Evolutionary causes and consequences of diversified CRISPR immune profiles in natural populations

Whitney E. England* **and Rachel J. Whitaker*[1]**

*Department of Microbiology, University of Illinois at Urbana-Champaign, 601 S. Goodwin Avenue, Urbana, IL 61801, U.S.A.

Abstract

Host–pathogen co-evolution is a significant force which shapes the ecology and evolution of all types of organisms, and such interactions are driven by resistance and immunity mechanisms of the host. Diversity of resistance and immunity can affect the co-evolutionary trajectory of both host and pathogen. The microbial CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR-associated) system is one host immunity mechanism which offers a tractable model for examining the dynamics of diversity in an immune system. In the present article, we review CRISPR variation observed in a variety of natural populations, examine the forces which can push CRISPRs towards high or low diversity, and investigate the consequences of various levels of diversity on microbial populations.

Background

Host–pathogen interactions are important in the ecology and evolution of all organisms. The traits that underlie these interactions are often under strong reciprocal selection between host and pathogen, resulting in co-evolution. One approach to studying host–pathogen co-evolution has been to focus on host defence traits such as immunity and resistance in model host systems [1–3]. Observing the evolution of these traits as they change over time and space can provide great insight into both past and future co-evolutionary dynamics [4–7]. Unfortunately, in many cases, it is difficult to measure the diversity and specificity of host resistance and immunity to a broad range of pathogens within natural systems without introducing experimental and culture biases. The repeat–spacer arrays of the CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR-associated) immune systems of bacteria and archaea provide a direct sequence-specific method to investigate polymorphism in host immunity in natural microbial populations because sequence variation in CRISPR arrays has a direct relationship with host immune profiles against specific pathogens [8,9]. We review differences in the structure of diversity of the CRISPR system that have been observed in natural populations and the evolutionary causes and consequences of different levels of polymorphism in the CRISPR system in natural microbial populations.

Key words: clustered regularly interspaced short palindromic repeats (CRISPR), co-evolution, diversity, immunity.

Abbreviations used: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; spoligotyping, spacer oligonucleotide typing.

[1]To whom correspondence should be addressed (email rwhitaker@life.illinois.edu).

CRISPR immune profiles in natural microbial populations

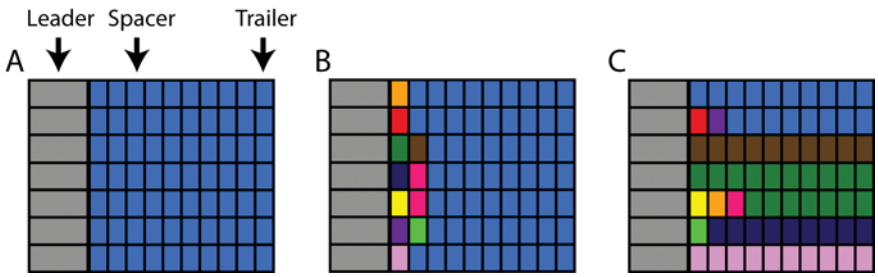
Polymorphism and rapid evolution of spacers between direct repeats within bacterial and archaeal genomes was observed long before the mechanism and function of CRISPR in adaptive immunity were recognized [10,11]. Polymorphism was assessed by looking at deletions of spacers between repeat sequences whose function in the cell was unknown [12]. For example, fingerprints of these repeated regions were used for typing of mycobacterial strains called spoligotyping (spacer oligonucleotide typing) [13] and to infer the structure of these pathogen populations [14]. Differences in diversity among populations were observed; for example, some mycobacterial outbreaks have a coexisting diverse set of strains (polyclonal) and others have identical alleles (monoclonal) [14].

