Evolution and Ecology of CRISPR

Edze R. Westra, Andrea J. Dowling, Jenny M. Broniewski, and Stineke van Houte

Environment and Sustainability Institute and Centre for Ecology and Conservation, Biosciences, University of Exeter, Tremough Campus, Penryn TR10 9FE, United Kingdom; email: E.R.Westra@exeter.ac.uk

Annu. Rev. Ecol. Evol. Syst. 2016. 47:307–31

First published online as a Review in Advance on August 24, 2016

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi: 10.1146/annurev-ecolsys-121415-032428

Keywords

bacteria, phage, virus, defense, adaptive immunity, CRISPR, Cas, evolution, ecology

Abstract

CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) systems are prokaryotic adaptive immune systems that provide protection against infection by parasitic mobile genetic elements, such as viruses and plasmids. CRISPR-Cas systems are found in approximately half of all sequenced bacterial genomes and in nearly all archaeal genomes. In this review, we summarize our current understanding of the evolutionary ecology of CRISPR-Cas systems, highlight their value as model systems to answer fundamental questions concerning host–parasite coevolution, and explain how CRISPR-Cas systems can be useful tools for scientists across virtually all disciplines.

1. INTRODUCTION

The discovery of CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) adaptive immune systems has revolutionized the study of life sciences. CRISPR-Cas systems, encoded on prokaryotic genomes, consist of a CRISPR array and *cas* genes. CRISPR arrays form the genetic memory of the prokaryotic adaptive immune system. They are composed of repeating sequences (repeats) that are interspersed by variable sequences (spacers) that match sequences from mobile genetic elements such as viruses (**Figure 1**). The number and length of CRISPR loci vary among organisms, with the highest number of 23 CRISPR loci found

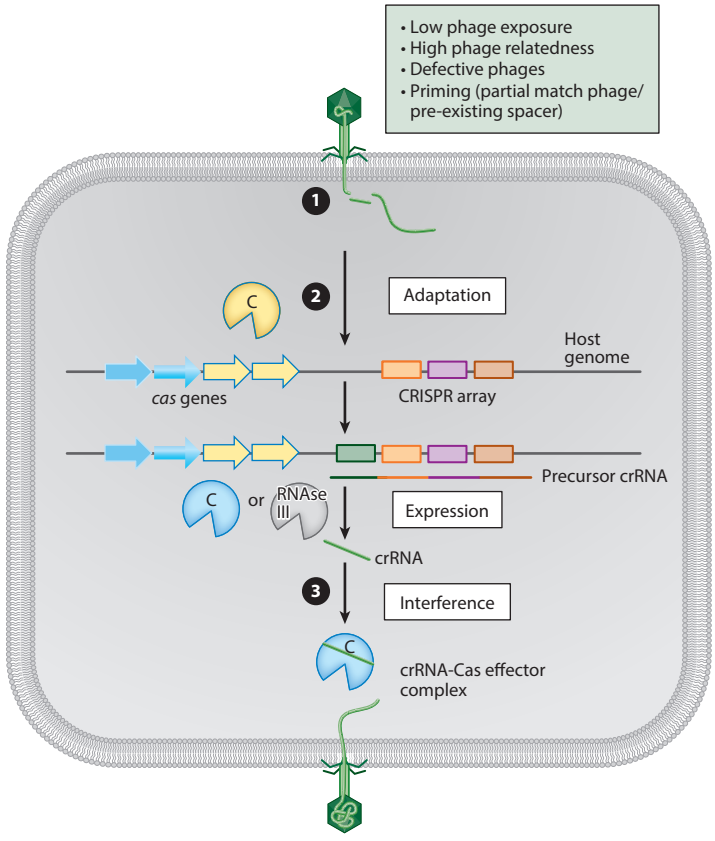

Figure 1

Mechanism of CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) systems and ecological factors that impact their evolution. During adaptation (upon phage infection), Cas1 and Cas2 capture a piece of phage DNA ❶ that is integrated into the host CRISPR locus. Cas9 is involved in protospacer adjacent motif (PAM) selectivity during this process in Class 2/Type II systems, whereas in Class 1/Type I systems, Cas1 and Cas2 have intrinsic PAM recognition ability. ❷ During expression, the CRISPR locus is transcribed into precursor CRISPR RNA (crRNA). This precursor crRNA is processed into crRNA either by Cas6-type endoribonucleases (Class 1 systems, with the exception of Type I-C, in which Cas5d carries out this role) or by RNase III (Class 2/Type II systems). ❸ During interference, crRNA–Cas complexes recognize complementary nucleic acids of a related phage followed by either recruitment of an effector Cas nuclease or target cleavage by the crRNA–Cas complex. Ecological factors that are predicted or shown to promote the evolution of CRISPR-Cas immunity are shown in the green box and include low phage exposure (which reduces the inducible fitness cost of CRISPR-Cas), high phage relatedness (which ensures that the CRISPR memory enables recognition upon secondary infection), and defective phages and priming (both of which increase the efficiency of adaptation).

in *Methanocaldococcus* sp. FS406-22 and the highest number of 600 individual spacers found in *Haliangium ochraceum* (Anderson et al. 2011). By comparison, a common model organism such as *Pseudomonas aeruginosa* UCBPP-PA14 has a single set of *cas* genes flanked by 2 CRISPR loci (with 21 and 14 spacers) (Cady et al. 2011). *Cas* genes encode the protein machinery that uses the information stored in the CRISPR array to launch an immune response against mobile genetic elements that carry a cognate sequence.

The first identification of a CRISPR locus dates back to 1987 (Ishino et al. 1987), but not until 2005 was it proposed that CRISPR loci act together with *cas* genes to provide adaptive immunity (Lillestol et al. 2006, Makarova et al. 2006, Mojica et al. 2005, Pourcel et al. 2005). The seminal paper by Barrangou and coworkers (2007) provided the first experimental evidence that confirmed this hypothesis. The authors exposed the lactic acid bacterium *Streptococcus thermophilus* strain DGCC7710 to phage and found that bacterial clones acquired resistance against phage by incorporating novel phage-derived sequences (spacers) into the CRISPR array. Spacers are thus acquired during the lifetime of an individual clone and the resultant phage resistance phenotype is inherited by future generations (Koonin & Wolf 2009, 2016).

During the past decade many mechanistic details of CRISPR-Cas have been elucidated, which has allowed these systems to be repurposed for sophisticated genome editing and gene regulation technologies. Although clear mechanistic differences exist between variants of the system, the general mode of action is that CRISPR transcripts are processed into small CRISPR RNA (crRNA) molecules that guide Cas proteins to bind and cleave complementary parasitic nucleic acids (see sidebar, The Three Stages of a CRISPR-Cas Immune Response). The mechanism of CRISPR-Cas is generally divided into three stages: adaptation, expression, and interference. During adaptation novel spacers are integrated into the CRISPR array [reviewed by Amitai & Sorek (2016)]. In the expression stage the CRISPR array is transcribed and processed into short crRNA molecules that are loaded onto Cas proteins [reviewed by Charpentier et al. (2015)]. During interference Cas–crRNA ribonucleoprotein complexes bind to complementary nucleic acids, followed by degradation of the target molecule (reviewed in van der Oost et al. 2014).

Although the mechanistic studies of CRISPR-Cas have raced ahead, we are only starting to understand the evolutionary ecology of CRISPR-Cas. In this review we first describe the scenario

THE THREE STAGES OF A CRISPR-Cas IMMUNE RESPONSE

CRISPR-Cas immune responses are generally divided into adaptation, expression, and interference. Adaptation always requires at least Cas1 and Cas2 (Yosef et al. 2012), which are the most conserved *cas* genes (Takeuchi et al. 2012) and are associated with nearly all CRISPR-Cas systems (however, sometimes Cas1 and Cas2 are supplied *in trans* by other CRISPR-Cas systems) (Makarova et al. 2015). Recent in vivo and in vitro studies demonstrated how Cas1 and Cas2 form a heterotetrameric complex that binds to the leader end of the CRISPR array, where Cas1 catalyzes the integration of spacers through a mechanism that shares similarities with retrovirus integration and DNA transposition [reviewed by Amitai & Sorek (2016)]. The polarity of spacer incorporation at the leader end of the CRISPR array (Barrangou et al. 2007, Yosef et al. 2012) results in a genetically encoded chronological record of previous host–parasite interactions (Vale & Little 2010). During expression, the CRISPR array is transcribed into a precursor crRNA, which is cleaved either by Cas6-like endoribonuclease (Class 1 systems) or by RNase III or Cpf1 (Class 2 systems) to yield mature crRNA that associate with Cas protein(s) (Charpentier et al. 2015, Fonfara et al. 2016). Each individual crRNA–Cas complex carries sequence information to bind a single complementary nucleic acid molecule (reviewed in van der Oost et al. 2014). During interference, crRNA–Cas complexes bind and cleave the complementary nucleic acid or mark it for destruction by an effector Cas nuclease.

for the evolution of CRISPR-Cas systems as proposed by Koonin and coworkers (Shmakov et al. 2015). The core of this review focuses on three important but as yet largely unanswered questions: (*a*) How important are CRISPR-Cas systems in shaping bacteria–phage interactions? (*b*) When do CRISPR-Cas systems matter in natural populations? (*c*) What are the long-term consequences of encoding CRISPR-Cas systems in microbial genomes? Finally, we highlight some of the various applications of CRISPR-Cas for evolutionary and ecological research. The mechanism of CRISPR-Cas systems has been thoroughly reviewed elsewhere; however, essential details concerning the mechanistic basis of the different CRISPR-Cas systems are given in the sidebars. For more detailed overviews, we refer to some of the many excellent reviews on this subject (Makarova et al. 2015, van der Oost et al. 2014, Wiedenheft et al. 2012).

2. THE EVOLUTION OF CRISPR-Cas SYSTEMS

CRISPR-Cas systems are extremely diverse (see sidebars) and are currently classified into 2 classes, 6 types, and 19 subtypes (**Figure 2**) (Makarova et al. 2015, Shmakov et al. 2015). Despite this diversity, all systems share the basic features of a CRISPR locus, consisting of repeats alternated by variable spacers, and a set of associated *cas* genes. The diverse nature of CRISPR-Cas, which likely results from rapid evolution and extensive horizontal gene transfer (HGT), poses considerable challenges for the evolutionary classification of CRISPR-Cas systems.

