Applied and Environmental Microbiology

Cas9-Based Tools for Targeted Genome Editing and Transcriptional Control

Tao Xu, Yongchao Li, Joy D. Van Nostrand, Zhili He and Jizhong Zhou

Appl. Environ. Microbiol. **2014, 80(5):1544. DOI: 10.1128/AEM.03786-13.**
Published Ahead of Print 3 January 2014.

AEM Journals.ASM.org

MINIREVIEW

Cas9-Based Tools for Targeted Genome Editing and Transcriptional Control

Tao Xu,[a,b] Yongchao Li,[a,b] Joy D. Van Nostrand,[a,b] Zhili He,[a,b] Jizhong Zhou[a,b,c,d]

Institute for Environmental Genomics[a] and Department of Microbiology and Plant Biology,[b] University of Oklahoma, Norman, Oklahoma, USA; Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA[c]; State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China[d]

Development of tools for targeted genome editing and regulation of gene expression has significantly expanded our ability to elucidate the mechanisms of interesting biological phenomena and to engineer desirable biological systems. Recent rapid progress in the study of a clustered, regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) protein system in bacteria has facilitated the development of newly facile and programmable platforms for genome editing and transcriptional control in a sequence-specific manner. The core RNA-guided Cas9 endonuclease in the type II CRISPR system has been harnessed to realize gene mutation and DNA deletion and insertion, as well as transcriptional activation and repression, with multiplex targeting ability, just by customizing 20-nucleotide RNA components. Here we describe the molecular basis of the type II CRISPR/Cas system and summarize applications and factors affecting its utilization in model organisms. We also discuss the advantages and disadvantages of Cas9-based tools in comparison with widely used customizable tools, such as Zinc finger nucleases and transcription activator-like effector nucleases.

In the postgenomic era, researchers are often overwhelmed by the enormous amount of genomic information available as a result of high-throughput sequencing technologies (1). Deciphering gene function and connecting genotype to phenotype have become primary challenges in utilizing these resources to engineer biological systems to relieve and address global challenges, such as environmental cleanup, clean energy production, and human disease treatment. To date, a variety of available tools have been applied to create genetic modifications in many organisms (2, 3). However, the demand for genetic engineering is transforming from targeting one site to targeting multiple sites in a single genome for efficient genome-scale engineering (2). The clustered, regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) protein system is an adaptive RNA-mediated immune system in approximately 40% of bacteria and ~90% of archaea (4). The CRISPR/Cas system can be reprogrammed to reject invading bacteriophages and conjugative plasmids (5, 6). Continued improvement in understanding the mechanisms of the type II CRISPR/Cas system launched the birth of novel programmable CRISPR/Cas9-based platforms, native Cas9 nuclease (Cas9)- or Cas9 nickase (Cas9n)-based targeted genome editing (7–16), and inactivated- or dead-Cas9 (dCas9)-based transcriptional control (11, 17–19). Cas9-based tools, thus far, have been successfully applied in diverse organisms and have shown great promise in realizing multiplex and efficient genome editing and regulation of gene expression without host dependence. Here, we review the molecular basis of the type II CRISPR/Cas system, the application of Cas9-based tools, and factors influencing their utilization. We also compare the advantages and limitations of Cas9-based tools with those of several widely used targeted tools, such as Zinc finger nucleases (ZFNs) (3, 20) and transcription activator-like effector nucleases (TALENs) (3, 21). In addition, we discuss possible strategies for further improvements.

Key abbreviations. All abbreviations are summarized in Table S1 in the supplemental material.

TYPE II CRISPR/Cas SYSTEM

The CRISPR/Cas system as an adaptive immune system (22) employs CRISPR RNA (crRNA)-guided Cas proteins to recognize target sites within the invader genome (known as protospacers) via base-pairing complementarity and then to cleave DNA within the protospacer sequences. It is classified into three types (I, II, and III) based on the sequence and structure of the Cas protein (23, 24). The crRNA-guided surveillance complexes in types I and III need multiple Cas subunits (25, 26); however, type II requires only Cas9 (27, 28). The type II system as a reduced system has been studied primarily in *Streptococcus* (27, 29) and *Neisseria* (30) (Fig. 1A and B) and also has been developed as a promising programmable tool. The native type II system requires at least three crucial components: RNA-guided Cas9 nuclease, crRNA, and a partially complementary *trans*-acting crRNA (tracrRNA) (27, 29, 30). Each of these components is discussed below.

Cas9 nuclease. Cas9 (formerly known as Csn1 or Csx12) is the first indispensable component of type II CRISPR/Cas systems and is able to cleave double-stranded DNA (dsDNA) in a sequence-specific manner (24, 27, 31). Although there are other *cas* genes (e.g., *cas1*, *cas2*, and *csn2*) present in a single genome, disruption of these other genes did not impair crRNA biogenesis (13, 27, 28). Cas9 is a large multidomain protein with two nuclease domains, a RuvC-like nuclease domain near the amino terminus and an HNH (or McrA-like) nuclease domain in the middle (5, 29). *In vitro* tests indicate that the endonuclease activity of *Streptococcus pyogenes* Cas9 creates blunt dsDNA breaks (DSBs) that are 3 bp upstream of the 3′-terminal complementarity region formed between the crRNA recognition sequence and the genomic proto-

Published ahead of print 3 January 2014

Address correspondence to Jizhong Zhou, jzhou@ou.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.03786-13.

