CRISPR-Based Typing and Next-Generation Tracking Technologies

Rodolphe Barrangou[1,2,*] and Edward G. Dudley[2]

[1]Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina 27695; email: rbarran@ncsu.edu

[2]Department of Food Science, The Pennsylvania State University, University Park, Pennsylvania 16802; email: egd100@psu.edu

Annu. Rev. Food Sci. Technol. 2016. 7:395–411

First published online as a Review in Advance on January 11, 2016

The *Annual Review of Food Science and Technology* is online at food.annualreviews.org

This article's doi: 10.1146/annurev-food-022814-015729

*Corresponding author.

Keywords

CRISPR, *cas*, typing, epidemiology

Abstract

Bacteria occur ubiquitously in nature and are broadly relevant throughout the food supply chain, with diverse and variable tolerance levels depending on their origin, biological role, and impact on the quality and safety of the product as well as on the health of the consumer. With increasing knowledge of and accessibility to the microbial composition of our environments, food supply, and host-associated microbiota, our understanding of and appreciation for the ratio of beneficial to undesirable bacteria are rapidly evolving. Therefore, there is a need for tools and technologies that allow definite, accurate, and high-resolution identification and typing of various groups of bacteria that include beneficial microbes such as starter cultures and probiotics, innocuous commensals, and undesirable pathogens and spoilage organisms. During the transition from the current molecular biology–based PFGE (pulsed-field gel electrophoresis) gold standard to the increasingly accessible omics-level whole-genome sequencing (WGS) N-gen standard, high-resolution technologies such as CRISPR-based genotyping constitute practical and powerful alternatives that provide valuable insights into genome microevolution and evolutionary trajectories. Indeed, several studies have shown potential for CRISPR-based typing of industrial starter cultures, health-promoting probiotic strains, animal commensal species, and problematic pathogens. Emerging CRISPR-based typing methods open new avenues for high-resolution typing of a broad range of bacteria and constitute a practical means for rapid tracking of a diversity of food-associated microbes.

INTRODUCTION

CRISPR: clustered regularly interspaced short palindromic repeats; a DNA repeat family

CRISPR-associated proteins (Cas): diverse family of proteins directing CRISPR-encoded immunization and immunity

Spacer: hypervariable sequence between palindromic DNA repeats of CRISPR loci; derived from invasive nucleic acid

Recent advances in microbiology have expanded our understanding of the critical roles that various bacteria play in human health and disease, and illustrate their potential for industrial exploitation in biotechnological, agricultural, and food applications (Fraser-Liggett 2005). The impacts of bacteria range widely from health-promoting beneficial probiotics that promote intestinal health of the host to disease-causing pathogenic bacteria that drive lethal processes in various organs and tissues of the host. Additionally, beneficial microbes also include bacterial species widely used as industrial workhorses that drive fermentation processes, enabling food processing and preservation of a broad range of products, including manufacturing of fermented vegetables, beverages, and various dairy products (Bourdichon et al. 2012). Some bacteria that are neither beneficial nor pathogenic can interfere with the food manufacturing and preservation process by driving spoilage, lowering quality, and shortening the shelf life of products.

In the past decade, technological advances in microbiology in general and genomics in particular have provided insights into the genetic underpinning of both pathogenic and beneficial traits in bacteria (Kyrpides 2009). For instance, the genetic basis for desirable probiotic functions such as modulation of the host immune system is being unraveled. Likewise, pathogenic islands encoding antibiotic resistance genes or toxin-manufacturing cassettes have been identified in several species (Relman 2011). Moreover, the genomic renaissance has also allowed the identification of genes that drive key industrial processes and features such as milk acidification and the genesis of desirable flavor and texture of fermented dairy products (Makarova et al. 2006a). The same applies to the biochemical pathways that drive the organoleptic subtleties of beer, wine, and champagne. Similarly, the biochemical processes that drive food spoilage and undesirable secondary fermentations have been characterized.

The functional genomics revolution has generated enough data to enable gazing into genome evolution and trajectories and has allowed scientists to expose the processes that drive genetic diversity and adaptation (Kyrpides 2009). In the past decade, one such milestone has been the discovery of clustered regularly interspaced short palindromic repeats (CRISPRs) (Bolotin et al. 2005, Ishino et al. 1987, Jansen et al. 2002a, Makarova et al. 2006b, Mojica et al. 2005, Pourcel et al. 2005) and CRISPR-associated proteins (Cas) (Haft et al. 2005, Jansen et al. 2002b), which constitute the CRISPR-Cas adaptive immune system in bacteria and archaea (Barrangou 2013, Barrangou & Marraffini 2014, Sontheimer & Barrangou 2015). This seminal biological discovery has given rise to a plethora of technologies and applications that range from antiviral vaccination of prokaryotes to gene therapies in eukaryotes (Barrangou & May 2015). Notwithstanding the biological function of these loci as adaptive immune systems or their potential as revolutionary genome-editing technologies (Cong et al. 2013, Ledford 2015, Pennisi 2013), there is much promise for their use as high-resolution genetic fingerprinting tools for a diversity of bacteria. Here, we discuss recent advances in CRISPR biology and genetics, assess their advantages and caveats, and highlight their potential as next-generation typing tools for bacteria.

CRISPR-CAS SYSTEMS

CRISPR arrays are constituted by alternating stretches of short, noncontiguous DNA repeats separated by variable spacer sequences that form peculiar loci in the sequenced genomes of many bacteria and most archaea (Barrangou & Marraffini 2014, Horvath & Barrangou 2010). The CRISPR repeat sequences are typically 20–38 nt and can occur up to ~600 times in a single genome (*Haliangium ochraceum* DSM 14365). CRISPR repeat sequences are often partially palindromic and predicted to form secondary structures (Kunin et al. 2007). Although multiple CRISPR arrays may

occur in a single chromosome, most organisms typically carry 1–2 CRISPR loci, as documented in the CRISPRdb (CRISPR database) (Grissa et al. 2007). In most cases, CRISPR arrays are flanked by CRISPR-associated (*cas*) genes (Makarova et al. 2011) that encode a diverse family of proteins that carry a wide range of functional domains involved in interaction with nucleic acids, notably nucleases (Makarova et al. 2006b, 2011, 2015). Although much diversity occurs between CRISPR-Cas systems with regard to *cas* gene content, arrangement, sequences, and functions, the *cas1* and *cas2* genes are nearly universal, and usually associated with signature *cas* genes that define the class and major CRISPR-Cas type, such as *cascade*, *cmr*, and *csm* for class I or *cas9* and *cpf1* for class II (Makarova et al. 2015). The nomenclature and classification of CRISPR-Cas systems is based on a robust polythetic system that has been refined and improved over time (Makarova et al. 2006b, 2011, 2015).