Two proteins that are present in almost all CRISPR-Cas systems (but see Makarova et al. 2015 for exceptions to this rule) are Cas1 and Cas2, which are involved in the adaptation phase (see sidebar, The Three Stages of a CRISPR-Cas Immune Response). Cas1 is the most highly conserved Cas protein and hence the most suitable marker to trace the evolutionary history of CRISPR-Cas. The phylogeny of Cas1 generally corresponds well to the subtype classification of CRISPR-Cas (see sidebars, Class 1 CRISPR-Cas Systems and Class 2 CRISPR-Cas Systems). Not all *cas1* genes, however, are CRISPR associated. Two clades of orphan *cas1* (coined Cas1-solo) were identified that predominantly occur in archaeal lineages, and these have been suggested to represent ancestral genes of CRISPR-associated *cas1* genes (Makarova et al. 2013). In some instances Cas1-solo genes are found within predicted transposable elements termed casposons, genomic islands flanked by terminal inverted repeats (Krupovic et al. 2014). Koonin and colleagues hypothesized that casposon-encoded Cas1 functions in the catalysis of casposon integration into the host genome, using a mechanism that may be akin to the spacer integration process, and speculated that casposon-encoded Cas1 may have played a key role in the early evolution of CRISPR-Cas (Krupovic et al. 2014, 2016). Specifically, they suggest that CRISPR-Cas systems may have emerged from a fusion of casposon and a *cas10*-like innate immunity gene.

Apart from Cas1, almost all CRISPR-Cas systems carry Cas2, which shares structural similarity with VapD, a toxin from *Haemophilus influenzae* (Kwon et al. 2012). Based on this finding, it has been suggested that Cas2 and other Cas proteins may act as a toxin to induce cell dormancy (He et al. 2014, Makarova et al. 2012), but this hypothesis awaits further experimental testing. Class 1 systems further encode a suite of genes that are essential for the expression and interference stages. Many of these genes belong to the RAMP (repeat associated mysterious protein) family. Makarova and coworkers (2011a, 2013) suggested a series of gene duplication events of an ancestral RAMP that carried a single RNA recognition motif (RRM) may have led to the relatively complex architecture of Class 1 systems of multiple RAMP genes, many of which carry multiple RRMs. Under this scenario, Class 1 systems evolved in thermophilic archaea and subsequently spread to bacteria (Makarova et al. 2011a). Class 2 systems lack much of the complexity observed for Class 1 systems and rely on an evolutionarily unrelated mechanism for target cleavage (see sidebar, Class 2 CRISPR-Cas Systems) in which the whole protein machinery necessary for the Class 1 interference stage is replaced by a single protein (Koonin & Makarova 2013) (*cas9* gene in Type II systems;

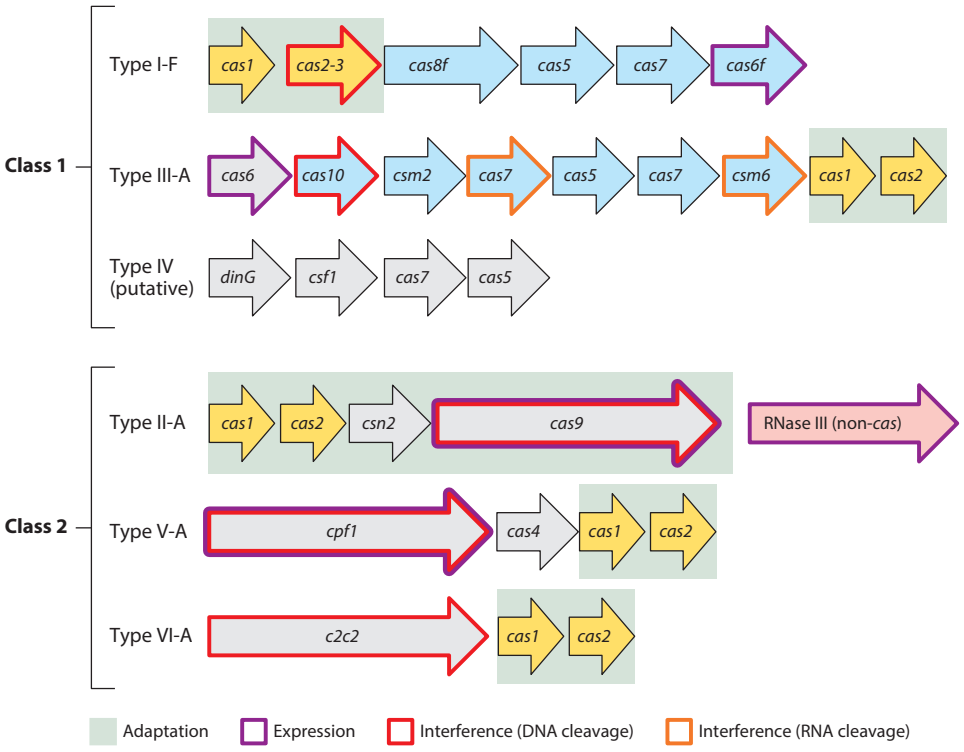

Figure 2

Examples of *cas* operon organization for CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) types belonging to each class. As shown in the color legend at the bottom of the figure, genes involved in adaptation, expression, and interference are color coded. Genes that encode multiprotein/CRISPR RNA (crRNA) complexes are shown in blue. Type IV systems await biochemical characterization. Adaptation in Type V and VI systems is likely to involve *cas1* and *cas2*, but experimental tests are lacking. Expression and interference in Type V systems involve *cpf1*, which encodes an enzyme that carries out both precursor crRNA and target DNA cleavage (Fonfara et al. 2016). The biochemistry of precursor crRNA processing in Type VI systems and the possible involvement of *cpf1* and *c2c2* in spacer acquisition (protospacer adjacent motif selection) in both Type V and VI systems has not yet been examined. In Type II-A systems, *cas9* is involved in both of these processes; precursor crRNA processing in Type II systems also requires the housekeeping enzyme RNase III and a *trans*-activating crRNA (Deltcheva et al. 2011). Some Type V systems also require *trans*-activating crRNA, but Type VI systems lack this requirement (Shmakov et al. 2015).

cpf1, *c2c1*, or *c2c3*, in Type V systems; and *c2c2* in Type VI systems; see **Figure 2** and sidebar, New Class 2 CRISPR-Cas Systems) (Makarova et al. 2015, Shmakov et al. 2015, Zetsche et al. 2015). It has been suggested by Shmakov et al. (2015) that the Class 2 effector nucleases may have acted as a stand-alone immune mechanism and may have been co-opted from a mobile genetic element by CRISPR-Cas to replace the typical Class 1 *cas* genes (Makarova et al. 2011a), resulting in the emergence of the two distinct classes of CRISPR-Cas. The sequence of events outlined in Shmakov et al. (2015) is shown in **Figure 3.**

3. HOW IMPORTANT ARE CRISPR-Cas SYSTEMS IN NATURE?

Although bacteria have many different immune mechanisms, CRISPR-Cas systems are—at least in terms of their mechanism—arguably their most sophisticated defense, and it is therefore tempting

CLASS 1 CRISPR-Cas SYSTEMS

Class 1 systems encode multisubunit crRNA–Cas complexes and are subdivided into Type I, III, and IV systems. Precursor crRNA processing is catalyzed by Cas6-type endoribonucleases, sometimes followed by further crRNA maturation steps by unidentified nucleases (reviewed in Charpentier et al. 2015). The model Type I system of *Escherichia coli* K12 encodes the crRNA–Cas complex known as Cascade, which consists of five different Cas proteins ($Cse1_1$, $Cse2_2$, $Cas5e_1$, $Cas7_6$, $Cas6e_1$), and a single crRNA (reviewed in van der Oost et al. 2014). The crRNA consists of a spacer sequence flanked by partial repeats that serve as conserved handles for the Cascade complex (van der Oost et al. 2014). Upon binding target DNA, Cascade recruits the Cas3 nuclease enzyme, which cleaves and digests the target (van der Oost et al. 2014). In agreement with distant phylogenetic relationships between Type I and Type III systems (Koonin & Makarova 2013; Makarova et al. 2011a,b, 2013, 2015), their associated multisubunit crRNA–Cas complexes share key structural features (reviewed in Jackson & Wiedenheft 2015). However, Type III complexes target both single-stranded RNA and transcriptionally active DNA (reviewed in Jackson & Wiedenheft 2015). The putative Type IV systems have been proposed only recently (Makarova et al. 2015) and await characterization.

to assume that CRISPR-Cas systems are of key importance for bacteria–virus interactions in nature. What evidence there is suggests that CRISPR loci can evolve rapidly in some environments, consistent with an important role in antagonistic coevolution.

First, CRISPR loci belonging to different populations of the same species are typically highly diverse regarding their spacer content (Andersson & Banfield 2008, DeBoy et al. 2006, Heidelberg et al. 2009, Held et al. 2010, Kunin et al. 2008, Pourcel et al. 2005, Rho et al. 2012, Tyson & Banfield 2008). Second, direct analysis of CRISPR loci over time reveals dynamic spacer content, for example, in microbial metapopulations from Lake Tyrell (Emerson et al. 2013) and in streptococci in saliva samples from healthy human individuals (Pride et al. 2011). Third, CRISPR spacers tend to match phage genomes from the same (sympatric) rather than other (allopatric) environments, demonstrating that CRISPR-based antagonistic coevolution occurs at a timescale

CLASS 2 CRISPR-Cas SYSTEMS

Despite being less common (Chylinski et al. 2014, Makarova et al. 2015), Class 2 systems received more attention owing to their application in genome editing. They are uniquely suited for this application, because a single protein carries out all functions of the multisubunit crRNA–Cas complexes of Class 1 systems. The Cas9 enzyme that is now widely used for genome editing (Sternberg & Doudna 2015) is encoded by Type II systems—one of the Class 2 types (Makarova et al. 2015). Type II systems encode a *trans*-activating crRNA (tracrRNA), which is essential for Cas9-dependent, RNase III–catalyzed cleavage of precursor crRNA (Deltcheva et al. 2011). Some Type II-C systems have a variant crRNA maturation pathway that involves transcription of short RNA molecules from promoters contained in CRISPR repeats (reviewed in Charpentier et al. 2015). The tracrRNA remains bound to the processed crRNA and forms an essential component of the tracrRNA–crRNA–Cas9 effector complex (Jinek et al. 2012). During interference, the effector complex binds and cleaves target dsDNA in a protospacer adjacent motif (PAM)–dependent manner (see sidebar, The Protospacer Adjacent Motif) (Sternberg et al. 2014), which results in a blunt-end cleavage product (Garneau et al. 2010). Recent high-resolution structures and fluorescence (Förster) resonance energy transfer (FRET) studies have propelled our understanding of the mechanism of recognition and cleavage of dsDNA target molecules by the tracrRNA–crRNA–Cas9 effector complex (reviewed in Wright et al. 2016).