doi:10.1128/AEM.03786-13

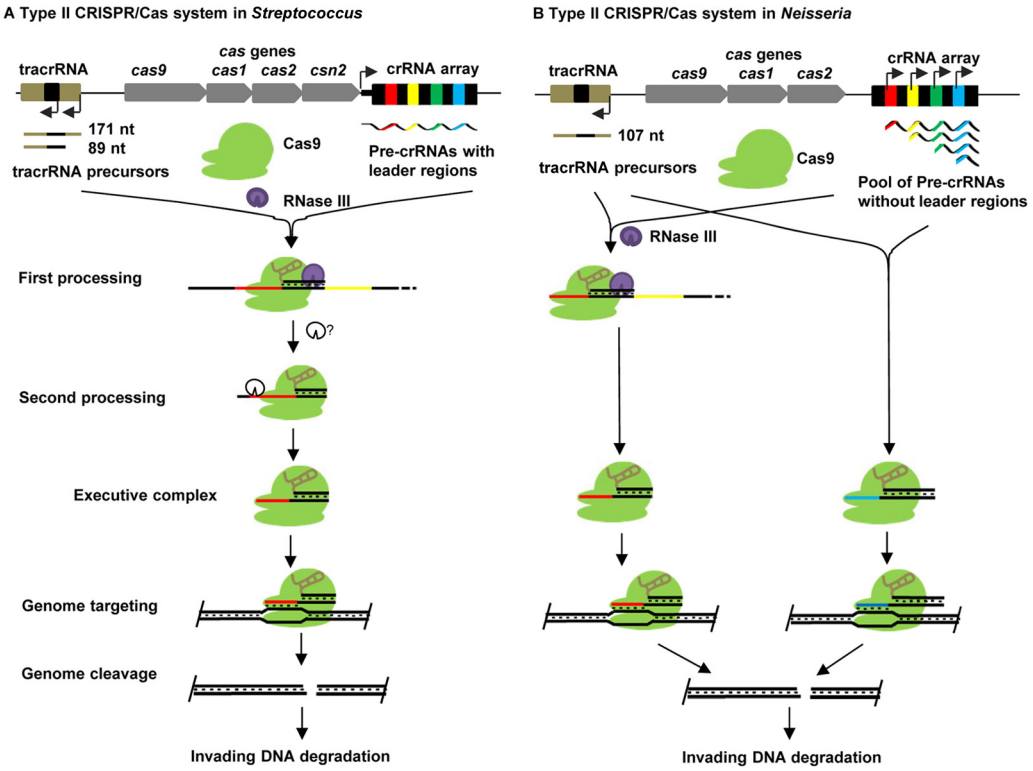

FIG 1 Type II CRISPR/Cas systems in *Streptococcus* (A) and *Neisseria* (B). (A) In *Streptococcus*, the type II system needs three major steps to accomplish target DNA cleavage. First, tracrRNA precursor and pre-crRNA transcripts are processed by RNase III (purple) in the presence of Cas9 (green) to split the crRNA array and shorten the tracrRNA precursor within the complementation regions formed between pre-crRNA repeats (black) and tracrRNA precursor antirepeats. Second, the spacer region (red, yellow, dark green, and blue) of crRNA is further trimmed by unknown RNases (white) to produce mature crRNA with a 20-nt target recognition region. Third, the tracrRNA-crRNA duplex is incorporated into Cas9, forming an executive complex to specify protospacers and create DSBs to degrade invading DNA. (B) In *Neisseria*, the type II system has two ways of degrading invading DNA. (Left) First, the pre-crRNAs are trimmed only at their 3′ ends, as with the first processing in streptococci, and then loaded into Cas9 for genome targeting. (Right) Second, the pre-crRNA pool interacting with tracrRNA precursors can be directly assembled with Cas9.

spacer (Fig. 2A) (5, 29). Mutagenesis of each catalytic site in the RuvC and HNH motifs abolishes the ability to create DSBs, leaving only nickase activity. Biochemically, the RuvC mutants (with D10A in *S. pyogenes* Cas9 and D31A in *Streptococcus thermophilus* Cas9) and HNH mutants (with H840A in *S. pyogenes* Cas9 and N891A in *S. thermophilus* Cas9) cut the noncomplementary and complementary strands, respectively, of the protospacer at the same positions as in the intact Cas9-crRNA complex (5, 16, 29), indicating that each active site acts on the opposite DNA strand to generate DSBs (5, 29). Intriguingly, mutations in these active sites did not alter the affinity of the CRISPR/Cas complex for binding the protospacer (29). Importantly, protospacer-adjacent motifs (PAMs) that are short conserved nucleotide stretches next to the protospacers, such as NGG (32), NGGNG (22), NAAR (32), and NNAGAAW (33), are absolutely necessary for Cas9 binding and cleavage (29). Orthogonal Cas9 nucleases from different microorganisms require different PAM sequences (31, 34). The recognition mechanism of the Cas9 protein for specific PAM sites is still unknown.

tracrRNA. tracrRNA is the second indispensable component of the type II CRISPR/Cas system and is a non-protein-coding RNA for crRNA maturation and subsequent DNA cleavage (35). In *S. pyogenes*, the tracrRNA gene is transcribed from two start sites producing two primary species of 171 nucleotides (nt) and 89 nt, both of which are processed into ~75-nt RNA species (27). In contrast, in *Neisseria*, the tracrRNA gene produces only one full-length 107-nt species, which is processed into a 91-nt form (30). The resulting tracrRNA precursors have a stretch of almost perfect (one mismatch) complementarity with each of the pre-crRNA repeats. The base-pairing RNA duplex is important for tracrRNA precursor trimming and crRNA maturation, as mentioned below (27, 36).

crRNA biogenesis in type II systems. Recent studies uncovered different crRNA maturation processes in type II systems (34, 37). For example, *S. pyogenes* and *Neisseria* spp. (Fig. 1A and B) present distinct pre-crRNA transcript features, pre-crRNA processing, and nucleoprotein complexes. *S. pyogenes* produces only one form of the full-length primary pre-crRNA of 511 nt, consisting of a leader region and a number of repeat-spacer-repeat units (27). In contrast, *Neisseria lactamica* produces a series of primary pre-crRNAs of different lengths (48 to 576 nt), because each CRISPR repeat carries a promoter element which can initiate independent transcription of the downstream array and none of the pre-crRNAs carry leader regions at their 5′ ends (30). Promoter-containing CRISPR repeats have also been found in *Neisseria meningitidis* and *Campylobacter jejuni* (30).

pre-crRNA processing varies as well. In *S. pyogenes*, a two-step crRNA biogenesis is used, with a first cleavage within the repeat regions and a second cleavage within the spacers (27). During the first cleavage, the base-pairing RNA duplex formed by the