The discovery that these repeat arrays within populations were specifically related to bacterial and archaeal immunity [15] transformed the significance of spoligotyping by directly relating diversity of repeat loci to variation in host immunity. Experimental work demonstrated that specific CRISPR spacer sequences in bacteria and archaea confer immunity to infection by viruses when their sequences match [15–17]. Even more astounding, the addition of new spacer sequences at the leader end is known to record a history of interactions from the most recent at the leader end to oldest at the trailer end of the repeat–spacer array [15,18] (Figure 1). Viruses have been shown to evade recognition by CRISPR immunity when they randomly acquire mutations in the position recognized by the CRISPR machinery [protospacer and/or PAM (protospacer-associated motif)] [19].

Using PCR to amplify, sequence and assemble CRISPR repeat–spacer alleles from isolated individuals or directly from environmental samples, several studies have investigated the diversity of CRISPRs within microbial populations. These studies have revealed a spectrum of population

Figure 1 | Schematic diagram of different populations of CRISPR alleles

CRISPR spacers are aligned relative to the leader sequences. In aligned vertical columns, different colours represent different CRISPR spacer sequences. Matching colours within horizontal rows do not signify identical spacers. (**A**) A monoclonal population where leader- and trailer-end alleles are identical. (**B**) A polyclonal population at the leader end with trailer-end clonality. (**C**) A polyclonal population where both leader- and trailer-end alleles have unique spacer sequences.



structures ranging from monoclonal to highly polyclonal. *Escherichia coli* isolates exhibit minimal variation; no new spacers are observed in strains which have diverged in the last 250 000 years, and most observed variation appears to be the result of spacer loss [20,21]. Similarly, *Salmonella* CRISPR loci show variation primarily due to spacer deletion rather than acquisition of new spacers and appear to be highly monoclonal [22,23]. Increasing in diversity, nearly clonal populations of *Leptospirillum* in acid mine drainage show extensive diversity at the leader end of the CRISPR array, but identity at the trailer end [17]. Populations of *Yersinia pestis*, known for extremely low sequence diversity at other genomic loci [24], show nearly clonal trailer-end spacers with leader-end variation [25,26]. At the other extreme, *Streptococcus thermophilus* exhibits hypervariability in its CRISPR loci; diversity is concentrated at the leader end, but multiple trailer types also exist [27]. Completely distinct CRISPR repeat spacer alleles (at both the leader and the trailer end) have been shown to coexist within a single population of the archaeon *Sulfolobus islandicus* [28]. A total of 39 isolates of *Sulfolobus islandicus* from a single hot spring sample collected in the year 2000 maintain extensive leader-end diversity, but also contain eight to ten completely different trailer-end alleles at each of three loci at relatively even abundance [28]. A later study involving 120 *S. islandicus* strains taken from the same population 10 years later found that leader- and trailer-end diversity persisted through time [29]. Why do some microbial populations exhibit extensive CRISPR immune diversity, whereas others do not? What does this difference in diversity between populations tell us about the ongoing co-evolutionary dynamic in these populations? What effect does diversity in CRISPR immunity have on evolution of host and pathogen populations?

Forces that lead to monoclonality in CRISPR immunity

On the basis of simple Lotka–Volterra dynamics [30], co-evolutionary models predict that at any one time and place, a population of hosts would have a single dominant CRISPR allele. Ongoing arms races between viruses and microbial hosts lead to periodic oscillations in immune host genotypes and subsequent selection of viral evasion mutants that can subvert the CRISPR immune surveillance [31]. The prediction is that each oscillation is driven by a selective sweep of an effective immune allele to fixation within a population. If these dynamics are actively ongoing, this model for diversity predicts that different populations of the same archaea or bacteria would have different monoclonal immune alleles, as they could be at different points in their co-evolutionary trajectories. Therefore if population structure is not well understood, this same evolutionary dynamic could result in apparent diversity of a particular allele when compared among populations that are in fact isolated.

Several other evolutionary, but not co-evolutionary, forces could result in monoclonality at the CRISPR locus. Other defence loci such as surface resistance could dominate the host–pathogen dynamic, resulting in clonal CRISPR alleles as the linked resistance locus is swept to fixation within the population. Clonality at a CRISPR locus could also result from demographic history such as bottleneck that would reduce diversity to a single individual allele by chance. Since new CRISPR spacers are added to the leader end, monoclonality at the leader end of the locus has been suggested to result from loss of function of the acquisition machinery in a particular system or the CRISPR system in general within a population [20].