NEW CLASS 2 CRISPR-Cas SYSTEMS

Two new Class 2 variants have been discovered only very recently: Type V and Type VI systems. Type V systems are further subdivided into Type V-A, V-B, and V-C systems (Shmakov et al. 2015). Type V systems encode Cpf1, C2c1, or C2c3 effector enzymes (Makarova et al. 2015), and Type VI systems encode the C2c2 effector enzyme (Shmakov et al. 2015). A recent study demonstrated that the Type V-A Cpf1 effector enzyme is a single crRNA-guided endonuclease (i.e., lacks a requirement for tracrRNA) and that crRNA processing is carried out by Cpf1 (Fonfara et al. 2016). Whereas mature Type II crRNA starts with a 5′ spacer sequence followed by a partial repeat (Charpentier et al. 2015, Deltcheva et al. 2011), mature crRNA of Type V-A starts with a 5′ partial repeat followed by a spacer sequence (Zetsche et al. 2015). Another clear difference with Type II systems is that Cpf1 introduces a staggered rather than a blunt double-stranded break in the complementary target DNA. Production of mature crRNA by the Type V-B C2c1 effector enzyme does depend on tracrRNA, and the tracrRNA molecule is also required during the interference stage (Shmakov et al. 2015). Type V-C systems have not yet been biochemically characterized. The Type VI C2c2 expression stage is again tracrRNA independent, and C2c2 cleavage activity (interference) has not yet been experimentally examined (Shmakov et al. 2015).

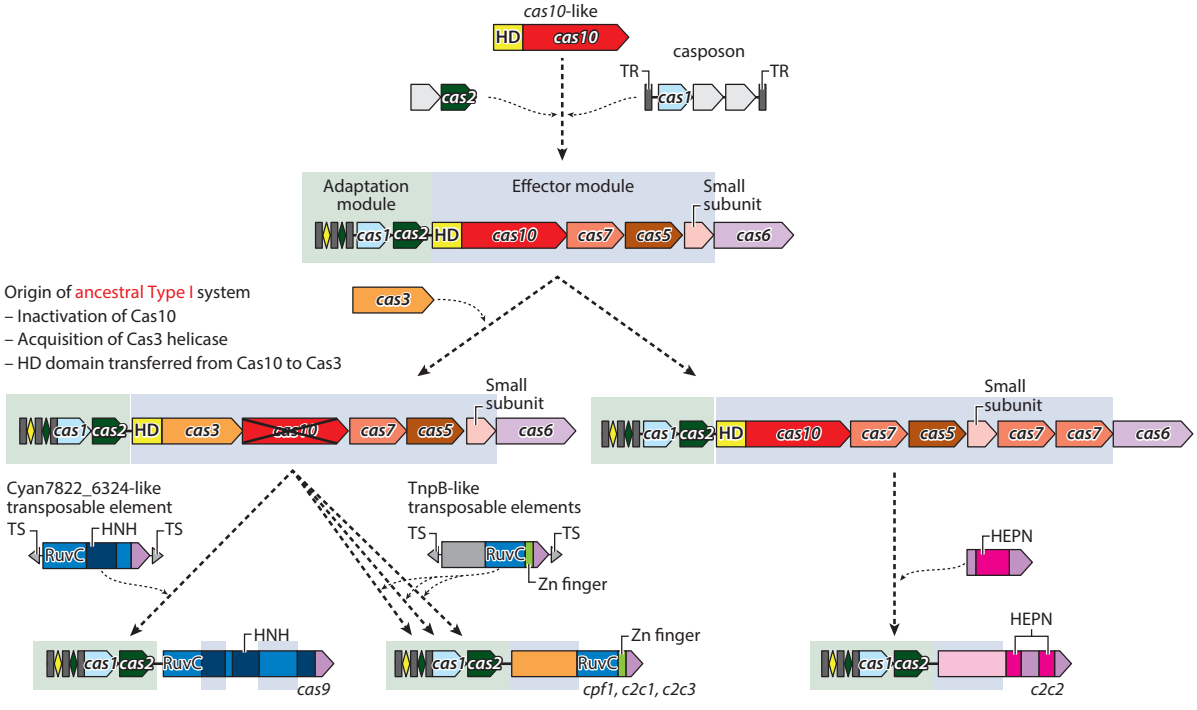

Figure 3

Scenario for the evolution of CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) systems. Abbreviations: HD, HD family endonuclease; HEPN, putative endoribonuclease of HEPN superfamily; HNH, HNH family endonuclease; RuvC, RuvC family endonuclease; TR, terminal repeats; TS, terminal sequences; Zn, zinc. Figure adapted with permission from Shmakov et al. (2015).

that is more rapid than bacterial dispersal (Berg Miller et al. 2012, Emerson et al. 2013, Held et al. 2010, Kunin et al. 2008, Sorokin et al. 2010, Vale & Little 2010). This effect was first reported by Kunin et al. (2008), showing that different isolates of a globally dispersed microbe (*Candidatus accumulibacter* phosphatis) carry unique CRISPR arrays that target sympatric rather than allopatric viruses. Analysis of available data sets from the Human Microbiome Project showed that CRISPR spacers are rarely shared between individual humans and some but not all spacers were shared between microbiomes resampled from the same individual through time (Rho et al. 2012). Another study found that many individuals shared bacteria with identical spacers, but the most recently acquired spacers were unique and matched co-occurring phage, in at least some cases (Stern et al. 2012). Streptococcal species from saliva were also shown to usually target sympatric phage, although in some instances they were equally likely to target sympatric and allopatric phage (Pride et al. 2012), perhaps due to global distribution of some phages. In the gut many phages are shared, even between geographically separated individuals (Stern et al. 2012), which may result from the fact that gut phages are often lysogenic (Breitbart et al. 2003) and therefore co-migrate with their host.

However, many studies have reported that spacer sequences from metagenomic data rarely match virus genomes (Anderson et al. 2011, Gogleva et al. 2014, Smedile et al. 2013). This finding has led to the suggestion that CRISPR may preferentially target rare viruses (Emerson et al. 2013). However, the lack of spacer matches is perhaps more likely due to low sequence availability of phage (Anderson et al. 2011, Berg Miller et al. 2012, Hatfull & Hendrix 2011, Reyes et al. 2010, Rohwer 2003), as illustrated by the fact that gut virome sequencing from different individuals still yields mostly novel viruses (Minot et al. 2011, 2012; Reyes et al. 2010, 2012). Indeed, when viromes were sequenced with bacterial CRISPR loci, spacers were typically found to match phage sequences from the same sample or geographical location (Andersson & Banfield 2008, Berg Miller et al. 2012, Emerson et al. 2013, Sorokin et al. 2010), with usually only the most recent spacers matching coexisting phage (Andersson & Banfield 2008, Stern et al. 2012).

Together, these studies have demonstrated that CRISPR loci can rapidly evolve in nature and are likely to play an important role in bacteria–phage interactions in these instances. However, from these correlational studies it is less clear when CRISPR-Cas systems are important fitness determinants of bacteria.

4. WHEN DO CRISPR–Cas SYSTEMS MATTER?

Despite the elegant mechanism of CRISPR-Cas–mediated defense, only half of all sequenced bacterial genomes encode CRISPR-Cas systems (Grissa et al. 2007), and different strains of the same species often differ with regard to the presence, number, and length of their CRISPR loci. Why do not all bacteria have CRISPR-Cas systems?

Theory and experimental studies suggest that ecological factors determine the relative benefit of CRISPR-Cas. Hence, these systems may be beneficial in some but not all environments. One example that suggests how ecology can impact the maintenance of CRISPR-Cas in bacterial genomes is provided by *Mycoplasma gallisepticum*, which lost its CRISPR-Cas system very rapidly following a host switch from chicken to wild finch (Delaney et al. 2012). Whether loss of CRISPR-Cas was adaptive is unclear (it may have hitchhiked along with an unrelated beneficial mutation), but other studies also suggest that the benefit of CRISPR-Cas is contingent on ecological factors (see Sections 4.2–4.6).

If we are to predict and manipulate the evolution of CRISPR-Cas immunity and the associated coevolution of CRISPR-Cas systems, we need to understand the selective forces that drive its evolution. Below we summarize the wide range of ecosystems where CRISPR-Cas systems have been found and then move on to discuss the role of four ecological factors that are predicted—and in

some cases confirmed—to be important: (*a*) virus genetic diversity (Iranzo et al. 2013), (*b*) the force of infection (Westra et al. 2015), (*c*) defective phages (Hynes et al. 2014), and (*d*) the presence of mutualists (Bikard et al. 2012, Gandon & Vale 2014, Jiang et al. 2013, Levin 2010). Finally, we discuss a sequence-specific interaction between nonimmune bacterial hosts and viruses, known as priming, which has received a lot of attention, as it is key for the evolution of CRISPR-based immunity.

4.1. CRISPR-Cas Systems Are Widespread but Enriched in Thermophilic Environments

CRISPR-Cas systems are widespread and have moved extensively by HGT between different species (Cui et al. 2008, Godde & Bickerton 2006, Haft et al. 2005, Heidelberg et al. 2009, Held et al. 2013, Horvath et al. 2009, Tyson & Banfield 2008). Metagenomics studies reveal that CRISPR-Cas systems can be found in bacteria and archaea living in diverse environments, such as the oral cavity (Pride et al. 2011, van der Ploeg 2009), gut (Gogleva et al. 2014, Rho et al. 2012, Stern et al. 2012), rumen of cows (Berg Miller et al. 2012), biofilms in an acid mine drainage (Andersson & Banfield 2008, Tyson & Banfield 2008), ocean (Sorokin et al. 2010) and deep-sea areas (Smedile et al. 2013), hydrothermal vents in the northeast Pacific (Anderson et al. 2011), sludge bioreactors (Kunin et al. 2008), hot springs (Heidelberg et al. 2009, Held & Whitaker 2009, Held et al. 2010, Snyder et al. 2010), hypersaline lakes (Emerson et al. 2013), and low-oxygen cyanobacterial mats in the Middle Island Sinkhole in Lake Huron (Voorhies et al. 2015). However, consistent with the idea that ecology matters, CRISPR-Cas systems are not evenly distributed across these environments. In particular, high-temperature environments are enriched for CRISPR-Cas systems, with thermophiles typically having both more and longer CRISPR arrays (Anderson et al. 2011, Makarova et al. 2006).

