
9 Bhaya, D. *et al.* (2011) CRISPR–Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.* 45, 273–297
10 Deltcheva, E. *et al.* (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607
11 Haurwitz, R.E. *et al.* (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329, 1355–1358
12 Sternberg, S.H. *et al.* (2012) Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* 18, 661–672
13 Brouns, S.J.J. *et al.* (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964
14 Garneau, J.E. *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71
15 Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845
16 Hale, C.R. *et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139, 945–956
17 Mojica, F.J.M. *et al.* (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740
18 Labrie, S.J. *et al.* (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8, 317–327
19 Samson, J.E. *et al.* (2013) Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* 11, 675–687
20 Stern, A. *et al.* (2012) CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* 22, 1985–1994
21 Rho, M. *et al.* (2012) Diverse CRISPRs evolving in human microbiomes. *PLoS Genet.* 8, e1002441
22 Reeks, J. *et al.* (2013) CRISPR interference: a structural perspective. *Biochem. J.* 453, 155–166
23 Westra, E.R. *et al.* (2012) The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annu. Rev. Genet.* 46, 311–339
24 Levin, B.R. (2010) Nasty viruses, costly plasmids, population dynamics, and the conditions for establishing and maintaining CRISPR-mediated adaptive immunity in bacteria. *PLoS Genet.* 6, e1001171
25 Navarre, W.W. *et al.* (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313, 236–238
26 Weinberger, A.D. and Gilmore, M.S. (2012) CRISPR–Cas: to take up DNA or not-that is the question. *Cell Host Microbe* 12, 125–126
27 Cady, K.C. *et al.* (2012) The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* 194, 5728–5738
28 Edgar, R. and Qimron, U. (2010) The *Escherichia coli* CRISPR system protects from lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192, 6291–6294
29 Vercoe, R.B. *et al.* (2013) Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet.* 9, e1003454
30 Yosef, I. *et al.* (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* 40, 5569–5576
31 Manica, A. *et al.* (2011) In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol. Microbiol.* 80, 481–491
32 Stern, A. *et al.* (2010) Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26, 335–340
33 Nozawa, T. *et al.* (2011) CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. *PLoS ONE* 6, e19543
34 Palmer, K.L. and Gilmore, M.S. (2010) Multidrug-resistant enterococci lack CRISPR–Cas. *MBio* 1, e00227–e310
35 Dugar, G. *et al.* (2013) High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* 9, e1003495
36 Jiang, W. *et al.* (2013) Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.* 9, e1003844
37 Deveau, H. *et al.* (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400
38 Avery, O.T. *et al.* (1944) Studies on the chemical nature of the substance inducing transformation of pneuomococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type III. *J. Exp. Med.* 79, 137–158
39 Bikard, D. *et al.* (2012) CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* 12, 177–186
40 Jorth, P. and Whiteley, M. (2012) An evolutionary link between natural transformation and CRISPR adaptive immunity. *MBio* 3, e00309–e312
41 Zhang, Y. *et al.* (2013) Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* 50, 488–503
42 Godde, J.S. and Bickerton, A. (2006) The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62, 718–729
43 Scholz, I. *et al.* (2013) CRISPR–Cas systems in the cyanobacterium *Synechocystis* sp. PCC6803 exhibit distinct processing pathways involving at least two Cas6 and a Cmr2 protein. *PLoS ONE* 8, e56470
44 Yang, Y. *et al.* (2011) pSLA2-M of *Streptomyces rochei* is a composite linear plasmid characterized by self-defense genes and homology with pSLA2-L. *Biosci. Biotechnol. Biochem.* 75, 1147–1153
45 Millen, A.M. *et al.* (2012) Mobile CRISPR/Cas-mediated bacteriophage resistance in *Lactococcus lactis*. *PLoS ONE* 7, e51663
46 Sebaihia, M. *et al.* (2006) The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat. Genet.* 38, 779–786

