REVIEW ARTICLE | SERIES
https://doi.org/10.1038/s41594-018-0173-y
nature structural & molecular biology

CRISPR RNA-guided autonomous delivery of Cas9

Royce A. Wilkinson, Coleman Martin, Artem A. Nemudryi and Blake Wiedenheft*

Cas9 is an endonuclease that can be programed to autonomously deliver diverse effectors to specified genetic addresses. High-resolution structures of this protein and its associated CRISPR RNA guide explain the molecular mechanisms of CRISPR-RNA-guided DNA recognition and provide a molecular blueprint that has facilitated structure-guided functional remodeling. Here we retrace events that led from early efforts to understand the central role of Cas9 in CRISPR-mediated adaptive immunity to contemporary efforts aimed at developing and deploying this enzyme for programmable genetic editing.

Recording information in DNA is the cornerstone of biology and the basis of genetic inheritance. In 2005, three groups independently reported that CRISPR (clustered regularly interspaced short palindromic repeats) loci in bacterial and archaeal genomes often contain short fragments of DNA derived from "transmissible genetic elements such as bacteriophages and conjugative plasmids"[1–3]. These observations led the authors to hypothesize that CRISPR loci were central components of a heritable, nucleic-acid-based immune system for protection from 'genetic aggressors'. In 2007, Barrangou et al. tested this hypothesis by challenging *Streptococcus thermophilus* with lytic phages. The authors showed that the surviving bacteria often contained new phage-derived 'spacer' sequences in the CRISPR locus. Deletion of specific spacers resulted in bacterial sensitivity to the corresponding phages. On rare occasions, phages were isolated that escaped CRISPR-mediated protection. These 'escape' phages contained mutations in the complementary target sequence (protospacer) or in a sequence adjacent to the protospacer, the protospacer-adjacent motif (PAM)[4–7]. Collectively, these experiments established a direct link between the DNA sequence at CRISPR loci and sequence-specific phage resistance (Box 1).

CRISPR loci are critical, but insufficient, for defense against phages[4]. They are often flanked by a diverse cassette of CRISPR-associated (*cas*) genes. Early experiments performed in *S. thermophilus* demonstrated that mutations in *cas9* (originally called *cas5*) result in loss of phage resistance, even when the CRISPR locus contains a spacer complementary to the invading phage[4]. This suggested that Cas9 is necessary for CRISPR-mediated phage defense. To investigate the mechanism of protection and the fate of phage DNA, Garneau et al. sequenced viral DNA isolated from infected bacterial cells and showed that both strands of the target DNA were cleaved within the spacer sequence, resulting in a blunt-ended cleavage product[8]. Although this work clarified the mechanism of defense, that is, cleavage of invading phage DNA, the mechanism for loading CRISPR RNAs (crRNAs) into Cas9 remained enigmatic. In 2011, Emmanuelle Charpentier's laboratory reported the identification of a trans-activating crRNA (tracrRNA) with sequence complementarity to repeat sequences of the crRNA[9]. The study showed that processing of the long primary CRISPR transcript into spacers for Cas9 delivery was dependent on the tracrRNA, Cas9, and an endogenous RNase III enzyme (Box 1). Subsequently, Jinek et al. purified Cas9 protein from *Streptococcus pyogenes* and showed that Cas9-mediated cleavage of double-stranded (ds)DNA relied on both the crRNA-guide and the tracrRNA[10]. To simplify this two-RNA system, Jinek et al. fused the 3′ end of the crRNA to the 5′ end of the tracrRNA to generate a single chimeric RNA and demonstrated that this single-guide RNA (sgRNA) could target Cas9 to cleave virtually any DNA sequence by design (Box 1 and Fig. 1). Similarly, Gasiunas et al. reported purification of the Cas9 protein from *S. thermophilus* and demonstrated programmable cleavage of dsDNA targets[11]. Together, this work demonstrated that Cas9 is a programmable nuclease, with potential for applications in genome engineering.

Science generally advances at a steady pace, one result leading to the next experiment, and years of work eventually guide the field to fundamental new understandings. However, work on CRISPR systems, and Cas9 in particular, has been anything but usual. In 2013, less than six months after the two reports on programmable cleavage of dsDNA by Cas9[10,11], three papers provided in vivo proof that RNA-guided Cas9 nucleases could be used to edit genes in both mouse and human cell lines[12–14]. Evidence of a new, easy-to-use genome-editing tool captured the attention of scientists across a wide range of different disciplines, and structural biologists raced to provide a molecular understanding of Cas9 function. In this review, we illustrate how mechanistic insights, particularly those gained from structures of *S. pyogenes* Cas9 (SpCas9), have facilitated engineering of the Cas9 ribonucleoprotein (RNP) to improve the genome editing and to impart new functions.

Structural insights into the Cas9 catalytic mechanism

The first structures of Cas9 were published in 2014, and currently there are 30 Cas9 structures from six different organisms deposited in the protein databank[15–22]. Cas9 proteins are diverse, ranging in length from 950 to 1,650 amino acids, and some share as little as 6% amino acid sequence identity (12% similarity)[23]. Despite this diversity, all Cas9 proteins share a bi-lobed structure, composed of nuclease (NUC) and recognition (REC) lobes (Fig. 1)[15,19]. The NUC lobe contains two nuclease domains (that is, HNH and RuvC) and a C-terminal PAM-interacting domain (PI). The REC lobe is composed of three α-helical domains (REC1–3) and a long arginine-rich helix (called the bridge helix, BH), which connects to the two lobes.

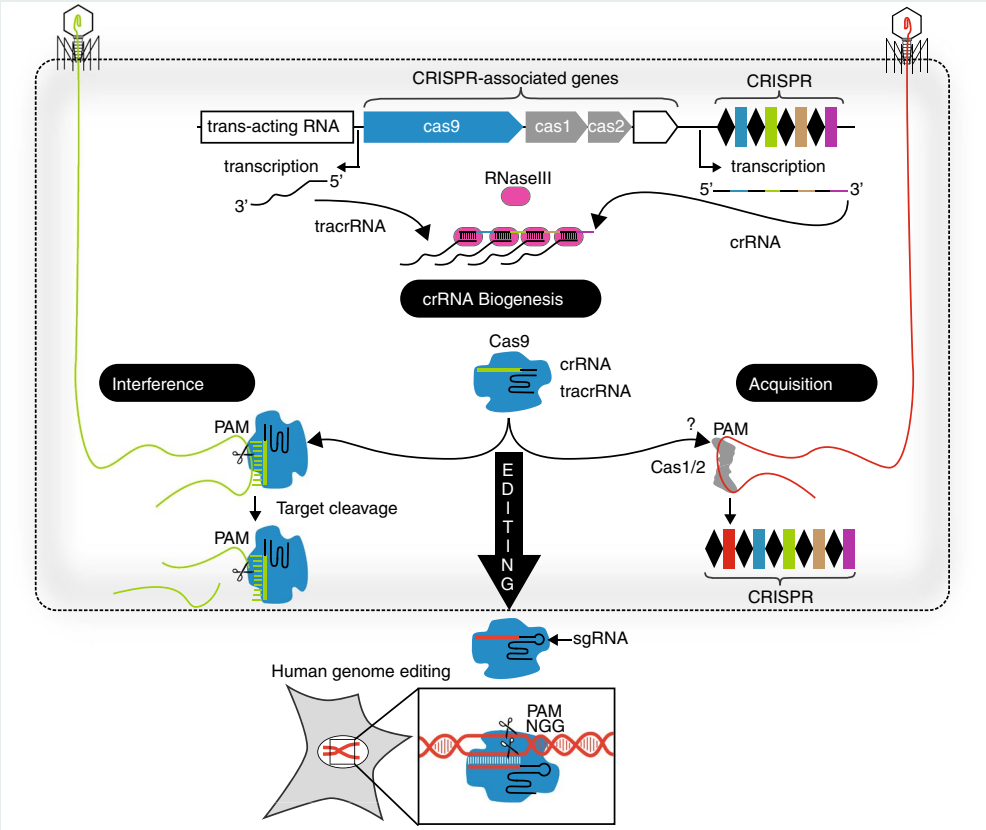
The structure of Cas9 alone was originally described as adopting an 'autoinhibited' conformation, such that the nuclease domains are inaccessible or inactive[15]. However, dynamic measurements of Cas9 performed using Förster resonance energy transfer (FRET), high-speed atomic force microscopy, and molecular dynamics simulations have shown that Cas9 is conformationally flexible before binding the RNA guide[24–26]. This flexibility may explain recent biochemical results suggesting that Cas9 from both *S. pyogenes* (SpCas9) and *Francisella novicida* (FnCas9) exhibit non-sequence-specific nuclease activity in the absence of an RNA guide[27].