4.2. Virus Genetic Diversity and Mutation Rates

Koonin and coworkers (Iranzo et al. 2013, Weinberger et al. 2012b) constructed theoretical models that suggested that differences between mesophiles and thermophiles in the rates of mutation fixation may explain why CRISPR-Cas systems may be more abundant in high-temperature environments. Specifically, both theoretical models predicted that the benefit of CRISPR-Cas adaptive immune systems decreases as virus genetic diversity increases, owing to the sequence specificity of virus recognition (Iranzo et al. 2013, Weinberger et al. 2012b). Virus genetic diversity covaries both with mutation rates and with population sizes (larger populations have more genetic diversity). The authors argue that high-temperature environments typically have lower mutation rates (Weinberger et al. 2012b) and lower microbial densities (Iranzo et al. 2013). Lower host densities directly cause a reduction in the density—and therefore genetic variation—of the virus population (Iranzo et al. 2013). Weinberger et al. (2012b) suggested that if carrying CRISPR-Cas systems was associated with a fitness cost (e.g., due to autoimmunity), the systems might be lost under conditions that correspond to high virus mutation rates as spacers are effective against only a small portion of the virus population. Such a constitutive cost of CRISPR-Cas was recently found to be associated with a Class 2 CRISPR-Cas system (Vale et al. 2015).

4.3. The Force of Infection

An alternative, mutually nonexclusive explanation why CRISPR-Cas systems are less beneficial and therefore less abundant at high microbial densities is that CRISPR-Cas–mediated immunity becomes increasingly costly as the frequency of infection increases (Westra et al. 2015). An inducible cost of resistance was independently observed in both *P. aeruginosa* (Westra et al.

THE PROTOSPACER ADJACENT MOTIF

In Type I and Type II systems, target-binding affinity of the crRNA–Cas complex strongly increases if target sequences (protospacers) are flanked by protospacer adjacent motifs (PAMs) (Rollins et al. 2015; Westra et al. 2012, 2013). In Type I systems, Cas1 and Cas2 take the PAM into account during adaptation (Swarts et al. 2012, Yosef et al. 2012). In Type II systems, the PAM specificity is provided by Cas9 during adaptation (Heler et al. 2015). As the PAM is absent from the CRISPR loci on the host genome, it serves to avoid autoimmunity problems (Deveau et al. 2008). Type III systems appear to lack PAMs. The Type III-A system of *Staphylococcus epidermidis* has a distinct mechanism to distinguish self from non-self. Rather than using a PAM-based mechanism, CRISPR interference is inhibited if the target sequence is flanked by repeat sequences (Marraffini & Sontheimer 2010), which are presumably sensed by the crRNA–Cas effector complex through extended base pairing over the crRNA repeat sequence. This mechanism also avoids self-targeting of CRISPR loci on the host genome.

2015) and *S. thermophilus* (Vale et al. 2015). These observations are consistent with studies that demonstrated that expression of many CRISPR-Cas adaptive immune systems is strictly regulated and specifically elicited upon infection (Agari et al. 2010, Quax et al. 2013, Young et al. 2012). In the *P. aeruginosa* system, high frequencies of infection (and therefore high inducible costs of resistance) were found to tip the balance from CRISPR-Cas–mediated immunity to surface modification–based defense, which is associated with a constitutive cost (Westra et al. 2015). Hence, these data suggest that CRISPR-Cas systems are more likely to confer a benefit in ecosystems associated with a low force of infection, such as high-temperature or low-resource environments, owing to the lower microbial and phage densities. At a high force of infection other immune mechanisms may be favored more than CRISPR-Cas.

At present the mechanistic basis of the observed fitness cost associated with CRISPR-Cas is unclear but may be related to autoimmunity (Bikard et al. 2012; Held & Whitaker 2009; Jiang et al. 2013; Paez-Espino et al. 2013, 2015; Stern et al. 2010; Vercoe et al. 2013) (see sidebar, The Protospacer Adjacent Motif) or allocating resources to defense that would otherwise be invested in growth.

4.4. Defective Phages

Recently, Hynes and coworkers (2014) provided compelling evidence that most of the spacer acquisition events in their *S. thermophilus* strain DGCC7710/phage 2972 experimental system occur in response to defective phages. This response is somewhat analogous to the role of defective viruses in triggering mammalian immune responses (Killip et al. 2015). The rate at which bacteria acquired CRISPR-based immunity also increased when the bacteria encoded both a restriction-modification (R-M)–based innate immune system and CRISPR-Cas (Dupuis et al. 2013, Hynes et al. 2014). In this case, R-M may supply CRISPR-Cas with phage genome cleavage products that can be integrated as novel spacers. These data also suggest that mounting a CRISPR-Cas adaptive immune response from scratch (i.e., bacteria going through the adaptation, expression, and interference stages; see sidebar, The Three Stages of a CRISPR-Cas Immune Response) may be too slow in the face of a highly virulent, rapidly replicating phage.

4.5. The Presence of Mutualists

In the presence of genetic parasites, such as phage, bacteria that carry CRISPR-Cas systems have a clear fitness advantage over strains that lack CRISPR-Cas (Westra et al. 2015). However, it is less

clear how the benefits of CRISPR-Cas are affected by mutualistic mobile genetic elements, such as plasmids that can confer antibiotics resistance or prophages that encode virulence factors. Because CRISPR-Cas immunity is adaptive, one would expect the system to discriminate between parasitic and mutualistic DNA elements. Yet, two lines of evidence indicate that CRISPR-Cas systems form a barrier for HGT. First, bacteria evolve CRISPR-based immunity not only against phage but also against mobile genetic elements that are not necessarily parasitic, such as conjugative plasmids and conjugative transposons (Erdmann & Garrett 2012, Lopez-Sanchez et al. 2012). Consistent with this, *Staphylococcus epidermidis* strain RP62a carries a spacer that targets the nickase gene found in all sequenced staphylococcal conjugative plasmids (Marraffini & Sontheimer 2008). Second, correlational studies indicate that CRISPR-Cas systems indeed limit HGT (but see Gophna et al. 2015). For example, in *Enterococcus faecium* and *Enterococcus faecalis*, which are currently among the most important causative agents of hospital infections (Paganelli et al. 2012), the presence of CRISPR-Cas inversely correlates with their antibiotics-resistance gene content (Palmer & Gilmore 2010). These studies therefore suggest that the presence of mutualistic mobile genetic elements and environmental factors that impact their benefit (e.g., heavy metals, antibiotics, etc.) may affect the benefit of CRISPR-Cas. It is interesting to note in this context that a Type III CRISPR-Cas system does not target a prophage unless the target gene is transcribed (Goldberg & Marraffini 2015), which may facilitate a host interaction with a mutualist despite CRISPR-Cas immunity.

4.6. Immune Priming

A key factor that has been experimentally shown as crucial for CRISPR evolution is priming. In the context of CRISPR-Cas, priming refers to the presence of a partial match between a pre-existing spacer and the genome of an invading phage or plasmid (Datsenko et al. 2012, Swarts et al. 2012). Despite the mismatches, the CRISPR-Cas surveillance complex can still recognize the partially complementary target sequence (Blosser et al. 2015), which triggers recruitment of the spacer acquisition machinery (Redding et al. 2015) and results in rapid primed spacer acquisition (see sidebar, The Molecular Characteristics of Priming). The number of mismatches can be relatively high (up to 13 in one study; see Fineran et al. 2014), indicating that priming allows for some level of promiscuity.

THE MOLECULAR CHARACTERISTICS OF PRIMING

Priming has been reported for only Type I systems (Class 1). Priming requires not only Cas1 and Cas2 (the core adaptation machinery; see sidebar, The Three Stages of a CRISPR-Cas Immune Response) but also the interference machinery Cascade and Cas3 (Datsenko et al. 2012). Cascade bound to an imperfect target has been suggested to adopt a different conformation (Blosser et al. 2015), which may result in recruitment of the adaptation machinery (Redding et al. 2015). Primed spacer acquisition causes a characteristic pattern of protospacer sampling. A primed *Escherichia coli* Type I-E CRISPR-Cas system acquired spacers almost exclusively from the same DNA strand of a plasmid target (Swarts et al. 2012). Primed Type I-F systems of *Pectobacterium atrosepticum* and *P. aeruginosa* acquired spacers from a constrained region around the priming site on the phage genome, resulting in a clustering of target sites in genomic regions of the phage genome (Richter et al. 2014, Westra et al. 2015). Furthermore, new spacers were found to be selected such that novel protospacers (i.e., the sequence complementary to crRNA) were located on the strand targeted by the priming crRNA when upstream of the priming site but on the nontarget strand when downstream of the target site (Richter et al. 2014, Westra et al. 2015). The same biased distribution was also observed when the *P. aeruginosa* Type I-F system was recombinantly expressed in *E. coli* (Vorontsova et al. 2015).

Why did the requirement for priming evolve? The efficiency of unprimed (or naive) spacer acquisition is very low, indicating that the requirement for priming imposes a severe constraint on the evolution of CRISPR-based immunity. One possibility is that priming helps to limit autoimmunity problems associated with CRISPR-Cas through self-targeting (Koonin & Wolf 2016). In addition, priming also helps to ensure that only parasitic DNA elements that resemble previous spacers and therefore known parasites are targeted. Hence, although indiscriminate spacer acquisition upon infection may increase the rate at which bacteria acquire immunity, it may trade off with higher levels of self-targeting and greater constraint on the uptake of beneficial DNA.

Taken together, the findings described above make it increasingly clear that CRISPR-Cas may be highly beneficial in some environments but potentially costly or ineffective in other environments. Teasing apart the relative importance of the ecological variables outlined above and their evolutionary consequences will be key to understanding and manipulating bacteria–virus interactions in natural environments.