Functionally, CRISPR-Cas systems provide adaptive immunity against exogenous genetic elements such as bacteriophages and plasmids in many bacteria and most archaea (Barrangou 2015, Barrangou & Marraffini 2014). These peculiar loci were originally discovered in *Escherichia coli* K12 in 1987 (Ishino et al. 1987) and repeatedly observed in prokaryotic genomes as draft sequences became increasingly available (Jansen et al. 2002a,b). Shortly after their involvement in adaptive immunity against bacteriophages was established (Barrangou et al. 2007), a series of seminal studies showed that CRISPR-Cas systems are DNA-encoded (Barrangou et al. 2007), RNA-mediated (Brouns et al. 2008), nucleic acid–targeting systems (Hale et al. 2009, Marraffini & Sontheimer 2008). Although many different CRISPR-Cas systems exist across two distinct classes and five major types (Makarova et al. 2015), there are three conserved stages that drive CRISPR-encoded immunization and CRISPR-mediated immunity, namely adaptation, expression, and interference (Barrangou & Marraffini 2014). During the adaptation stage, DNA sequences derived from invasive elements are sampled by a copy-paste process, which yields the integrase-driven insertion of a new repeat-spacer unit at the leader end of the locus (Nuñez et al. 2015). This process defines the adaptive nature of CRISPR-Cas immune systems and constitutes the vaccination portion of immunization (Barrangou et al. 2013). Once a strain has been vaccinated, during the expression stage, the repeat-spacer array is transcribed into pre-crRNA (CRISPR RNA), which is processed into small interfering crRNAs that carry a portion of a spacer sequence and a section of a CRISPR repeat (Brouns et al. 2008). Subsequently, during the interference stage, the crRNAs guide Cas nucleases toward complementary sequences for sequence-specific recognition and degradation through bona fide nucleic acid targeting (Garneau et al. 2010, Gasiunas et al. 2012). The nature of the ribonucleoprotein complexes that drive interference differs across classes, types, and subtypes, and the best-characterized models involve Cascade and the Cas3 nuclease in Type I systems; the Cas9 endonuclease in Type II systems; and the Cmr or Csm complex in Type III systems. In some cases, targeting also requires ancillary elements such as the protospacer adjacent motif (PAM), RNaseIII, or the *trans*-encoded crRNA (tracrRNA) (Deltcheva et al. 2011, Deveau et al. 2008, Horvath et al. 2008, Mojica et al. 2009). In Type II systems, the widely popular Cas9 endonuclease is a large multidomain protein that generates double-stranded breaks in target DNA (Garneau et al. 2010, Gasiunas et al. 2012, Jinek et al. 2012) using two nickase domains (RuvC and HNH) that each nick one target DNA strand within an R-loop structure at a precise distance from the PAM.

With regard to applications, CRISPR-Cas systems have been used for various purposes in their native hosts, including genotyping (Barrangou & Horvath 2012), vaccination against phages (Barrangou et al. 2013), interference against plasmid uptake (Marraffini & Sontheimer 2008), use as antimicrobials (Gomaa et al. 2014), and genome editing and remodeling (Jiang et al. 2013, Oh & van Pijkeren 2014, Selle et al. 2015, Selle & Barrangou 2015). Notwithstanding their implementation in prokaryotes, the recent CRISPR craze (Ledford 2015, Pennisi 2013) has focused on the

CRISPR database (CRISPRdb): website offering several tools to identify and analyze CRISPR loci and Cas sequences

Bacteriophage: virus that infects bacteria, also known as phage

Interference: process by which an invasive nucleic acid is targeted by crRNA-guided CASCADE

CRISPR RNA (crRNA): small noncoding ribonucleic acid mediating CRISPR-encoded immunity

exploitation of CRISPR-Cas9 technologies (Jinek et al. 2012) for a plethora of genome-editing applications in eukaryotes, including genetic engineering (Cong et al. 2013, Jiang et al. 2013, Mali et al. 2013), transcriptional control (Qi et al. 2013), methylation and epigenetics (Hilton et al. 2015), imaging (Chen et al. 2013), and large-scale screens (Shalem et al. 2014), with implications for translational medicine, biotechnology, and agriculture (Doudna & Charpentier 2014). With regard to applications in food science and technology, pioneering work and seminal studies in lactic acid bacteria have already been exploited in industrially relevant bacteria for various purposes. With regard to phage resistance, CRISPR-driven immunization of industrial *Streptococcus thermophilus* dairy starter cultures has been successfully implemented for yogurt and cheese manufacturing on a global scale (Barrangou et al. 2013). Similar approaches may be underway to vaccinate other dairy cultures against lytic phages. It is noteworthy that there are natural means to harness this technology that would alleviate any concerns about genetic engineering and manipulation, as screening of natural vaccination events can be readily implemented in the laboratory to replicate natural phenomena. One advantage of this technology is the ability to run iterative vaccination rounds to enhance the breadth and depth of phage resistance and combine resulting isogenic strains in the formulation of starter culture rotations. Furthermore, the exploitation of CRISPR-based genotyping has been implemented for various dairy cultures and probiotic strains and for tracking, typing, and tagging purposes in *Streptococcus*, *Lactobacillus*, and *Bifidobacterium* species (Briner et al. 2015, Horvath et al. 2009, Ventura et al. 2009). Looking at future applications for the food and agricultural industries, we anticipate that CRISPR-based genome editing holds much potential in animal and plant breeding for husbandry and major crop trait enhancement, respectively.

CRISPR-BASED GENOTYPING OF BACTERIA

A core feature of CRISPR-Cas systems is the ability to acquire novel spacers in an ordinal manner within the CRISPR array, which essentially captures exposure to invasive genetic elements over time in a sequential manner (**Figure 1**). This peculiar feature thus constitutes a unique set of DNA sequences that provide a valuable basis for genotyping of bacteria using a single hypervariable region, which includes information about the origin, evolutionary trajectory, and recent path of an isolate (**Figure 1**) (Barrangou & Horvath 2012).