Loading of either an sgRNA or the tracrRNA + crRNA pair induces large-scale conformational changes that create extensive contacts between the RNA and the enzyme (Fig. 1b,c and Supplementary Video 1). These interactions coincide with a reduction in Cas9 flexibility and a substantial reduction in non-sequence-specific

Department of Microbiology and Immunology, Montana State University, Bozeman, Montana, USA. *e-mail: bwiedenheft@gmail.com

Box 1 | Overview of CRISPR-mediated immunity and applications in editing

In type-II CRISPR-Cas systems, acquisition of new spacers requires Cas1 and Cas2 proteins, as well as Cas9. Although the specific function of Cas9 in the new sequence acquisition has not been well established, it does have a role in selecting prespacers that contain PAMs[110–112]. During interference, CRISPR loci are transcribed and each repeat in the CRISPR RNA base-pairs with a small portion of the tracrRNA. The resulting RNA duplexes are cleaved by RNaseIII, and the mature crRNA-tracrRNA hybrids are loaded into Cas9[9]. Cas9 scans dsDNA for PAMs, and PAM binding destabilizes the DNA duplex for crRNA-guided sampling of the adjacent sequence[29]. Hybridization between the crRNA-guide and target DNA triggers a conformational change in Cas9 that activates the RuvC and HNH nuclease domains[24,25,31,35,101], leading to the cleavage of the invading phage DNA. Jinek et al. showed that the tracrRNA and the crRNA hybrid could be covalently connected with a short linker and that the resulting sgRNA could be used to target cleavage of any dsDNA adjacent to a PAM by design[10]. This programmable nuclease has now been used for a wide variety of genome-engineering applications in cells from all domains of life.



nuclease activity[17,27]. RNA binding by Cas9 results in a 65-Å rigid-body migration of the REC3 domain from one end of the REC lobe to the other and a 60-Å rigid-body migration of the entire NUC lobe (Fig. 1c)[15,17]. Movement of the NUC lobe orients the PI domain for PAM detection, and the BH is wedged between the tracrRNA and the crRNA (Fig. 1b,c). This positions the BH perpendicular to the RNA duplex formed by complementary regions of the tracrRNA + crRNA pair and spatially separates the tracrRNA from the crRNA guide sequence. One side of the BH interacts with the first nine nucleotides of the crRNA guide (that is, the seed region), whereas the other side of the BH makes contacts with the first stem-loop of the tracrRNA (Fig. 1b). Collectively, the RNA forms a triangular structure that completely encases the BH. The guide, tetraloop, and stem-loop 1 of the RNA bind the REC lobe (K_d = 0.7 nM), whereas stem-loops 2 and 3 have high affinity (K_d = 0.2 nM) for the NUC lobe. RNA binding brings the two lobes together, restricting conformational flexibility[28].

To understand the mechanism of crRNA-guided DNA interrogation and nuclease activation, several labs have determined structures of the SpCas9 RNP bound to short fragments of DNA[16,19,20]. A structure of the Cas9 bound to a 23-nucleotide fragment of single-strand (ss)DNA complementary to the crRNA-guide reveals rigid-body rotations in both REC2 and REC3 domains that are necessary to accommodate the complementary target strand while avoiding steric clashes. These rotations are accompanied by a 25-Å movement of the HNH domain toward the DNA-RNA duplex, although the active site is oriented away from the DNA (Fig. 1c)[17,19]. However, this ssDNA target does not contain a PAM. Anders et al. subsequently determined the crystal structure of SpCas9 bound to a partially duplexed DNA target and a duplexed TGG PAM sequence, which explained the molecular mechanism of PAM recognition. The guanine bases of the PAM are recognized by nucleotide-specific hydrogen bonds to a pair of arginine residues (R1333 and R1335) that extend into the major groove (Figs. 1c and 2a)[20]. Notably, FnCas9 also recognizes an NGG PAM, and the structure reveals a similar mechanism that relies on a pair of arginines, although the arginine residues are on distinct secondary structural elements[21]. Although the PI domains of SpCas9 and FnCas9 share a conserved fold, the arginines responsible for PAM recognition in FnCas9 are located on two different loops

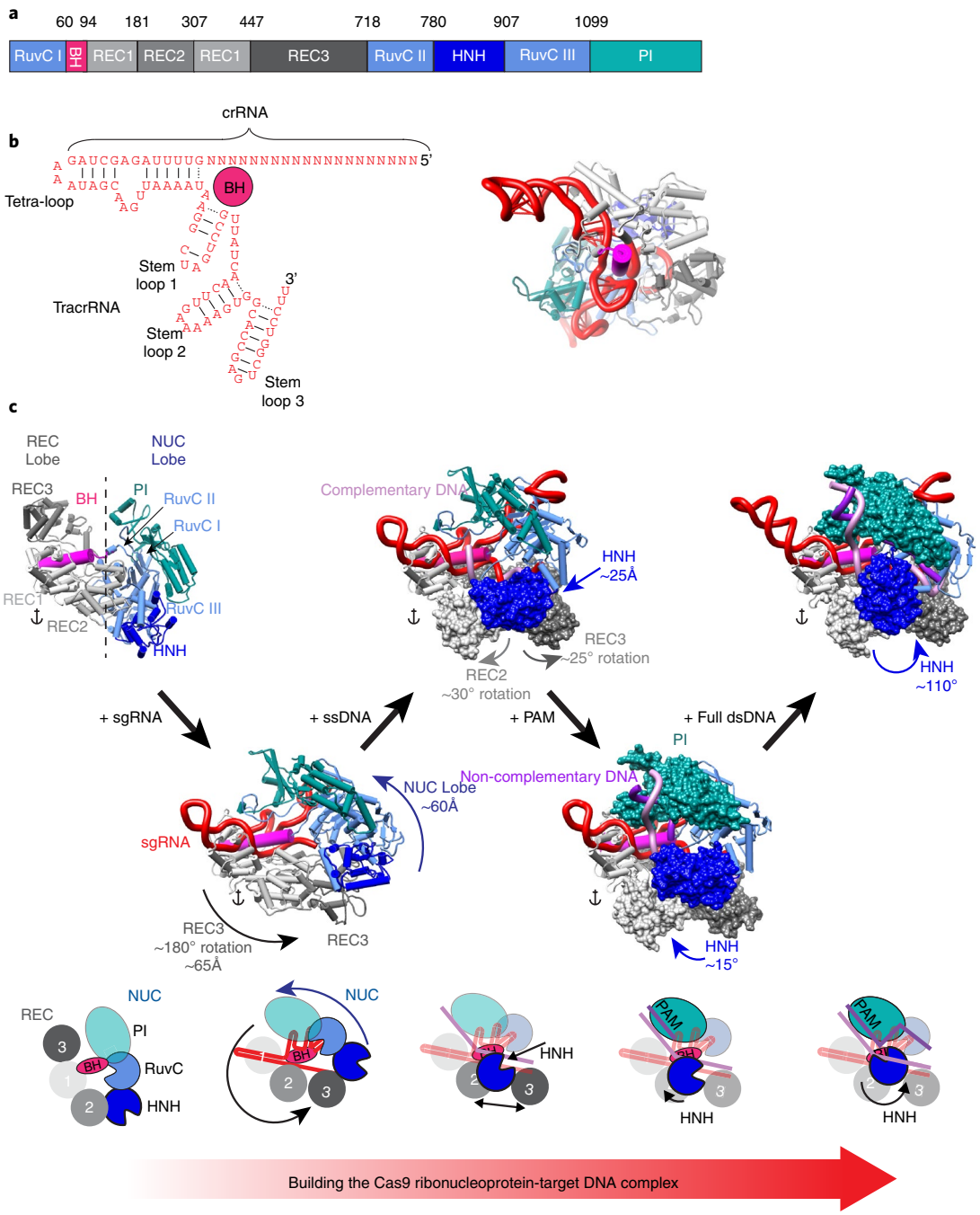


Fig. 1 | The impact of sgRNA-loading and target DNA-binding on Cas9 structure. a, Domain organization of the Cas9 protein. **b**, Left, the sequence and secondary structure of the sgRNA. The location of the BH is highlighted in magenta. Right, a ribbon diagram of the Cas9 RNP oriented to highlight RNA interactions with the BH (magenta). **c**, The Cas9 protein has a crescent shape composed of two lobes. Three recognition domains (REC1–3, gray) and the BH (magenta) form the REC-lobe. The NUC lobe is composed of two nuclease domains (RuvC, light blue, and HNH, deep blue) and a PI domain (green). Five crystal structures (Cas9 alone (4CMP); with sgRNA (4ZT0); with sgRNA and complementary ssDNA (4OO8); with sgRNA, complementary ssDNA, and dsDNA PAM (4UN3); and with sgRNA and dsDNA containing a PAM and protospacer (5F9R)) highlight major conformational changes at each stage of assembly.

that are separated by 29 residues, whereas the arginines in SpCas9 are separated by a single amino acid[20,21].