5. CRISPR-Cas–MEDIATED BACTERIA–PHAGE COEVOLUTION

What are the short-term coevolutionary consequences of CRISPR immunity against viruses? Several experimental studies have demonstrated that phage can overcome CRISPR immunity by mutating the target sequence or the protospacer adjacent motif (PAM) (see sidebar, The Protospacer Adjacent Motif) (Deveau et al. 2008). The high specificity of CRISPR–virus interactions leads to predictions of persistent coevolution (Agrawal & Lively 2002, Iranzo et al. 2013, Vale & Little 2010). Yet, CRISPR-specific models and recent coevolution studies reveal conditions in which CRISPR–virus coevolution is short-lived due to virus extinction (Iranzo et al. 2013, van Houte et al. 2016). What factors determine the type and duration of coevolution? Below we discuss the impact of: (*a*) spacer diversity, (*b*) viral mutation and host spacer acquisition rates, (*c*) the number of different viruses, and (*d*) anti-CRISPR genes.

5.1. Spacer Diversity

Theory predicts (Childs et al. 2014) and in vitro coevolution experiments show that CRISPR–virus coevolution is short-lived if the bacterial host population generates high levels of spacer diversity, which results in rapid extinction of the virus (van Houte et al. 2016). Crucially, although viruses rapidly evolved infectivity against nearly all individual clones in monoculture, viruses lost this ability when the same bacterial clones were mixed (van Houte et al. 2016). In addition to this evolutionary effect, diversity is also predicted to have an epidemiological effect. First, densities of the matching host genotype are reduced (dilution effect). Second, resistance of nonmatching host genotypes reduces the number of successful secondary infections (Lively 2010). This latter effect is amplified once an escape virus emerges, since the virus drives the frequency of its matching host genotype down (Lively 2010). The propensity to generate spacer diversity is thus an important fitness determinant of CRISPR-Cas and is dependent on priming because this accelerates spacer acquisition.

5.2. Viral Mutation and Host Spacer Acquisition Rates

As explained above, the virus mutation and spacer acquisition rates are predicted to be important for CRISPR–virus coevolution (Iranzo et al. 2013). Specifically, higher rates of mutation fixation are predicted to work to the advantage of the virus. Hence, increased mutation supply rates can move viruses away from extinction toward persistent coevolution or even to host extinction (Iranzo et al. 2013). Spacer acquisition rates work in the opposite direction and could lead to phage

extinction (Childs et al. 2014, Iranzo et al. 2013). The reasons why the rate of spacer acquisition is important are twofold. First, it allows the host to renew immunity against escape viruses. Second, it increases spacer diversity in the population, which—as explained above—increases overall population resistance (van Houte et al. 2016).

5.3. Multiple Viruses

In many natural environments, a bacterium is likely to interact with multiple viruses, which may have a strong impact on CRISPR–virus coevolution. Theory and metagenomics data suggest that CRISPR-Cas systems may cause selective sweeps in the host populations if a single host genotype acquires resistance against two phages (Tyson & Banfield 2008, Weinberger et al. 2012a). Another effect of multiple phages, observed in coevolution studies using *S. thermophilus*, was that a phage may persist longer compared with a single-phage infection (Paez-Espino et al. 2015). Interestingly, it was found that the two phages recombined to escape CRISPR-Cas (Paez-Espino et al. 2015), which is consistent with previous metagenomics analyses (Andersson & Banfield 2008) and may help to explain commonly observed mosaic genomes of phage (Paez-Espino et al. 2015, Pedulla et al. 2003). The dynamics of CRISPR-Cas–mediated, single-host, multiphage coevolution awaits further study.

5.4. Anti-CRISPR Genes

The rapid extinction of viruses by the *P. aeruginosa* CRISPR-Cas system may have provided strong selective pressure on phages to evolve more sophisticated escape mechanisms. Many *Pseudomonas* phages encode so-called anti-CRISPR genes (Bondy-Denomy et al. 2013, Pawluk et al. 2014), which bind either the crRNA–Cas complex of *P. aeruginosa* (known as the Csy complex) to interfere with target DNA recognition or the effector nuclease (Cas3) to block target DNA destruction (Bondy-Denomy et al. 2015). Anti-CRISPR proteins are encoded by an extremely diverse set of genes often located in a conserved locus on phage genomes (Bondy-Denomy et al. 2013, Pawluk et al. 2014) and other mobile genetic elements (Pawluk et al. 2014, van Belkum et al. 2015). The evolution of anti-CRISPR genes likely occurs over much longer timescales compared with phage evolution to escape CRISPR-Cas by point mutation. The dynamics of these long-term coevolutionary processes have not been investigated, and if and how hosts can evolve to overcome the anti-CRISPR genes remains an open question.

6. CRISPR AS A BARRIER OF HORIZONTAL GENE TRANSFER: IMPLICATIONS FOR MICROBIAL ADAPTATION

Apart from short-term coevolutionary consequences, CRISPR-Cas can also impact long-term microbial adaptation. As explained above, CRISPR-Cas systems can form a barrier for HGT. Several correlational studies indicate that CRISPR-Cas limits gene transfer and that this results in important differences between genomes of CRISPR$^+$ and CRISPR$^-$ strains. For example, strains of the opportunistic pathogen *P. aeruginosa*, which often causes lung infections in cystic fibrosis patients, have significantly smaller genomes if they encode CRISPR-Cas systems compared with strains that lack CRISPR-Cas (van Belkum et al. 2015). Furthermore, CRISPR-Cas systems are absent in the genomes of species for which gene transfer is an important fitness determinant, such as *Streptococcus pneumoniae*, the causative agent of pneumonia that relies heavily on natural transformation for capsule switching during infection (Hatoum-Aslan & Marraffini 2014). CRISPR-Cas systems have been proposed to interfere with natural transformation (Bikard et al. 2012).

Similarly, genomes of *Streptococcus pyogenes*, a major human pathogen that can cause, among other symptoms, pharyngitis, sepsis, and necrotizing fasciitis, often carry a range of prophage-encoded virulence factors that are key for *S. pyogenes* virulence, and a higher number of prophages negatively correlates with CRISPR-Cas (Hatoum-Aslan & Marraffini 2014).

CRISPR-Cas systems may be lost during evolution if they acquire resistance against beneficial mobile genetic elements. For example, when exposed to antibiotics, CRISPR-Cas mutants readily emerged in a clonal population of *S. epidermidis* that was CRISPR-resistant against an antibiotics-resistant plasmid (Jiang et al. 2013). Similarly, CRISPR-Cas mutants emerged in *S. pneumoniae* engineered to carry a CRISPR-Cas system from *S. pyogenes* that targets capsule genes (Bikard et al. 2012). Although in both studies the strains were engineered to target a mobile genetic element that was essential for survival, these studies highlight how selection may favor loss of CRISPR-Cas systems in microbial populations.

Taken together, accumulating evidence shows that CRISPR-Cas can impact microbial adaptation via HGT, and the need for HGT may select against CRISPR-Cas systems.

7. NONCANONICAL FUNCTIONS OF CRISPR-Cas SYSTEMS

In addition to adaptive immunity against invading genetic elements, there is increasing experimental evidence that CRISPR-Cas systems play roles in other cellular functions, including transcriptional control, stress response, and pathogenicity (see Westra et al. 2014 for a review).

The most commonly identified function of CRISPR-Cas systems other than adaptive immunity is the regulation of gene expression. One of the best-characterized examples of a directly selected noncanonical function of a CRISPR-Cas system is that of virulence regulation in the highly infectious zoonotic pathogen *Francisella novicida*. In *F. novicida* (strain U112), a Type II-B CRISPR-Cas system is involved in repressing production of the cell surface–associated bacterial lipoprotein (BLP). Because the host innate immune system recognizes BLP and is activated, reduction in BLP enables *F. novicida* to evade host detection and therefore increases virulence (Sampson et al. 2013). The repression of BLP involves the Cas9 nuclease, *trans*-activating crRNA (tracrRNA), and small CRISPR-Cas associated RNA, which degrade BLP messenger RNA (mRNA). Mutants that lack any of these CRISPR-Cas components demonstrate marked attenuation in virulence in mouse models (Sampson et al. 2013, 2014), and mutations in these BLP regulators correlate with mutations in the *blp* gene (Sampson & Weiss 2013).

Type II CRISPR-Cas systems have also been observed to modulate virulence in a number of other bacterial pathogens. Some strains of the enteric pathogen *Campylobacter jejuni* encode a Type II-C CRISPR-Cas system. Expression of Cas9 protein in *C. jejuni* strains in which this CRISPR-Cas system is absent leads to an increase in virulence. Further, mutated strains lacking *cas9* showed reduced adherence, invasion, and attenuated cytotoxicity toward human gut cell lines (Louwen et al. 2013). Cas9 has also been implicated in the virulence of *Neisseria meningitidis*, which requires Cas9 for attachment to host lung epithelial cells and for invasion and intracellular replication (Sampson & Weiss 2013). However, the mechanism of action of Cas9 in these bacteria is unknown, but they may act in concert with other CRISPR components, as in *F. novicida*, to regulate virulence-associated genes.

In *Legionella pneumophila*, the causative agent of Legionnaires' disease, Cas2 (present within a Type II CRISPR-Cas system) promotes intracellular infection of host amoebae (Gunderson & Cianciotto 2013). Mutants in *cas2* are impaired in their ability to cause infection, whereas mutants in any other part of the CRISPR-Cas system display no change. *L. pneumophila* Cas2 has been shown to have DNase and RNase activity, which is important in the establishment of intracellular infection in amoebae (Gunderson et al. 2015). The introduction of *cas2* into a

noncarrying *L. pneumophila* strain increased infectivity, and this introduction has been proposed to provide an advantage for acquisition and maintenance of *cas2* within strains. The *L. pneumophila* Cas2 is currently the only Cas2 protein with a function exclusive from a role in adaptive immunity.

Another interesting case is that of the pathogen *Listeria monocytogenes*, a facultative intracellular bacterium and the causative agent of listeriosis. All strains of *Listeria* sequenced to date carry a CRISPR locus termed RliB. Notably, this locus is not adjacent to any *cas* genes and is also present in *L. monocytogenes* strains that do not carry any other *cas* genes. Overexpression of the *rliB* CRISPR transcript upregulates expression of a ferrous iron transporter (Mandin et al. 2007). Studies show that a mutant strain lacking *rliB* colonizes the liver of infected mice more effectively than wild-type strains, implicating a role in virulence regulation (Toledo-Arana et al. 2009).