Forces that lead to polyclonality in CRISPR immunity

Three basic theoretical models of host–pathogen co-evolution predict that polymorphism can be generated and maintained within populations [32]. In the first, explicit trade-offs between resistance or immunity and fitness are required for diversity to emerge [33]. In experimental evolutionary models, such trade-offs have been observed to promote the coexistence of multiple host or viral genotypes over time [33–38]. Associating a cost with CRISPR

immunity, these models predict that the maintenance of a CRISPR system is only adaptive if viral diversity is limited. A second model to promote diversity is negative frequency-dependent selection, where the adaptive benefit of a novel allele decreases as it increases in frequency within a population [39–42]. In this model, diversity is maintained in microbial populations through negative frequency dependence without explicit trade-offs between resistance alleles. Finally, spatial structure within populations has been shown to promote diversity. Spatial models have been applied to simulations on the basis of CRISPR immunity to predict that they are essential to maintain diversity [43,44].

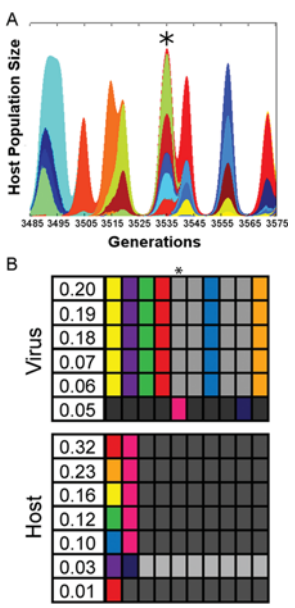
In applying these co-evolutionary models to explain the CRISPR diversity, it is important to consider two crucial elements of the CRISPR system that distinguish it in mechanism and model from previously described systems for host–pathogen co-evolution. First, the addition of new spacers to the CRISPR system is 'Lamarkian' in that new genotypes are created upon a viral encounter and can be passed on to the next generation [45]. Although the frequency at which new spacers are acquired by hosts has not been well defined, it is believed to be higher than the genome mutation rate, leading to the potential for competing mutations to exist within a population at one time [28]. Secondly, host cells are not subject to a large fitness drag as a consequence of investment in new immune phenotypes. The number of potential immune phenotypes for a population of infected cells is limited only by the number of protospacer sites within each virus, and each is likely to have equal fitness consequences to the host cell.

Integrating this mechanistic knowledge of CRISPR immunity, several new co-evolutionary models have been proposed (reviewed in [46]). Many of these models predict CRISPR immune profiles where there is leader-end diversity, but trailer-end clonality. Some models predict that, owing to the rapid acquisition of new spacers, neutral variation persists at the leader end until selection for a particular spacer causes a selective sweep [47]. An additional model suggests that, since each different spacer confers equal immunity to a given virus, diversity is maintained within a population because each distinct genotype has the same immune phenotype [48]. We have demonstrated that this phenomenon emerges within simulated populations (Figure 2) and quantified the extent to which this mechanism promotes polymorphism over time. As shown in Figure 2, host density peaks exhibit both clonal and polyclonal peaks within a single simulation. We have demonstrated that varying host and virus biological parameters such as rates of viral mutation, the number of potential protospacers in the virus genome, the number of protospacers in the host genome and the rate of acquisition of new spacers by the host will alter the relative frequency of these two evolutionary dynamics within a simulation. For the parameters tested to date most often result in trailer-end clonality even in simulations where distributed immunity maintains leader-end diversity. Additionally, we have shown that the impact of this type of 'distributed immunity' results in stable and increased host populations and unstable viral populations since each escape mutation has very little advantage to each viral mutant (L.M. Childs, W.E. England, M.J. Young, J.S. Weitz and R.J. Whitaker, unpublished work). Although not explicitly investigated as the basis through which diversity is maintained, similar dynamics have been observed in other mathematical models [49]. Together these models suggest that biological parameters for both the host and virus may result in differences in population structures between different host–pathogen pairs, where some stably maintain a diversity over time [49].

