Coevolution between bacterial CRISPR-Cas systems and their bacteriophages

Bridget N.J. Watson[1*], Jurre A. Steens[2], Raymond H.J. Staals[2], Edze R. Westra[1*] and Stineke van Houte[1*]

[1]ESI, Biosciences, University of Exeter, Cornwall Campus, Penryn TR10 9FE, UK.

[2]Laboratory of Microbiology, Wageningen University and Research, 6708 WE Wageningen, The Netherlands

*For correspondence: B.Watson3@exeter.ac.uk; E.R.Westra@exeter.ac.uk; C.van-Houte@exeter.ac.uk

Abstract (150 words max)

CRISPR-Cas systems provide bacteria and archaea with adaptive, heritable immunity against their viruses (bacteriophages; phages) and other parasitic genetic elements. CRISPR-Cas systems are highly diverse, and we are only beginning to understand their relative importance in phage defense. In this review we will discuss when and why CRISPR-Cas immunity against phages evolves, and how this in turn selects for the evolution of immune evasion by phages. Finally, we will discuss our current understanding on if, and when, we observe coevolution between CRISPR-Cas systems and phages, and how this may be influenced by the mechanism of CRISPR-Cas immunity.

22 **Introduction**

23 Bacteriophages (phages) are genetically and morphologically diverse and outnumber their microbial
24 hosts in most environments (Suttle, 2007). They shape microbial evolution, community structure, and
25 their ecological functions, such as carbon and nutrient cycling (Koskella and Brockhurst, 2014).

26 Bacteria use different strategies to limit infection by phages. Of these, CRISPR-Cas (clustered regularly
27 interspaced short palindromic repeats; CRISPR-associated) immunity is the only adaptive and heritable
28 defense system known to date (Hampton et al., 2020; Hille et al., 2018). Generally, the CRISPR-Cas
29 immune response is thought of as a three-stage process (Figure 1). During adaptation, the acquisition
30 machinery captures a fragment of the invading phage genome (known as a 'protospacer') and
31 incorporates it into a CRISPR array in the bacterial chromosome as a 'spacer' sequence between repeat
32 sequences. In the expression phase, the array is transcribed and processed into crRNAs, that contain a
33 single spacer sequence. CrRNAs associate with an effector protein or complex, which, during the
34 interference stage, surveys the cell for genetic material complementary to the spacer. If detected, the
35 invading nucleic acid will be cleaved or degraded.

36 CRISPR-Cas systems have been identified in ~40% of bacterial and ~90% of archaeal genomes and can
37 be grouped into two Classes containing six types (Makarova et al., 2020) (Figure 2). In general, Class
38 1 are more abundant than Class 2, with types I, II and III being the most abundant, comprising 60, 13
39 and 25 percent of bacterial CRISPR systems, respectively, (Makarova et al., 2015; Weissman et al.,
40 2019). In archaea, type I systems make up 64 percent, and type III systems, 34 percent, while type II
41 systems are rare (Burstein et al., 2017). In agreement, most of our knowledge of the ecology and
42 evolution of CRISPR-Cas systems exists for type I, II and III systems, which will therefore be the focus
43 of this review.

44 The weak association between the phylogenies of bacterial and archaeal hosts and the phylogenies of
45 the CRISPR-Cas systems they carry suggests that horizontal gene transfer has mediated their spread
46 (Makarova et al., 2015). Further, the distribution and prevalence of CRISPR-Cas systems is not uniform
47 across bacterial and archaeal taxa or environments, and some lineages that are found across a wide
48 range of environments appear to lack CRISPR-Cas systems altogether (Burstein et al., 2016). While the
49 reasons for variation in distribution and prevalence are largely unknown, they are likely related to the
50 ecology of the host, as the balance of the costs of carrying or expressing the system and the benefits of
51 adaptive immunity will ultimately determine whether CRISPR-Cas is lost or retained. To date, several
52 ecological factors have been correlated with CRISPR-Cas prevalence, including oxygen requirement
53 and temperature, as more systems are found in thermophilic bacteria and those associated with an
54 anaerobic lifestyle (Weissman et al., 2019). CRISPR-Cas systems are also more commonly found in
55 free-living than host-associated microbes (Burstein et al., 2016; Weissman et al., 2019). In addition, the
56 presence or absence of various dsDNA repair mechanisms in the host has been linked with CRISPR-

57 Cas prevalence, as some systems have been shown to require the activity of repair mechanisms
58 (Bernheim et al., 2019) or impair the mechanism function (Bernheim et al., 2017). Explanations for the
59 roles of these environmental and genetic factors have been proposed, but the mechanisms through which
60 they act are in many cases not yet determined.

61

62 Diversity of CRISPR-Cas systems

63 CRISPR-Cas systems are diverse and can be categorised into two Classes, six types and over 30
64 subtypes based on differences between the Cas proteins encoded (Makarova et al., 2020)(Figure 2).
65 The adaptation machinery, Cas1 and Cas2, is conserved across almost all CRISPR-Cas variants. Class
66 I systems (consisting of types I, III and IV) encode multi-protein effector complexes, whereas Class II
67 systems (types II, V and VI) consist of single protein effectors. Across the different CRISPR-Cas types,
68 activation of the effector components occurs when different types of nucleic acids are detected and
69 results in different outcomes for the cells and invading genetic materials (Figure 2). In Type I systems,
70 the Cas protein effector complex (known as Cascade) recognises double strand (ds) DNA and recruits
71 Cas3 to progressively degrade the foreign DNA (Brouns et al., 2008; Westra et al., 2012). Interference
72 relies on recognising a short sequence, called the protospacer adjacent motif (PAM) and a perfect
73 spacer-protospacer match in the 8 nt adjacent to the PAM, called the 'seed' sequence (Semenova et al.,
74 2011; Wiedenheft et al., 2011). Clearing the phage DNA will typically result in host survival (Westra
75 et al., 2015). However, if phage infection progresses before the infection can be cleared, cell death may
76 occur, which is likely due to the irreversible damage inflicted on bacterial processes in the cell for phage
77 replication, resulting in abortive infection (Watson et al., 2019b). In both instances, phages will be
78 removed from the population to protect vulnerable cells.

79 Type II systems also require the recognition of a (different) PAM and seed sequence complementarity
80 in the dsDNA target sequence (Jinek et al., 2012). The effector protein, Cas9, will generate a double
81 strand break in the genome of the invading genetic material and bacteria survive infection (Garneau et
82 al., 2010). In type III systems, target RNAs resulting from transcription of the foreign DNA will bind
83 to the effector complex and subsequently activate the Cas10 effector, initiating different pathways of
84 defence, including target RNA and DNA cleavage (Deng et al., 2013; Hale et al., 2009; Samai et al.,
85 2015; Staals et al., 2014; Tamulaitis et al., 2014; Zhang et al., 2012). Cas10 also generates cyclic
86 oligoadenylates (cOAs), which activate CARF effector proteins to induce non-specific degradation of
87 RNA in the cell (Kazlauskiene et al., 2017; Niewoehner et al., 2017). While immunity mediated by type
88 III systems can result in host survival (Pyenson et al., 2017), the collateral RNA damage can induce
89 dormancy and possibly cell death. Further, no canonical PAM is recognised by type III systems,
90 although some systems recognise motifs flanking the target RNA (called rPAMs) (Elmore et al., 2016).

Type IV systems are typically found on plasmids or plasmid-like elements, where they are thought to be involved in plasmid competition (Pinilla-Redondo et al., 2020). As they typically lack an adaptation module, it is thought they may associate with compatible host-encoded CRISPR-Cas systems (Pinilla-Redondo et al., 2020). Like type II systems, type V cuts both strands of the target DNA, albeit asymmetrically (Zetsche et al., 2015). Additionally, activation of the Cas12 protein can result in non-target ssDNA degradation (Chen et al., 2018), although more work is needed to understand the consequences of this collateral damage on host outcome (Varble and Marraffini, 2019). Finally, type VI systems uniquely recognise and cleave RNA. The effector, Cas13, degrades the target RNA, as well as non-targeted RNA in the cell. Due to the high levels of RNA degradation in the cell, type VI-mediated immunity induces dormancy, but cells act as 'sinks' to remove phages from the population (Meeske et al., 2019).

When does CRISPR immunity evolve?

While CRISPR-based immunity can protect against phages and other parasitic mobile genetic elements, bacteria do not always rely on CRISPR-Cas immune systems for phage defence. First, CRISPR-Cas systems are not present in all prokaryotic genomes, as discussed above (Burstein et al., 2016; Makarova et al., 2020). Second, even if bacteria do carry a CRISPR-Cas immune system, they may rely on other defenses instead. Indeed, cells often have multiple defence options, such as surface mutation or modification (*sm*) and Restriction-Modification (R-M) systems (Hampton et al., 2020). Like CRISPR-Cas, R-M systems function following phage DNA injection and cleave unmodified DNA at certain sequence motifs. On the other hand, *sm* prevents phage binding and entry by altering or masking the receptor, so phages remain in the population. Recent studies have demonstrated that the defense strategies that bacteria evolve in response to phage strongly depends on the environment. When we consider an initially sensitive bacterial population that is infected with phages, the most commonly observed mechanisms to acquire resistance in the short-term are mutation or loss of the receptor that is used by the phage to attach to the cell surface and CRISPR-Cas. Initially, the rates at which both types of resistance evolve (i.e., rates of spacer acquisition and rates of receptor mutation) will be one of the key determinants of the type of defence that dominates in the bacterial population, especially if different competing genotypes have similar relative fitness. In the longer term, which defense ultimately prevails will depend on whether hosts with CRISPR immunity have higher fitness than those with mutated receptors or vice versa, which in turn will depend on the environment (Westra and Levin, 2020). In addition, other defense mechanisms may be acquired horizontally over these longer timescales that compete or combine with CRISPR immunity or *sm* to provide more robust or lower cost defense (Dimitriu et al., 2020). It is becoming increasingly clear that both in the short- and long-term, the

125 ecological context is a major determinant for the evolution of bacterial defenses (van Houte et al.,
126 2016a).