Finally, binding a dsDNA target containing a PAM and a protospacer results in a 110° rotation of the HNH active site toward the scissile phosphate on the complementary DNA, suggesting that crRNA-guided hybridization allosterically regulates the HNH domain (Fig. 1c)[16]. However, in this structure, active sites of the HNH and RuvC nuclease domains remain 10 Å and 5.5 Å from their respective scissile phosphates[16]. Notably, both HNH and RuvC are metal-ion-dependent phosphodiesterase domains (one-metal and two-metal mechanisms, respectively), and EDTA was used to prevent DNA cleavage during complex reconstitution and subsequent crystallization. This locked the complex in a precleavage state. However, the metal ions may also facilitate active site engagement with the DNA target, which cannot be observed under these conditions. A structure of Cas9 with engaged active sites will help clarify mechanisms of cleavage, and we anticipate that non-

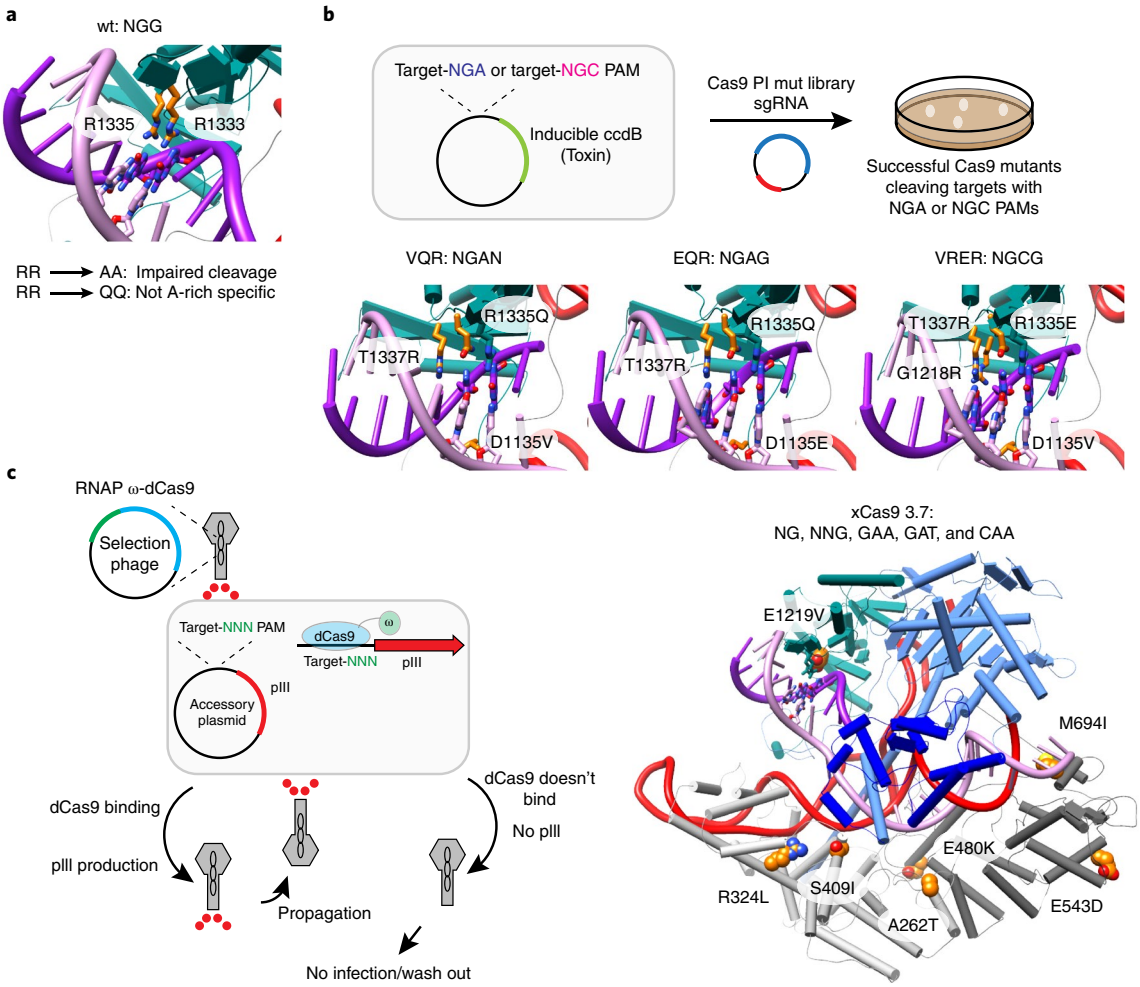


Fig. 2 | Engineering Cas9 to change PAM specificity. a, A pair of arginine residues (R1333 and R1335) in the PI domain of SpCas9 extend into the major groove, making nucleotide-specific hydrogen bonds with guanine bases on the noncomplementary strand (PDB: 4UN3). **b**, A screen for Cas9 mutants with altered PAM specificity using a library of Cas9 PI domain mutants identified the VQR, EQR, and VRER mutants, which recognize the NGAN, NGAG, and NGCG PAMs, respectively. Crystal structures 5B2R (VQR), 5B2S (EQR), and 5B2T (VRER) show the interactions between the mutant Cas9s and altered PAMs. **c**, The ω-dCas9 fusion protein was programmed to bind a protospacer upstream of the phage pIII gene, and a phage-assisted continuous-evolution protocol was used to isolate a family of Cas9 mutants that bind to diverse PAM sequences (NNN, green). Cas9 mutants that bind the protospacer activate expression of pIII, which encodes an essential phage tail fiber protein (G3P). Phages expressing G3P will package mutant *cas9* and amplify it by reinfection. Absence of G3P results in phage particles with attachment defects, which are lost from the library. Mutations with broadened PAM specificities can induce pIII expression from a greater number of the NNN PAM sequences. Amino acid mutations in the xCas9 3.7 variant, which has the broadest PAM binding while maintaining efficient DNA cleavage rates, are shown as spheres (PDB: 4UN3).

cleavable modifications to the phosphate backbone may help trap the metal-dependent nuclease domains in action.

Target identification and nuclease activation proceed through a series of sequential steps. Cas9 searches the genomic landscape for PAMs via a combination of both one-dimensional sliding and three-dimensional diffusion[29,30]. PAM binding destabilizes the DNA duplex and facilitates directional crRNA-guided strand invasion that initiates at the PAM and proceeds toward the 5′ end of the guide[29]. Mismatches between the crRNA-guide and the complementary DNA result in binding and cleavage defects, which are position-specific. PAM-proximal mismatches in the seed region lead to rapid dissociation, while mismatches at the opposite end (PAM-distal) have much less of an impact on binding. However, binding alone is insufficient for nuclease activation, and recent work indicates that complete (or near-complete) base pairing is necessary for nuclease activation[31–34]. Target hybridization positions the HNH nuclease domain in the active conformation, which then allosterically activates the RuvC nuclease domain[31,35]. Allosteric activation is transmitted by formation of an extended α-helix between the two domains (S909–N940), and mutations that inhibit helix formation block RuvC cleavage[31]. Following cleavage of both DNA strands, Cas9 remains tightly bound to the two ends of the DNA. Thus, SpCas9 is a single-turnover enzyme[29,35], although this may not be true for all Cas9 enzymes. In fact, Yourik et al. recently reported that the Cas9 from *Staphylococcus aureus* is a multiturnover enzyme[36].