One additional biological factor that has attracted less attention in terms of its impact on maintaining diversity in natural populations has been the effect of reassortment and recombination of CRISPR loci among individuals within a population. This effect is not dependent upon the CRISPR mechanism of mutation and action, but has been shown to be important for any traits that are under strong selection

Figure 2 | Distributed immunity in a simulated microbial population

(**A**) Host population dynamics over time in simulated populations as in [48]. Each colour represents a different CRISPR genotype; colours can repeat when not touching. The host and viral populations at generation 3535 (*) are detailed in (**B**). (**B**) Host population (lower panel) contains multiple unique spacers matching a large portion of the viral population (upper panel). Each row represents a host or virus strain; the first column indicates that strain's proportion of the population. Each column is a spacer or protospacer position. Coloured boxes indicate spacer–protospacer matches; grey boxes represent spacers or protospacers with no match. Closely related viral strains making up the majority of the population are differentiated by polymorphism at the non-matching protospacer position marked with an asterisk. Hosts making up <1% and viruses making up <5% of the population are omitted for clarity.



such as those involved in resistance and immunity. The Red Queen hypothesis [50] states that antagonistically co-evolving organisms must continually adapt simply to survive against their ever-changing antagonists. This hypothesis has been applied to explain the maintenance of sexual reproduction; sexual exchange of genetic information increases the emergence of novel genotypes and consequently new immunity, resistance and virulence mechanisms [51–56]. In a microbial population co-evolving with lytic viruses, the viral threat provides strong selection pressure for immunity or resistance in the microbial host. The immunity provided by the spacer of the CRISPR–Cas system makes it likely that horizontal transfer of repeat–spacer arrays and *cas* genes would occur in such populations.

A recent study of a single natural *S. islandicus* population found evidence for rapid recombinatorial reassortment of entire CRISPR loci among strains [29]. Out of a set of 53 natural *S. islandicus* isolates with CRISPR loci where leader-end spacers were observed multiple times, leader-end alleles were found linked to the same trailer-end alleles in the majority of cases. By contrast, linkage among the three CRISPR loci present in this population was found to be low, indicating that complete repeat–spacer arrays are reassorted throughout this population [29]. These findings are in line with observations from other studies outside of a single natural population. A comparison of natural *E. coli* isolates found that phylogenetically close strains typically have very similar spacers, but, in some cases, these strains harbour completely different spacer sets; in over half of cases where the spacers were completely different, the spacers match those of very distant strains, implicating horizontal transfer of spacer arrays [21]. A broad comparison of bacterial 16S rRNA, *cas1* and direct repeat sequences also found evidence for transfer of the CRISPR locus in its entirety [57]. This shuffling of spacer arrays could benefit the host's battle against its viral antagonists. Reassortment of CRISPR loci can redistribute beneficial antiviral spacers through a population and provides another avenue for hosts to acquire different spacers rather than relying solely on leader-end addition. In addition, the reassortment of these loci will also prevent selective sweeps from removing polymorphism within a population and can promote diversity by maintaining trailer-end diversity. Horizontal transfer of *cas* genes from divergent sources has been observed frequently in microbial populations [20,29,58–60]. This could result from a similar mechanism in which horizontal gene flow increases the efficacy of selection on these essential pieces of the CRISPR immunity [61,62]. This observation suggests that variation in the efficiency of recombination and horizontal gene transfer between species might affect the level of polyclonality observed within a microbial population.