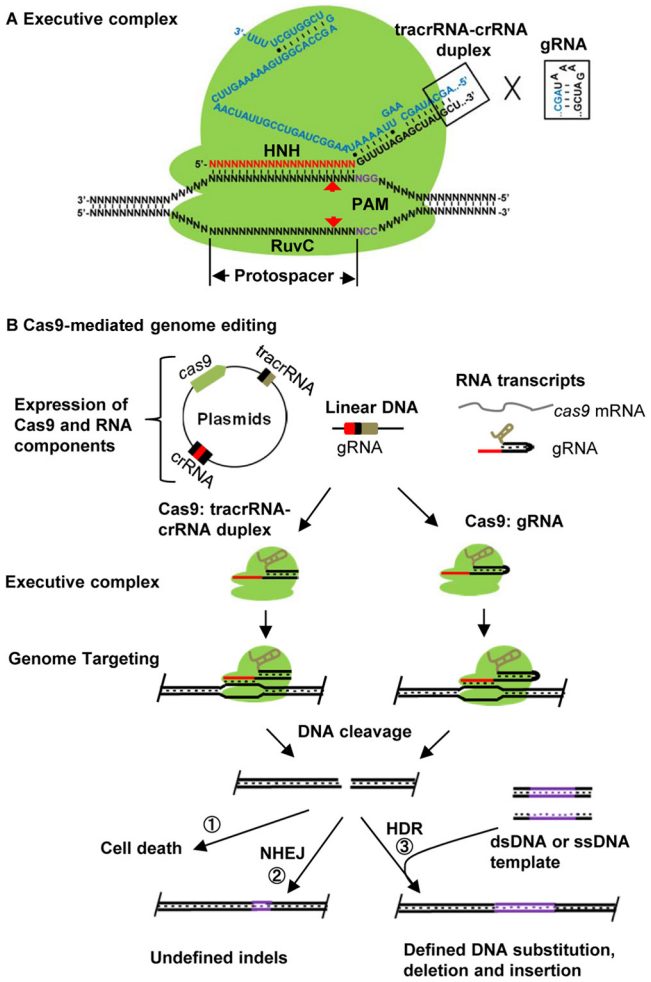

FIG 2 Application of CRISPR/Cas9 in targeted genome editing. (A) Widely used *S. pyogenes* Cas9 with the HNH and RuvC domains is directed by tracrRNA-crRNA duplexes or gRNA to cut the complementary or non-complementary strand, respectively. Cuts are made at the positions (indicated by red arrows) that are 3 bp upstream of PAM sites (purple characters). All components required for RNA-guided genome editing in foreign hosts are expressed by delivering coexpression plasmids, DNA expression cassette fragments, or sole RNA transcripts (B). Expressed tracrRNA-crRNA duplexes or gRNA is assembled with Cas9, generating executive complexes. These complexes generate breaks in the genome that may lead to cell death if the DSBs are not removed (①), induce error-prone nonhomologous end joining (NHEJ) to rejoin the ends and introduce undefined small deletions and additions (indels) (②), or trigger homology-directed repair (HDR) when homology-containing dsDNA or ssDNA templates are given (③) to confer precise DNA substitution, deletion, or insertion.

tracrRNA precursor and the pre-crRNA is attacked by the housekeeping RNase III within the repeats, generating a 75-nt tracrRNA and a 66-nt intermediate crRNA species. The coordinated action of RNase III and Cas9 is necessary to process the duplex, and the complementarity of the duplex is a prerequisite for RNase III-mediated coprocessing (27, 35). The second cleavage is assumed to depend on the Cas9-mediated ruler-type mechanism whereby the spacers are cleaved at a fixed distance using the first processing site as an anchor, generating 39- to 42-nt mature species carrying a unique 20-nt spacer sequence and a 19- to 22-nt repeat sequence (27). These spacers specifically determine cleavage sites in invading genomes. Strikingly, in *N. meningitidis*, pre-crRNAs are trimmed once at their 3′ end to produce 48-nt mature crRNAs by the coordinated action of RNase III and Cas9, as with the first cleavage in *S. pyogenes* but without the second cleavage within the spacers (30).

These processed RNA components are assembled with Cas9, forming executive nucleoprotein complexes that target and cleave the protospacer recognized by 20-nt spacer sequences in crRNAs. Obviously, the mature crRNAs employed in *S. pyogenes* and *Neisseria* spp. present different features. *S. pyogenes* crRNAs are trimmed at both their 5′ and 3′ ends, leaving 5′ phosphate and 3′ hydroxyl residues; however, *Neisseria* crRNAs are trimmed only at their 3′ end, resulting in 5′ triphosphate and 3′ hydroxyl ends. Interestingly, loss of *N. meningitidis* RNase III did not affect Cas9-mediated DNA cleavage *in vivo*, even though pre-crRNA and tracrRNA processing was abolished (30). This means that RNase III-mediated pre-crRNA processing is not required for interference activity in *Neisseria*. It appears that the *Neisseria* type II system is simpler than that in *Streptococcus*. However, recent applications of type II systems were derived mainly from *S. pyogenes* because fundamental studies in *Neisseria* are very recent.

APPLICATION OF TYPE II CRISPR/Cas SYSTEM

Due to the simplicity and customizability of type II CRISPR systems, host-independent gene-targeting platforms have been developed for genome editing and transcriptional control in both eukaryotes (7, 10, 12, 14–16, 38–42) and prokaryotes (13, 17, 19, 31) (see Table S2 in the supplemental material). In general, current applications of type II systems can be classified into three categories: native Cas9-mediated genome editing, Cas9 nickase-mediated genome editing, and inactivated Cas9-mediated transcriptional control. Promisingly, type II systems can also be engineered for high-throughput genome editing and silencing.

Native Cas9-mediated genome editing. Cas9-mediated genome editing depends on two sequential steps (Fig. 2B). First, genomic DNA is cleaved by Cas9 at a specific site determined by the 20-nt target recognition sequence in crRNA (5, 27). Second, DSBs are ligated by native DNA repair systems (43), native non-homologous end joining (NHEJ) (44), or template-dependent homology-directed repair (HDR) (45). NHEJ, as an error-prone process, often generates undefined small insertions and deletions (indels) during the repair process (12, 14, 16), presumably resulting in malfunction of targeted genes. When an editing template with homologous flanking arms was used, the DSBs could be precisely repaired by HDR, generating defined deletions, insertions (7, 16, 41), and nucleotide substitutions (8, 13).

To utilize type II CRISPR/Cas systems, three components, including the Cas9 protein, tracrRNA, and customized crRNA, need to be expressed in foreign hosts. Even though *S. pyogenes* RNase III has been reported to be an indispensable component (27), it was not necessary in a number of diverse heterogeneous systems (12–14, 16). Transfer of the natural type II CRISPR/Cas system by genomic DNA transformation realized targeted genome editing in *Streptococcus pneumoniae*, and then a plasmid-based CRISPR/Cas system was established to edit the *Escherichia coli* genome using two plasmids: (i) pCas9 expressing tracrRNA and Cas9 and (ii) pCRISPR expressing the crRNA array (13). Through use of customized 20-nt target recognition sequences in a crRNA array, double deletion and/or multiplexed editing has been achieved in *E. coli* (13) and human (16) genomes in a single step. Concurrent DSBs using two crRNAs against the *EMX1* gene deleted 119-bp

human genome regions (16). Thus, three-component CRISPR/Cas9 systems are convenient for realizing targeted multiplexed editing in a genome simply by programming the crRNA array.