In *P. aeruginosa* lysogens, disruption of a Type I-F CRISPR-Cas system affects both biofilm formation and swarming motility (Zegans et al. 2009). However, this effect is a by-product of cytotoxicity of CRISPR-mediated self-targeting (cleavage of the lysogenic phage, which is integrated in the host genome) (Heussler et al. 2015). In *P. aeruginosa* the primary function of this CRISPR-Cas system is in immunity, as shown by, among other indications, high levels of evolved CRISPR immunity in response to phage (Westra et al. 2015) and the evolution of anti-CRISPR genes in *P. aeruginosa* phage genomes (Bondy-Denomy et al. 2013). In brief, it appears from the examples above and others (Westra et al. 2014) that CRISPR transcripts and Cas proteins are both independently able to influence bacterial virulence via gene regulation. However, the majority of the mechanisms remain to be elucidated. A central question relating to the evolution of noncanonical functions of CRISPR-Cas systems asks if these are simply by-products of their role in immunity or if they are indeed selected functions. From the studies described above it appears that both occur.

8. APPLICATIONS OF CRISPR-Cas IN EVOLUTION AND ECOLOGY

CRISPR-Cas systems have emerged as tools across life sciences, and their applications range from strain typing to engineering genomes or regulating their gene expression. In addition to these applications, CRISPR-Cas are also emerging as model systems to examine host–parasite coevolution (Vale & Little 2010).

8.1. Ecological and Epidemiological Studies

CRISPR diversity results both from rapid acquisition of new spacers (Andersson & Banfield 2008) and from spacer loss (Held et al. 2010, Lopez-Sanchez et al. 2012, Pourcel et al. 2005, Schouls et al. 2003, Tyson & Banfield 2008). Because spacer acquisition is polar (i.e., new spacers are added at one end of the array) (Barrangou et al. 2007, Lillestol et al. 2006, Pourcel et al. 2005, Tyson & Banfield 2008), diversity in CRISPR loci is mostly localized in the area where novel spacers are integrated, the leader-proximal end of the CRISPR array (Horvath et al. 2008, Tyson & Banfield 2008, Weinberger et al. 2012a). As a consequence, trailer ends of CRISPR loci may be conserved between different microbial populations, whereas the middle may be population specific and the leader end unique at the individual level (Tyson & Banfield 2008). The highly variable sequence content of CRISPR loci can and has been exploited to distinguish closely related bacterial strains (Pourcel et al. 2005, Vergnaud et al. 2007). CRISPR-based typing can be done using hybridization-based methods, such as spoligotyping, which is based on hybridization of a CRISPR amplicon to known spacer probes, or using sequencing-based methods (for a review, see Shariat & Dudley 2014). Spoligotyping was developed as an early, rapid, and cost-effective CRISPR-based method to discriminate *Mycobacterium tuberculosis* strains and has been widely used

over the past decades both for strain typing and to examine pathogen evolution and population structuring. *M. tuberculosis* has ceased to acquire novel spacers, and differences between strains are therefore predominantly due to spacer loss (van Embden et al. 2000). Spoligotyping has also been used for other bacterial pathogens, such as *Corynebacterium diphtheriae*, *L. pneumophila*, and *Salmonella enterica* (reviewed in Shariat & Dudley 2014). Pourcel and colleagues (2005) have directly used the sequence information stored in CRISPR loci for *Yersinia pestis* typing, and in a follow-up study the authors were able to link different strains to distinct geographical locations (Cui et al. 2008). CRISPR sequences have also been used to distinguish *Erwinia amylovora* strains, which are plant pathogens that are indistinguishable using other common strain typing techniques, and to study *Salmonella* outbreaks or phylogeny (reviewed in Shariat & Dudley 2014). CRISPR sequence–based typing coupled to multi-virulence-locus sequence typing (CRISPR-MVLST) has been used to identify outbreak isolates in patients (reviewed in Shariat & Dudley 2014). Finally, sensitive and specific real-time polymerase chain reaction–based methods using primers that anneal to CRISPR arrays have been developed to identify Shiga-toxin-producing *Escherichia coli* serogroups (Delannoy et al. 2012a, 2012b).

Apart from studying the host, CRISPR sequences can also be used to study mobile genetic elements. One key application of CRISPR spacer sequences has been their use in identifying virus and plasmid sequences in metagenomics sequencing data and linking the phage/virus to a specific host. The Banfield laboratory was the first to use CRISPR from metagenomics sequences of biofilms in an acid mine drainage to identify host–virus interactions (Andersson & Banfield 2008). Over recent years, this method has been used extensively to identify viruses and link them to a specific host (Anderson et al. 2011, Garrett et al. 2010, Minot et al. 2013, Sanguino et al. 2015, Stern et al. 2012). Furthermore, spacer sequences have also been used as probes in microarrays to monitor the abundance of specific viruses in environmental samples (Snyder et al. 2010).

8.2. Genome Editing

The requirement for a single Cas effector protein, rather than multisubunit crRNA–Cas complexes, makes Class 2 systems uniquely suited for genome editing technologies (see sidebar, Class 2 CRISPR-Cas Systems, for details). Nearly all genome editing and related techniques have been developed using the Cas9 enzymes (Class 2 Type II), but as explained in the sidebar, New Class 2 CRISPR-Cas Systems, novel Class 2 enzymes are currently being examined as well (Shmakov et al. 2015, Zetsche et al. 2015). At present, the Cas9 genome editing technology has been used across many different organisms, for example, to make transgenic crops and mutant mouse models, to modify human embryos, and to prevent or treat disease in animal models (reviewed in Sternberg & Doudna 2015). Cas9 cleavage sites are repaired either using the error-prone nonhomologous end joining DNA repair pathway, which results in loss of function due to sequence deletions or insertions, or using the homology-directed repair (HDR) pathway if a homologous DNA molecule with mutations of interest, which is used by the HDR machinery during repair, is supplied. HDR provides much greater control over the mutations that are introduced into the genome. Apart from editing, a catalytically inactive Cas9 mutant can be used to bind promoter regions to inhibit expression of target genes, or Cas9 can also be fused to other proteins, such as transcriptional regulators or fluorescent proteins to image specific genetic loci (reviewed in Sternberg & Doudna 2015). This technology is now widely used to examine gene function, but it can also be applied to examine specific evolutionary questions, such as the adaptive value of specific mutations and their interactions (e.g., epistasis). Furthermore, in the future the application of CRISPR-Cas9 in genome editing may prove to be very important in evolutionary genetics to confirm candidate

gene function and to test genes' effects in ecological interactions. As such, CRISPR-Cas9 has clear potential as a tool in evolutionary ecology studies.

8.3. Manipulating Community Composition

Apart from sequence-specific editing of genomes or their transcriptional regulation, CRISPR-Cas can also be used to specifically manipulate species communities. For example, CRISPR-Cas9 has been used to target specific viruses (Kennedy & Cullen 2015), although rapid virus evolution can result in escape from CRISPR-Cas9 (Wang et al. 2016). In a microbial population, a CRISPR-Cas system encoded by a mobile genetic element, such as a virus or conjugative plasmid, can be programmed such that it kills one specific host genotype. In the in vitro laboratory environment, an sgRNA-Cas9–encoding phagemid (plasmid packaged in a phage capsid) could effectively target virulent *Staphylococcus aureus* genotypes while leaving nonvirulent genotypes unaffected (Bikard et al. 2014). A similar approach of using a virus-encoded Class 1 (Yosef et al. 2015) or Class 2 (Citorik et al. 2014) CRISPR-Cas system was used to target antibiotic resistance genes in *E. coli*. These proof-of-concept studies clearly demonstrate that CRISPR-Cas–based approaches can in principle be used to manipulate microbial community composition, although this will be much harder to achieve in real environments, where reaching sufficiently high infection rates of the phage is more difficult. Although these studies focused on important human pathogens and their virulence genes, the approach may be extended to examine specific functions of individual strains or species in complex microbial communities.

Apart from manipulating microbial community composition, it has been speculated that CRISPR-Cas can also be used as a tool to manipulate communities of sexually reproducing organisms. Specifically, Esvelt and colleagues (2014) proposed that sgRNA-Cas9 enzymes can be repurposed as gene drives to spread engineered traits through a population, which can in effect be used to alter community composition. The use of gene drives has been considered to control or reduce the spread of insect-borne diseases, but progress on this front has been hindered by technical difficulties associated with genome engineering. Esvelt et al. (2014) built a strong case that the use of sgRNA-Cas9 may overcome many of these problems and potentially open the way to ecological engineering. Indeed, Cas9-based gene drives have been developed that efficiently spread genomic alterations in populations of a number of organisms. Two proof-of-principle studies have demonstrated highly efficient Cas9-based gene drive activity in yeast (DiCarlo et al. 2015) and in fruit flies (DiCarlo et al. 2015, Gantz & Bier 2015), with homing efficiencies (the rate at which the drive gene is copied onto the opposite chromosome) of the gene drives reaching 99% and 97%, respectively. In two recent studies Cas9-based gene drives were successfully employed to spread traits in mosquitos that could potentially limit the spread of vector-borne diseases (Gantz et al. 2015, Hammond et al. 2016). Gantz et al. (2015) used a gene drive to spread antimalarial genes in *Anopheles stephensi*. Despite the large size of the gene drive used, efficiency reached 99%, although the effect was partially lost in the progeny of females due to instability of the homing element. A different approach to limit malaria parasite transmission was used in a study by Hammond et al. (2016), in which a Cas9 gene drive was used to target genes that are necessary for female reproduction in *Anopheles gambiae* populations. This gene drive was also efficiently transmitted to offspring but suffered from inactivation of the target genes in heterozygous females, thereby greatly reducing female fertility and hence gene drive transmission (Hammond et al. 2016). However, these studies show that Cas9-based gene drives can be highly efficient systems for ecological engineering of populations (Champer et al. 2016).