Concluding remarks

Diversity in the CRISPR system has major implications for the trajectory of co-evolution between microbial hosts and viruses. Of course, it is only one side of the co-evolutionary equation. Viral mechanisms of action and biology will also greatly affect the diversity observed in the host population, although they are less well understood and more difficult to assess. Co-evolutionary dynamics are affected by the mechanism of host–virus interaction and host and viral biology, as well as by demographic and evolutionary factors on both the host and virus. Links between mechanism, model and observation are still not strong in the co-evolutionary field. As biochemists and molecular biologists work at breakneck speed to illuminate the molecular mechanisms of the CRISPR–Cas system, these discoveries will need to be integrated into models and predictions about the impact on the co-evolutionary process. As we move forward, we must understand that population structure changes in space and time are important in order to provide the link between molecular mechanism and co-evolutionary model. Depending on when you look, the structure of diversity might change; it is the way that structure changes over time and space that is most informative. Because of the elegant link between genotype and immune phenotype within the CRISPR repeat–spacer loci, these molecular markers provide a tractable model for understanding co-evolution dynamics in natural populations. What we learn from microbial models may more generally apply to models of host–pathogen co-evolution in the near future.

Funding

This work was supported by the National Aeronautics and Space Administration (NASA) [grant number NNX09AM92G], the National Science Foundation [grant number DEB-0816885], a National Institutes of Health training grant [grant number 5T32AI078876] to the School of Molecular and Cellular Biology at the University of Illinois, and a fellowship from the Institute for Advanced Computing Applications and Technologies at the University of Illinois.

References

1 Lively, C.M. and Dybdahl, M.F. (2000) Parasite adaptation to locally common host genotypes. Nature **405**, 679–681
2 Mutikainen, P., Salonen, V., Puustinen, S. and Koskela, T. (2000) Local adaptation, resistance, and virulence in a hemiparasitic plant–host plant interaction. Evol. Int. J. Org. Evol. **54**, 433–440
3 Thrall, P.H. and Burdon, J.J. (2000) Effect of resistance variation in a natural plant host–pathogen metapopulation on disease dynamics. Plant Pathol. **49**, 767–773
4 May, R. and Anderson, R. (1983) Epidemiology and genetics in the coevolution of parasites and hosts. Proc. R. Soc. London Ser. B **219**, 281–313
5 Frank, S.A. (2000) Polymorphism of attack and defense. Trends Ecol. Evol. **15**, 167–171
6 Lively, C.M. (2010) A review of Red Queen models for the persistence of obligate sexual reproduction. J. Hered. **101** (Suppl. 1), S13–S20
7 Duffy, S., Burch, C.L. and Turner, P.E. (2007) Evolution of host specificity drives reproductive isolation among RNA viruses. Evolution **61**, 2614–2622
8 Horvath, P., Romero, D.A., Coûté-Monvoisin, A.C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C. and Barrangou, R. (2008) Diversity, activity and evolution of CRISPR loci in *Streptococcus thermophilus*. J. Bacteriol. **190**, 1401–1412