127

128 ***The rates of spacer acquisition***

129 Following phage infection, the initial rise in the frequency of CRISPR-immune bacteria in a phage-
130 sensitive population will strongly depend on both the rates of spacer acquisition and the rates at which
131 surface-based resistance mutations are generated (Gurney et al., 2019). If the host has a higher mutation
132 rate, then more mutations in the receptor are generated and this will result in a higher proportion of
133 surface-based resistance in the population (Chevallereau et al., 2019). Three factors have been identified
134 that determine the rates of spacer acquisition by CRISPR-Cas immune systems: defective phages, *cas*
135 gene expression and priming. The first is related to the challenge of any CRISPR-Cas system to capture
136 a protospacer from the phage genome and carry out interference before phage replication and lysis of
137 the bacterial cell. It turns out that high rates of spacer acquisition occur from infection with defective
138 phages that naturally exist in phage populations (Hynes et al., 2014). This form of vaccination, whereby
139 the CRISPR-Cas can generate immunity against the phage from an attenuated form, is analogous to the
140 formation of defective interfering particles (DIPs) in eukaryotic viruses and the immune response they
141 trigger (Yang et al., 2019). DIPs are formed by error-prone viral replication and while they lack viral
142 genetic material, they can still infect a host cell. Consequently, these attenuated viral particles can be
143 used as a viral vaccine. Spacer adaptation is also enhanced in cells that also carry R-M systems, as these
144 limit phage replication through phage DNA degradation and these cleavage products can be captured
145 by the Cas adaptation machinery (Hynes et al., 2014). Generally, immunity is more readily generated
146 when the CRISPR-Cas system is exposed to free DNA ends, through processes including RecBCD-
147 mediated DNA repair, degradation or processing of DNA breaks occurring at stalled replication forks
148 (Levy et al., 2015) or when linear phage DNA first enters the cell (Modell et al., 2017).

149 The rates at which spacers are being acquired, whether it be from defective or intact phages, is
150 also determined by the expression levels of the CRISPR-Cas immune system (Patterson et al., 2017;
151 Shivram et al., 2021). Expression of CRISPR-Cas is often tightly regulated, presumably because their
152 expression may carry costs in the absence of phage infection, for example due to acquisition of spacers
153 from the host genome (reviewed in (Weissman et al., 2020)). Two key factors that determine the
154 infection risk of an individual bacterium are the density of bacterial hosts in the population, since phages
155 spread more effectively when their density is high, and the expression levels of phage receptors on the
156 cell surface. Consistent with this, the type I-E, I-F and III-A CRISPR-Cas systems of *Serratia* sp. ATCC
157 39006 are all regulated by quorum sensing (QS) (Patterson et al., 2016), and a single regulator, the Rcs
158 stress response, regulates both CRISPR-Cas immunity and the expression of cell surface proteins (Smith
159 et al., 2021). In *Pseudomonas aeruginosa*, expression of both the Type IV pilus, which is a key phage

160 receptor for this species, and the CRISPR-Cas immune system are positively regulated by quorum
161 sensing (QS) (Hoyland-Kroghsbo et al., 2016); Broniewski, unpublished data). The alginate
162 biosynthesis pathway represses CRISPR-Cas expression in *P. aeruginosa* cells growing on surfaces,
163 which also ensures that the immune system is expressed at levels necessary for the risk of infection.
164 Further, CRISPR-Cas expression levels can be induced by phage infection (Agari et al., 2010; He et al.,
165 2017; Quax et al., 2013), as well as membrane stress (Ratner et al., 2015) and metabolic status (Agari
166 et al., 2010; Patterson et al., 2015).

167 Priming (or primed adaptation) is a third factor that determines the rate of spacer acquisition.
168 This mechanism of type I systems increases the rates at which new spacers are acquired from invading
169 elements. Perfectly matched spacers, or those with mismatches in the protospacer adjacent motif (PAM)
170 or PAM-adjacent 'seed' sequence in the targeted region, can activate the incorporation of new spacers
171 into the CRISPR array (Staals et al., 2016). Consequently, in short-term evolution experiments, the
172 observed levels of spacer acquisition are higher when bacteria are primed (Datsenko et al., 2012; Staals
173 et al., 2016). Another potential benefit of primed adaptation is that it introduces a bias towards spacer
174 acquisition from parasitic DNA with sequence similarity to pre-existing spacers relative to the
175 acquisition of spacers from the bacterial genome, which would result in autoimmunity (Weissman et
176 al., 2020). Primed adaptation has also been observed in some type II systems, and similarly to type I
177 systems (Staals et al., 2016), it occurs more often with perfectly matched spacers than with those
178 carrying mismatches, since Cas9 must cleave the phage DNA to generate substrates for adaptation
179 (Nussenzweig et al., 2019). As a result, immunity may not be as rapidly generated in type II systems
180 and the system is less able to quickly acquire new spacers in response to escape phages.

181

182 ***Fitness costs and benefits of CRISPR immunity***

183 In the long term, the type of resistance that prevails in a population is determined by the fitness costs
184 and benefits of each strategy, as the one with the highest net benefits will dominate (van Houte et al.,
185 2016a; Westra et al., 2015). Several factors have been identified that influence the relative fitness of
186 bacteria with mutated surface receptors and those with CRISPR immunity. The type II-A CRISPR-Cas
187 system of *Streptococcus thermophilus* has been shown to be costly to maintain and express (Vale et al.,
188 2015), but the evolution of alternative defence such as *sm* is almost never observed in *S. thermophilus*,
189 explaining why bacteria almost exclusively rely on CRISPR immunity (discussed in (Westra and Levin,
190 2020)). For *P. aeruginosa*, CRISPR-Cas immunity is also associated with a fitness cost, which was
191 found to be strictly infection-dependent (i.e. CRISPR immune bacteria that are infected by phages have
192 a lower fitness relative to uninfected cells) (Westra et al., 2015). Recent mechanistic studies have shown
193 that these phage infection-induced fitness costs are at least partially due to the production of phage-
194 encoded proteins prior to phage removal by CRISPR-Cas (Meaden et al., 2020). Due to this cost, which

195 is paid every time a cell gets infected, CRISPR-Cas immunity is favoured over *sm* at low phage
196 densities, when infections are relatively rare, whereas high phage densities select for *sm*, which has a
197 constitutive cost independent of the number of infections (Meaden et al., 2020; Westra et al., 2015).
198 Hence, in environments where there is a constant immigration of phages (Westra et al., 2015) or where
199 high phage densities are maintained due to the immigration of naïve hosts into the population (Chabas
200 et al., 2016), bacteria with surface-based resistance ultimately dominate the population.

201 The magnitude of the fitness costs of CRISPR immunity and surface resistance depend on the
202 ecological context beyond phage densities. For example, it was found that the cost of surface-based
203 resistance relative to those of CRISPR immunity are amplified in the presence of a microbial community
204 (Alseth et al., 2019). Moreover, these amplified costs manifested in the presence of some, but not other
205 bacterial species. Understanding the mechanistic basis of how microbial community composition
206 determines the fitness costs and benefits of alternative phage defence strategies will be key to
207 understanding their ecological distribution.

208 The fitness benefits of spacer acquisition by a CRISPR-Cas immune system rely on the ability
209 to protect against infection by genetically similar phage in the future. Since CRISPR-Cas immune
210 systems rely on sequence identity between the phage from which a spacer is acquired and subsequent
211 infecting phages, high genetic diversity in the phage population is predicted to reduce the benefits of
212 acquired spacers, and therefore lead to lower levels of CRISPR immunity (Iranzo et al., 2013;
213 Weinberger et al., 2012b). Indeed, experimental tests show that increasing phage genetic diversity
214 promotes the evolution of more generalist defense through mutation of the phage receptor instead of
215 CRISPR immunity (Broniewski et al., 2020). Those bacteria that do evolve CRISPR immunity often
216 carry multiple spacers, which leads to higher protection as phages are less likely to carry mutations in
217 multiple targets (Broniewski et al., 2020). The presence of genetically identical phages in the population
218 that encode genes that block the CRISPR-Cas immune system also favours the evolution of surface-
219 based resistance (Chevallereau et al., 2020).

220 CRISPR-based immunity can also have fitness costs due to the incorporation of self-targeting
221 spacers (Wimmer and Beisel, 2019). Spacers targeting bacterial genomes are widespread, although
222 there are often mutations in either the protospacers or the repeat sequences in the CRISPR array (Stern
223 et al., 2010), since self-targeting spacers are cytotoxic (Gomaa et al., 2014; Vercoe et al., 2013). While
224 mutations might make these spacers non-functional, the mismatches can activate primed adaptation,
225 resulting in autoimmunity (Staals et al., 2016; Vercoe et al., 2013). In some type I systems, almost half
226 of the self-targeting spacers map to prophages in the genome but targeting may be prevented by phage-
227 encoded anti-CRISPR proteins (Nobrega et al., 2020). The consequences of prophage-targeting vary
228 between CRISPR-Cas types, as in type I systems, prophage-targeting spacers can be cytotoxic and the
229 CRISPR-Cas system may be lost to remove the cost and maintain the prophage (Rollie et al., 2020). On

230 the other hand, in type III systems, the requirement of target transcription for targeting allows prophages
231 to be tolerated (Goldberg et al., 2014). However, fitness costs due to cytotoxicity can still occur if
232 spacers target prophage genes that are expressed (Goldberg et al., 2018).