Engineering altered PAM specificity. The requirement for PAM binding increases target specificity, which limits off-target binding. However, strict PAM requirements also reduce the programmable versatility of the enzyme. Several strategies have been used to alter PAM specificity. Different Cas9 proteins have distinct PAM recognition profiles, and grafting the PI domain from *S. thermophilus* Cas9 (NGGNG PAM) onto SpCas9 (NGG PAM) resulted in a chimera with specificity for the desired NGGNG PAM[19]. However, these two proteins are closely related (60% identity), and a PAM-swap strategy may not be generalizable to more diverse Cas9 proteins. To expand

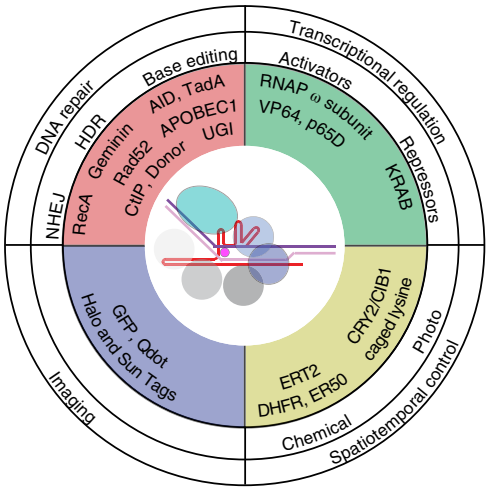


Fig. 3 | Cas9 for autonomous delivery of diverse effectors. Cas9 activity can be enhanced or modified by association with various effectors. Effectors are fused directly to Cas9 or tethered to Cas9 through RNA aptamers integrated into the sgRNA. In general, effectors have been used to influence DNA repair, control the timing and location of Cas9 activity, regulate transcription, or image chromosomes.

the PAM-binding repertoire of SpCas9 from NGG to NAA, Anders et al. borrowed a trick previously used to alter binding sites of homing endonucleases[20]. Given that the major groove recognition of adenosines often involves bidentate hydrogen bonding with either asparagine or glutamine, the authors mutated the SpCas9 arginines (R1333 and R1335) to glutamines in an attempt to broaden targeting to an A-rich PAM (Fig. 2a)[20]. However, this structure-guided engineering effort failed, resulting in Cas9 proteins with DNA-binding defects.

As an alternative approach to expand the PAM-binding repertoire, Kleinstiver et al. used a library encoding Cas9 proteins with random mutations in the PI domain to identify Cas9 mutants with distinct PAM preferences[37]. The Cas9 library was programed to target a protospacer in the *ccdB* gene, which encodes a toxin targeting DNA gyrase, flanked by either an NGA or NGC PAM. Cells with a Cas9 mutant capable of cleaving the toxic DNA survived, and sequencing *cas9* alleles in the survivors revealed three predominate mutations (that is, VQR, EQR, and VRER), which recognize NGAN, NGAG, and NGCG PAMs, respectively (Fig. 2b). DNA cleavage efficiencies for these mutants were similar to wild-type Cas9 cleavage rates (5–53%), and off-target double-strand breaks (DSBs) were comparable to those reported for wild-type Cas9. Notably, a D1135E mutation of Cas9 (a mutation present in the EQR variant) also increases the preference of Cas9 for NGG over NGA PAMs. Structures of these Cas9 mutants bound to their respective PAMs reveal multiple changes that synergistically displace the phosphodiester backbone of the PAM duplex in a way that reorients bases for sequence-specific recognition in the major groove[38]. A similar screening approach was used to identify Cas9 mutants from *S. aureus* with relaxed PAM specificity (NNGRRT to NNNRRT)[39].

To evolve SpCas9 variants with promiscuous PAM compatibility and high DNA specificity, Hu et al. recently performed a phage-assisted continuous evolution experiment designed to derive variants of SpCas9 with an expanded PAM (xCas9; Fig. 2c)[40]. In this setup, *Escherichia coli* phages encode a mutant library of catalytically dead Cas9 proteins that are fused to the ω subunit of bacterial RNA polymerase (ω-dCas9). The ω-dCas9 is programmed to target a protospacer upstream of a gene encoding a phage-attachment protein (that is, G3P), which is necessary to form infectious virions. However, the PAM sequence is intentionally scrambled (NNN), so only phages carrying ω-dCas9 mutants capable of promiscuous PAM recognition are able to drive expression of the phage-attachment protein. This approach identified a mutant (xCas9 3.7) with broadened PAM specificity (NGG, NG, GAA, and GAT). The E1219V mutation carried by this variant is near the PAM binding site and arose in the first round of evolution. However, it is difficult to rationalize the impact of many of the other mutations on PAM recognition, although some are near the RNA-DNA duplex and may have an impact on target binding.

Programmable repair. Cas9-induced DNA DSBs are often repaired by nonhomologous end joining (NHEJ), an efficient repair mechanism that 'stitches' the free ends of the DNA back together. However, NHEJ is error-prone, often resulting in insertions or deletions, which can perturb gene function and generate functional gene knockouts. Knockouts are critical for determining the biological function of a specific gene, but the next major challenge is to create methods that enable efficient and precise sequence modification of prescribed genes. Homology-directed repair (HDR) of double-stranded DNA breaks, using DNA repair templates flanked by sequences homologous to regions on either side of the break, has proven effective for introducing new sequences at specified locations, but the efficiency of HDR is low relative to NHEJ.

To alter the cellular NHEJ/HDR DNA repair ratio, Cas9 has been fused to various effector proteins (Fig. 3). Genetically tethering Cas9 to *E. coli* Rec A, an ssDNA binding protein with a central role in homologous recombination, resulted in an unexpected increase in NHEJ in human cells[41]. Although Rec A from *E. coli* did not have the desired outcome, fusions to eukaryotic enzymes that have a critical role in HDR (that is, Rad52 or CtIP) enhances HDR by two- to threefold[42,43]. Given that many of the proteins necessary for HDR are only active in late S and G2 phases of the cell cycle, synchronizing Cas9 expression with activation of the HDR machinery also enhances templated repair. This has been accomplished by fusing Cas9 to the N-terminal region of geminin, a protein that is subject to ubiquitination and degradation in phases of the cell cycle during which the HDR machinery is not active[44].

In addition to fusing Cas9 to proteins involved in HDR, DNA repair templates have also been hitched to Cas9 or sgRNA for delivery to the cleavage location. Single-stranded DNA donors have been linked to synthetic crRNAs using 'click' chemistry[45]. In addition, sgRNA containing a streptavidin-binding aptamer designed to bind biotinylated DNA donors has been used[46]. Alternatively, Cas9 has been fused to SNAP-tags, which covalently react with O^6-benzylguanine-labeled DNA repair templates[47]. This results in a 20-fold increase in the HDR/NHEJ ratio, with a typical increase of two- to fourfold in total HDR levels.

Base editing is an elegant alternative to efforts that rely on HDR (Fig. 3)[48–50]. Fusing dCas9 to cytidine deaminase results in irreversible conversion of C to U, which is read as T by DNA polymerase[48,49]. However, cytidine deamination can occur spontaneously in cells, and G:U mismatches are efficiently repaired by cellular uracil DNA glycosylase (UDG), which restores the original G:C pair. To improve G:C to A:T editing efficiencies, researchers fused cytidine deaminase and a UDG inhibitor protein (UGI) to a single dCas9 molecule, leading to a fourfold improvement in C-to-T conversion[48,49]. Base-editing efficiency is further improved by switching from dCas9 to a Cas9 nickase (nCas9), which only cleaves the complementary strand. Eukaryotic mismatch repair uses nicks in newly synthesized DNA to determine which strand is used as the repair template. Nicking the complementary (G-containing) strand forces the cell to use the edited U as the template and further increases the desired C:G to T:A conversion by two- to fourfold[48,49]. To broaden the applicability and versatility of base editing, Gaudelli et al. created an adenosine deaminase–nCas9 fusion for base-editing A:T pairs to G:C[50]. Using directed evolution of a naturally occurring

bacterial tRNA adenosine deaminase, they generated a novel adenosine deaminase that works on DNA substates[50].

Cas9 as a governor of gene expression. Catalytically dead Cas9 (dCas9) is routinely used to control expression of specific genes without permanently altering the DNA sequence (Fig. 3). Hitching transcriptional activators, such as VP6, p65AD, or the ω subunit of RNAP, to dCas9 has been successful in promoting gene expression in both eukaryotes[51–53] and bacteria[54]. As an alternative approach, integrating RNA aptamers into the sgRNA facilitates functionalization of the Cas9 RNP without requiring protein engineering. RNA aptamers have been inserted in the tetraloop, stem-loop 2, and the 3′ end of sgRNA with little impact on RNA loading into Cas9 (Fig. 1b). Fusing aptamer-binding proteins to transcriptional activators then allows activation of targets specified by the guide RNA[55].