9 Bhaya, D., Davison, M. and Barrangou, R. (2011) CRISPR–Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. **45**, 273–297
10 Mojica, F.J., Ferrer, C., Juez, G. and Rodríguez-Valera, F. (1995) Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. Mol. Microbiol. **17**, 85–93
11 Jansen, R., Embden, J.D. A., van, Gaastra, W. and Schouls, L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. **43**, 1565–1575
12 Hermans, P.W., van Soolingen, D., Bik, E.M., de Haas, P.E., Dale, J.W. and van Embden, J.D. (1991) Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. Infect. Immun. **59**, 2695–2705
13 Groenen, P.M., Bunschoten, A.E., van Soolingen, D. and van Embden, J.D. (1993) Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. Mol. Microbiol. **10**, 1057–1065
14 Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M. et al. (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. **35**, 907–914
15 Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science **315**, 1709–1712
16 Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V. and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. Science **321**, 960–964
17 Tyson, G.W. and Banfield, J.F. (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Environ. Microbiol. **10**, 200–207
18 Pourcel, C., Salvignol, G. and 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**, 653–663
19 Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P. and Moineau, S. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. **190**, 1390–1400
20 Touchon, M. and Rocha, E.P.C. (2010) The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. PLoS ONE **5**, e11126
21 Touchon, M., Charpentier, S., Clermont, O., Rocha, E.P. C., Denamur, E. and Branger, C. (2011) CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. J. Bacteriol. **193**, 2460–2467
22 Shariat, N., DiMarzio, M.J., Yin, S., Dettinger, L., Sandt, C.H., Lute, J.R., Barrangou, R. and Dudley, E.G. (2013) The combination of CRISPR-MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. Food Microbiol. **34**, 164–173
23 Fabre, L., Zhang, J., Guigon, G., Le Hello, S., Guibert, V., Accou-Demartin, M., de Romans, S., Lim, C., Roux, C., Passet, V. et al. (2012) CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. PLoS ONE **7**, e36995
24 Achtman, M. (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. **62**, 53–70
25 Cui, Y., Li, Y., Gorgé, O., Platonov, M.E., Yan, Y., Guo, Z., Pourcel, C., Dentovskaya, S.V., Balakhonov, S.V., Wang, X. et al. (2008) Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. PLoS ONE **3**, e2652
26 Riehm, J.M., Vergnaud, G., Kiefer, D., Damdindorj, T., Dashdavaa, O., Khurelsukh, T., Zöller, L., Wölfel, R., Le Flèche, P. and Scholz, H.C. (2012) *Yersinia pestis* lineages in Mongolia. PLoS ONE **7**, e30624
27 Horvath, P., Romero, D.A., Coûté-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C. and Barrangou, R. (2008) Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. J. Bacteriol. **190**, 1401–1412
28 Held, N.L., Herrera, A., Cadillo-Quiroz, H. and Whitaker, R.J. (2010) CRISPR associated diversity within a population of *Sulfolobus islandicus*. PLoS ONE **5**, e12988
29 Held, N.L., Herrera, A. and Whitaker, R.J. (2013) Reassortment of CRISPR repeat–spacer loci in *Sulfolobus islandicus*. Environ. Microbiol. **15**, 3065–3076
30 Lotka, A.J. (1920) Undamped oscillations derived from the law of mass action. J. Am. Chem. Soc. **42**, 1595–1599
31 Bohannan, B.J. and Lenski, R.E. (2000) Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecol. Lett. **3**, 362–377
32 Buckling, A. and Hodgson, D.J. (2007) Short-term rates of parasite evolution predict the evolution of host diversity. J. Evol. Biol. **20**, 1682–1688
33 Winter, C., Bouvier, T., Weinbauer, M.G. and Thingstad, T.F. (2010) Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. Microbiol. Mol. Biol. Rev. **74**, 42–57
34 Lenski, R.E. (1988) Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. Evolution **42**, 425–432
35 Bohannan, B.J. M. and Lenski, R.E. (1997) Effect of resource enrichment on a chemostat community of bacteria and bacteriophage. Ecology **78**, 2303–2315
36 Buckling, A. and Rainey, P.B. (2002) Antagonistic coevolution between a bacterium and a bacteriophage. Proc. R. Soc. London Ser. B **269**, 931–936
37 Weitz, J.S., Hartman, H. and Levin, S.A. (2005) Coevolutionary arms races between bacteria and bacteriophage. Proc. Natl. Acad. Sci. U.S.A. **102**, 9535–9540
38 Buckling, A., Wei, Y., Massey, R.C., Brockhurst, M.A. and Hochberg, M.E. (2006) Antagonistic coevolution with parasites increases the cost of host deleterious mutations. Proc. R. Soc. London Ser. B **273**, 45–49
39 Nee, S. (1989) Antagonistic coevolution and the evolution of genotypic randomization. J. Theor. Biol. **140**, 499–518
40 Clarke, B. (1976) The ecological relationships of host–parasite relationships. In Genetic Aspects of Host–Parasite Relationships (Taylor, A.E.R. and Muller, R., eds), pp. 87–103, Blackwell, Oxford
41 Hamilton, W.D. (1982) Pathogens as causes of genetic diversity in their host populations. In Population Biology of Infectious Diseases (Anderson, R.M. and May, R.M., eds), pp. 269–296, Springer Verlag, Berlin
42 Koskella, B. and Lively, C.M. (2009) Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. Evol. Int. J. Org. Evol. **63**, 2213–2221
43 Haerter, J.O., Trusina, A. and Sneppen, K. (2011) Bacterial CRISPR defense system buffers phage diversity. J. Virol. **85**, 10554–10560
44 Haerter, J.O. and Sneppen, K. (2012) Spatial structure and Lamarckian adaptation explain extreme genetic diversity at CRISPR locus. mBio **3**, e00126–12
45 Koonin, E.V. and Wolf, Y.I. (2009) Is evolution Darwinian or/and Lamarckian? Biol. Direct **4**, 42
46 Held, N.L., Childs, L.M., Davison, M., Weitz, J.S., Whitaker, R.J. and Bhaya, D. (2013) CRISPR–Cas systems to probe ecological diversity and host–viral interactions. In In CRISPR–Cas Systems (Barrangou, R. and van der Oost, J., eds), pp. 221–250, Springer Verlag, Berlin
47 He, J. and Deem, M.W. (2010) Heterogeneous diversity of spacers within CRISPR (clustered regularly interspaced short palindromic repeats). Phys. Rev. Lett. **105**, 128102
48 Childs, L.M., Held, N.L., Young, M.J., Whitaker, R.J. and Weitz, J.S. (2012) Multiscale model of CRISPR-induced coevolutionary dynamics: diversification at the interface of Lamarck and Darwin. Evolution **66**, 2015–2029
49 Iranzo, J., Lobkovsky, A.E., Wolf, Y.I. and Koonin, E.V. (2013) Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR–Cas in an explicit ecological context. J. Bacteriol. **195**, 3834–3844
50 Van Valen, L. (1973) A new evolutionary law. Evol. Theory **1**, 1–30
51 Jaenike, J. (1978) An hypothesis to account for the maintenance of sex within populations. Evol. Theory **3**, 191–194
52 Bell, G. (1982) The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Cambridge University Press, Cambridge
53 Hamilton, W.D., Axelrod, R. and Tanese, R. (1990) Sexual reproduction as an adaptation to resist parasites (a review). Proc. Natl. Acad. Sci. U.S.A. **87**, 3566–3573
54 Peters, A.D. and Lively, C.M. (1999) The Red Queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution. Am. Nat. **154**, 393–405