The tracrRNA:crRNA duplex has been engineered into one molecule, called a chimeric guide RNA (gRNA), that is 39 to 40 nt long and contains a 20-nt target recognition sequence at its 5′ end followed by a hairpin structure (or gRNA scaffold) that retains the base-pairing interactions within the tracrRNA:crRNA duplex (Fig. 2A) (5, 10, 14). This progress further simplified the application of type II CRISPR systems in genome editing. By coexpressing Cas9 nuclease and custom gRNAs, people have successfully engineered and edited the genomes of humans (10, 15, 16, 39), mice (8, 39), *Drosophila melanogaster* (41), zebrafish (40, 42), *Saccharomyces cerevisiae* (14), *Arabidopsis thaliana* (12, 46), tobacco plants (12, 46, 47), and wheat and rice plants (48) (see Table S2 in the supplemental material). Construction of a mammalian expression system with codon-optimized Cas9 and gRNA generated targeting rates of 2 to 25% in various human cells (10) and 36 to 48% in mouse embryonic stem cells (8). Cotransformation of a gRNA plasmid and editing template DNA into a yeast cell constitutively expressing Cas9 generated a near 100% donor DNA recombination frequency at the target loci (14). Intriguingly, coinjection of Cas9 mRNA and gRNA transcripts into mouse zygotes generated mutants with an efficiency of 80%, and when the zygotes were given mutant oligonucleotides, precise point mutations were introduced simultaneously in two target genes (8). By designing DNA donor templates, multiple point mutations (49), site-specific recombination sites (*loxP* and *attP*) (41, 42, 49), endogenous protein tagging (49), and expression cassettes of green fluorescent protein (10) have been successfully introduced into the targeted genome loci. The Cas9-gRNA complex has been used to simultaneously disrupt five genes in a single genome (8). Therefore, the CRISPR/Cas9 system is an efficient tool for editing genomes with wide applications in a broad range of hosts.

Cas9 nickase-mediated genome editing. gRNA-guided Cas9n with a RuvC or HNH mutation has the ability to create a nick instead of a DSB at the target site (9, 11, 16, 29). Although individual nicks are predominantly repaired by the high-fidelity base excision pathway (50), the combination of nick generation and HDR has successfully edited genomes at the intended site (16). Introduction of a double nick using a pair of gRNA-directed Cas9ns targeting the opposite strands of the target site has been successfully applied to generate DSBs and NHEJ-induced mutations (9, 11). A paired-nicking strategy was reported to facilitate high-efficiency HDR at levels comparable to those of native Cas9-mediated HDR and at significantly higher rates than single Cas9n-mediated HDR (9). Interestingly, this paired nicking significantly reduced off-target cleavages by 50- to 1,500-fold in human cells, but without sacrificing on-target cleavage efficiency (9). Because double nicking has the ability to create predictably defined overhangs, NHEJ-mediated ligation, by offering double-stranded repair templates with compatible overhangs, has successfully facilitated HDR-independent fragment integration at specific sites (9). Additionally, creating a pair of double nicks at two sites by four customized gRNAs successfully deleted genomic fragments of up to 6 kb in HEK 293FT cells (9). Thus, multiplex nicking created by Cas9n has the ability to create high-precision genome editing.

Inactivated Cas9-based transcriptional control. CRISPR/Cas systems have also been developed as an innovative facile and multiplexable approach for transcriptional control without altering the target gene sequence, called CRISPR interference (CRISPRi) (Fig. 3) (17). It consists of a completely inactive dCas9 and a custom gRNA (or tracrRNA:crRNA duplex). As mentioned before, dCas9 loses its endonuclease activity, but its ability to incorporate gRNA and bind to targets is not affected. Like RNA interference (RNAi), CRISPRi depends on base-pairing complementarity to recognize target sites. However, they apply different mechanisms to control gene expression. RNAi causes mainly transcript degradation and/or translation blocking (51), but CRISPRi blocks transcription initiation and elongation (17). Qi et al. reported the mechanism of CRISPRi and its initial applications in efficiently repressing the expression of targeted genes in *E. coli* and human cells (17). Through cocustomization of several gRNAs, simultaneous regulation of multiple genes became possible. dCas9-mediated transcriptional control also has been tested in *S. pneumoniae* (19), and silencing effects can be induced and reversed using an anhydrotetracycline-inducible promoter to drive dCas9 and gRNA expression (17). The repression efficiency varied (10- to 300-fold) depending on several major factors, which will be discussed below. Combining two gRNAs targeting the same gene could produce up to 1,000-fold repression (17). Therefore, the CRISPRi targeting platform holds promise as a general approach for modulating gene expression at the transcriptional level.

Like a variety of ZFNs and TALENs that were generated by coupling specific DNA binding domains with different nonspecific effectors (52–54), dCas9 has been fused with transcription effectors, generating chimeric dCas9 effector proteins (Fig. 3) (18, 55). The consequence caused by the chimera depends on effector functions, since the major role of gRNA-guided dCas9 is just to recognize and localize the chimera. KRAB, a repressive chromatin modifier domain, was grafted onto dCas9 and presented significantly higher repression efficiency than dCas9 by itself in HEK293 cells (18). Apart from silencing gene expression, dCas9 has been successfully applied to activate gene expression. dCas9-activator proteins, like dCas9-VP64 and dCas9-p65AD, exhibited an up to 25-fold increase in gene expression (18). In *E. coli*, activation of gene expression was realized by fusing dCas9 to the ω subunit of RNA polymerase (19). Also, by tethering customized gRNA with the MS2 bacteriophage coat protein-binding RNA stem-loop, an MS2-VP64 fusion protein was localized to the target site by the dCas9 complex, and then it stimulated gene expression (Fig. 3) (11). Thus, the dCas9-gRNA complex has a large potential for the design of sequence-specific transcriptional regulation in different organisms and potentially for diverse epigenetic investigation.

Cas9-based high-throughput forward genetic screen. To date, all Cas9-based tools have shown a powerful ability to enhance reverse genetics and synthetic biology. We are also interested in developing CRISPR/Cas9-derived platforms for forward genetic studies. It is highly possible to use multifunctional Cas9 variants to create mutant libraries for screening and identifying genome-scale phenotype-related genetic elements (55, 56). A general flowchart to reach this goal is illustrated in Fig. 4. For high-throughput targeting, the key is to construct high-specificity gRNA libraries. The rules applied to select genome-wide targetable sites have been discussed (14, 38, 56). Coexpression of native Cas9 and gRNA libraries in a host will generate loss-of-function mutant libraries. Very recently, this strategy was successfully applied in genetic screening in human cells (57, 58). If dCas9 or a dCas9-effector chimera is used, knockdown or activation mutant libraries will be generated. Then, mutants of interest can be