233 The potential for CRISPR-Cas systems to target and exclude prophages can be a disadvantage
234 to the host as they can provide beneficial traits (Westra and Levin, 2020). Similarly, CRISPR-Cas
235 systems can limit the uptake of novel genetic material which may restrict their evolutionary potential
236 (Bikard et al., 2012; Marraffini and Sontheimer, 2008; Watson et al., 2018). Consequently, CRISPR-
237 Cas systems may be lost or inactivated to enable the uptake of plasmids (Jiang et al., 2013) which is the
238 case for several bacterial pathogens (Hatoum-Aslan and Marraffini, 2014). In accordance, in both
239 *Enterococcus faecalis* and *P. aeruginosa* a negative correlation was found between CRISPR-Cas
240 systems and horizontally acquired elements (Palmer and Gilmore, 2010; Wheatley and MacLean,
241 2020). However, a global analysis of all sequenced genomes did not detect an interaction between
242 CRISPR-Cas and signatures of horizontal gene transfer, suggesting that this effect may vary across taxa
243 or ecosystems (Gophna et al., 2015).

244

245 What limits the durability of CRISPR immunity?

246 Evolution of CRISPR immunity in bacteria can lead to rapid phage extinction. One way for phages to
247 persist is through immune evasion, which can be achieved through various different mechanisms
248 (extensively reviewed in (Hampton et al., 2020; Malone et al., 2020a), including, through mutation of
249 target sequences, production of anti-CRISPR proteins, and physical barriers that shield phage DNA
250 from cleavage (Figure 2). However, in addition to phage evolution, bacteria can also evolve to lose their
251 CRISPR immunity, which can play an important role in the coexistence of bacteria with CRISPR
252 immune systems and their phages.

253

254 *Evolution of phage infectivity on CRISPR immune bacteria*

255 Phages can overcome CRISPR-Cas targeting by altering the protospacer sequence through mutations,
256 deletions, or gene rearrangements. For type I and type II systems, single point mutations in the PAM or
257 the seed sequence can prevent CRISPR-Cas interference (Deveau et al., 2008). The probability of
258 phages acquiring mutations to overcome type I or II CRISPR-Cas immunity depends on the composition
259 of the host population (Chabas et al., 2018). Escape phages are most likely to evolve in a population
260 with an intermediate level of resistant hosts, since there needs to be enough hosts on which phages can
261 replicate, but also enough selection for acquiring infectivity. However, having hosts with multiple or
262 different phage-targeting spacers decreases the likelihood of escape phages emerging. For escape

263 phages that do evolve, the accumulation of point mutations that enable CRISPR escape can have
264 negative fitness consequences (Chabas et al., 2019; Common et al., 2019; Watson et al., 2019a), which
265 may select against those phages in the long-term. It is more difficult for phages to overcome multiple
266 spacers within a single host and deletion or recombination may be a more effective way to do this (Han
267 et al., 2013; Paez-Espino et al., 2015; Watson et al., 2019a), although large deletions can alter the phage
268 structure and reduce phage infectivity (Watson et al., 2019a). In contrast to type I and II, type III
269 CRISPR-Cas systems are much more tolerant of mismatches due to a flexible seed sequence (Steens et
270 al., 2021) and phages can only escape through deletions (Pyenson et al., 2017).

271 Phages can also evade CRISPR by providing physical barriers that shield the phage DNA from
272 CRISPR-Cas immune complexes (Malone et al., 2020a). Examples of these include modification of the
273 phage DNA to avoid detection by CRISPR-Cas (Vlot et al., 2018), or, in the case of jumbo phages, the
274 production of a nucleus-like structure within the host cell that surrounds the phage DNA but excludes
275 CRISPR-Cas machinery (Malone et al., 2020b; Mendoza et al., 2020). Such mechanisms may have a
276 fitness cost, although this has not been studied to date. While these physical barriers protect phages
277 from CRISPR-Cas systems that recognize DNA, they do not protect against type III immunity as phage
278 RNA will still be targeted in the cytoplasm (Malone et al., 2020b). Finally, phages may encode anti-
279 CRISPR proteins (Acrs) that inhibit CRISPR-Cas activity. Different Acrs have been found to act at
280 each phase of immunity, but most impair interference by interacting with effector proteins or complexes
281 (See (Li and Bondy-Denomy, 2021) from this special issue).

282

283 ***Loss of CRISPR immunity through bacterial mutation***

284 Another mechanism that can enable the coexistence of bacteria with CRISPR immune systems and the
285 phages they target is through the loss of spacers or mutation of *cas* genes (Bradde et al., 2017; Han and
286 Deem, 2017; Levin et al., 2013; Weissman et al., 2018b). Such spontaneous loss of immunity can
287 provide phages with sensitive hosts that they can use to replicate, hence avoiding extinction. Both the
288 mutation of *cas* genes and the loss of spacers from CRISPR arrays has been observed experimentally
289 (Jiang et al., 2013; Weinberger et al., 2012a). The idea that CRISPR immunity can be lost at a high rate
290 is further supported by experiments where *Staphylococcus epidermidis* was transformed with an
291 antibiotic resistance plasmid, resulting in estimated rates of between one in a thousand and one in ten
292 thousand loss events per individual per generation (Jiang et al., 2013). In an observational study, spacer
293 loss events correlated with the resurgence of a phage population that was no longer targeted
294 (Weinberger et al., 2012a), and an experimental study with *S. thermophilus* DGCC7710 and its lytic
295 phage 2972 found that the stable coexistence of bacteria and phages, that occurred despite the high
296 frequency of CRISPR immunity in their experiments, dynamics that were best captured in a

297 mathematical model that assumed high rates of loss of CRISPR-Cas immunity (Weissman et al.,
298 2018b).

299 Another mechanism through which pre-existing levels of CRISPR immunity can decrease is
300 through the acquisition of new spacers that target other genetic elements, which are inserted at the
301 leader-proximal end of the CRISPR array. Spacers that are closer to the leader-end of the CRISPR array
302 provide higher levels of immunity than those towards the trailer-end of the array (McGinn and
303 Marraffini, 2016), since expression of spacers decreases with distance from the leader end of the array,
304 where transcription is initiated (Zoephel and Randau, 2013). This may help to explain why most
305 CRISPR arrays contain only between ~10-40 spacers, even if they can in theory contain hundreds of
306 spacers (Bradde et al., 2020; Grissa et al., 2007). Indeed, theoretical studies that explore the trade-offs
307 between effectiveness against a specific phage and coverage of as many phages as possible, find that
308 the common ranges of ~10-40 spacers provides optimal defence under a broad range of realistic
309 parameter estimates (Bradde et al., 2020; Martynov et al., 2017). The acquisition of multiple CRISPR
310 arrays within a single host, along with sufficiently high levels of *cas* gene expression (Watson et al.,
311 2019b), may occur to overcome these challenges by maximising novel spacer acquisition and memory
312 span (Weissman et al., 2018a).

313

314 <u>**What does CRISPR-mediated co-evolution look like?**</u>

315 The observations that the CRISPR-Cas system acquires a "memory" of past infections through spacer
316 integration (Barrangou et al., 2007), and that phages can readily overcome CRISPR-Cas immunity
317 through protospacer or PAM mutation (Deveau et al., 2008; Malone et al., 2020a) when using a clonal
318 population of CRISPR-immune bacteria, has led to the idea that CRISPR-immune bacteria and their
319 phages are engaged in an ongoing coevolutionary arms race, in which hosts accumulate spacers, and
320 phage accumulate point mutations (Levin, 2010). This model of CRISPR-phage coevolution has been
321 revised in recent years, based on more refined models (Childs et al., 2014; van Houte et al., 2016a;
322 Weissman et al., 2018b), and an appreciation of the mechanistic differences that exist between type I,
323 type II and type III systems (Hille et al., 2018).

324

325 ***Type I CRISPR-phage coevolution***

326 Theoretical studies on CRISPR-phage coevolution predicted that over time, different spacers that each
327 target the phage at different positions in its genome, appear in the population, resulting in high levels
328 of spacer diversity at the population level. Meanwhile, phages with mutations to evade CRISPR
329 targeting appear (Childs et al., 2014; Iranzo et al., 2013; Levin, 2010; Weissman et al., 2018b). The

330 ultimate outcome (i.e. ongoing coevolution or extinction of host and/or phage) was predicted to be
331 influenced by the spacer acquisition rate, as well as the number of potential unique spacers (as
332 determined by the requirement for PAM sequences) (Childs et al., 2014; Iranzo et al., 2013).
333 Specifically, with different phage-targeting spacers in different bacterial clones in the population (i.e.
334 population-level spacer diversity), phages may not be able to overcome all of the spacers and might
335 become eliminated (Childs et al., 2014). Experimental evolution studies where a virulent mutant of
336 phage DMS3 was used to infect *P. aeruginosa* strain PA14 found high population-level spacer diversity
337 in the CRISPR array due to the acquisition of new spacers targeting the phage. Moreover, the majority
338 of hosts in the population acquired a single, unique spacer (Meaden et al., 2020; Westra et al., 2015).
339 This bacterial strain carries a type I CRISPR-Cas system that is primed against phage DMS3, which
340 promotes rapid spacer acquisition and therefore generation of spacer diversity at the population level,
341 as discussed above (Figure 3). As predicted by theory (Childs et al., 2014), the likelihood that phages
342 are driven extinct is positively correlated with the level of CRISPR spacer diversity, and the level of
343 spacer diversity naturally generated in a population drives phage extinct rapidly and consistently
344 (Morley et al., 2016; van Houte et al., 2016b). The immigration of susceptible hosts into the population
345 can provide permissive hosts and enable phages to replicate (Chabas et al., 2016). However, these
346 phages will still not evolve to be able to replicate on the CRISPR clones and will just coexist in the
347 population, rather than coevolve with CRISPR. Hence, for CRISPR-Cas systems that generate high
348 levels of population-level diversity, CRISPR-phage coevolution is likely only very short lived, and in
349 accordance experimental studies have not found evidence to date for CRISPR-phage coevolution when
350 bacteria carry a primed type I CRISPR-Cas system.