47 Soutourina, O.A. *et al.* (2013) Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet.* 9, e1003493
48 Minot, S. *et al.* (2011) The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* 21, 1616–1625
49 Minot, S. *et al.* (2013) Rapid evolution of the human gut virome. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12450–12455
50 Seed, K.D. *et al.* (2013) A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* 494, 489–491
51 Pul, U. *et al.* (2010) Identification and characterization of *E. coli* CRISPR–Cas promoters and their silencing by H-NS. *Mol. Microbiol.* 75, 1495–1512
52 Touchon, M. *et al.* (2011) CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J. Bacteriol.* 193, 2460–2467
53 Touchon, M. and Rocha, E.P.C. (2010) The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS ONE* 5, e11126
54 Touchon, M. *et al.* (2012) Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. *Microbiology* 158, 2997–3004
55 Westra, E.R. *et al.* (2010) H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77, 1380–1393
56 Medina-Aparicio, L. *et al.* (2011) The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 193, 2396–2407
57 Bondy-Denomy, J. *et al.* (2013) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493, 429–432
58 Makarova, K.S. *et al.* (2006) 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
59 Schunder, E. *et al.* (2013) First indication for a functional CRISPR/Cas system in *Francisella tularensis*. *Int. J. Med. Microbiol.* 303, 51–60
60 Sampson, T.R. *et al.* (2013) A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497, 254–257
61 Mandin, P. *et al.* (2007) Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.* 35, 962–974
62 Zegans, M.E. *et al.* (2009) Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191, 210–219
63 Cady, K.C. and O'toole, G.A. (2011) Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* 193, 3433–3445
64 Semenova, E. *et al.* (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10098–10103
65 Biswas, A. *et al.* (2013) CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol.* 10, 817–827
66 Babu, M. *et al.* (2011) A dual function of the CRISPR–Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* 79, 484–502
67 Gunderson, F.F. and Cianciotto, N.P. (2013) The CRISPR-associated gene *cas2* of *Legionella pneumophila* is required for intracellular infection of amoebae. *MBio* 4, e00074–13 (http://dx.doi.org/10.1128/mBio.00074-13)
68 Lopez-Sanchez, M-J. *et al.* (2012) The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol. Microbiol.* 85, 1057–1071
69 Gudbergsdottir, S. *et al.* (2011) Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* 79, 35–49
70 Erdmann, S. and Garrett, R.A. (2012) Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol.* 85, 1044–1056
71 Fischer, S. *et al.* (2012) An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J. Biol. Chem.* 287, 33351–33363
72 Elmore, J.R. *et al.* (2013) Programmable plasmid interference by the CRISPR–Cas system in *Thermococcus kodakarensis*. *RNA Biol.* 10, 828–840
73 Almendros, C. *et al.* (2012) Target motifs affecting natural immunity by a constitutive CRISPR–Cas system in *Escherichia coli*. *PLoS ONE* 7, e50797
74 Delaney, N.F. *et al.* (2012) Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLoS Genet.* 8, e1002511
75 Semenova, E. *et al.* (2009) Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. *FEMS Microbiol. Lett.* 296, 110–116
76 Tyson, G.W. and Banfield, J.F. (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ. Microbiol.* 10, 200–207
77 Andersson, A.F. and Banfield, J.F. (2008) Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320, 1047–1050
78 Cui, Y. *et al.* (2008) Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLoS ONE* 3, e2652
79 Held, N.L. *et al.* (2013) Reassortment of CRISPR repeat-spacer loci in *Sulfolobus islandicus*. *Environ. Microbiol.* http://dx.doi.org/10.1111/1462-2920.12146
80 Rezzonico, F. *et al.* (2011) Diversity, evolution, and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. *Appl. Environ. Microbiol.* 77, 3819–3829
81 Watanabe, T. *et al.* (2013) CRISPR regulation of intraspecies diversification by limiting IS transposition and intercellular recombination. *Genome Biol. Evol.* 5, 1099–1114
82 Cady, K.C. *et al.* (2011) Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* 157, 430–437