This CRISPR-based typing of bacteria has also been expanded to the analysis of populations, with the ability to truly assess genetic diversity even within relatively clonal sets of bacteria (Paez-Espino et al. 2013). In some cases, these loci can provide insights into host-virus population interactions and dynamics in natural habitats, with potential for resolution in both space (geographical distribution) and time (evolution over the sampling period), as established in several studies of environmental samples (Anderson et al. 2011, Andersson & Banfield 2008, Heidelberg et al. 2009, Held & Whitaker 2009, Held et al. 2010, Tyson & Banfield 2008). In these studies, deep-sequencing analysis of environmental samples has allowed probing of host-virus interactions. For instance, using metagenomics to analyze acid mine drainage biofilm samples over space and time, the Banfield group was able to reconstruct viral genomes and match them to their archaeal and bacterial hosts by using homology to CRISPR spacer sequences; in *Leptospirillum*, CRISPR genotypes were able to distinguish two subpopulations based on conserved ancestral spacers (Andersson & Banfield 2008, Tyson & Banfield 2008). Likewise, in a study of hydrothermal vents, reconstruction of CRISPR loci was used to match CRISPR spacers of *Methanocaldococcus* strains to viral sequences and to gaze into the interplay between the co-occuring host and viral populations by using sequence matches between CRISPR spacers and viral sequences (Anderson et al. 2011). Similar analyses in *Sulfolobus* have allowed typing of strains across continents, and the subtyping of *Sulfolobus islandicus* strains isolated from a single hotspring in Russia, with phylogenetic insights

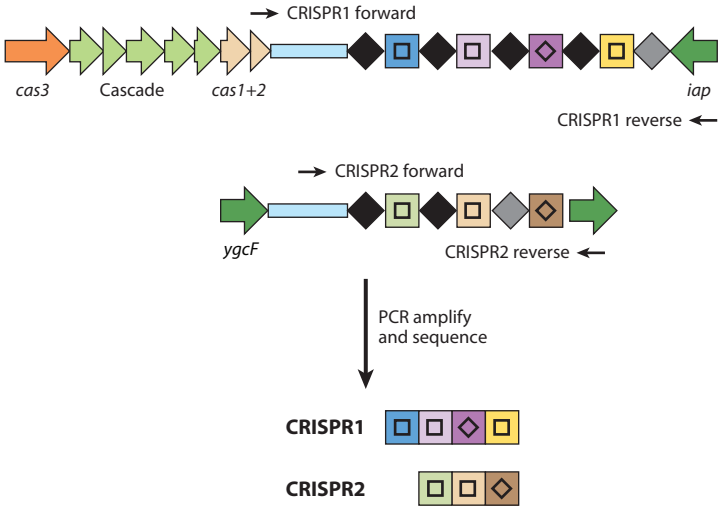

Figure 1

CRISPR locus typing scheme. Schematic of CRISPR typing protocol for *Salmonella enterica*. The top diagram represents a typical CRISPR1 locus, with the array shown as repeats (*black diamonds*) and unique spacer sequences (*colored boxes*). Different symbols inside the boxes indicate the length of the spacer, which is most commonly 32 bp for *Salmonella*. Upstream of the first repeat is the leader region (*light blue box*), which contains genes responsible for spacer acquisition (*cas1*+2) and interference (*cas3* and the Cascade complex). Using polymerase chain reaction (PCR) primers that target conserved sequences, the arrays from CRISPR1 and CRISPR2 can be amplified and sequenced. For analysis, repeat sequences are removed, and the spacer sequences are concatenated. Typically, comparable CRISPR loci belong to the same subtype and share CRISPR repeat sequence identity but need not share the same genomic context.

provided by conserved ancestral spacers (Held & Whitaker 2009, Held et al. 2010). This approach has also been implemented in the analysis of a hotspring microbial mat community to determine the host-virus interplay in *Synechococcus* (Heidelberg et al. 2009), with distinct CRISPR-Cas systems targeting different viral genotypes. Most of these studies used CRISPR spacer diversity to distinguish strains from various samples or to assess genetic diversity within a sample. This approach was successful because CRISPR-Cas systems occurred frequently in these data sets (high enough coverage in deep-sequencing outputs to assemble reads with overlapping spacer sequences and reconstruct CRISPR arrays), and spacer polymorphism was sufficient to assemble distinct loci bioinformatically (distinct contigs could be generated and compared and contrasted using differential spacer content). However, this method is of limited value when CRISPR-Cas systems occur at a low frequency and in cases in which novel spacer acquisition does not allow distinction between closely related strains that share identical spacer content.

Although much interest was initially focused on CRISPR array diversity in *Mycobacterium* and *Yersinia*, the potential of CRISPR-based genotyping in bacteria quickly expanded to industrial dairy cultures and has recently been applied across a range of clinically relevant isolates. Early work in *Mycobacterium tuberculosis* (Groenen et al. 1993) has led to several studies showing potential for widespread typing in this species (Abadia et al. 2010, 2011; Borile et al. 2011; Brudey et al. 2006; Zhang et al. 2010). The same applies to *Yersinia pestis*, for which early work (Cui et al. 2008, Pourcel et al. 2005) established a basis for recent studies (**Table 1**). This also applies to *Corynebacterium diphtheriae* (Mokrousov et al. 2007, 2009), *Pseudomonas aeruginosa* (Cady et al. 2011), *E. coli* (Diez-Villasenor et al. 2010, Toro et al. 2014, Touchon & Rocha 2010, Touchon et al. 2011,

Table 1 CRISPR-based genotyping of bacteria

Organism	Type	Reference
Bifidobacterium animalis	Probiotic	Barrangou et al. 2009
Campylobacter jejuni	Pathogen	Kovanen et al. 2014
Clostridium difficile	Pathogen	Hargreaves et al. 2014
Corynebacterium diphtheriae	Pathogen	Mokrousov et al. 2009
Enterococcus faecalis	Pathogen	Lindenstrauss et al. 2011
Enterococcus faecium	Pathogen	Tremblay et al. 2013
Erwinia amylovora	Pathogen	Rezzonico et al. 2011
Escherichia coli	Pathogen	Yin et al. 2013
Lactobacillus buchneri	Spoilage	Briner & Barrangou 2014
Lactobacillus casei	Probiotic	Broadbent et al. 2012
Legionella pneumophila	Pathogen	D'Auria et al. 2010
Microcystis aeruginosa	Pathogen	Kuno et al. 2012
Mycobacterium tuberculosis	Pathogen	Groenen et al. 1993
Propionibacterium acnes	Pathogen	Brüggemann et al. 2012
Salmonella enterica	Pathogen	Shariat et al. 2014
Streptococcus agalactiae	Pathogen	Lier et al. 2015
Streptococcus thermophilus	Starter culture	Horvath et al. 2008
Staphylococcus aureus	Pathogen	Kinnevey et al. 2013
Vibrio parahemolyticus	Pathogen	Sun et al. 2015
Xanthomonas aeruginosa	Pathogen	Semenova et al. 2009
Yersinia pestis	Pathogen	Riehm et al. 2012

Yin et al. 2013), *Legionella* (D'Auria et al. 2010), and *Streptococcus pyogenes* (McShan et al. 2008). Although multiple groups have shown ad hoc potential of CRISPR-based diversity and hypervariability for typing of various bacteria (**Table 1**), there is a need to assess the potential of these loci for typing across a broad phylogenetic set of organisms. Of course, the applicability of this approach depends on the frequency at which CRISPR-Cas systems occur in these genomes in general and on the occurrence of identical CRISPR repeats in particular. Furthermore, the novel CRISPR spacer acquisition propensity must be sufficient to distinguish closely related strains, whereas ancestral spacer conservation is desirable to anchor phylogenetic relationships. Therefore, the potential of CRISPR-based genotyping must be assessed on an individual basis for each candidate species, and the sole presence of a CRISPR array in a draft genome is only a starting point.