55 Otto, S.P. and Nuismer, S.L. (2004) Species interactions and the evolution of sex. Science **304**, 1018–1020
56 Becks, L. and Agrawal, A.F. (2010) Higher rates of sex evolve in spatially heterogeneous environments. Nature **468**, 89–92
57 Chakraborty, S., Waise, T.M. Z., Hassan, F., Kabir, Y., Smith, M.A. and Arif, M. (2009) Assessment of the evolutionary origin and possibility of CRISPR–Cas (CASS) mediated RNA interference pathway in *Vibrio cholerae* O395. In Silico Biol. **9**, 245–254
58 Haft, D.H., Selengut, J., Mongodin, E.F. and Nelson, K.E. (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. **1**, e60
59 Godde, J.S. and Bickerton, A. (2006) The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. J. Mol. Evol. **62**, 718–729
60 Horvath, P., Coûté-Monvoisin, A.-C., Romero, D.A., Boyaval, P., Fremaux, C. and Barrangou, R. (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int. J. Food Microbiol. **131**, 62–70
61 Makarova, K.S. and Koonin, E.V. (2013) Evolution and classification of CRISPR–Cas systems and Cas protein families. In CRISPR–Cas Systems (Barrangou, R. and van der Oost, J., eds), pp. 61–91, Springer Verlag, Berlin
62 Takeuchi, N., Wolf, Y.I., Makarova, K.S. and Koonin, E.V. (2012) Nature and intensity of selection pressure on CRISPR-associated genes. J. Bacteriol. **194**, 1216–1225

Received 8 October 2013
doi:10.1042/BST20130243