Methods for repressing gene expression have also been developed. In fact, simply targeting dCas9 to the 5′ untranslated region or coding sequence of a gene of interest inhibits transcription[56]. Linking a transcriptional repressor domain, such as the KRAB (Krüppel-associated box) domain, to dCas9 improves gene silencing further and can be used to target promoter regions as well as the 5′ untranslated region or coding sequence[51].

Engineering complex genetic circuits requires sophisticated control methods that permit simultaneous repression, activation, or DNA cleavage at multiple locations in a single cell. For this purpose, several guides can be used to deliver Cas9 to multiple genetic loci in a single cell (multiplexing). Additional layers of complexity can be built into the sgRNA by including different aptamer sequences in sgRNA loops as described above and by altering the guide length. Complete complementarity between the RNA-guide and the DNA target is necessary to activate the Cas9 nuclease domains, while short guides (14 or 15 nucleotides) result in specific binding without cleavage. Catalytically active Cas9 can thus be sent to several genetic addresses, and the length of the RNA-guide can be used to control Cas9-mediated gene activation or repression versus DNA cleavage[57,58].

Finally, epigenetic control of gene expression is a key factor in fundamental biological processes such as development and cellular differentiation. It also has an important function in human diseases, including cancer, and has potential roles in aging and age-related diseases[59,60]. dCas9-guided epigenetic modifications (for example, cytosine methylation[61,62], histone demethylation[63], and histone acetylation or deacetylation[64,65]) can be used to experimentally test cause-and-effect relationships associated with these modifications. That said, recent evidence suggests that at least the dCas9-methylase enzymes have significant off-target effects at sites with as little as a five-nucleotide seed match. They can even cause hypermethylation on a global level as a result of high expression levels of the methylase and its non-sequence-specific interaction with DNA or other DNA-binding proteins[66,67]. The substantial off-target binding observed agrees with FRET and early dCas9 chromatin-immunoprecipitation sequencing experiments observing stable dCas9 binding to sites with up to 11 PAM-distal mismatches[31,32,34,68]. Conversely, the Cas9-KRAB repressor showed no substantial off-target transcriptional silencing by RNA-sequencing[51]. This might be a result of the fact that repression requires binding near the promoter or transcription start site, and off-target binding at other loci may have no impact on RNA levels[69]. Pflueger et al. recently reported a modular, noncovalent coupling of dCas9 to DNA methyltransferase using the SunTag system, a multimeric peptide display tag that can be bound by single-chain antibodies fused to the desired effector protein (for example, DNA methyltransferase)[70]. This permits independent control of expression levels for Cas9 and the methylase and can be used to substantially reduce off-target methylation by limiting methylase expression levels. High on-target methylation is retained through multimeric recruitment of the methylase to the dCas9-SunTag[70]. Collectively, these results caution that off-target impacts of dCas9-guided enzymatic effectors should be examined on a genome-wide scale[71].

Box 2 | CRISPR-anti-CRISPR conflict

Bacteria and archaea have evolved CRISPR–Cas-based adaptive immune systems for protection from phage predation, and phages have evolved anti-CRISPRs that neutralize these immune systems[113,114]. The competing interests of host fitness and viral replication creates a dynamic landscape of selective pressures that 'incentivizes' genetic variation and innovation[115]. The diversity of currently recognized CRISPR–Cas systems have been partitioned into 33 different subtypes[116], and according to a recent review by Bondy-Denomy et al., there are "36 distinct families of anti-CRISPRs described in the literature that block seven subtypes of CRISPR-Cas systems"[117]. Nine of these anti-CRISPRs target Cas9 proteins, and seven have been used to control Cas9 activity for applications in editing[113,118–120].

Spatiotemporal control of Cas9. Programmable delivery of Cas9 to genetic targets in the cell is relatively robust, but in vivo delivery of Cas9 to a specific subset of cells for therapeutic intervention remains a major challenge. A recent review by Lino et al. provides an overview of the various delivery vehicles reported for CRISPR–Cas9[72]. Below we highlight a few methods that impart spatial or temporal control of Cas9 activity.

The CRISPR-Cas9 genome editing system has been modified for use in conjunction with adeno-associated virus (AAV). AAV is highly infectious, even in nondividing cells. AAV serotypes preferentially infect specific cell types, providing a natural platform for delivery to major organs, including the heart, kidney, liver, lung, pancreas, muscle, and central nervous system. Another advantage of AAV is its ssDNA genome, as single-stranded DNA donors result in higher levels of template-directed repair[73]. Although there are advantages to using AAV, cargo size is a clear limitation. The total foreign DNA cargo is limited to less than 4.4 kb. Although SpCas9 and a chimeric sgRNA (~4.2 kb) are just below the limit, this leaves little room for additional regulatory elements or an HDR donor. To overcome the AAV packaging limit, Cas9 can be split into two intein-tagged polypeptides for delivery by two viral constructs. Inteins are protein sequences analogous to mRNA introns that are able to excise and then join two polypeptides (for example, two halves of Cas9) with a peptide bond[74–76]. As an alternative to inteins, Wright et al. designed a split-Cas9 system in which the NUC and REC lobes are expressed as separate polypeptides that can be delivered in separate AAVs[28]. Here the sgRNA serves as a regulatory agent that 'glues' the two halves together, creating a functional RNP. In contrast to efforts aimed at controlling assembly of split Cas9, the activity of Cas9 also has been regulated by controlling subcellular localization or by controlling Cas9 stability[77–81]. Fusing Cas9 to the hormone-binding domain of the estrogen receptor enables spatial control through ligand-induced nuclear translocation of Cas9 in the presence of 4-hydroxytamoxifen[81]. Similarly, Cas9 has been fused to destabilization domains that are stabilized in the presence of small molecules (for example, trimethoprim), leading to small-molecule-induced Cas9 activity[79,80].

Another approach for controlling Cas9 activity is fusing Cas9 to proteins that dimerize upon exposure to light[77,78]. Most photoregulators of Cas9 reported to date rely on blue light, which has phototoxic effects on mammalian cells and limited penetration through turbid human tissues[77,78,82]. However, Shao et al. recently developed a far-red-light-activated CRISPR-dCas9 effector system that has deep