Potential of CRISPR-Based Typing for Beneficial Bacteria: Lessons from *Streptococcus thermophilus* Industrial Starter Cultures

As CRISPR-based typing evolves, several studies have shown that there is potential for typing of industrially relevant organisms such as starter cultures and probiotics. Indeed, several groups have shown potential for CRISPR-based typing in lactobacilli and streptococci (Guinane et al. 2011; Horvath et al. 2008, 2009). The implementation of CRISPR-based genotyping in the ubiquitously used dairy starter culture *S. thermophilus* species has been very useful, and it was demonstrated that the occurrence and diversity of CRISPR-Cas systems in many strains within this species provide valuable insights into the origin and genetic type of particular strains (Horvath et al. 2008). The presence of the universal CRISPR-Cas1 system in highly divergent strains has been instrumental

for PCR-based amplification and sequencing of a single locus for high-resolution genotyping of even clonal strains. The occasionally conserved and shared ancestral spacers at the trailer end provide phylogenetic anchors, whereas the spacers recently acquired at the leader end provide insights into recent phage exposures. Indeed, the ability of this locus to readily acquire novel spacers following phage exposure yields enough activity to generate spacer-based diversity even within isolates derived from the same ancestral strain. One notable advantage of CRISPR-based typing in *S. thermophilus* is the differential presence of four distinct CRISPR-Cas systems (encompassing the universally occurring and variable CRISPR-Cas1), including two systems that have been shown to actively acquire novel spacers, enabling distinction of strains isolated within a short period of time (a few months for some genotypes) (Barrangou et al. 2013, Horvath et al. 2008, Horvath & Barrangou 2010). In laboratory settings with phage evolutionary pressure, it has been shown that active CRISPR loci can acquire novel spacers on a daily scale (Paez-Espino et al. 2013). This CRISPR-based typing methodology has been broadly implemented for the genotyping of dairy cultures derived from various public and industrial collections, and isolated from a plethora of industrial cultures and dairy products across the globe. Such methods also provide unequivocal and easy to interpret genetic fingerprints that can prove instrumental in tracking and monitoring the use of proprietary industrial strains that have unique sets of vaccination events that have been selectively screened, yielding a series of spacer combinations that are extremely unlikely to randomly occur in nature. Actually, selection of variants following iterative rounds of vaccination against phages can yield unique profiles that enable efficient and affordable monitoring of proprietary strains (Barrangou et al. 2013). Furthermore, in some cases, CRISPR typing has enabled the detection of natural vaccination events that likely occurred in industrial settings and yielded an expanded CRISPR spacer repertoire that can distinguish a naturally generated phage-resistant variant from the parent strain that was originally formulated in a commercial starter culture.

Exploiting CRISPR Diversity for Pathogen Tracking: Insights from Studies of *Salmonella* Genomes and CRISPR Diversity

Notwithstanding early success using spacer-oligonucleotide typing (spoligotyping) in *M. tuberculosis* (Groenen et al. 1993, Zhang et al. 2010), one key species for which an extensive set of CRISPR-based genotyping studies is available is *Salmonella*, with efforts from several laboratories establishing a framework for the exploitation of CRISPR loci for in-depth comparative analysis of clinical isolates with high-resolution insights into microevolution (DiMarzio et al. 2013; Fabre et al. 2012; Fricke et al. 2011; Liu et al. 2011a,b; Pettengill et al. 2014; Shariat et al. 2013a,b,c, 2014; Timme et al. 2013; Wehnes et al. 2014). The occurrence and diversity of CRISPR-Cas systems and loci in a broad diversity of *Salmonella enterica* strains, isolates, genomes, and phylotypes is well documented. A number of studies have focused on *Salmonella* strains derived from reference strains and genomes, as well as a wide set of isolates sampled from actual outbreaks involving a broad range of food products and patients. Indeed, determining the CRISPR spacer content of *Salmonella* serovar Typhimurium as well as *Salmonella* serovar Enteritidis from 10 distinct outbreaks (Fabre et al. 2012) has proven epidemiologically valuable (**Figure 2**). Furthermore, comparative analysis of CRISPR sequences in *S. enterica* 6,7:c1.5 isolates that are closely related has provided results consistent with multilocus sequence typing (MLST) (Fabre et al. 2012), illustrating epidemiological concordance between distinct typing methods that target different chromosomal regions and types of sequences. Similar findings also apply to *Salmonella* serovar Typhi versus *Salmonella* serovar Paratyphi. This technology has also been applied to more than 1,500 isolates, representing more than 100 serovars, and has proven to yield valuable and insightful results (Shariat & Dudley 2014).

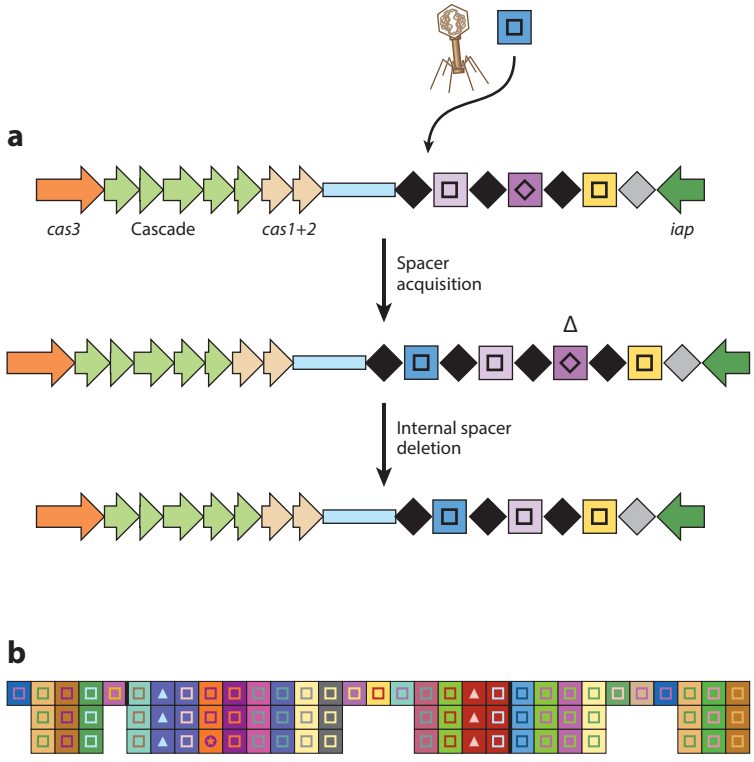

Figure 2

CRISPR locus evolution. (*a*) Two mechanisms of CRISPR spacer array diversification. As a defense mechanism, *cas1* and *cas2* (*tan arrows*) insert new spacer (*light blue box*) sequences derived from foreign DNA into the leader-proximal end of a repeat-spacer array. Additionally, spacer sequences can be deleted (*indicated by Δ above purple spacer*), presumably through homologous recombination between repeat sequences. (*b*) CRISPR1 from three selected *Salmonella* Typhimurium isolates, highlighting the two mechanisms of diversification.