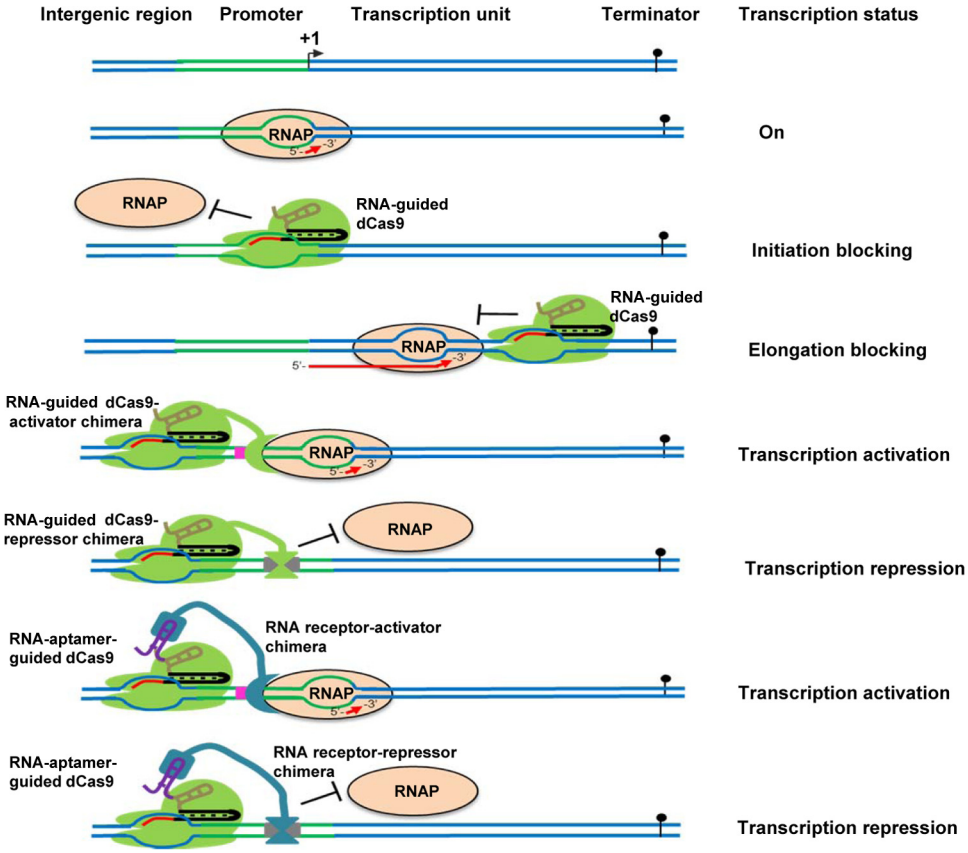

FIG 3 Application of engineered dCas9 and/or RNA components in transcriptional control. RNA polymerase (RNAP) initiates transcription within the promoter region; however, the binding of RNA-guided dCas9 to the promoter region and the encoding region may block transcription initiation and transcription elongation, respectively, leading to the repression of gene expression at the transcriptional level. Through fusion of dCas9 with transcriptional activators or repressors, the positioning function of gRNA or crRNA molecules will direct the dCas9-effector chimera to bind in the promoter's vicinity, and then the effector modules will stimulate or repress gene transcription by interacting with DNA motifs or RNAP. Also, gRNA or crRNA might be fused with RNA aptamers, generating chimeric RNA that will direct dCas9 to bind to specific sites, allowing localization of specific RNA receptors. Generation of an RNA receptor-activator or -repressor chimera will lead to activator or repressor localization, followed by expression activation or repression of neighboring genes.

screened with suitable methodologies. Compared to loss-of-function mutant libraries, the knockdown or activation mutant libraries have an unmatched advantage in the study of lethal genes.

INFLUENTIAL FACTORS OF CRISPR/Cas APPLICATION

Thus far, numerous studies have examined the diverse factors that impact the efficiency and/or specificity of Cas9-based tools, such as Cas9 activity, RNA component lengths and structures, Cas9/RNA ratio, and RNA target complementarity extent and complementary position. Discussion of these factors will help direct future experiments using CRISPR and improve performance.

Cas9 is a pivotal component. Mutation of catalytic sites, incorrect subcellular localization, or inappropriate Cas9 dosage all affect genome editing. In eukaryotic cells, prokaryote-derived Cas9 is generally fused with a nuclear location signal (NLS) at the N or C terminus, or both, to direct protein translocation into the nucleus (8, 10, 14, 17, 40, 41, 59). Recent studies found that the addition of a 32-amino-acid linker between the NLS and Cas9 enhanced genomic DNA cleavage activity. This might be caused by increased subcellular localization since the NLS peptide would be buried or shielded during Cas9 folding without the linker (39). Codon optimization is also necessary for producing functional Cas9 in heterogeneous expression systems (10, 12, 60). Addition of exogenous gRNA substantially increased DNA cleavage activity (59). The ratio of Cas9 to gRNA greatly affected mutagenesis efficiency (8, 12, 60). Theoretically, the more complexes are formed, the higher editing efficiency is expected to be. However, a potential risk that accompanies excessive executive complex availability is the off-target effect due to the unavoidably low complementarity of nonspecific regions in the genome (61). To address the above issues, we need to optimize the CIRSPR/Cas9 system by controlling component expression, improving target selection criteria, and engineering the Cas9 protein to provide higher specificity.

Another major class of determinants is the RNA components. For tracrRNAs, *S. pyogenes* generated two kinds of precursor tracrRNA species. Tests in heterogeneous cells showed that the smaller tracrRNA was more effective (16). The gRNA chimera exhibits efficiency comparable to that of the tracrRNA:crRNA duplex in *in vitro* plasmid cleavage assays (5). To mutate the rice genome, gRNAs presented higher efficiency than RNA duplexes (62). However, studies using human and mouse cells showed that cleavage efficiencies of gRNAs were either lower than those of RNA duplexes or undetectable when the same protospacer was targeted (16). This suggests that some undetermined cellular factors or RNA features might influence editing efficiency. In addi-

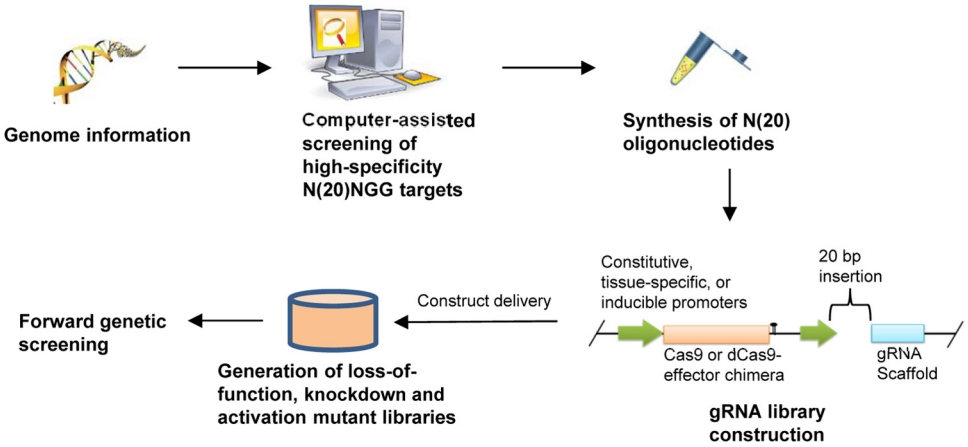

FIG 4 Proposed flowchart for Cas9-based high-throughput forward genetic screening. Valid and specific N_{20}NGG targetable sites across the whole genome are selected by computer-assisted programs. DNA oligonucleotides with N_{20} sequences are synthesized and modified for ligation with a construct, generating a gRNA library able to coexpress Cas9 or a Cas9-effector chimera and diverse gRNAs driven by constitutive, tissue-specific, or inducible promoters. Then the gRNA library is delivered to host cells to generate diverse mutant libraries for forward genetic screening.