351

352 ***Type II CRISPR-phage coevolution***

353 The most important model system for studying type II CRISPR-phage coevolution is that of the lactic
354 acid bacterium *S. thermophilus* DGCC7710 and its phage 2972. Using this experimental system,
355 CRISPR-Cas systems were first demonstrated to provide adaptive immunity against phages (Barrangou
356 et al., 2007) and phages were shown to evolve to overcome CRISPR immunity through point mutation
357 (Deveau et al., 2008). Co-culture studies in milk and in defined media have demonstrated that spacer
358 acquisition occurs at a much lower rate than those observed in type I systems, despite some form of a
359 priming mechanism in type II systems (Nussenzweig et al., 2019) (Figure 3). As a result of this, bacterial
360 populations during the early stages of a phage epidemic are virtually clonal, which is the result of a
361 single bacterium that acquired a spacer sweeping to fixation (Common et al., 2019; Paez-Espino et al.,
362 2013; Paez-Espino et al., 2015; Weissman et al., 2018b). Although initially, spacer abundance is
363 determined by acquisition rates, rather than selection (Heler et al., 2019). Several studies found that the
364 bacteria and phage can coexist for extensive periods of time, ranging from tens to hundreds of

365 generations (Common et al., 2019; Paez-Espino et al., 2015). In accordance with theory that predicts
366 that lower rates of spacer acquisition increases the probability for CRISPR-phage coevolution (Childs
367 et al., 2014; Iranzo et al., 2013), escape phages were found to emerge in these experiments, and an arms
368 race where bacteria accumulate spacers and phage accumulate escape mutations ensues (Common et
369 al., 2019). However, the arms race is asymmetrical, with the host acquiring cost-free spacers and slowly
370 increasing the population-levels of spacer diversity, whereas the phage accumulates costly point
371 mutations and is unable to keep up with the increase in spacers at the population level, ultimately
372 resulting in phage extinction (Common et al., 2019). Interestingly, a recent study examined how the
373 environment shapes the coevolutionary interaction, with a focus on the role of spatial structure, which
374 limits host and phage mobility and therefore could affect the effects of spacer diversity in the bacterial
375 population. Data from this study suggest that escape phages emerge more readily in structured
376 environments compared to well-mixed broth, leading to a greater number of coevolutionary cycles and
377 hence a greater number of spacers that are acquired by the bacterial hosts (Pyenson and Marraffini,
378 2020).

379

380 ***Type III CRISPR-phage coevolution.***

381 Apart from diversity at the population level, diversity can also be generated at the individual level
382 through the acquisition of multiple different spacers within the CRISPR locus of an individual
383 bacterium, which make it much harder to overcome CRISPR immunity by point mutation than single
384 spacers. Recent studies reveal that type III systems have a distinct and unique mechanism to generate
385 individual-level diversity, explaining why escape from this CRISPR-Cas type is so rare, requiring
386 deletions in targeted phage sequences (Pyenson et al., 2017). Mechanistic studies have demonstrated
387 that type III immune complexes vary in size, with smaller complexes carrying crRNA of reduced size
388 (Hale et al., 2009; Hatoum-Aslan et al., 2011). The heterogeneity in the size of the crRNAs in type III
389 systems is due to a secondary maturation step at the 3' end of the crRNAs, resulting in crRNAs with a
390 variable 3' sequences. Crucially, a recent study showed that the type III-B system of *T. thermophilus*
391 carries a 3' seed region that is critical for target RNA cleavage (Steens et al., 2021). Where phages can
392 overcome CRISPR immunity of type I and II systems through a single point mutation in the seed
393 sequence (Deveau et al., 2008; Watson et al., 2019a), this is much harder in the case of this type III
394 system, due to its variable 3' end, which defines different seed sequences (Steens et al., 2021). By
395 employing this unique strategy, type III complexes are able to create within-host diversity with just a
396 single unique spacer (Figure 3) and are much more robust against rapidly evolving phages (Pyenson et
397 al., 2017). In addition to RNA cleavage, base pairing between the target RNA and the crRNA of type
398 III systems activates the Cas10 subunit of the immune complex, leading to the production of cOA
399 signaling molecules and sequence non-specific ssDNA cleavage activity (Kazlauskiene et al., 2017).

The activation of Cas10 is regulated by the Cas10 Activating Region (CAR) at the 5' end of the crRNA (Steens et al., 2021). Mutations in the target at the 5' end of the crRNA affect the production of cOA but do not affect the sequence-specific target RNA cleavage (Steens et al., 2021). The different effects of mutations in the 5' CAR or 3' seed sequence of the crRNA may help to explain why previous studies have suggested seed regions at both ends of the crRNA (Cao et al., 2016; Manica et al., 2013; Peng et al., 2015; Wang et al., 2019). Taken together, variable 3' processing and the resulting complex composition, which defines the location of the seed sequence, along with activation of Cas10 through the 5' end of the crRNA, creates significant challenges for phage to overcome type III CRISPR immunity by mutation.

Conclusions and Outlook

While we have made progress in understanding the fundamental concepts of CRISPR-phage coevolution using lab-based experiments, we are far from understanding the extent and mode of coevolution in natural environments. Metagenomics, comparative genomics and experimental studies in semi-natural environments, including biofilms in a mine drainage systems (Andersson and Banfield, 2008), fish farms (Laanto et al., 2017) and recolonised mouse gut (Cornuault et al., 2020), have provided evidence for the evolution of CRISPR immunity outside of the lab (reviewed in (Westra and Levin, 2020)). Moreover, these studies have identified CRISPR spacers that match co-sampled phage genomes and found that phages that persist over time often carry mutations that are predicted to enable escape from spacers found in earlier time points (Andersson and Banfield, 2008; Martinez Arbas et al., 2021; Sun et al., 2016; Weinberger et al., 2012a). These studies suggest that in natural environments, phages and bacteria with CRISPR-Cas immune systems can coexist and coevolve. It is likely that the differences seen between natural environments and lab evolution experiments exist due to various ecological and evolutionary factors that have not been captured in lab experiments. These likely include the microbial community context in which these interactions take place, the spatial structure of the environment, as well as the levels of phage diversity. While some observational, theoretical and experimental studies have started to explore how these and other factors impact CRISPR-phage coevolution (Paez-Espino et al., 2015; Weinberger et al., 2012a), the generality of these observations is not yet clear, and further work with a greater number of model systems and different CRISPR-Cas types is needed to address the many open questions that remain.

Acknowledgements

R.H.J.S. was supported by a VENI grant (016.Veni.171.047), E.R.W. acknowledges funding from the Natural Environment Research Council (Grant ref. NE/S001921/1 and NE/M018350/1), and the

434 European Research Council (ERC-STG-2016-714478 - EVOIMMECH). S.v.H. acknowledges funding
435 from the Biotechnology and Biological Sciences Research Council (BB/S017674/1 and
436 BB/R010781/10).

437

438 **References**

439 Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. (2010).
440 Transcription profile of Thermus thermophilus CRISPR systems after phage infection. J Mol Biol
441 *395*, 270-281.
442 Alseth, E.O., Pursey, E., Lujan, A.M., McLeod, I., Rollie, C., and Westra, E.R. (2019).
443 Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. Nature *574*, 549-552.
444 Andersson, A.F., and Banfield, J.F. (2008). Virus population dynamics and acquired virus
445 resistance in natural microbial communities. Science *320*, 1047-1050.
446 Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero,
447 D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes.
448 Science *315*, 1709-1712.
449 Bernheim, A., Bikard, D., Touchon, M., and Rocha, E.P.C. (2019). A matter of background:
450 DNA repair pathways as a possible cause for the sparse distribution of CRISPR-Cas systems in
451 bacteria. Philos Trans R Soc Lond B Biol Sci *374*, 20180088.
452 Bernheim, A., Calvo-Villamanan, A., Basier, C., Cui, L., Rocha, E.P.C., Touchon, M., and
453 Bikard, D. (2017). Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria. Nature
454 communications *8*, 2094.
455 Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L.A. (2012). CRISPR interference
456 can prevent natural transformation and virulence acquisition during in vivo bacterial infection. Cell
457 host & microbe *12*, 177-186.
458 Bradde, S., Nourmohammad, A., Goyal, S., and Balasubramanian, V. (2020). The size of the
459 immune repertoire of bacteria. Proc Natl Acad Sci U S A *117*, 5144-5151.
460 Bradde, S., Vucelja, M., Tesileanu, T., and Balasubramanian, V. (2017). Dynamics of
461 adaptive immunity against phage in bacterial populations. PLoS computational biology *13*, e1005486.
462 Broniewski, J.M., Meaden, S., Paterson, S., Buckling, A., and Westra, E.R. (2020). The effect
463 of phage genetic diversity on bacterial resistance evolution. The ISME journal *14*, 828-836.
464 Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P.,
465 Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs
466 guide antiviral defense in prokaryotes. Science *321*, 960-964.
467 Burstein, D., Harrington, L.B., Strutt, S.C., Probst, A.J., Anantharaman, K., Thomas, B.C.,
468 Doudna, J.A., and Banfield, J.F. (2017). New CRISPR-Cas systems from uncultivated microbes.
469 Nature *542*, 237-241.
470 Burstein, D., Sun, C.L., Brown, C.T., Sharon, I., Anantharaman, K., Probst, A.J., Thomas,
471 B.C., and Banfield, J.F. (2016). Major bacterial lineages are essentially devoid of CRISPR-Cas viral
472 defence systems. Nature communications *7*, 10613.
473 Cao, L., Gao, C.-H., Zhu, J., Zhao, L., Wu, Q., Li, M., and Sun, B. (2016). Identification and
474 functional study of type III-A CRISPR-Cas systems in clinical isolates of Staphylococcus aureus.
475 International Journal of Medical Microbiology *306*, 686-696.
476 Chabas, H., Lion, S., Nicot, A., Meaden, S., van Houte, S., Moineau, S., Wahl, L.M., Westra,
477 E.R., and Gandon, S. (2018). Evolutionary emergence of infectious diseases in heterogeneous host
478 populations. PLoS Biol *16*, e2006738.
479 Chabas, H., Nicot, A., Meaden, S., Westra, E.R., Tremblay, D.M., Pradier, L., Lion, S.,
480 Moineau, S., and Gandon, S. (2019). Variability in the durability of CRISPR-Cas immunity. Philos
481 Trans R Soc Lond B Biol Sci *374*, 20180097.