8.4. Model to Study Host–Parasite Coevolution

Many studies on host–parasite coevolutionary processes are constrained by a lack of knowledge about the mechanisms of resistance and infectivity. Focusing on the CRISPR-Cas system to study host–parasite coevolution greatly overcomes this problem because we have a deep understanding about the molecular mechanism by which it confers immunity (Vale & Little 2010). As such, CRISPR-Cas can be used as a model system to investigate general questions concerning the evolutionary ecology of host–parasite interactions. For example, bacteria–phage interactions have been used recently to examine when an induced defense, such as CRISPR-Cas, is favored over a constitutive defense, such as surface modification, and vice versa (Westra et al. 2015). Using this tractable experimental system, it was demonstrated that the force of infection determines the relative investment in the two arms of defense, which is in agreement with theoretical predictions (Hamilton et al. 2008). More recently, the impact of host resistance allele diversity on parasite persistence and evolution was examined. This study found that parasites (viruses) rapidly evolved infectivity against monocultures of bacterial hosts that all carried the same CRISPR spacer (targeting the phage), but phage was unable to evolve infectivity on the same clones when they were mixed together (van Houte et al. 2016). This study helps us understand how host diversity can limit the spread of infectious disease. We envisage that CRISPR–virus interactions will further emerge as an important model system for experimental evolution to answer key questions concerning host–parasite interactions.

9. CONCLUSION AND OUTLOOK

The field of CRISPR-Cas biochemistry has raced ahead, leading to groundbreaking applications in genome editing and beyond. However, many questions concerning the ecology and evolution of CRISPR-Cas remain. For example, it is unclear when CRISPR-Cas systems provide a selective advantage. There are many examples in which these sophisticated adaptive immune systems have been lost (Delaney et al. 2012, Sampson & Weiss 2013) or they appear inactive (Touchon & Rocha 2010). It is becoming clear that the benefit of CRISPR-Cas depends on a range of ecological variables, but much more work needs to be done if we are to understand and manipulate the evolution of CRISPR-Cas immunity in the lab, let alone in nature. Moreover, relatively little experimental work has been done on the coevolutionary consequences of CRISPR-Cas–virus interactions and the long-term consequences of CRISPR-Cas on microbial adaptation (reviewed in Hatoum-Aslan & Marraffini 2014). As the mechanistic details of CRISPR-Cas become better characterized, future research is likely to focus on these outstanding questions.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Angus Buckling for discussions. E.R.W. acknowledges the Natural Environment Research Council, the British Biological Sciences Research Council, and the Wellcome Trust for funding. S.V.H. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 660039.

LITERATURE CITED

Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* 395:270–81

Agrawal A, Lively CM. 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evol. Ecol. Res.* 4:79–90

Amitai G, Sorek R. 2016. CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.* 14:67–76

Anderson RE, Brazelton WJ, Baross JA. 2011. Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol. Ecol.* 77:120–33

Andersson AF, Banfield JF. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320:1047–50

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–12

Berg Miller ME, Yeoman CJ, Chia N, Tringe SG, Angly FE, et al. 2012. Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. *Environ. Microbiol.* 14:207–27

Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, et al. 2014. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* 32:1146–50

Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. 2012. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* 12:177–86

Blosser TR, Loeff L, Westra ER, Vlot M, Kunne T, et al. 2015. Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. *Mol. Cell* 58:60–70

Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, et al. 2015. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* 526:136–39

Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. 2013. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493:429–32

Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, et al. 2003. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 185:6220–23

Cady KC, White AS, Hammond JH, Abendroth MD, Karthikeyan RS, et al. 2011. Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* 157:430–37

Champer J, Buchman A, Akbari OS. 2016. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat. Rev. Genet.* 17:146–59

Charpentier E, Richter H, van der Oost J, White MF. 2015. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.* 39:428–41

Childs LM, England WE, Young MJ, Weitz JS, Whitaker RJ. 2014. CRISPR-induced distributed immunity in microbial populations. *PLOS ONE* 9:e101710

Chylinski K, Makarova KS, Charpentier E, Koonin EV. 2014. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 42:6091–105

Citorik RJ, Mimee M, Lu TK. 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* 32:1141–45

Cui Y, Li Y, Gorge O, Platonov ME, Yan Y, et al. 2008. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLOS ONE* 3:e2652

Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3:945

DeBoy RT, Mongodin EF, Emerson JB, Nelson KE. 2006. Chromosome evolution in the *Thermotogales*: large-scale inversions and strain diversification of CRISPR sequences. *J. Bacteriol.* 188:2364–74

Delaney NF, Balenger S, Bonneaud C, Marx CJ, Hill GE, et al. 2012. Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLOS Genet.* 8:e1002511

Delannoy S, Beutin L, Burgos Y, Fach P. 2012a. Specific detection of enteroaggregative hemorrhagic *Escherichia coli* O104:H4 strains by use of the CRISPR locus as a target for a diagnostic real-time PCR. *J. Clin. Microbiol.* 50:3485–92

Delannoy S, Beutin L, Fach P. 2012b. Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic *Escherichia coli* strains of serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7 by real-time PCR. *J. Clin. Microbiol.* 50:4035–40

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471:602–7

Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, et al. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1390–400

DiCarlo JE, Chavez A, Dietz SL, Esvelt KM, Church GM. 2015. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat. Biotechnol.* 33:1250–55

Dupuis ME, Villion M, Magadan AH, Moineau S. 2013. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat. Commun.* 4:2087

Emerson JB, Andrade K, Thomas BC, Norman A, Allen EE, et al. 2013. Virus-host and CRISPR dynamics in Archaea-dominated hypersaline Lake Tyrrell, Victoria, Australia. *Archaea* 2013:370871

Erdmann S, Garrett RA. 2012. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol.* 85:1044–56

Esvelt KM, Smidler AL, Catteruccia F, Church GM. 2014. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife* 3:e03401

Fineran PC, Gerritzen MJ, Suarez-Diez M, Kunne T, Boekhorst J, et al. 2014. Degenerate target sites mediate rapid primed CRISPR adaptation. *PNAS* 111:E1629–38

Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. 2016. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532:517–21

Gandon S, Vale PF. 2014. The evolution of resistance against good and bad infections. *J. Evol. Biol.* 27:303–12

Gantz VM, Bier E. 2015. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* 348:442–44

Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, et al. 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *PNAS* 112:E6736–43

Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, et al. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468:67–71

Garrett RA, Prangishvili D, Shah SA, Reuter M, Stetter KO, Peng X. 2010. Metagenomic analyses of novel viruses and plasmids from a cultured environmental sample of hyperthermophilic neutrophiles. *Environ. Microbiol.* 12:2918–30

Godde JS, Bickerton A. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62:718–29

Gogleva AA, Gelfand MS, Artamonova II. 2014. Comparative analysis of CRISPR cassettes from the human gut metagenomic contigs. *BMC Genom.* 15:202

Goldberg GW, Marraffini LA. 2015. Resistance and tolerance to foreign elements by prokaryotic immune systems—curating the genome. *Nat. Rev. Immunol.* 15:717–24

Gophna U, Kristensen DM, Wolf YI, Popa O, Drevet C, Koonin EV. 2015. No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *ISME J.* 9:2021–27

Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinform.* 8:172

Gunderson FF, Cianciotto NP. 2013. The CRISPR-associated gene *cas2* of *Legionella pneumophila* is required for intracellular infection of amoebae. *mBio* 4:e00074–13

Gunderson FF, Mallama CA, Fairbairn SG, Cianciotto NP. 2015. Nuclease activity of *Legionella pneumophila* Cas2 promotes intracellular infection of amoebal host cells. *Infect. Immun.* 83:1008–18

Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLOS Comput. Biol.* 1:e60

Hamilton R, Siva-Jothy M, Boots M. 2008. Two arms are better than one: Parasite variation leads to combined inducible and constitutive innate immune responses. *Proc. R. Soc. B* 275:937–45

Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, et al. 2016. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat. Biotechnol.* 34:78–83

Hatfull GF, Hendrix RW. 2011. Bacteriophages and their genomes. *Curr. Opin. Virol.* 1:298–303
Hatoum-Aslan A, Marraffini LA. 2014. Impact of CRISPR immunity on the emergence and virulence of bacterial pathogens. *Curr. Opin. Microbiol.* 17C:82–90
He F, Chen L, Peng X. 2014. First experimental evidence for the presence of a CRISPR toxin in *Sulfolobus*. *J. Mol. Biol.* 426:3683–88
Heidelberg JF, Nelson WC, Schoenfeld T, Bhaya D. 2009. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLOS ONE* 4:e4169
Held NL, Herrera A, Cadillo-Quiroz H, Whitaker RJ. 2010. CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLOS ONE* 5:e12988
Held NL, Herrera A, Whitaker RJ. 2013. Reassortment of CRISPR repeat-spacer loci in *Sulfolobus islandicus*. *Environ. Microbiol.* 15:3065–76
Held NL, Whitaker RJ. 2009. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ. Microbiol.* 11:457–66
Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, et al. 2015. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519:199–202
Heussler GE, Cady KC, Koeppen K, Bhuju S, Stanton BA, O'Toole GA. 2015. Clustered regularly interspaced short palindromic repeat-dependent, biofilm-specific death of *Pseudomonas aeruginosa* mediated by increased expression of phage-related genes. *mBio* 6:e00129
Horvath P, Coute-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* 131:62–70
Horvath P, Romero DA, Coute-Monvoisin AC, Richards M, Deveau H, et al. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1401–12
Hynes AP, Villion M, Moineau S. 2014. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat. Commun.* 5:4399
Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV. 2013. Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *J. Bacteriol.* 195:3834–44
Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169:5429–33
Jackson RN, Wiedenheft B. 2015. A conserved structural chassis for mounting versatile CRISPR RNA-guided immune responses. *Mol. Cell* 58:722–28
Jiang W, Maniv I, Arain F, Wang Y, Levin BR, Marraffini LA. 2013. Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLOS Genet.* 9:e1003844
Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–21
Kennedy EM, Cullen BR. 2015. Bacterial CRISPR/Cas DNA endonucleases: A revolutionary technology that could dramatically impact viral research and treatment. *Virology* 479–480:213–20
Killip MJ, Fodor E, Randall RE. 2015. Influenza virus activation of the interferon system. *Virus Res.* 209:11–22
Koonin EV, Makarova KS. 2013. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol.* 10:679–86
Koonin EV, Wolf YI. 2009. Is evolution Darwinian or/and Lamarckian? *Biol. Direct.* 4:42
Koonin EV, Wolf YI. 2016. Just how Lamarckian is CRISPR-Cas immunity: The continuum of evolvability mechanisms. *Biol. Direct.* 11:9
Krupovic M, Makarova KS, Forterre P, Prangishvili D, Koonin EV. 2014. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.* 12:36
Krupovic M, Shmakov S, Makarova KS, Forterre P, Koonin EV. 2016. Recent mobility of casposons, self-synthesizing transposons at the origin of the CRISPR-Cas immunity. *Genome Biol. Evol.* 8:375–86
Kunin V, He S, Warnecke F, Peterson SB, Garcia Martin H, et al. 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res.* 18:293–97
Kwon AR, Kim JH, Park SJ, Lee KY, Min YH, et al. 2012. Structural and biochemical characterization of HP0315 from *Helicobacter pylori* as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res.* 40:4216–28