tion, base pairing is critical to the folding structures of gRNAs. Elongation of the self-complementation region in gRNAs enhanced site-specific NHEJ-mediated mutagenesis (59).

crRNA and gRNA molecules harboring target recognition sequences determine target specificity; as a result, the selection of target protospacers is a critical issue. A protospacer within an N_{21}GG (or N_{20}NGG) format is widely used for *S. pyogenes* Cas9 targeting. This protospacer contains a 20-nt base-pairing region immediately followed by a PAM (NGG). Other Cas9 orthologs requiring longer PAM sequences would reduce our choices on targetable sites in a given gene or genome. The amount of base pair complementarity between target recognition sequences and protospacers is of importance to Cas9-based editing efficiency and dCas9-based transcriptional control (13). Extension of the 5′ end of the gRNA target recognition region to increase base-pairing complementarity with a protospacer did not improve either editing efficiency or targeting specificity because all gRNA transcripts with >20 nt at their 5′ ends were trimmed to 20 nt (9). Several studies reported that mismatches occurring in the 3′ half of the gRNA severely affected Cas9-mediated cleavage (5, 13, 17, 63). Recently, the effects of mismatches on Cas9-gRNA functionality were demonstrated to be target site dependent (61). The same position within different targeting sequences presented varying importance, and not all mismatches in the 5′ half of the gRNA were well tolerated (61). For the double-nicking strategy, the relative positions of the gRNA pairs with offsets from −4 to 20 bp were most efficient in inducing NHEJ (9), and introduction of 5′ overhangs created by offset nicks stimulated more-robust NHEJ and HDR events than introduction of 3′ overhangs (9, 11). For CRISPRi, dCas9 also presented similar rules for maintenance of silencing efficiency (17).

The above discussion focuses on the determinants of DNA cleavage, which is the most critical step in introducing frameshift mutations to a specific genome site by error-prone NHEJ. Another way to resolve DSBs is to stimulate HDR by providing editing templates, which are single-stranded DNAs (ssDNAs) or dsDNA fragments with homologous flanking arms. DSB generation can increase rates of homologous recombination of ssDNA and dsDNA donors by 5-fold and 130-fold, respectively (14). During recombination, editing templates should not overlap crRNA/gRNA target recognition sequences, which might decrease editing efficiency (14). If multiple template DNAs are cotransformed with plasmids expressing Cas9 and gRNAs targeting multiple sites, a single-step double deletion or more deletions could be generated as desired (13). However, some factors potentially affecting HDR, including the sizes and positions of the homologous flanking arms and the stability of the given templates before HDR occurs, remain to be evaluated.

CRISPRi has been systematically studied (17), and several factors, in addition to the ones aforementioned, have been identified as influencing dCas9-based transcriptional control. First, CRISPRi-mediated blocking of transcriptional elongation presents strand specificity (17, 19). gRNAs targeting a nontemplate DNA strand presented much higher repression efficiency than those targeting the template strand. Second, the silencing efficiency is inversely correlated with the distance of the target from the translation start codon. Third, an augmentative silencing effect may be observed when two or more gRNAs bind to separate target sites on the same gene (11, 17); however, if they bind to overlapping regions, repression is suppressed. Finally, CRISPRi-mediated blocking of transcriptional initiation has site effects but not strand specificity. To block transcriptional initiation in *E. coli*, the −35 box-containing regions chosen as gRNA targets are more efficient than other adjacent regions. Interestingly, a recent study has found that orthogonal dCas9 proteins exhibited quite different repression abilities (31). For dCas9 effector-dependent transcriptional regulation, performance also presented position and accumulation effects (11). Using the dCas9-VP64 chimera, gRNAs proximal to the transcriptional start site significantly triggered gene expression, and multiple gRNAs working together induced multifold transcriptional enhancement (11). Several orthogonal dCas9s fused with VP64 also presented robust transcriptional activation in human cells at levels similar to that of the corresponding transcription activator-like (TAL)-VP64 activator (31).

COMPARISON OF TARGETED GENETIC-ENGINEERING TOOLS

A wide variety of tools are available for editing targeted genomes and regulating gene expression. Based on target recognition mechanisms, they can be grouped into two major classes, protein-

directed and nucleotide-directed specificity (2). Recombinases, integrases, ZFNs, and TALENs are well-known approaches that depend on protein-directed specificity; RNA interference (RNAi), group II intron retrotransposition, and the innovative Cas9-based platforms rely on nucleotide-directed specificity. All of these have been widely used in prokaryotes and/or eukaryotes. This section will discuss the advantages and limitations of these widely used tools in terms of their flexibility, multiplex targeting potential, and targeting efficiency and specificity.

Generally, protein-directed specificity is comparatively harder to customize than nucleotide-directed specificity. Recombinases and integrases require suitable preexisting recognition sites in the genome and often have some inherent application limitations (2, 64, 65). Both ZFNs and TALENs are generated by coupling a customized DNA binding domain with a nonspecific nuclease domain (3, 21). The DNA binding domain of ZFNs and TALENs is a tandem array of zinc finger (ZF) motifs and TAL repeats, respectively (21, 66). However, it is difficult and expensive to customize ZFs or TALs by protein engineering (3, 20), and if the FokI nuclease domain is used, two ZFNs or TALENs must be customized for each new target site (67). Also, ZFN and TALEN activities are affected by numbers of factors (3, 21). Even though ZFNs and TALENs have already generated extensive modifications (3, 21), they are difficult to apply to the creation of multiple mutations in a single genome via stepwise mutagenesis.