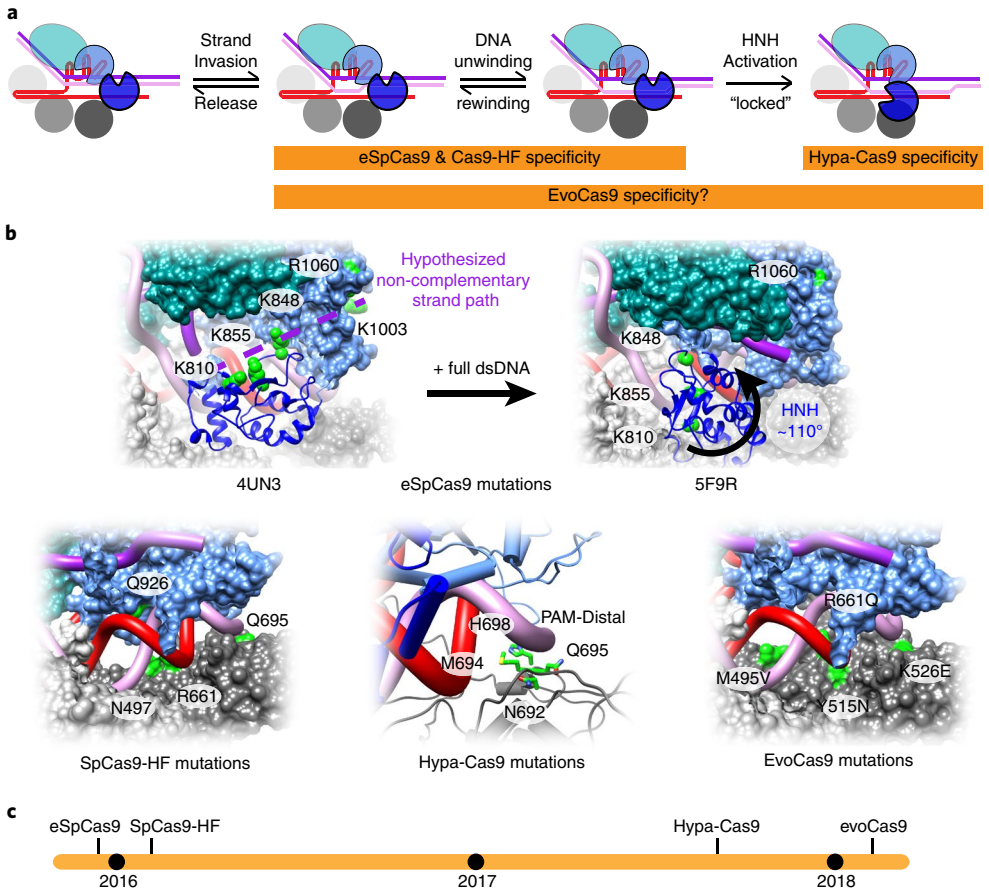


Fig. 4 | Structure-guided design of Cas9 for enhanced target-specificity. a, Cas9 target verification is a multistep process. After binding to a double-stranded PAM, the guide RNA invades the target duplex and begins base-pairing with the complementary strand at the PAM-proximal end (seed region). Hybridization proceeds toward the PAM-distal end, checking for mismatches. Once the DNA is fully unwound, the HNH domain can enter the active conformation and allosterically activate the RuvC domain for cleavage. Several Cas9 mutants have been created to improve the specificity of Cas9. **b**, Positively charged residues on the nuclease domains of Cas9 were predicted to interact with the noncomplementary DNA strand (PDB: 4UN3). A series of alanine substitutions were designed to limit non-sequence-specific interactions with the displaced strand and enhanced-specificity Cas9 (eSpCas9). However, a subsequent structure of SpCas9 bound to dsDNA revealed a conformational change that precludes many of these predicted interactions (PDB: 5F9R). The 4UN3 structure was also used to create Cas9-HF1, with mutations that eliminate interactions with the phosphate backbone of the complementary DNA strand. Mutations in eSpCas9 and Cas9-HF alter the balance between unwinding and rewinding of the DNA duplex in the presence of mismatched targets. Amino acids in the REC3 domain at the PAM-distal end of the RNA-DNA duplex act as an allosteric regulator of HNH nuclease activation. Mutation of these residues to alanine produced hyperaccurate Cas9 (HypaCas9). Screening for mutants with increased specificity using a library of Cas9s with random mutations in REC3 resulted in evolved Cas9 (evoCas9), which may improve specificity by reducing stabilizing interactions with the complementary strand (that is, increasing the off-rate similarly to Cas9-HF1) and by altering allosteric regulation of the HNH nuclease domain (similarly to mutations in HypaCas9). **c**, Timeline of Cas9 mutants.

tissue-penetration capacity and negligible phototoxicity, which may improve precision therapies[83].

Finally, virally encoded anti-CRISPRs (Acr) have also been used to control Cas9 (Box 2). Typically, Acr expression is used to suppress Cas9 activity. However, Bubeck et al. recently fused the anti-CRISPR protein AcrIIA4 to the LOV2 photosensor domain[84]. The N- and C- termini of LOV2 are in close proximity in the dark, but photoexcitation unfolds the terminal helix. By taking advantage of recently published structure of AcrIIA4 bound to SpCas9[85,86], the authors strategically inserted LOV2 in a loop of AcrIIA4 such that light-induced unfolding of LOV2 perturbs the fold of AcrIIA4 and inhibits anti-CRISPR function, which results in Cas9 activation.

In an effort to target Cas9 to the liver, Rouet et al. decorated Cas9 with synthetic ligands that bind asialoglycoprotein receptors expressed on hepatocytes[87]. This approach results in preferential Cas9 uptake in hepatocytes, but the protein is frequently trapped in the endosome, which limits editing efficiencies. To liberate Cas9 from the endosome, the ligand-decorated Cas9 was co-transfected with endosomolytic peptides, which improved editing, but the efficiency nevertheless remains low (5%). Collectively, the emerging repertoire of delivery systems that impart spatiotemporal control on Cas9 is progressing toward minimizing risks associated with editing DNA in specific cell types in vivo.

Chromosomal imaging and manipulation. There is growing interest in understanding the three-dimensional organization of the cellular genome and its impact on gene expression[88]. Cas9 has been used to label genetic loci either by fusion of fluorescent proteins directly to Cas9 or by using RNA aptamers to attach fluorescent proteins or quantum dots (Fig. 3)[46,89,90]. Chen et al. used this approach to monitor telomere dynamics following DNA damage[90]. Beyond imaging, Morgan et al. developed a method for inducing chromosomal looping by tethering dCas9 from *S. aureus* and *S. pyogenes* to two proteins that heterodimerize in the presence of S-(+)-abscisic acid (ABA)[91].

The dCas9 proteins were programmed to bind spatially separated regions of the chromosome, and looping was induced by ABA. Using this approach, the authors demonstrated that the promoter for a gene could be 'recruited' to another regulatory sequence for activation.

Designing Cas9 and the RNA-guide for higher specificity. The clinical success of Cas9 will depend on the specificity of these enzymes. One of the first examples of enhanced specificity was the use of Cas9 nickases. Wild-type Cas9 cleaves both strands of the DNA duplex, while HNH or RuvC mutants act as nickases that cleave a single DNA strand. Thus, generating DSBs with Cas9 nickases requires two paired enzymes that cleave adjacent locations on opposite strands of the genome. The requirement for two independent nicks at adjacent locations reduces on-target efficiency, but it also minimizes the risk of off-target DSBs[55]. Similarly, specificity can be increased by fusing dCas9 to a dimerization-dependent FokI nuclease. Similar to Cas9 nickases, DNA cleavage requires colocalization of two independently programmed dCas9 molecules to facilitate dimerization and activation of the FokI nuclease[92,93].

The 'two-hit' requirement for introducing DSBs (for example, Cas9 nickases or dCas9-FokI fusions) is not practical in many situations where delivery of one Cas9 is already a challenge, particularly in the context of using Cas9 RNPs. To improve specificity of the nuclease-active Cas9, Slaymaker et al. created a panel of enhanced-specificity Cas9s by making a series of structure-guided mutations predicted to destabilize interactions with the noncomplementary DNA strand (Fig. 4b)[94]. Protein-mediated interactions with the displaced strand facilitate crRNA-guided strand invasion by limiting the backreaction (i.e., re-annealing of the DNA duplex). Thus, mutations that make the displaced strand more accessible for re-annealing to the complementary DNA strand were predicted to increase stringency for on-target interactions by promoting DNA rewinding. Similar engineering strategies have been used to create a series of structure-guided mutations that eliminate interactions with the complementary strand. Eliminating these protein contacts results in higher-fidelity Cas9 (Cas9-HF1; Fig. 4b)[95]. Chen et al. hypothesized that there must be a mechanism for detecting complete RNA–DNA heteroduplex formation resulting in transition of the HNH domain to the active conformation. They identified residues in the REC3 domain near the PAM-distal end of the RNA-DNA duplex that could interact with the fully hybridized duplex. Mutation of four of these residues to alanine resulted in hyperaccurate Cas9, with improvements in specificity similar to what has been reported for enhanced-specificity Cas9s and SpCas9-HF1 (Fig. 4b)[96]. Finally, screening for mutants with increased specificity using a Cas9 library with random mutations in REC3 resulted in evolved Cas9 (evoCas9) (Fig. 4b)[97]. EvoCas9 may improve specificity through both mechanisms, since it combines mutations that are predicted to restrict HNH activation and mutations to residues that interact with the complementary DNA strand, similarly to SpCas9-HF1 (Fig. 4a).

In addition to protein modifications, truncated RNA guides (17 or 18 nucleotides rather than 20) have also been used to decreased off-target nuclease activity with minimal impact on cleavage at the intended site[98,99]. Complete base pairing is necessary to activate the nuclease domains and single-molecule FRET (smFRET) experiments indicate that Cas9 spends threefold less time in the catalytically active conformation when loaded with a 17-nucleotide guide. A single PAM-distal mismatch with a 17-nucleotide guide prevents HNH activation and cleavage entirely[100,101]. Further, structural insights into the interactions between Cas9 and the 2'-OH of the RNA ribose have been used to identify crRNA or tracrRNA positions amenable to incorporation of DNA nucleotides or other 2'-OH modifications into the guide RNA that result in decreased off-target activity[102–105]. Incorporation of DNA residues may increase specificity due to the reduced stability of DNA-DNA duplexes compared to DNA-RNA duplexes[102,103]. However, the mechanism of improvement seen with other modifications is unclear, and the level of improvement varies at different off-target sites[104,105]. Finally, incorporating bicyclic bridged or 'locked' nucleic acids increases mismatch-discrimination in nucleic acid duplexes. These modified bases can be incorporated into synthetic guide RNAs, particularly at key positions of known off-target sites or single-nucleotide polymorphisms[106].