Investigations of foodborne outbreaks primarily use PFGE (pulsed-field gel electrophoresis) for strain identification and tracking, and CRISPR technology can be useful as a complement to traditional subtyping methods (Very et al. 2015) or as an alternative to PFGE. Although CRISPR diversity may provide strain-level discrimination, it should also be noted that conservation of certain spacers within closely related strains can be used for serogroup/serotype identification of *Salmonella* (Fricke et al. 2011). The concordance between CRISPR-based genotyping and serotyping has been valuable in showcasing the potential of this methodology. Even within *Salmonella* serovars, spacer microevolution (including evolution by duplication or removal of internal spacers; see **Figure 2** and **3**) has proven useful for typing of very closely related isolates.

Monitoring Spoilage Organisms Using CRISPR: Rapid Assessment of *Lactobacillus buchneri* Diversity in Cucumber Fermentation Tanks

Notwithstanding the focus on beneficial and pathogenic bacteria, the quality and shelf life of a food product heavily depend on the presence, activity, and growth of naturally occurring bacteria that can turn into undesirable spoilage organisms. Therefore, protecting and maintaining the

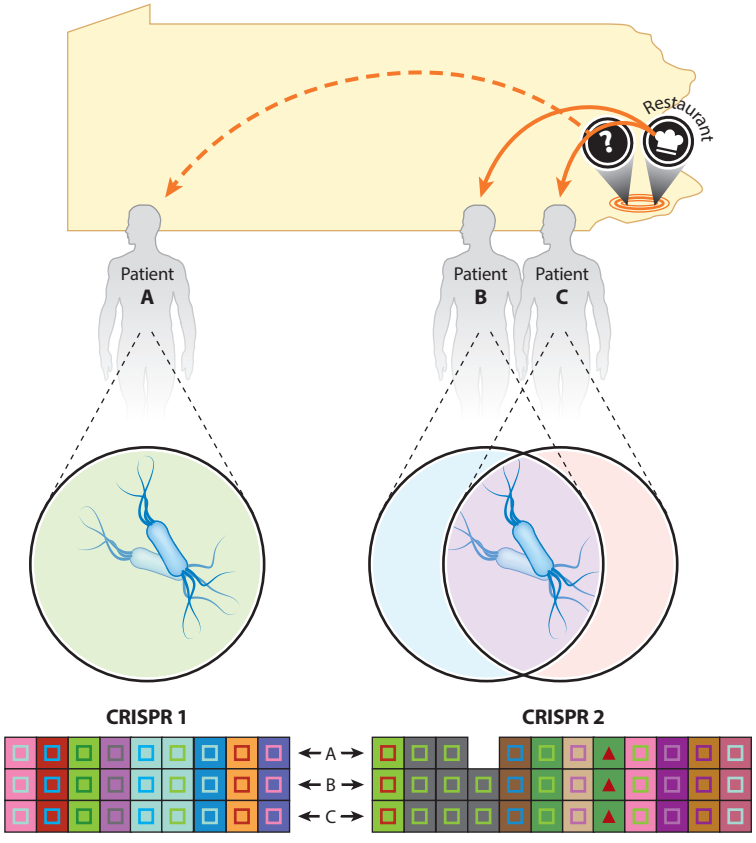

Figure 3

Application of CRISPR typing in epidemiologic investigations. Three *Salmonella* Enteritidis isolates were received by the Pennsylvania Department of Health during the same time frame, and all three were typed as pattern .0004 by PFGE (pulsed-field gel electrophoresis). Food questionnaires indicated that patients B and C ate at the same restaurant; however, no epidemiologic data linked patient A to the outbreak. CRISPR sequencing revealed a one-spacer difference between this latter isolate and those from the other two patients, confirming that two separate incidents were responsible for these illnesses (N. Shariat & E.G. Dudley, unpublished data). Adapted from illustration by Joselyn N. Allen (Pennsylvania State University) with permission.

quality of food products partially hinge on our ability to monitor and control spoilage organisms. In fermented foods, especially in complex products that undergo extended and/or several rounds of fermentations, such as fermented vegetables (sauerkraut, kimchi, pickles), much focus ought to be placed on tracking the occurrence and load of spoilage organisms.

A recent study of CRISPR-based typing of *Lactobacillus buchneri*, a beneficial organism for silage inoculation and bioethanol manufacturing but an undesirable species involved in cucumber fermentation spoilage, has established a basis for assessing the occurrence and diversity of spoilage organisms (Briner & Barrangou 2014). Specifically, it was shown that a Type II CRISPR-Cas system universally occurs in this species, and a comparative analysis of CRISPR sequences in pickle fermentation isolates revealed that multiple genotypes can coexist. In particular, spacers conserved at the ancestral end reflect common origin, whereas diversity at the leader end illustrates

recent divergence. Subsequent analysis of homologous sequences in *L. buchneri* genomes confirms the broad applicability of this approach in other strains and illustrates the potential of similar approaches in other spoilage organisms, especially lactic acid bacteria, in which CRISPR-Cas systems are highly enriched.

As surveys and studies of microbial content and composition of environmental and clinical samples increasingly delve into the true composition of populations on a comprehensive basis, there is a need to ascertain diversity at the strain level to assess within-sample heterogeneity and strain microevolution between closely related clones. This allows for tracking of lineages within monomorphic clonal populations, and this is where CRISPR-based typing of clones is a practical means to dig into complex situations in which multiple variants derived from common ancestors must be resolved. This can be done using classical microbiology and sequencing approaches (Sanger-based sequencing of a PCR amplicon of a CRISPR locus from a colony) or next-generation deep-sequencing technologies (in silico assembly and reconstitution of CRISPR arrays from deep Illumina sequencing or resequencing). Such approaches have been successfully implemented for the analysis of complex environmental samples in acid mine drainage (Andersson & Banfield 2008, Tyson & Banfield 2008), microbial mats (Heidelberg et al. 2009), hyperthermophilic environments (Held & Whitaker 2009, Held et al. 2010), and the human oral cavity (Pride et al. 2011). These studies have indeed established that reconstruction of CRISPR loci from metagenomic data can establish ancestral relationships of phylogenetically linked strains and provide insights into evolutionary events that have shaped microbial composition at the population level, such as sweeps and bottlenecks that constrain population constitution. These data sets can also be mathematically modeled to quantitatively assess genetic diversity with high-level resolution (He & Deem 2010, Levin et al. 2013, Weinberger et al. 2012). In some studies, it is also possible to infer and/or reconstitute the interplay between hosts and their viruses via sequence homologies between CRISPR spacers (derived from host contigs) and matching bacteriophage sequences (derived from viral contigs). Further analyses can also unravel the genomic strategies and trajectories employed by viral populations to escape and circumvent CRISPR-encoded immunity (Sun et al. 2013).