482 Chabas, H., van Houte, S., Hoyland-Kroghsbo, N.M., Buckling, A., and Westra, E.R. (2016).
483 Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. Proc
484 Biol Sci *283*.
485 Chen, J.S., Ma, E., Harrington, L.B., Da Costa, M., Tian, X., Palefsky, J.M., and Doudna, J.A.
486 (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity.
487 Science *360*, 436.
488 Chevallereau, A., Meaden, S., Fradet, O., Landsberger, M., Maestri, A., Biswas, A., Gandon,
489 S., van Houte, S., and Westra, E.R. (2020). Exploitation of the Cooperative Behaviors of Anti-
490 CRISPR Phages. Cell host & microbe *27*, 189-198 e186.
491 Chevallereau, A., Meaden, S., van Houte, S., Westra, E.R., and Rollie, C. (2019). The effect
492 of bacterial mutation rate on the evolution of CRISPR-Cas adaptive immunity. Philos Trans R Soc
493 Lond B Biol Sci *374*, 20180094.
494 Childs, L.M., England, W.E., Young, M.J., Weitz, J.S., and Whitaker, R.J. (2014). CRISPR-
495 induced distributed immunity in microbial populations. PloS one *9*, e101710.
496 Common, J., Morley, D., Westra, E.R., and van Houte, S. (2019). CRISPR-Cas immunity
497 leads to a coevolutionary arms race between Streptococcus thermophilus and lytic phage. Philos Trans
498 R Soc Lond B Biol Sci *374*, 20180098.
499 Cornuault, J.K., Moncaut, E., Loux, V., Mathieu, A., Sokol, H., Petit, M.A., and De Paepe,
500 M. (2020). The enemy from within: a prophage of Roseburia intestinalis systematically turns lytic in
501 the mouse gut, driving bacterial adaptation by CRISPR spacer acquisition. The ISME journal *14*, 771-
502 787.
503 Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E.
504 (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity
505 system. Nature communications *3*, 945.
506 Deng, L., Garrett, R.A., Shah, S.A., Peng, X., and She, Q. (2013). A novel interference
507 mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. Mol Microbiol *87*, 1088-1099.
508 Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero,
509 D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in
510 *Streptococcus thermophilus*. J Bacteriol *190*, 1390-1400.
511 Dimitriu, T., Szczelkun, M.D., and Westra, E.R. (2020). Evolutionary Ecology and Interplay
512 of Prokaryotic Innate and Adaptive Immune Systems. Current biology : CB *30*, R1189-R1202.
513 Elmore, J.R., Sheppard, N.F., Ramia, N., Deighan, T., Li, H., Terns, R.M., and Terns, M.P.
514 (2016). Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas
515 system. Genes Dev *30*, 447-459.
516 Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P.,
517 Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial
518 immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67-71.
519 Goldberg, G.W., Jiang, W., Bikard, D., and Marraffini, L.A. (2014). Conditional tolerance of
520 temperate phages via transcription-dependent CRISPR-Cas targeting. Nature *514*, 633-637.
521 Goldberg, G.W., McMillan, E.A., Varble, A., Modell, J.W., Samai, P., Jiang, W., and
522 Marraffini, L.A. (2018). Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces
523 the fitness of lysogenic hosts. Nature communications *9*, 61.
524 Gomaa, A.A., Klumpe, H.E., Luo, M.L., Selle, K., Barrangou, R., and Beisel, C.L. (2014).
525 Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. mBio *5*,
526 e00928-00913.
527 Gophna, U., Kristensen, D.M., Wolf, Y.I., Popa, O., Drevet, C., and Koonin, E.V. (2015). No
528 evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. The
529 ISME journal.
530 Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display
531 CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics *8*, 172.
532 Gurney, J., Pleska, M., and Levin, B.R. (2019). Why put up with immunity when there is
533 resistance: an excursion into the population and evolutionary dynamics of restriction-modification and
534 CRISPR-Cas. Philos Trans R Soc Lond B Biol Sci *374*, 20180096.

535 Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and
536 Terns, M.P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell *139*,
537 945-956.
538 Hampton, H.G., Watson, B.N.J., and Fineran, P.C. (2020). The arms race between bacteria
539 and their phage foes. Nature *577*, 327-336.
540 Han, P., and Deem, M.W. (2017). Non-classical phase diagram for virus bacterial coevolution
541 mediated by clustered regularly interspaced short palindromic repeats. J R Soc Interface *14*.
542 Han, P., Niestemski, L.R., Barrick, J.E., and Deem, M.W. (2013). Physical model of the
543 immune response of bacteria against bacteriophage through the adaptive CRISPR-Cas immune
544 system. Phys Biol *10*, 025004.
545 Hatoum-Aslan, A., Maniv, I., and Marraffini, L.A. (2011). Mature clustered, regularly
546 interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism
547 anchored at the precursor processing site. Proc Natl Acad Sci U S A *108*, 21218-21222.
548 Hatoum-Aslan, A., and Marraffini, L.A. (2014). Impact of CRISPR immunity on the
549 emergence and virulence of bacterial pathogens. Curr Opin Microbiol *17*, 82-90.
550 He, F., Vestergaard, G., Peng, W., She, Q., and Peng, X. (2017). CRISPR-Cas type I-A
551 Cascade complex couples viral infection surveillance to host transcriptional regulation in the
552 dependence of Csa3b. Nucleic acids research *45*, 1902-1913.
553 Heler, R., Wright, A.V., Vucelja, M., Doudna, J.A., and Marraffini, L.A. (2019). Spacer
554 Acquisition Rates Determine the Immunological Diversity of the Type II CRISPR-Cas Immune
555 Response. Cell host & microbe *25*, 242-249 e243.
556 Hille, F., Richter, H., Wong, S.P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The
557 Biology of CRISPR-Cas: Backward and Forward. Cell *172*, 1239-1259.
558 Hoyland-Kroghsbo, N.M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-
559 Denomy, J., and Bassler, B.L. (2016). Quorum sensing controls the Pseudomonas aeruginosa
560 CRISPR-Cas adaptive immune system. Proc Natl Acad Sci U S A.
561 Hynes, A.P., Villion, M., and Moineau, S. (2014). Adaptation in bacterial CRISPR-Cas
562 immunity can be driven by defective phages. Nature communications *5*, 4399.
563 Iranzo, J., Lobkovsky, A.E., Wolf, Y.I., and Koonin, E.V. (2013). Evolutionary Dynamics of
564 the Prokaryotic Adaptive Immunity System CRISPR-Cas in an Explicit Ecological Context. Journal
565 of Bacteriology *195*, 3834-3844.
566 Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B.R., and Marraffini, L.A. (2013). Dealing
567 with the Evolutionary Downside of CRISPR Immunity: Bacteria and Beneficial Plasmids. PLoS
568 Genet *9*, e1003844.
569 Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A
570 programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*,
571 816-821.
572 Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A
573 cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science *357*, 605.
574 Koskella, B., and Brockhurst, M.A. (2014). Bacteria-phage coevolution as a driver of
575 ecological and evolutionary processes in microbial communities. FEMS microbiology reviews *38*,
576 916-931.
577 Laanto, E., Hoikkala, V., Ravantti, J., and Sundberg, L.R. (2017). Long-term genomic
578 coevolution of host-parasite interaction in the natural environment. Nature communications *8*, 111.
579 Levin, B.R. (2010). Nasty viruses, costly plasmids, population dynamics, and the conditions
580 for establishing and maintaining CRISPR-mediated adaptive immunity in bacteria. PLoS Genet *6*,
581 e1001171.
582 Levin, B.R., Moineau, S., Bushman, M., and Barrangou, R. (2013). The population and
583 evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. PLoS Genet *9*,
584 e1003312.
585 Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U.,
586 and Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA.
587 Nature *520*, 505-510.
588 Li, Y., and Bondy-Denomy, J. (2021). Anti-CRISPRs go viral: the infection biology of
589 CRISPR-Cas inhibitors. Cell host & microbe.