As tools based on nucleotide-directed specificity, RNA-directed RNAi, group II intron retrotransposition, and Cas9-based methods require only DNA synthesis or PCR amplification to retarget, so obviously these methods are more convenient and economical. RNAi is used mostly to repress gene expression in both prokaryotes and eukaryotes instead of knocking them out. Although RNAi also can be used to target multiple genes, the need for long target sites and amplification of small interference RNAs can sometimes result in severe off-target effects (68, 69). Group II intron retrotransposition is widely applied to inactivate genes in bacterial genomes (2). Cas9-based tools can be used in diverse applications, as mentioned above. All of the essential components required by these tools can be expressed by delivering plasmids (10, 12, 13, 16, 48), linear DNA expression cassettes (14), or RNA transcripts (7, 8, 40, 70). In addition, bioinformatic analysis of genome-wide target sites ($N_{21}GG$) revealed that most genes or exons can be targeted specifically in *Arabidopsis* (12), rice (38), and yeast (14). Therefore, Cas9-based genome editing provides a highly flexible and programmable method.

The ability to multiplex target is another notable advantage that Cas9-based tools have. Efficient methods enabling multiplex genome editing are urgently needed for genome-scale engineering. Several reports demonstrated the creation of simultaneous multiple mutations with Cas9-based tools (8, 16). To realize multiplexed editing, the only thing required is the construction of crRNA arrays that produce various crRNAs or the construction of several different chimeric gRNAs to direct Cas9 to edit multiple targets at the same time. In this way, as many as five gene mutations have been generated simultaneously in mouse embryonic stem cells with high efficiency (8). In addition, with gRNAs to direct mutated dCas9 to specifically target the transcriptional regions of two different genes, the expression of both targeted genes was simultaneously decreased (17). Then, multiple genes were activated or repressed at the transcriptional level by coupling dCas9 with transcriptional effectors or fusing gRNA with recognizable RNA aptamers (11, 17, 19). Thus, versatile Cas9-based tools hold promise for realizing both multiplexed genome editing and transcriptional control, while avoiding tedious stepwise genetic manipulations.

Targeting efficiency and specificity greatly impact the application potential of targeted tools. The editing efficiencies of Cas9-based tools vary greatly among different organisms, cell types, and mutation types and even target sites (see Table S2 in the supplemental material). NHEJ-induced indels occurred at efficiencies of 2 to 33% in human HEK293T cells (15, 42, 59), 66% in *Drosophila* (41), and >35% in zebrafish embryos (42). Notably, the efficiency of indel occurrence in tobacco protoplast reached 37.7 to 38.5% (12), much higher than the 5.6% observed for *Arabidopsis* protoplasts (12). However, when HDR-mediated insertion was chosen, the editing efficiency in *Arabidopsis* protoplasts increased to 10.7% (12). HDR-mediated insertion occurred at efficiencies of 100% in *S. pneumoniae*, 64% in *E. coli* (13), and 100% in *S. cerevisiae* (14). Cas9-mediated genome editing in human cells and zebrafish embryos produced efficiencies similar to those obtained using ZFNs and/or TALENs (16, 40), although its efficiency in other organisms needs to be evaluated in the future. To date, Cas9-based tools have presented the ability to delete 6-kb genomic fragments (9) and insert up to 3 kb of DNA into the intended genomic locus (7). However, for application in synthetic biology, the potential of delivering larger DNA fragments still needs to be evaluated. Off-target activity, which potentially produces misleading conclusions, is a big challenge for all targeted tools. Cas9-based tools face the same problem (38, 61). TALENs appear to have lower off-target activities than ZFNs (71). Cas9-gRNA complexes and 18-mer TAL effectors can potentially tolerate 1 to 3 and 1 to 2 target mismatches, respectively (11, 61). Further studies with Cas9-gRNA complexes revealed that the frequency of off-target cleavage was sometimes the same as for on-target frequency (61). Cas9n was reported to greatly reduce off-target effects without sacrificing the efficiency of HDR induction (16). To improve the efficiency and specificity of Cas9-based tools, much effort needs to be put into Cas9 engineering, optimizing gRNA selection rules, and further elucidating Cas9-gRNA recognition features.

CONCLUDING REMARKS

Cas9-based tools possess notable advantages that current, widely used targeted tools cannot match. These tools will greatly enhance our ability to engineer and edit genomes and regulate gene expression in diverse organisms. These technologies also pave the way for easily dissecting individual gene functions and are expected to accelerate the *in vivo* study of functionally redundant genes and epigenetic investigations. They will enable a broad range of research and applications in diverse biological fields: biotechnology, metabolic engineering, and medicine. The ability to do multiplex targeting will revolutionize genome-scale engineering by providing a method for multiple disruptions, insertions, and deletions at high efficiency and low cost (2). However, even though Cas9-based tools have been applied in many model organisms (see Table S2 in the supplemental material), several fundamental attributes are still unclear, including the molecular structure and catalytic mechanism of Cas9, PAM dependence, and the gRNA loading mechanism. Understanding these issues will assist in engineering Cas9 to be PAM independent, broaden our choices of targets, and generate highly precise Cas9-gRNA complexes, especially for use as human therapeutic agents. The major challenges

to current Cas9-based tools is off-target disruption caused by the tolerance of the Cas9-gRNA complex and imperfect base-pairing complementation between gRNA and target sites (61). The aforementioned Cas9 engineering is a promising way to address this issue. Additionally, comprehensive profiling of off-target events by high-throughput methods will aid in establishing rules for target selection and even in composing programs for customized targeting. Other aspects will need further evaluation in the future: (i) the impact of the GC content and secondary structure of customized gRNA and of the Cas9/gRNA molecular ratio on editing and silencing efficiency and specificity (61), (ii) the efficacy of diverse delivery systems for generating active Cas9 and gRNA components in different cells (e.g., plasmid-, DNA fragment-, and RNA-based delivery) (8, 14, 16), (iii) the universality and efficacy of NHEJ- and HDR-mediated DSB repair in diverse organisms, especially in bacteria, and (iv) the potential application of Cas9-based high-throughput forward genetic studies. In summary, much effort needs to be expended to fully understand Cas9-based tools and exploit their potential.

ACKNOWLEDGMENTS

We thank Chengcheng Zhang for the helpful discussions.

This work was supported by the NSF EPSCoR award EPS 0814361.