Challenges regarding the widespread use of CRISPR loci for bacterial genotyping include their phylogeny-independent occurrence in genera and species of interest, and sometimes variable distribution patterns within a species. Indeed, although CRISPR loci occur in many bacteria, they are currently documented in only 47% of sequenced bacterial genomes (Grissa et al. 2007; Makarova et al. 2011, 2015). Furthermore, various classes, types, and subtypes are differentially distributed in phylogenetic clades, with Type I systems most frequently widespread in bacteria, Type II systems least frequent and exclusive to bacteria, and Type III systems more common in archaea. Another element compounding this issue is the propensity of CRISPR-Cas systems for horizontal gene transfer (Godde & Bickerton 2006). Nevertheless, they are widespread in *Firmicutes* and occur rather frequently in pathogenic bacteria of clinical interest.

Another limitation to the widespread use of CRISPR loci for genotyping applications is the paucity of biological and genetic insights into their activity, as defined by the ability to acquire novel spacers and thus have an evolutionarily relevant polymorphic and evolving genotype. Indeed, the extent of spacer polymorphism (both spacer number and sequence) is directly correlated with the activity of the CRISPR locus. Unfortunately, only a few systems have been characterized to date with the ability to acquire novel CRISPR spacers following exposure to phages and/or plasmids. Although inactive CRISPR arrays evolve via decay, accumulation of mutations and internal deletion of spacers, their optimal phylogenetic value relies on an active system that expands by novel spacer acquisition over time. Thus, there is a need for more advanced and systematic studies of spacer acquisition in genera and species of clinical relevance.

CONCLUSIONS AND PERSPECTIVE FOR FUTURE APPLICATIONS

There is an increasing need for technologies that allow for rapid, definite, and practical identification of bacterial isolates at strain-level resolution. This is particularly important for the monitoring and analysis of complex microbial environments such as food products that concurrently contain beneficial, commensal, and pathogenic bacteria. Although genomic analysis of food samples can provide ecological insights into the dominant microflora, there are limits to depth of sequencing that are coverage constrained, and resulting data are often fragmented, difficult to interpret, and sometimes phylogenetically orphaned. In many cases, CRISPR-based typing technologies deliver unique DNA fingerprints that provide valuable insights into the origin and path of a strain as well as into the unequivocal relationships between even clonal isolates. These technologies can be customized to target CRISPR loci that are unique to genera or species of interest, and to selectively and specifically monitor genotypes or ribotypes of interest. Several studies have provided proof of concept in health-promoting genera (and species) such as *Lactobacillus* (e.g., *Lactobacillus casei*) and *Bifidobacterium* (e.g., *Bifidobacterium animalis* subsp. *lactis*), as well as starter cultures (e.g., *S. thermophilus*), food spoilage organisms (e.g., *L. buchneri*), and a plethora of human pathogens that pose a food safety threat (i.e., *Salmonella*). Furthermore, the presence of CRISPR loci in other genera of interest and their occurrence and distribution in food microbiome data sets open new avenues for investigation of CRISPR-based typing technologies for next-generation microbial tracking.

Nevertheless, there are some limitations to the widespread exploitation of CRISPR loci for high-resolution genotyping of bacteria, such as their absence in some species of interest and their low variability in some cases, which may preclude high-resolution typing. In select cases, this technology may be practically best used in combination with other rapid technologies that selectively target variable genetic regions, such as MLST, and variants such as multi-virulence-locus sequence typing (MVLST). Although whole-genome sequencing has become increasingly accessible and available, it is still a technology constrained by our ability to rapidly and easily interpret large data sets and for which interpretation can be quality dependent, thus rendering CRISPR-based technologies useful in the interim.

SUMMARY POINTS

1. CRISPR-Cas systems provide DNA-encoded, RNA-mediated, DNA-targeting immunity against phages and plasmids.
2. CRISPR loci are widespread in many bacteria and most archaea, with variable occurrence and diversity in beneficial, commensal, and pathogenic species.
3. During the immunization process, CRISPR-Cas systems acquire exogenous DNA as novel spacers that constitute a time-resolved record of vaccination events.
4. CRISPR-based typing provides valuable targets for high-resolution phylogenetic studies of select food pathogens and clinical isolates.

FUTURE ISSUES

1. Notwithstanding current knowledge in *Escherichia*, *Salmonella*, and *Streptococcus*, CRISPR-based typing holds potential for high-resolution typing of *Clostridium* and other clinically relevant organisms.

2. The potential of CRISPR-based typing technologies is constrained by their limited distribution and activity.
3. As our understanding of CRISPR biology, genetics, and functionality advances in a broad set of organisms, new avenues will open for enhanced use of CRISPR-based technologies.
4. In species in which CRISPR systems are active, there is a need to assess the rate at which spacers are acquired to measure microevolution quantitatively.

DISCLOSURE STATEMENT

R.B. is on the Board of Directors of Caribou Biosciences, a co-founder and member of the Scientific Advisory Board of Intellia Therapeutics, and a co-inventor on several patents related to various uses of CRISPR-Cas systems. E.G.D. is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank our colleagues and collaborators for their support of our work and many contributions to the ever-expanding CRISPR field, as well as the United States Army Research Office (grant number W911NF-11-1-0442). We also thank Joselyn N. Allen (Pennsylvania State University) for creating **Figure 3**.

LITERATURE CITED

Abadia E, Zhang J, dos Vultos T, Ritacco V, Kremer K, et al. 2010. Resolving lineage assignation on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method. *Infect. Genet. Evol.* 10:1066–74

Abadia E, Zhang J, Ritacco V, Kremer K, Tuimy R, et al. 2011. The use of microbead-based spoligotyping for *Mycobacterium tuberculosis* complex to evaluate the quality of the conventional method: providing guidelines for quality assurance when working on membranes. *BMC Infect. Dis.* 11:110

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

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

Barrangou R. 2013. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip. Rev. RNA* 4:267–78

Barrangou R. 2015. The roles of CRISPR-Cas systems in adaptive immunity and beyond. *Curr. Opin. Immunol.* 32:36–41

Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, et al. 2009. Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J. Bacteriol.* 191:4144–51

Barrangou R, Coûté-Monvoisin AC, Stahl B, Chavichvily I, Damange F, et al. 2013. Genomic impact of CRISPR immunization against bacteriophages. *Biochem. Soc. Trans.* 41:1383–91

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

Barrangou R, Horvath P. 2012. CRISPR: new horizons in phage resistance and strain identification. *Annu. Rev. Food Sci. Technol.* 3:143–62

Barrangou R, Marraffini LA. 2014. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54:234–44