590 Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J.,
591 Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H.*, et al.* (2015). An updated evolutionary
592 classification of CRISPR-Cas systems. Nature reviews Microbiology *13*, 722-736.
593 Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J.,
594 Charpentier, E., Cheng, D., Haft, D.H., Horvath, P.*, et al.* (2020). Evolutionary classification of
595 CRISPR-Cas systems: a burst of class 2 and derived variants. Nature reviews Microbiology *18*, 67-83.
596 Malone, L.M., Birkholz, N., and Fineran, P.C. (2020a). Conquering CRISPR: how phages
597 overcome bacterial adaptive immunity. Curr Opin Biotechnol *68*, 30-36.
598 Malone, L.M., Warring, S.L., Jackson, S.A., Warnecke, C., Gardner, P.P., Gumy, L.F., and
599 Fineran, P.C. (2020b). A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA
600 targeting but is vulnerable to type III RNA-based immunity. Nat Microbiol *5*, 48-55.
601 Manica, A., Zebec, Z., Steinkellner, J., and Schleper, C. (2013). Unexpectedly broad target
602 recognition of the CRISPR-mediated virus defence system in the archaeon Sulfolobus solfataricus.
603 Nucleic acids research *41*, 10509-10517.
604 Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene
605 transfer in staphylococci by targeting DNA. Science *322*, 1843-1845.
606 Martinez Arbas, S., Narayanasamy, S., Herold, M., Lebrun, L.A., Hoopmann, M.R., Li, S.,
607 Lam, T.J., Kunath, B.J., Hicks, N.D., Liu, C.M.*, et al.* (2021). Roles of bacteriophages, plasmids and
608 CRISPR immunity in microbial community dynamics revealed using time-series integrated meta-
609 omics. Nat Microbiol *6*, 123-135.
610 Martynov, A., Severinov, K., and Ispolatov, I. (2017). Optimal number of spacers in CRISPR
611 arrays. PLoS computational biology *13*, e1005891.
612 McGinn, J., and Marraffini, L.A. (2016). CRISPR-Cas systems optimize their immune
613 response by specifying the site of spacer integration. Molecular cell *64*, 616-623.
614 Meaden, S., Capria, L., Alseth, E., Gandon, S., Biswas, A., Lenzi, L., van Houte, S., and
615 Westra, E.R. (2020). Phage gene expression and host responses lead to infection-dependent costs of
616 CRISPR immunity. The ISME journal.
617 Meeske, A.J., Nakandakari-Higa, S., and Marraffini, L.A. (2019). Cas13-induced cellular
618 dormancy prevents the rise of CRISPR-resistant bacteriophage. Nature.
619 Mendoza, S.D., Nieweglowska, E.S., Govindarajan, S., Leon, L.M., Berry, J.D., Tiwari, A.,
620 Chaikeeratisak, V., Pogliano, J., Agard, D.A., and Bondy-Denomy, J. (2020). A bacteriophage
621 nucleus-like compartment shields DNA from CRISPR nucleases. Nature *577*, 244-248.
622 Modell, J.W., Jiang, W., and Marraffini, L.A. (2017). CRISPR-Cas systems exploit viral
623 DNA injection to establish and maintain adaptive immunity. Nature *544*, 101-104.
624 Morley, D., Broniewski, J.M., Westra, E.R., Buckling, A., and van Houte, S. (2016). Host
625 diversity limits the evolution of parasite local adaptation. Mol Ecol.
626 Niewoehner, O., Garcia-Doval, C., Rostol, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J.,
627 Marraffini, L.A., and Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate
628 second messengers. Nature *548*, 543-548.
629 Nobrega, F.L., Walinga, H., Dutilh, B.E., and Brouns, S.J.J. (2020). Prophages are associated
630 with extensive CRISPR-Cas auto-immunity. Nucleic acids research *48*, 12074-12084.
631 Nussenzweig, P.M., McGinn, J., and Marraffini, L.A. (2019). Cas9 Cleavage of Viral
632 Genomes Primes the Acquisition of New Immunological Memories. Cell host & microbe *26*, 515-526
633 e516.
634 Paez-Espino, D., Morovic, W., Sun, C.L., Thomas, B.C., Ueda, K., Stahl, B., Barrangou, R.,
635 and Banfield, J.F. (2013). Strong bias in the bacterial CRISPR elements that confer immunity to
636 phage. Nature communications *4*, 1430.
637 Paez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B.C., Barrangou, R., and
638 Banfield, J.F. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus*
639 *thermophilus*. mBio *6*, e00262-00215.
640 Palmer, K.L., and Gilmore, M.S. (2010). Multidrug-resistant enterococci lack CRISPR-*cas*.
641 mBio *1*, e00227-00210.
642 Patterson, A.G., Chang, J.T., Taylor, C., and Fineran, P.C. (2015). Regulation of the Type I-F
643 CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. Nucleic
644 acids research *43*, 6038-6048.

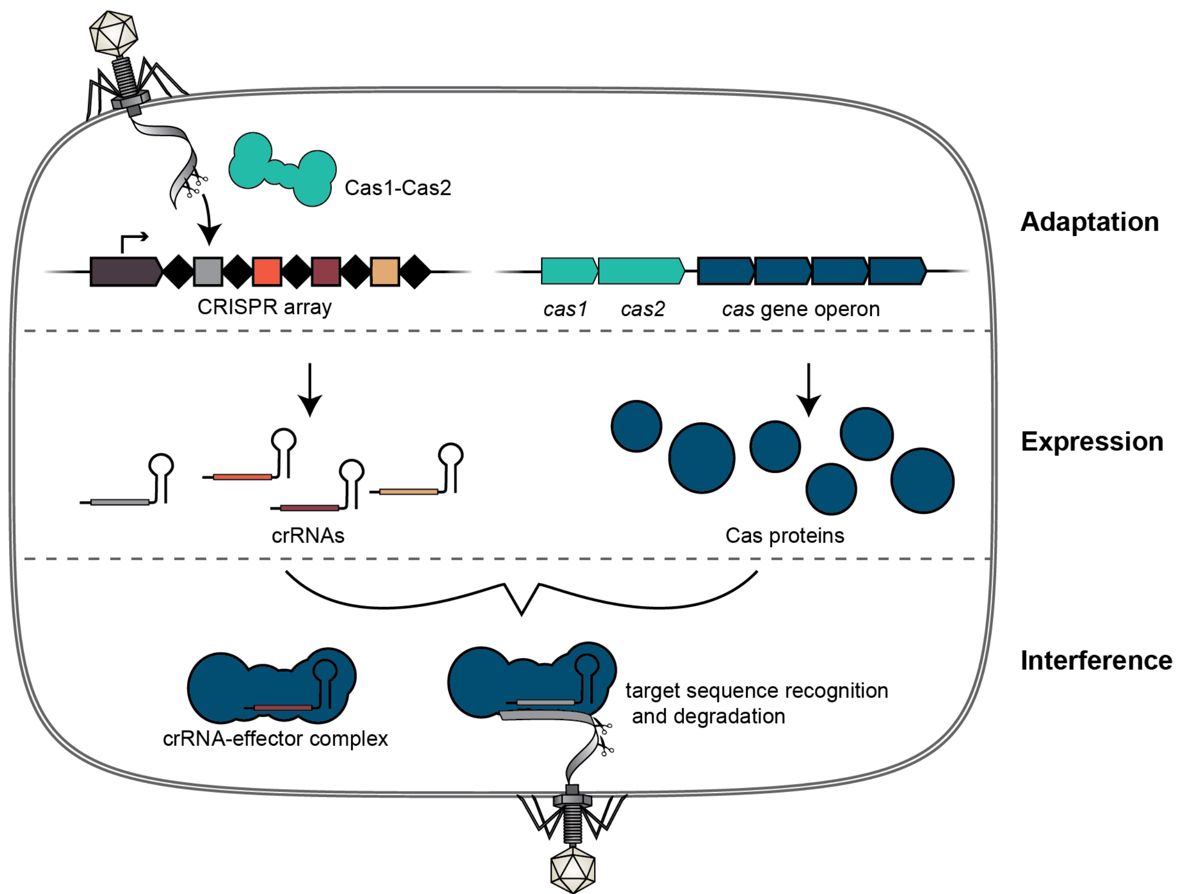
645 Patterson, Adrian G., Jackson, Simon A., Taylor, C., Evans, Gary B., Salmond, George P.C.,
646 Przybilski, R., Staals, Raymond H.J., and Fineran, Peter C. (2016). Quorum Sensing Controls
647 Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. Molecular cell.
648 Patterson, A.G., Yevstigneyeva, M.S., and Fineran, P.C. (2017). Regulation of CRISPR-Cas
649 adaptive immune systems. Curr Opin Microbiol *37*, 1-7.
650 Peng, W., Feng, M., Feng, X., Liang, Y.X., and She, Q. (2015). An archaeal CRISPR type III-
651 B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA
652 interference. Nucleic acids research *43*, 406-417.
653 Pinilla-Redondo, R., Mayo-Munoz, D., Russel, J., Garrett, R.A., Randau, L., Sorensen, S.J.,
654 and Shah, S.A. (2020). Type IV CRISPR-Cas systems are highly diverse and involved in competition
655 between plasmids. Nucleic acids research *48*, 2000-2012.
656 Pyenson, N.C., Gayvert, K., Varble, A., Elemento, O., and Marraffini, L.A. (2017). Broad
657 targeting specificity during bacterial type III CRISPR-Cas immunity constrains viral escape. Cell host
658 & microbe *22*, 343-353 e343.
659 Pyenson, N.C., and Marraffini, L.A. (2020). Co-evolution within structured bacterial
660 communities results in multiple expansion of CRISPR loci and enhanced immunity. Elife *9*.
661 Quax, T.E., Voet, M., Sismeiro, O., Dillies, M.A., Jagla, B., Coppee, J.Y., Sezonov, G.,
662 Forterre, P., van der Oost, J., Lavigne, R*., et al.* (2013). Massive activation of archaeal defense genes
663 during viral infection. Journal of virology *87*, 8419-8428.
664 Ratner, H.K., Sampson, T.R., and Weiss, D.S. (2015). I can see CRISPR now, even when
665 phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Current
666 opinion in infectious diseases *28*, 267-274.
667 Rollie, C., Chevallereau, A., Watson, B.N.J., Chyou, T.Y., Fradet, O., McLeod, I., Fineran,
668 P.C., Brown, C.M., Gandon, S., and Westra, E.R. (2020). Targeting of temperate phages drives loss of
669 type I CRISPR-Cas systems. Nature *578*, 149-153.
670 Samai, P., Pyenson, N.C., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini,
671 L.A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity.
672 Cell *161*, 1164-1174.
673 Semenova, E.V., Jore, M.M., Westra, E.R., Oost, v.d.J., and Brouns, S.J.J. (2011).
674 Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed
675 by a seed sequence. Proc Natl Acad Sci U S A *108*, 10098-18424.
676 Shivram, H., Cress, B.F., Knott, G.J., and Doudna, J.A. (2021). Controlling and enhancing
677 CRISPR systems. Nat Chem Biol *17*, 10-19.
678 Smith, L.M., Jackson, S.A., Malone, L.M., Ussher, J.E., Gardner, P.P., and Fineran, P.C.
679 (2021). The Rcs stress response inversely controls surface and CRISPR-Cas adaptive immunity to
680 discriminate plasmids and phages. Nat Microbiol.
681 Staals, R.H., Jackson, S.A., Biswas, A., Brouns, S.J., Brown, C.M., and Fineran, P.C. (2016).
682 Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native
683 CRISPR-Cas system. Nature communications *7*, 12853.
684 Staals, Raymond H.J., Zhu, Y., Taylor, David W., Kornfeld, Jack E., Sharma, K., Barendregt,
685 A., Koehorst, Jasper J., Vlot, M., Neupane, N., Varossieau, K*., et al.* (2014). RNA Targeting by the
686 Type III-A CRISPR-Cas Csm Complex of *Thermus thermophilus*. Molecular cell *56*, 518-530.
687 Steens, J.A., Zhu, Y., Taylor, D.W., Bravo, J.P.K., Prinsen, S.H.P., Schoen, C.D., Keijser,
688 B.J.F., Ossendrijver, M., Hofstra, L.M., Brouns, S.J.J*., et al.* (2021). SCOPE: Flexible targeting and
689 stringent CARF activation enables type III CRISPR-Cas diagnostics. bioRxiv,
690 2021.2002.2001.429135.
691 Stern, A., Keren, L., Wurtzel, O., Amitai, G., and Sorek, R. (2010). Self-targeting by
692 CRISPR: gene regulation or autoimmunity? Trends in genetics : TIG *26*, 335-340.
693 Sun, C.L., Thomas, B.C., Barrangou, R., and Banfield, J.F. (2016). Metagenomic
694 reconstructions of bacterial CRISPR loci constrain population histories. The ISME journal *10*, 858-
695 870.
696 Suttle, C.A. (2007). Marine viruses--major players in the global ecosystem. Nature reviews
697 Microbiology *5*, 801-812.