Levin BR. 2010. Nasty viruses, costly plasmids, population dynamics, and the conditions for establishing and maintaining CRISPR-mediated adaptive immunity in bacteria. *PLOS Genet.* 6:e1001171

Lillestol RK, Redder P, Garrett RA, Brugger K. 2006. A putative viral defence mechanism in archaeal cells. *Archaea* 2:59–72

Lively CM. 2010. The effect of host genetic diversity on disease spread. *Am. Nat.* 175:E149–52

Lopez-Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina E, et al. 2012. The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol. Microbiol.* 85:1057–71

Louwen R, Horst-Kreft D, de Boer AG, van der Graaf L, de Knegt G, et al. 2013. A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barre syndrome. *Eur. J. Clin. Microbiol. Infect. Dis.* 32:207–26

Makarova KS, Anantharaman V, Aravind L, Koonin EV. 2012. Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes. *Biol. Direct.* 7:40

Makarova KS, Aravind L, Wolf YI, Koonin EV. 2011a. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct.* 6:38

Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, et al. 2011b. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9:467–77

Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, et al. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13:722–36

Makarova KS, Wolf YI, Koonin EV. 2013. The basic building blocks and evolution of CRISPR-CAS systems. *Biochem. Soc. Trans.* 41:1392–400

Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P. 2007. Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.* 35:962–74

Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322:1843–45

Marraffini LA, Sontheimer EJ. 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463:568–71

Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD. 2013. Rapid evolution of the human gut virome. *PNAS* 110:12450–55

Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, et al. 2011. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* 21:1616–25

Minot S, Wu GD, Lewis JD, Bushman FD. 2012. Conservation of gene cassettes among diverse viruses of the human gut. *PLOS ONE* 7:e42342

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60:174–82

Paez-Espino D, Morovic W, Sun CL, Thomas BC, Ueda K, et al. 2013. Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat. Commun.* 4:1430

Paez-Espino D, Sharon I, Morovic W, Stahl B, Thomas BC, et al. 2015. CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *mBio* 6:e00262

Paganelli FL, Willems RJ, Leavis HL. 2012. Optimizing future treatment of enterococcal infections: attacking the biofilm? *Trends Microbiol.* 20:40–49

Palmer KL, Gilmore MS. 2010. Multidrug-resistant enterococci lack CRISPR-*cas*. *mBio* 1:e00227

Pawluk A, Bondy-Denomy J, Cheung VH, Maxwell KL, Davidson AR. 2014. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* 5:e00896

Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, et al. 2003. Origins of highly mosaic mycobacteriophage genomes. *Cell* 113:171–82

Pourcel C, Salvignol G, Vergnaud G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151:653–63

Pride DT, Salzman J, Relman DA. 2012. Comparisons of clustered regularly interspaced short palindromic repeats and viromes in human saliva reveal bacterial adaptations to salivary viruses. *Environ. Microbiol.* 14:2564–76

Pride DT, Sun CL, Salzman J, Rao N, Loomer P, et al. 2011. Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. *Genome Res.* 21:126–36

Quax TE, Voet M, Sismeiro O, Dillies MA, Jagla B, et al. 2013. Massive activation of archaeal defense genes during viral infection. *J. Virol.* 87:8419–28

Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, et al. 2015. Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell* 163:854–65

Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, et al. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466:334–38

Reyes A, Semenkovich NP, Whiteson K, Rohwer F, Gordon JI. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. *Nat. Rev. Microbiol.* 10:607–17

Rho M, Wu YW, Tang H, Doak TG, Ye Y. 2012. Diverse CRISPRs evolving in human microbiomes. *PLOS Genet.* 8:e1002441

Richter C, Dy RL, McKenzie RE, Watson BN, Taylor C, et al. 2014. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.* 42:8516–26

Rohwer F. 2003. Global phage diversity. *Cell* 113:141

Rollins MF, Schuman JT, Paulus K, Bukhari HS, Wiedenheft B. 2015. Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 43:2216–22

Sampson TR, Napier BA, Schroeder MR, Louwen R, Zhao J, et al. 2014. A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. *PNAS* 111:11163–68

Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, Weiss DS. 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497:254–57

Sampson TR, Weiss DS. 2013. Degeneration of a CRISPR/Cas system and its regulatory target during the evolution of a pathogen. *RNA Biol.* 10:1618–22

Sanguino L, Franqueville L, Vogel TM, Larose C. 2015. Linking environmental prokaryotic viruses and their host through CRISPRs. *FEMS Microbiol. Ecol.* 91:fiv046

Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJ, et al. 2003. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *J. Clin. Microbiol.* 41:15–26

Shariat N, Dudley EG. 2014. CRISPRs: molecular signatures used for pathogen subtyping. *Appl. Environ. Microbiol.* 80:430–39

Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, et al. 2015. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60:385–97

Smedile F, Messina E, La Cono V, Tsoy O, Monticelli LS, et al. 2013. Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan-Vavilov Deep. *Environ. Microbiol.* 15:167–82

Snyder JC, Bateson MM, Lavin M, Young MJ. 2010. Use of cellular CRISPR (clusters of regularly interspaced short palindromic repeats) spacer-based microarrays for detection of viruses in environmental samples. *Appl. Environ. Microbiol.* 76:7251–58

Sorokin VA, Gelfand MS, Artamonova II. 2010. Evolutionary dynamics of clustered irregularly interspaced short palindromic repeat systems in the ocean metagenome. *Appl. Environ. Microbiol.* 76:2136–44

Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26:335–40

Stern A, Mick E, Tirosh I, Sagy O, Sorek R. 2012. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* 22:1985–94

Sternberg SH, Doudna JA. 2015. Expanding the biologist's toolkit with CRISPR-Cas9. *Mol. Cell* 58:568–74

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507:62–67

Swarts DC, Mosterd C, van Passel MW, Brouns SJ. 2012. CRISPR interference directs strand specific spacer acquisition. *PLOS ONE* 7:e35888

Takeuchi N, Wolf YI, Makarova KS, Koonin EV. 2012. Nature and intensity of selection pressure on CRISPR-associated genes. *J. Bacteriol.* 194:1216–25

Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, et al. 2009. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459:950–96

Touchon M, Rocha EP. 2010. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLOS ONE* 5:e11126

Tyson GW, Banfield JF. 2008. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ. Microbiol.* 10:200–7

Vale PF, Lafforgue G, Gatchitch F, Gardan R, Moineau S, Gandon S. 2015. Costs of CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. *Proc. R. Soc. B* 282:20151270

Vale PF, Little TJ. 2010. CRISPR-mediated phage resistance and the ghost of coevolution past. *Proc. R. Soc. B* 277:2097–103

van Belkum A, Soriaga LB, LaFave MC, Akella S, Veyrieras JB, et al. 2015. Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. *mBio* 6:e01796

van der Oost J, Westra ER, Jackson RN, Wiedenheft B. 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 12:479–92

van der Ploeg JR. 2009. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology* 155:1966–76

van Embden JD, van Gorkom T, Kremer K, Jansen R, van Der Zeijst BA, Schouls LM. 2000. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J. Bacteriol.* 182:2393–401

van Houte S, Ekroth AK, Broniewski JM, Chabas H, Ben A, et al. 2016. The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* 532:385–88

Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, et al. 2013. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLOS Genet*. 9:e1003454

Vergnaud G, Li Y, Gorge O, Cui Y, Song Y, et al. 2007. Analysis of the three *Yersinia pestis* CRISPR loci provides new tools for phylogenetic studies and possibly for the investigation of ancient DNA. *Adv. Exp. Med. Biol*. 603:327–38

Voorhies AA, Eisenlord SD, Marcus DN, Duhaime MB, Biddanda BA, et al. 2015. Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. *Environ. Microbiol.* 18:358–71

Vorontsova D, Datsenko KA, Medvedeva S, Bondy-Denomy J, Savitskaya EE, et al. 2015. Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. *Nucleic Acids Res*. 43:10848–60

Wang Z, Pan Q, Gendron P, Zhu W, Guo F, et al. 2016. CRISPR/Cas9-derived mutations both inhibit HIV-1 replication and accelerate viral escape. *Cell Rep*. 15:481–89

Weinberger AD, Sun CL, Plucinski MM, Denef VJ, Thomas BC, et al. 2012a. Persisting viral sequences shape microbial CRISPR-based immunity. *PLOS Comput. Biol.* 8:e1002475

Weinberger AD, Wolf YI, Lobkovsky AE, Gilmore MS, Koonin EV. 2012b. Viral diversity threshold for adaptive immunity in prokaryotes. *mBio* 3:e00456

Westra ER, Buckling A, Fineran PC. 2014. CRISPR-Cas systems: beyond adaptive immunity. *Nat. Rev. Microbiol.* 12:317–26

Westra ER, Semenova E, Datsenko KA, Jackson RN, Wiedenheft B, et al. 2013. Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLOS Genet*. 9:e1003742

Westra ER, van Erp PB, Kunne T, Wong SP, Staals RH, et al. 2012. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46:595–605

Westra ER, van Houte S, Oyesiku-Blakemore S, Makin B, Broniewski JM, et al. 2015. Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr. Biol.* 25:1043–49

Wiedenheft B, Sternberg SH, Doudna JA. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482:331–38

Wright AV, Nunez JK, Doudna JA. 2016. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164:29–44

Yosef I, Goren MG, Qimron U. 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* 40:5569–76

Yosef I, Manor M, Kiro R, Qimron U. 2015. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *PNAS* 112:7267–72

Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, et al. 2012. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLOS ONE* 7:e38077

Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA. 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191:210–19

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, et al. 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163:759–71