Barrangou R, May AP. 2015. Unraveling the potential of CRISPR-Cas9 for gene therapy. *Expert Opin. Biol. Ther.* 15:311–14

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151:2551–61
Borile C, Labarre M, Franz S, Sola C, Refregier G. 2011. Using affinity propagation for identifying subspecies among clonal organisms: lessons from *M. tuberculosis*. *BCM Bioinform.* 12:224
Bourdichon F, Casaregola S, Farrokh C, Frisvad JC, Gerds ML, et al. 2012. Food fermentations: microorganisms with technological beneficial use. *Int. J. Food Microbiol.* 154:87–97
Briner AE, Barrangou R. 2014. *Lactobacillus buchneri* genotyping on the basis of clustered regularly interspaced short palindromic repeat (CRISPR) locus diversity. *Appl. Environ. Microbiol.* 80:994–1001
Briner AE, Lugli GA, Milani C, Duranti S, Turroni F, et al. 2015. Occurrence and diversity of CRISPR-Cas systems in the genus *Bifidobacterium*. *PLOS ONE* 10:e0133661
Broadbent JR, Neeno-Eckwall EC, Stahl B, Tandee K, Cai H, et al. 2012. Analysis of the *Lactobacillus casei* supragenome and its influence in species evolution and lifestyle adaptation. *BMC Genomics* 13:533
Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–64
Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, et al. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 6:23
Brüggemann H, Lomholt HB, Tettelin H, Kilian M. 2012. CRISPR/Cas loci of type II *Propionibacterium acnes* confer immunity against acquisition of mobile elements present in type I *P. acnes*. *PLOS ONE* 7:e34171
Cady KC, White AS, Hammond JH, Abendroth MD, Karthikeyan RS, et al. 2011. Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* 157:430–37
Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, et al. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155:1479–91
Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–23
Cui Y, Li Y, Gorge O, Platonov ME, Yan Y, et al. 2008. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLOS ONE* 3:e2652
D'Auria G, Jimenez-Hernandez N, Peris-Bondia F, Moya A, Latorre A. 2010. *Legionella pneumophila* pangenome reveals strain-specific virulence factors. *BMC Genomics* 11:181
Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. 2011. CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* 471:602–7
Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, et al. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1390–400
Diez-Villasenor C, Almendros C, Garcia-Martinez J, Mojica FJ. 2010. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 156:1351–61
DiMarzio M, Shariat N, Kariyawasam S, Barrangou R, Dudley EG. 2013. Antibiotic resistance in *Salmonella* Typhimurium associates with CRISPR sequence type. *Antimicrob. Agents Chemother.* 9:4282–89
Doudna JA, Charpentier E. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:6213
Fabre L, Zhang J, Guigon G, Le Hello S, Guibert V, et al. 2012. CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. *PLOS ONE* 7:e36995
Fraser-Liggett CM. 2005. Insights on biology and evolution from microbial genome sequencing. *Genome Res.* 15:1603–10
Fricke WF, Mammel MK, McDermott PF, Tartera C, White DG, et al. 2011. Comparative genomics of 28 *Salmonella enterica* isolates: evidence for CRISPR-mediated adaptive sublineage evolution. *J. Bacteriol.* 193:3556–68
Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, et al. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468:67–71
Gasiunas G, Barrangou R, Horvath P, Siksnys V. 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *PNAS* 109:E2579–86
Godde JS, Bickerton A. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62:718–29

Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. 2014. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *mBio* 5:e00928–13

Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinform.* 23:172

Groenen PM, Bunschoten AE, van Soolingen D, van Embden JD. 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10:1057–65

Guinane C, Kent RM, Norberg S, Hill C, Fitzgerald GF, et al. 2011. Host specific diversity in *Lactobacillus johnsonii* as evidenced by a major chromosomal inversion and phage resistance mechanisms. *PLOS ONE* 6:e18740

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

Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, et al. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139:945–56

Hargreaves KR, Flores CO, Lawley TD, Clokie MR. 2014. Abundant and diverse clustered regularly interspaced short palindromic repeat spacers in *Clostridium difficile* strains and prophages target multiple phage types within this pathogen. *mBio* 5:e01045-13

He J, Deem MW. 2010. Heterogeneous diversity of spacers within CRISPR (clustered regularly interspaced short palindromic repeats). *Phys. Rev. Lett.* 105:128102

Heidelberg JF, Nelson WC, Schoenfeld T, Bhaya D. 2009. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLOS ONE* 4:e4169

Held NL, Herrera A, Cadillo-Quiroz H, Whitaker RJ. 2010. CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLOS ONE* 5:e12988

Held NL, Whitaker RJ. 2009. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ. Microbiol.* 11:457–66

Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, et al. 2015. Epigenome editing by a CRISPR-Cas9–based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33:510–17

Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–70

Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* 131:62–70

Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, et al. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1401–12

Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169:5429–33

Jansen R, Embden JD, Gaastra W, Schouls LM. 2002a. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43:1565–75

Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002b. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* 6:23–33

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31:233–39

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–21

Kinnevey PM, Shore AC, Brennan GI, Sullivan DJ, Ehricht R, et al. 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome mec (SCCmec)-SCC-SCCCRISPR composite element in Irish hospitals. *Antimicrob. Agents Chemother.* 57:524–31

Kovanen SM, Kivistö RI, Rossi M, Hänninen ML. 2014. A combination of MLST and CRISPR typing reveals dominant *Campylobacter jejuni* types in organically farmed laying hens. *J. Appl. Microbiol.* 117:249–57

Kunin V, Sorek R, Hugenholtz P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8:R61

Kuno S, Yoshida T, Kaneko T, Sako Y. 2012. Intricate interactions between the bloom-forming cyanobacterium *Microcystis aeruginosa* and foreign genetic elements, revealed by diversified clustered regularly interspaced short palindromic repeat (CRISPR) signatures. *Appl. Environ. Microbiol.* 78:5353–60

Kyrpides NC. 2009. Fifteen years of microbial genomics: meeting the challenges and fulfilling the dream. *Nat. Biotechnol.* 27:627–32

Ledford H. 2015. CRISPR, the disruptor. *Nature* 522:20–24

Levin BR, Moineau S, Bushman M, Barrangou R. 2013. The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. *PLOS Genet.* 9:e1003312

Lier C, Baticle E, Horvath P, Haguenoer E, Valentin AS, et al. 2015. Analysis of the type II-A CRISPR-Cas system of *Streptococcus agalactiae* reveals distinctive features according to genetic lineages. *Front. Genet.* 6:214

Lindenstrauss AG, Pavlovic M, Bringmann A, Behr J, Ehrmann MA, Vogel RF. 2011. Comparison of genotypic and phenotypic cluster analyses of virulence determinants and possible role of CRISPR elements towards their incidence in *Enterococcus faecalis* and *Enterococcus faecium*. *Syst. Appl. Microbiol.* 34:553–60