698 Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, C., Nwokeoji, A.O., Dickman,
699 M.J., Horvath, P., and Siksnys, V. (2014). Programmable RNA shredding by the type III-A CRISPR-
700 Cas system of Streptococcus thermophilus. Molecular cell *56*, 506-517.
701 Vale, P.F., Lafforgue, G., Gatchitch, F., Gardan, R., Moineau, S., and Gandon, S. (2015).
702 Costs of CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. Proc Biol Sci *282*,
703 20151270.
704 van Houte, S., Buckling, A., and Westra, E.R. (2016a). Evolutionary Ecology of Prokaryotic
705 Immune Mechanisms. Microbiology and molecular biology reviews : MMBR *80*, 745-763.
706 van Houte, S., Ekroth, A.K., Broniewski, J.M., Chabas, H., Ashby, B., Bondy-Denomy, J.,
707 Gandon, S., Boots, M., Paterson, S., Buckling, A.*, et al.* (2016b). The diversity-generating benefits of
708 a prokaryotic adaptive immune system. Nature *532*, 385-388.
709 Varble, A., and Marraffini, L.A. (2019). Three New Cs for CRISPR: Collateral,
710 Communicate, Cooperate. Trends in genetics : TIG *35*, 446-456.
711 Vercoe, R.B., Chang, J.T., Dy, R.L., Taylor, C., Gristwood, T., Clulow, J.S., Richter, C.,
712 Przybilski, R., Pitman, A.R., and Fineran, P.C. (2013). Cytotoxic chromosomal targeting by
713 CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands.
714 PLoS Genet *9*, e1003454.
715 Vlot, M., Houkes, J., Lochs, S.J.A., Swarts, D.C., Zheng, P., Kunne, T., Mohanraju, P.,
716 Anders, C., Jinek, M., van der Oost, J.*, et al.* (2018). Bacteriophage DNA glucosylation impairs target
717 DNA binding by type I and II but not by type V CRISPR-Cas effector complexes. Nucleic acids
718 research *46*, 873-885.
719 Wang, L., Mo, C.Y., Wasserman, M.R., Rostol, J.T., Marraffini, L.A., and Liu, S. (2019).
720 Dynamics of Cas10 Govern Discrimination between Self and Non-self in Type III CRISPR-Cas
721 Immunity. Molecular cell *73*, 278-290 e274.
722 Watson, B.N.J., Easingwood, R.A., Tong, B., Wolf, M., Salmond, G.P.C., Staals, R.H.J.,
723 Bostina, M., and Fineran, P.C. (2019a). Different genetic and morphological outcomes for phages
724 targeted by single or multiple CRISPR-Cas spacers. Philos Trans R Soc Lond B Biol Sci *374*,
725 20180090.
726 Watson, B.N.J., Staals, R.H.J., and Fineran, P.C. (2018). CRISPR-Cas-mediated phage
727 resistance enhances horizontal gene transfer by transduction. mBio *9*, e02406-02417.
728 Watson, B.N.J., Vercoe, R.B., Salmond, G.P.C., Westra, E.R., Staals, R.H.J., and Fineran,
729 P.C. (2019b). Type I-F CRISPR-Cas resistance against virulent phages results in abortive infection
730 and provides population-level immunity. Nature communications *10*, 5526.
731 Weinberger, A.D., Sun, C.L., Plucinski, M.M., Denef, V.J., Thomas, B.C., Horvath, P.,
732 Barrangou, R., Gilmore, M.S., Getz, W.M., and Banfield, J.F. (2012a). Persisting viral sequences
733 shape microbial CRISPR-based immunity. PLoS computational biology *8*, e1002475.
734 Weinberger, A.D., Wolf, Y.I., Lobkovsky, A.E., Gilmore, M.S., and Koonin, E.V. (2012b).
735 Viral diversity threshold for adaptive immunity in prokaryotes. mBio *3*, e00456-00412.
736 Weissman, J.L., Fagan, W.F., and Johnson, P.L.F. (2018a). Selective Maintenance of
737 Multiple CRISPR Arrays Across Prokaryotes. CRISPR J *1*, 405-413.
738 Weissman, J.L., Holmes, R., Barrangou, R., Moineau, S., Fagan, W.F., Levin, B., and
739 Johnson, P.L.F. (2018b). Immune loss as a driver of coexistence during host-phage coevolution. The
740 ISME journal *12*, 585-597.
741 Weissman, J.L., Laljani, R.M.R., Fagan, W.F., and Johnson, P.L.F. (2019). Visualization and
742 prediction of CRISPR incidence in microbial trait-space to identify drivers of antiviral immune
743 strategy. The ISME journal *13*, 2589-2602.
744 Weissman, J.L., Stoltzfus, A., Westra, E.R., and Johnson, P.L.F. (2020). Avoidance of Self
745 during CRISPR Immunization. Trends Microbiol *28*, 543-553.
746 Westra, E.R., and Levin, B.R. (2020). It is unclear how important CRISPR-Cas systems are
747 for protecting natural populations of bacteria against infections by mobile genetic elements. Proc Natl
748 Acad Sci U S A *117*, 27777-27785.
749 Westra, E.R., van Erp, P.B., Kunne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S.,
750 Jore, M.M., Semenova, E., Severinov, K.*, et al.* (2012). CRISPR immunity relies on the consecutive
751 binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Molecular cell
752 *46*, 595-605.

753 Westra, E.R., van Houte, S., Oyesiku-Blakemore, S., Makin, B., Broniewski, J.M., Best, A.,
754 Bondy-Denomy, J., Davidson, A., Boots, M., and Buckling, A. (2015). Parasite exposure drives
755 selective evolution of constitutive versus inducible defense. Current biology : CB *25*, 1043-1049.
756 Wheatley, R.M., and MacLean, R.C. (2020). CRISPR-Cas systems restrict horizontal gene
757 transfer in Pseudomonas aeruginosa. The ISME journal.
758 Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A.,
759 Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J.*, et al.* (2011). RNA-guided complex from a
760 bacterial immune system enhances target recognition through seed sequence interactions. Proceedings
761 of the National Academy of Sciences *108*, 10092-10097.
762 Wimmer, F., and Beisel, C.L. (2019). CRISPR-Cas Systems and the Paradox of Self-
763 Targeting Spacers. Front Microbiol *10*, 3078.
764 Yang, Y., Lyu, T., Zhou, R., He, X., Ye, K., Xie, Q., Zhu, L., Chen, T., Shen, C., Wu, Q.*, et*
765 *al.* (2019). The Antiviral and Antitumor Effects of Defective Interfering Particles/Genomes and Their
766 Mechanisms. Front Microbiol *10*, 1852.
767 Zetsche, B., Gootenberg, J S., Abudayyeh, O O., Slaymaker, I M., Makarova, K S.,
768 Essletzbichler, P., Volz, S E., Joung, J., van der Oost, J., Regev, A.*, et al.* (2015). Cpf1 Is a Single
769 RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell *163*, 759-771.
770 Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Graham, S., Reimann, J., Cannone,
771 G., Liu, H., Albers, S.V.*, et al.* (2012). Structure and mechanism of the CMR complex for CRISPR-
772 mediated antiviral immunity. Molecular cell *45*, 303-313.
773 Zoephel, J., and Randau, L. (2013). RNA-Seq analyses reveal CRISPR RNA processing and
774 regulation patterns. Biochem Soc Trans *41*, 1459-1463.