Future directions

Structural insights will continue to have a fundamental role in understanding Cas9 function. For example, a high-resolution structure of Cas9 with the active sites engaged on both strands of the DNA target is likely to provide additional mechanistic insights that might aid in the design of more efficient Cas9 tools with fewer off-target effects. Most Cas9 proteins studied thus far remain bound to the DNA target after cleavage, which may be advantageous in the context of an immune response, because it prevents access of cellular DNA repair and replication enzymes to the viral genome. However, Clarke et al. recently demonstrated that RNA polymerase ejects Cas9 from DSBs in human cells, turning the single-turnover Cas9 into a multiturnover enzyme, which improves the efficiency of editing[107]. A similar effect was observed in bacteria[56], where a twofold enrichment of crRNAs targeting the template versus nontemplate strand of invading phage genomes was noted in *S. thermophilus* CRISPR arrays[107]. This suggests that crRNA targeting of the template strand may provide a protective advantage to the host.

In addition to enhancing editing efficiencies, there may be additional incentives for targeting the template strand in genome-engineering experiments. Cas9-generated DSBs have recently been shown to stall the eukaryotic cell cycle in G1 in a p53-dependent manner[108]. Long-lived Cas9 binding following DNA cleavage may reduce HDR by delaying DSB repair until the cells arrest in G1 when the HDR machinery is not available. This means that Cas9-mediated DSBs will be repaired primarily by NHEJ in cells that have a functional p53 pathway. Thus, targeting the template strand may reduce the DNA-bound half-life of Cas9, allowing template-dependent editing before the cells can arrest in G1. Alternatively, we hypothesize that Cas9 mutants that destabilize DNA interactions after cleavage may also increase HDR in untransformed ($p53^+$) cells by allowing repair before G1 arrest. A structure of the post-Cas9 cleavage complex may help guide the design of these 'quick-release' mutants.

Although current base-editors provide a method for editing without creating a cellular DSB, the deaminase is prone to target multiple adenines or cytosines within a small window on the displaced strand (the R-loop), which limits specificity. Modifications that alter placement of the deaminase or the conformational flexibility may further restrict the deaminase to a more specific location. Alternatively, engineered deaminases with sequence-specific activity can also reduce 'bystander' activity[109]. In addition, further improvements in base-editing may expand the technology from transitions (C-to-T or A-to-G) to transversion mutations.

Applications of Cas9-based genome editing have considerable implications for bioengineering, agriculture, and healthcare, although many challenges remain. Although our understanding of the rules that govern the autonomous delivery of Cas9 to specified genetic addresses continue to improve, delivery of Cas9 to specific cells in multicellular organisms, potentially including human patients, and switching repair from unpredictable indels to programmable modification at specific loci remains difficult. We anticipate that continued improvements that limit off-target activity will allow for programmable genetic manipulations, and delivery of Cas9 to specific cell types will lead to a technology that will reshape our genetic landscape.

Received: 22 August 2018; Accepted: 13 November 2018;
Published online: 31 December 2018

References

1. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).
2. Pourcel, C., Salvignol, G. & Vergnaud, G. 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 (2005).
3. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).
4. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
5. Deveau, H. et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390–1400 (2008).
6. Horvath, P. et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1401–1412 (2008).
7. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).
8. Garneau, J. E. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
9. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
10. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
Cas9 is a dual-RNA-guided endonuclease that creates dsDNA-breaks, and the two RNAs can be linked by a tetraloop into a chimeric single-guide RNA.
11. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **109**, E2579–E2586 (2012).
12. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
13. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
14. Jinek, M. et al. RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
15. Jinek, M. et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014).
This paper provides the first crystal structure of SpCas9 without guide RNA or target DNA.
16. Jiang, F. et al. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867–871 (2016).
This crystal structure of active Cas9 with sgRNA and 30-bp double-strand target DNA reveals the rotation of the NHN domain toward the DNA.
17. Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477–1481 (2015).
The crystal structure of sgRNA-bound active Cas9 shows extensive protein rearrangement upon RNA binding and preordering of the 10-nt seed region.
18. Nishimasu, H. et al. Crystal structure of *Staphylococcus aureus* Cas9. *Cell* **162**, 1113–1126 (2015).
19. Nishimasu, H. et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
This study provides the crystal structure of Cas9 in complex with an sgRNA and a 23-nt single-strand target DNA.
20. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).
This study provides the crystal structure of sgRNA-Cas9 bound to a complementary DNA strand and a duplexed PAM. This structure explains the mechanism of PAM recognition.
21. Hirano, H. et al. Structure and engineering of *Francisella novicida* Cas9. *Cell* **164**, 950–961 (2016).
22. Yamada, M. et al. Crystal structure of the minimal Cas9 from *Campylobacter jejuni* reveals the molecular diversity in the CRISPR-Cas9 systems. *Mol. Cell* **65**, 1109–1121.e3 (2017).
23. Fonfara, I. et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* **42**, 2577–2590 (2014).
24. Shibata, M. et al. Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nat. Commun.* **8**, 1430 (2017).
25. Osuka, S. et al. Real-time observation of flexible domain movements in CRISPR-Cas9. *EMBO J.* **37**, e96941 (2018).
26. Palermo, G., Miao, Y., Walker, R. C., Jinek, M. & McCammon, J. A. CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. *Proc. Natl. Acad. Sci. USA* **114**, 7260–7265 (2017).
27. Sundaresan, R., Parameshwaran, H. P., Yogesha, S. D., Keilbarth, M. W. & Rajan, R. RNA-independent DNA cleavage activities of Cas9 and Cas12a. *Cell Reports* **21**, 3728–3739 (2017).
28. Wright, A. V. et al. Rational design of a split-Cas9 enzyme complex. *Proc. Natl. Acad. Sci. USA* **112**, 2984–2989 (2015).
29. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
This paper clarifies the mechanism target identification, showing that the search process is primarily governed by three-dimensional diffusion.
30. Globyte, V. et al. CRISPR Cas9 searches for a protospacer adjacent motif by one-dimensional diffusion. *EMBO J.* (in the press).
31. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* **527**, 110–113 (2015).
These authors demonstrate that complete base pairing to the target is an allosteric regulator of the HNH and RuvC nuclease domains.
32. Singh, D., Sternberg, S. H., Fei, J., Doudna, J. A. & Ha, T. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* **7**, 12778 (2016).
33. Yang, M. et al. The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET. *Cell Reports* **22**, 372–382 (2018).
34. Singh, D. et al. Mechanisms of improved specificity of engineered Cas9s revealed by single-molecule FRET analysis. *Nat. Struct. Mol. Biol.* **25**, 347–354 (2018).
35. Gong, S., Yu, H. H., Johnson, K. A. & Taylor, D. W. DNA unwinding is the primary determinant of CRISPR-Cas9 activity. *Cell Reports* **22**, 359–371 (2018).
36. Yourik, P., Fuchs, R.T., Mabuchi, M., Curcuru, J.L. & Robb, G.B. *Staphylococcus aureus* Cas9 is a multiple-turnover enzyme. *RNA* rna.067355.118 (2018).
37. Kleinstiver, B. P. et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
38. Hirano, S., Nishimasu, H., Ishitani, R. & Nureki, O. Structural basis for the altered PAM specificities of engineered CRISPR-Cas9. *Mol. Cell* **61**, 886–894 (2016).
39. Kleinstiver, B. P. et al. Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **33**, 1293–1298 (2015).
40. Hu, J. H. et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
41. Lin, L. et al. Fusion of SpCas9 to *E. coli* Rec A protein enhances CRISPR-Cas9 mediated gene knockout in mammalian cells. *J. Biotechnol.* **247**, 42–49 (2017).
42. Shao, S. et al. Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing *Saccharomyces cerevisiae* Rad52. *Int. J. Biochem. Cell Biol.* **92**, 43–52 (2017).
43. Charpentier, M. et al. CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nat. Commun.* **9**, 1133 (2018).
44. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Reports* **14**, 1555–1566 (2016).
45. Lee, K. et al. Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *eLife* **6**, e25312 (2017).
46. Carlson-Stevermer, J. et al. Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711 (2017).
47. Savic, N. et al. Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *eLife* **7**, e33761 (2018).
48. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
Cas9 fusion to rat APOBEC1 cytidine deaminase and the uracil glycosylase inhibitor enables programmable base editing in cells without dsDNA cleavage.
49. Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016).
Fusion of Cas9 to activation-induced cytidine deaminase and the uracil glycosylase inhibitor enables programmable base editing in cells without dsDNA cleavage.
50. Gaudelli, N. M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

51. Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
52. Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).
53. Maeder, M. L. et al. CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).
54. Bikard, D. et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013).
55. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).
There are two main points in this paper: insertion of aptamer sequences into the guide-RNA enables site-specific recruitment of transcriptional activators; and paired Cas9 nickases reduce off-target cleavage activity.
56. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
This study provides the first demonstration of using nuclease inactive Cas9 (dCas9) for transcriptional repression in bacteria.
57. Dahlman, J. E. et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat. Biotechnol.* **33**, 1159–1161 (2015).
58. Kiani, S. et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* **12**, 1051–1054 (2015).
59. Paradise, B. D., Barham, W. & Fernandez-Zapico, M. E. Targeting epigenetic aberrations in pancreatic cancer, a new path to improve patient outcomes?. *Cancers (Basel)* **10**, E128 (2018).
60. Song, S. & Johnson, F. B. Epigenetic mechanisms impacting aging: a focus on histone levels and telomeres. *Genes (Basel)* **9**, E201 (2018).
61. Lei, Y. et al. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat. Commun.* **8**, 16026 (2017).
62. Liu, X. S. et al. Editing DNA methylation in the mammalian genome. *Cell* **167**, 233–247.e17 (2016).
63. Kearns, N. A. et al. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).
This study provides the first demonstration of dCas9-targeted epigenetic modification using the LSD1 histone demethylase to define enhancer elements in mouse embryonic stem cells.
64. Hilton, I. B. et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).
This paper demonstrated that dCas9 can be used control cellular phenotypes by targeted manipulation of histone acetylation.
65. Kwon, D. Y., Zhao, Y. T., Lamonica, J. M. & Zhou, Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat. Commun.* **8**, 15315 (2017).
66. Lin, L. et al. Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. *Gigascience* **7**, 1–19 (2018).
67. Galonska, C. et al. Genome-wide tracking of dCas9-methyltransferase footprints. *Nat. Commun.* **9**, 597 (2018).
68. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32**, 677–683 (2014).
69. Gilbert, L. A. et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014).
70. Pflueger, C. et al. A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res.* **28**, 1193–1206 (2018).
71. Xu, X. & Qi, L. S. A CRISPR-dCas toolbox for genetic engineering and synthetic biology. *J. Mol. Biol.* **S0022-2836**(18), 30666–1 (2018).
72. Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257 (2018).
73. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).
74. Fine, E. J. et al. Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes. *Sci. Rep.* **5**, 10777 (2015).
75. Truong, D. J. et al. Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res.* **43**, 6450–6458 (2015).
76. Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A. & Liu, D. R. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat. Chem. Biol.* **11**, 316–318 (2015).
77. Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H. & Sato, M. CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* **22**, 169–174 (2015).
78. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198–200 (2015).
79. Balboa, D. et al. Conditionally stabilized dCas9 activator for controlling gene expression in human cell reprogramming and differentiation. *Stem Cell Reports* **5**, 448–459 (2015).
80. Maji, B. et al. Multidimensional chemical control of CRISPR-Cas9. *Nat. Chem. Biol.* **13**, 9–11 (2017).
81. Liu, K. I. et al. A chemical-inducible CRISPR-Cas9 system for rapid control of genome editing. *Nat. Chem. Biol.* **12**, 980–987 (2016).
82. Müller, K. et al. A red/far-red light-responsive bi-stable toggle switch to control gene expression in mammalian cells. *Nucleic Acids Res.* **41**, e77 (2013).
83. Shao, J. et al. Synthetic far-red light-mediated CRISPR-dCas9 device for inducing functional neuronal differentiation. *Proc. Natl. Acad. Sci. USA* **115**, E6722–E6730 (2018).
84. Bubeck, F. et al. Engineered anti-CRISPR proteins for optogenetic control of CRISPR-Cas9. *Nat. Methods* **15**, 924–927 (2018).
85. Shin, J. et al. Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* **3**, e1701620 (2017).
86. Yang, H. & Patel, D. J. Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9. *Mol. Cell* **67**, 117–127.e5 (2017).
87. Rouet, R. et al. Receptor-mediated delivery of CRISPR-Cas9 endonuclease for cell-type-specific gene editing. *J. Am. Chem. Soc.* **140**, 6596–6603 (2018).
88. Davies, J. O., Oudelaar, A. M., Higgs, D. R. & Hughes, J. R. How best to identify chromosomal interactions: a comparison of approaches. *Nat. Methods* **14**, 125–134 (2017).
89. Wang, S., Su, J. H., Zhang, F. & Zhuang, X. An RNA-aptamer-based two-color CRISPR labeling system. *Sci. Rep.* **6**, 26857 (2016).
90. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
91. Morgan, S. L. et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* **8**, 15993 (2017).
92. Tsai, S. Q. et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32**, 569–576 (2014).
93. Guilinger, J. P., Thompson, D. B. & Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **32**, 577–582 (2014).
94. Slaymaker, I. M. et al. Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
95. Kleinstiver, B. P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
96. Chen, J. S. et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).
97. Casini, A. et al. A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* **36**, 265–271 (2018).
98. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
99. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
100. Lim, Y. et al. Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. *Nat. Commun.* **7**, 13350 (2016).
101. Dagdas, Y. S., Chen, J. S., Sternberg, S. H., Doudna, J. A. & Yildiz, A. A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Sci. Adv.* **3**, o0027 (2017).
102. Rueda, F. O. et al. Mapping the sugar dependency for rational generation of a DNA-RNA hybrid-guided Cas9 endonuclease. *Nat. Commun.* **8**, 1610 (2017).
103. Yin, H. et al. Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat. Chem. Biol.* **14**, 311–316 (2018).
104. Rahdar, M. et al. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc. Natl. Acad. Sci. USA* **112**, E7110–E7117 (2015).
105. Hendel, A. et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015).
106. Cromwell, C. R. et al. Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. *Nat. Commun.* **9**, 1448 (2018).
107. Clarke, R. et al. Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. *Mol. Cell* **71**, 42–55.e8 (2018).
108. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).
109. Gehrke, J. M. et al. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.***36**, 977–982 (2018).
110. Heler, R. et al. Mutations in Cas9 enhance the rate of acquisition of viral spacer sequences during the CRISPR-Cas immune response. *Mol. Cell* **65**, 168–175 (2017).

111. Heler, R. et al. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* **519**, 199–202 (2015).
112. Wei, Y., Terns, R. M. & Terns, M. P. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev.* **29**, 356–361 (2015).
113. Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The discovery, mechanisms, and evolutionary impact of anti-CRISPRs. *Annu. Rev. Virol.* **4**, 37–59 (2017).
114. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: discovery, mechanism and function. *Nat. Rev. Microbiol.* **16**, 12–17 (2018).
115. Carter, J., Hoffman, C. & Wiedenheft, B. The interfaces of genetic conflict are hot spots for innovation. *Cell* **168**, 9–11 (2017).
116. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
This paper provides an evolutionary road map for CRISPR-Cas classification.
117. Bondy-Denomy, J. et al. A unified resource for tracking anti-CRISPR names. *The CRISPR Journal* **1**, 304–305 (2018).
118. Pawluk, A. et al. Naturally occurring off-switches for CRISPR-Cas9. *Cell* **167**, 1829–1838.e9 (2016).
119. Rauch, B. J. et al. Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell* **168**, 150–158.e10 (2017).
120. Hynes, A. P. et al. Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. *Nat. Commun.* **9**, 2919 (2018).

Acknowledgements

Research in the Wiedenheft lab is supported by the National Institutes of Health (P20GM103500, P30GM110732, R01GM110270, R01GM108888, and R21 AI130670), the National Science Foundation EPSCoR (EPS-110134), the M. J. Murdock Charitable Trust, a young investigator award from Amgen, and the Montana State University Agricultural Experimental Station (USDA NIFA).

Competing interests

B.W. is the founder of SurGene LLC and is an inventor on patent applications related to CRISPR-Cas systems and applications thereof.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41594-018-0173-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to B.W.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.