Liu F, Barrangou R, Gerner-Smidt P, Ribot EM, Knabel SJ, et al. 2011a. Novel virulence gene and clustered regularly interspaced short palindromic repeat (CRISPR) multilocus sequence typing scheme for subtyping of the major serovars of *Salmonella enterica* subsp. *enterica*. *Appl. Environ. Microbiol.* 77:1946–56

Liu F, Kariyawasam S, Jayarao B, Barrangou R, Gerner-Smidt P, et al. 2011b. Subtyping *Salmonella* serovar Enteritidis isolates from different sources using sequence typing based on virulence genes and CRISPRs. *App. Environ. Microbiol.* 77:4520–26

Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, et al. 2006a. Comparative genomics of the lactic acid bacteria. *PNAS* 103:15611–16

Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 2006b. A putative RNA-interference–based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1:7

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

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

Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. 2013. RNA-guided human genome engineering via Cas9. *Science* 339:823–26

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

McShan WM, Ferretti JJ, Karasawa T, Suvorov AN, Lin S, et al. 2008. Genome sequence of a nephritogenic and highly transformable M49 strain of *Streptococcus pyogenes*. *J. Bacteriol.* 190:7773–85

Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–40

Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60:174–82

Mokrousov I, Limeschenko E, Vyazovaya A, Narvskaya O. 2007. *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. *Biotechnol. J.* 2:901–6

Mokrousov I, Vyazovaya A, Kolodkina V, Limeschenko E, Titov L, Narvskaya O. 2009. Novel macroarray-based method of *Corynebacterium diphtheriae* genotyping: evaluation in a field study in Belarus. *Eur. J. Clin. Microbiol. Infect. Dis.* 28:701–3

Nuñez JK, Lee AS, Engelman A, Doudna JA. 2015. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519:193–98

Oh JH, van Pijkeren JP. 2014. CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* 42:e131

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

Pennisi E. 2013. The CRISPR craze. *Science* 341:833–36

Pettengill JB, Timme RE, Barrangou R, Toro M, Allard MW, et al. 2014. The evolutionary history and diagnostic utility of the CRISPR-Cas system within *Salmonella enterica* ssp. *enterica*. *Peer J.* 2:e340

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

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

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, et al. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–83

Relman DA. 2011. Microbial genomics and infectious diseases. *N. Engl. J. Med.* 365:347–57

Rezzonico F, Smits THM, Duffy B. 2011. Diversity, evolution and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. *Appl. Environ. Microbiol.* 77:3819–29

Riehm JM, Vergnaud G, Kiefer D, Damdindorj T, Dashdavaa O, et al. 2012. *Yersinia pestis* lineages in Mongolia. *PLOS ONE* 7:e30624

Selle K, Barrangou R. 2015. Harnessing CRISPR-Cas systems for bacterial genome editing. *Trends Microbiol.* 23:225–32

Selle K, Klaenhammer TR, Barrangou R. 2015. CRISPR-based screening of genomic island excision events in bacteria. *PNAS* 112:8076–81

Semenova E, Nagornykh M, Pyatnitskiy M, Artamonova, II, Severinov K. 2009. Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. *FEMS Microbiol. Lett.* 296:110–16

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, et al. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343:84–87

Shariat N, DiMarzio MJ, Yin S, Dettinger L, Sandt CH, et al. 2013a. 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.* 34:164–73

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

Shariat N, Kirchner MK, Sandt CH, Trees E, Barrangou R, Dudley EG. 2013b. Subtyping of *Salmonella enterica* serovar Newport outbreak isolates by CRISPR-MVLST and determination of the relationship between CRISPR-MVLST and PFGE results. *J. Clin. Microbiol.* 51:2328–36

Shariat N, Sandt CH, DiMarzio MJ, Barrangou R, Dudley EG. 2013c. CRISPR-MVLST subtyping of *Salmonella enterica* subsp. *enterica* serovars Typhimurium and Heidelberg and application in identifying outbreak isolates. *BMC Microbiol.* 13:254

Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley E. 2014. Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology* 161:374–86

Sontheimer E, Barrangou R. 2015. The bacterial origins of the CRISPR genome editing revolution. *Hum. Gene Ther.* 26:413–24

Sun CL, Barrangou R, Thomas BC, Horvath P, Fremaux C, Banfield JF. 2013. Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol.* 15:463–70

Sun H, Li Y, Shi X, Lin Y, Qiu Y, et al. 2015. Association of CRISPR/Cas evolution with *Vibrio parahaemolyticus* virulence factors and genotypes. *Foodborne Pathog. Dis.* 12:68–73

Timme RE, Pettengill JB, Allard MW, Strain E, Barrangou R, et al. 2013. Phylogenetic diversity of the enteric pathogen *Salmonella enterica* subsp. *enterica* inferred from genome-wide reference-free SNP characters. *Genome Biol. Evol.* 5:2109–23

Toro M, Cao G, Ju W, Allard M, Barrangou R, et al. 2014. Association of clustered regularly interspaced short palindromic repeat (CRISPR) elements with specific serotypes and virulence potential of shiga toxin–producing *Escherichia coli*. *Appl. Environ. Microbiol.* 80:1411–20

Touchon M, Charpentier S, Clermont O, Rocha EPC, Denamur E, Branger C. 2011. CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J. Bacteriol.* 193:2460–67

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

Tremblay CL, Charlebois A, Masson L, Archambault M. 2013. Characterization of hospital-associated lineages of ampicillin-resistant *Enterococcus faecium* from clinical cases in dogs and humans. *Front. Microbiol.* 4:245

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

Ventura M, Turroni F, Lima-Mendez G, Foroni E, Zomer A, et al. 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. *Appl. Environ. Microbiol.* 75:6929–36

Very KJ, Kirchner MF, Shariat N, Cottrell W, Sandt CH, et al. 2015. Prevalence and spacial distribution of *Salmonella* infections in the Pennsylvania common raccoon (*Procyon lotor*). *Zoonoses Pub. Health* doi: 10.1111/zph.12222

Wehnes CA, Rehberger TG, Barrangou R, Smith AH. 2014. Determination of *Salmonella* clustered regularly interspaced short palindromic repeats (CRISPR) diversity on dairy farms in Wisconsin and Minnesota. *J. Dairy Sci.* 97:6370–77

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

Yin S, Jensen MA, Bai J, Debroy C, Barrangou R, Dudley EG. 2013. The evolutionary divergence of shiga toxin–producing *Escherichia coli* is reflected in clustered regularly interspaced short palindromic repeat (CRISPR) spacer composition. *Appl. Environ. Microbiol.* 79:5710–20

Zhang J, Abadia E, Refregier G, Tafaj S, Boschiroli ML, et al. 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of "spoligotyping" with new spacers and a microbead-based hybridization assay. *J. Med. Microbiol.* 59:285–94