775

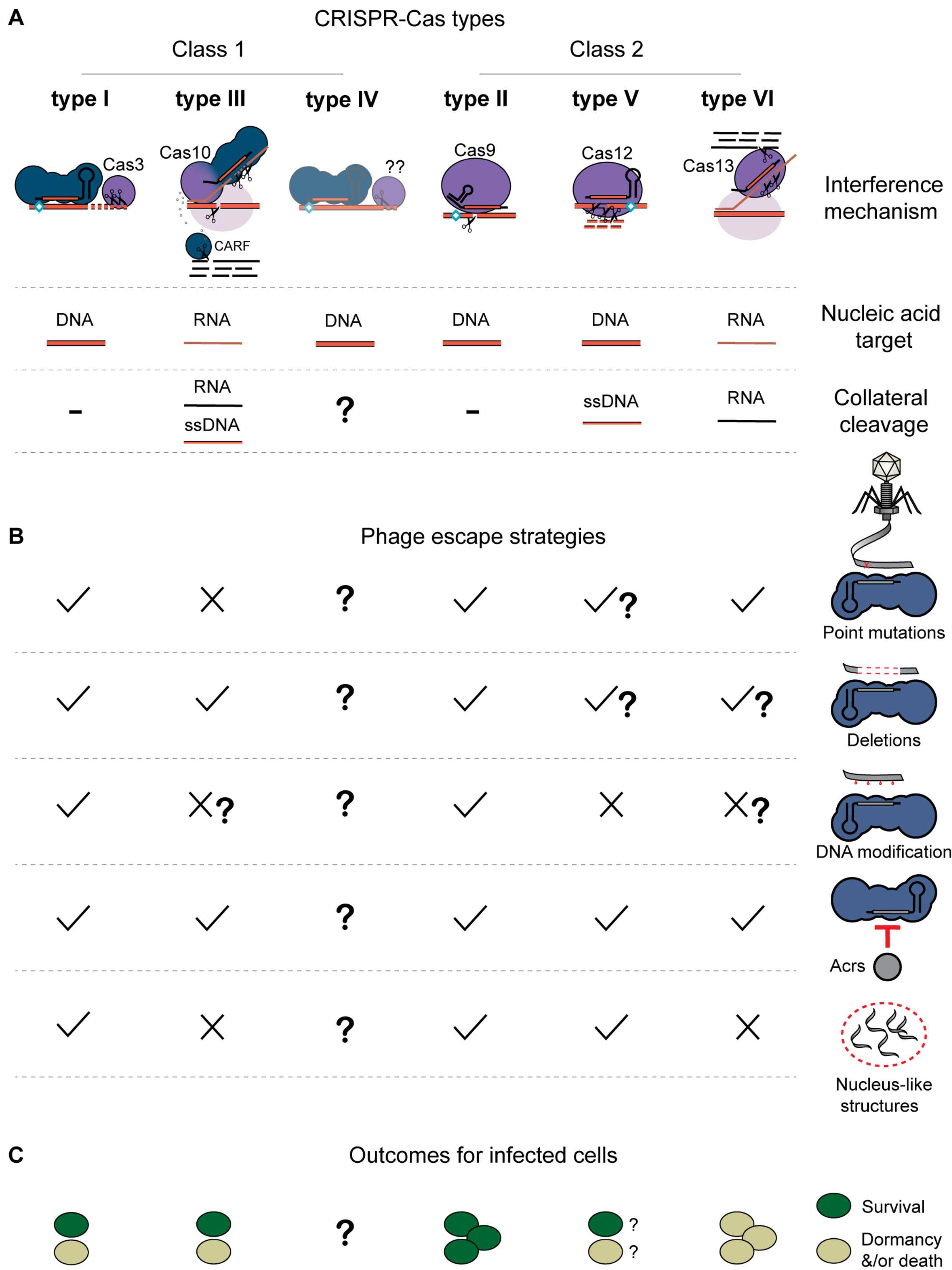
776



777

778 Figure 1 Overview of CRISPR-Cas immunity. **Adaptation**: Cas1 and Cas2 capture a fragment of
779 invading genetic material and incorporate it into the CRISPR array as a spacer (squares), between repeat
780 sequences (black diamonds). **Expression**: the CRISPR array is transcribed and processed into smaller
781 CRISPR RNAs (crRNAs) and the Cas effector protein(s) are produced. **Interference**: the crRNAs
782 associate with the effector protein(s) and any sequence detected that is complementary to the crRNA is
783 degraded.

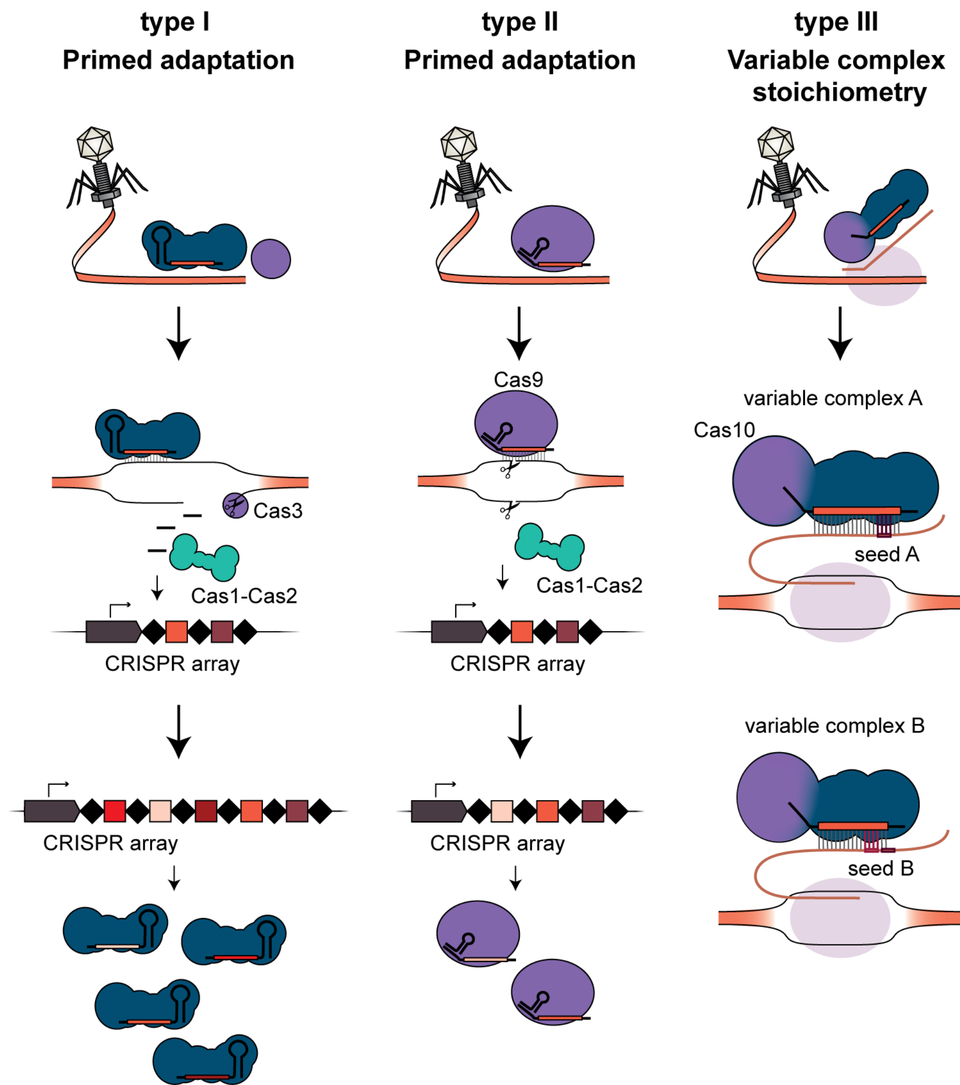
784



785

786 Figure 2 Summary of CRISPR-Cas types, phage escape mechanisms and the outcomes for infected
787 cells. **A** The components involved in interference, the nucleic acids that are targeted and affected by
788 collateral cleavage are shown for the different CRISPR-Cas types. The labelled, purple proteins are the
789 signature proteins and effector components for each type. The light-blue diamonds represent the
790 requirement for a PAM sequence. Grey circles (types III and VI) represent replication bubbles. **B** Phages
791 can escape targeting through different strategies but these vary in their efficacy against different
792 CRISPR-Cas types. Strategies involving alterations to the phage DNA to prevent crRNA recognition
793 include point mutations in the PAM or seed sequence of the protospacer, deletions of the protospacer

794 and protospacer DNA modifications. Anti-CRISPR (Acr) proteins can inactivate CRISPR-Cas immune
795 proteins. A nucleus-like structure in the bacterial cell that occludes CRISPR-Cas immune proteins is
796 produced by some phages, including jumbo phages. ✓ = phage escape, ✕ = CRISPR-Cas is still
797 effective, (? indicates the predicted outcome where no evidence is available). **C** Type I and II systems
798 promote survival of infected cells by cleaving invading DNA (green circle) but in some type I systems,
799 slow phage clearance can result in cell death (light-grey circle) Type III systems can result in survival
800 of infected cells but activation of non-specific RNA cleavage through cOA signalling induces cell
801 dormancy. Little is known about the outcome of type V immunity, but ssDNA degradation may induce
802 dormancy or cell death. The collateral damage induced by type VI systems through RNA degradation
803 induces dormancy. It is not yet clear whether type IV systems provide phage resistance.



804

805 Figure 3 Diversity-generating mechanisms for the different CRISPR-Cas types. **Type I** systems
806 generate population-level spacer diversity through primed adaptation when exposed to phages with
807 perfect sequences or mutations that prevent interference. The effector complex (dark blue) will
808 recognise and bind the target sequence. Cas3 (purple circle) is then recruited and will degrade the target
809 sequence, generating substrates for spacer acquisition into the CRISPR array (squares represent spacers,
810 black diamonds represent repeats) by Cas1-Cas2 (green). **Type II** systems can generate low levels of
811 population-level spacer diversity, including through primed adaptation with spacers that perfectly target
812 phages. Cleavage of the target DNA by Cas9 (purple circle) generates substrates for spacer acquisition
813 by Cas1-Cas2 (green) **Type III** effector complexes vary in their size due to variability in the number of
814 Cas7-Cas11 (Cmr4-Cmr5) backbone segments. The variation in the length of the corresponding crRNA
815 creates seed sequence flexibility, which then gives rise to individual-level diversity in CRISPR-Cas
816 immune complexes. Type III effector complexes bind to target RNA (grey circles represent replication
817 bubbles). In variable complex A, seed sequence A (purple lines) represents sequence that must have
818 perfectly complementary for targeting to occur. Variable complex B represents a smaller complex, with
819 fewer backbone segments. Hence, seed B (pink lines) is in a different position to seed A.

820

821