REFERENCES

1. **Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M.** 2012. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. **2012:**251364. http://dx.doi.org/10.1155/2012/251364.
2. **Esvelt KM, Wang HH.** 2013. Genome-scale engineering for systems and synthetic biology. Mol. Syst. Biol. **9:**641. http://dx.doi.org/10.1038/msb.2012.66.
3. **Gaj T, Gersbach CA, Barbas CF.** 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. **31:**397–405. http://dx.doi.org/10.1016/j.tibtech.2013.04.004.
4. **Marraffini LA, Sontheimer EJ.** 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat. Rev. Genet. **11:**181–190. http://dx.doi.org/10.1038/nrg2749.
5. **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:**816–821. http://dx.doi.org/10.1126/science.1225829.
6. **Carroll D.** 2012. A CRISPR approach to gene targeting. Mol. Ther. **20:**1658–1660. http://dx.doi.org/10.1038/mt.2012.171.
7. **Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R.** 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell **154:**1370–1379. http://dx.doi.org/10.1016/j.cell.2013.08.022.
8. **Wang HY, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R.** 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell **153:**910–918. http://dx.doi.org/10.1016/j.cell.2013.04.025.
9. **Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F.** 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell **154:**1380–1389. http://dx.doi.org/10.1016/j.cell.2013.08.021.
10. **Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM.** 2013. RNA-guided human genome engineering via Cas9. Science **339:**823–826. http://dx.doi.org/10.1126/science.1232033.
11. **Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM.** 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. **31:**833–838. http://dx.doi.org/10.1038/nbt.2675.
12. **Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J.** 2013. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat. Biotechnol. **31:**688–691. http://dx.doi.org/10.1038/nbt.2654.
13. **Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA.** 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. **31:**233–239. http://dx.doi.org/10.1038/nbt.2508.
14. **DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM.** 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. **41:**4336–4343. http://dx.doi.org/10.1093/nar/gkt135.
15. **Cho SW, Kim S, Kim JM, Kim JS.** 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. **31:**230–232. http://dx.doi.org/10.1038/nbt.2507.
16. **Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F.** 2013. Multiplex genome engineering using CRISPR/Cas systems. Science **339:**819–823. http://dx.doi.org/10.1126/science.1231143.
17. **Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA.** 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell **152:**1173–1183. http://dx.doi.org/10.1016/j.cell.2013.02.022.
18. **Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS.** 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell **154:**442–451. http://dx.doi.org/10.1016/j.cell.2013.06.044.
19. **Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA.** 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. **41:**7429–7437. http://dx.doi.org/10.1093/nar/gkt520.
20. **Pabo CO, Peisach E, Grant RA.** 2001. Design and selection of novel Cys(2)His(2) zinc finger proteins. Annu. Rev. Biochem. **70:**313–340. http://dx.doi.org/10.1146/annurev.biochem.70.1.313.
21. **Mussolino C, Cathomen T.** 2012. TALE nucleases: tailored genome engineering made easy. Curr. Opin. Biotechnol. **23:**644–650. http://dx.doi.org/10.1016/j.copbio.2012.01.013.
22. **Horvath P, Barrangou R.** 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science **327:**167–170. http://dx.doi.org/10.1126/science.1179555.
23. **Makarova KS, Aravind L, Wolf YI, Koonin EV.** 2011. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol. Direct **6:**38. http://dx.doi.org/10.1186/1745-6150-6-38.
24. **Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV.** 2011. Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. **9:**467–477. http://dx.doi.org/10.1038/nrmicro2577.
25. **Sinkunas T, Gasiunas G, Waghmare SP, Dickman MJ, Barrangou R, Horvath P, Siksnys V.** 2013. In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. EMBO J. **32:**385–394. http://dx.doi.org/10.1038/emboj.2012.352.
26. **Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, Graham S, Reimann J, Cannone G, Liu H, Albers SV, Naismith JH, Spagnolo L, White MF.** 2012. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. Mol. Cell **45:**303–313. http://dx.doi.org/10.1016/j.molcel.2011.12.013.
27. **Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E.** 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature **471:**602–607. http://dx.doi.org/10.1038/nature09886.
28. **Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V.** 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Res. **39:**9275–9282. http://dx.doi.org/10.1093/nar/gkr606.
29. **Gasiunas G, Barrangou R, Horvath P, Siksnys V.** 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. **109:**E2579–E2586. http://dx.doi.org/10.1073/pnas.1208507109.
30. **Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, Vogel J, Sontheimer EJ.** 2013. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. Mol. Cell **50:**488–503. http://dx.doi.org/10.1016/j.molcel.2013.05.001.
31. **Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM.** 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods **10:**1116–1121. http://dx.doi.org/10.1038/nmeth.2681.
32. **van der Ploeg JR.** 2009. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by

M102-like bacteriophages. Microbiology **155:**1966–1976. http://dx.doi.org/10.1099/mic.0.027508-0.
33. **Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S.** 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. **190:**1390–1400. http://dx.doi.org/10.1128/JB.01412-07.
34. **Bhaya D, Davison M, Barrangou R.** 2011. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. **45:**273–297. http://dx.doi.org/10.1146/annurev-genet-110410-132430.
35. **Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V.** 2013. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. RNA Biol. **10:**841–851. http://dx.doi.org/10.4161/rna.24203.
36. **Chylinski K, Le Rhun A, Charpentier E.** 2013. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol. **10:**726–737. http://dx.doi.org/10.4161/rna.24321.
37. **Sorek R, Lawrence CM, Wiedenheft B.** 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu. Rev. Biochem. **82:**237–266. http://dx.doi.org/10.1146/annurev-biochem-072911-172315.
38. **Xie K, Yang Y.** 2013. RNA-guided genome editing in plants using a CRISPR-Cas system. Mol. Plant **6:**1975–1983. http://dx.doi.org/10.1093/mp/sst119.
39. **Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P, Huang X.** 2013. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. Cell Res. **23:**720–723. http://dx.doi.org/10.1038/cr.2013.46.
40. **Hwang WY, Fu YF, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JRJ, Joung JK.** 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. **31:**227–229. http://dx.doi.org/10.1038/nbt.2501.
41. **Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM.** 2013. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics **113:**152710. http://dx.doi.org/10.1534/genetics.113.152710.
42. **Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong JW, Xi JJ.** 2013. Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. Cell Res. **23:**465–472. http://dx.doi.org/10.1038/cr.2013.45.
43. **Wyman C, Kanaar R.** 2006. DNA double-strand break repair: all's well that ends well. Annu. Rev. Genet. **40:**363–383. http://dx.doi.org/10.1146/annurev.genet.40.110405.090451.
44. **Shuman S, Glickman MS.** 2007. Bacterial DNA repair by non-homologous end joining. Nat. Rev. Microbiol. **5:**852–861. http://dx.doi.org/10.1038/nrmicro1768.
45. **Smith GR.** 2001. Homologous recombination near and far from DNA breaks: alternative roles and contrasting views. Annu. Rev. Genet. **35:**243–274. http://dx.doi.org/10.1146/annurev.genet.35.102401.090509.
46. **Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP.** 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res. **41:**e188. http://dx.doi.org/10.1093/nar/gkt780.
47. **Nekrasov V, Staskawicz B, Weigel D, Jones JD, Kamoun S.** 2013. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. Nat. Biotechnol. **31:**691–693. http://dx.doi.org/10.1038/nbt.2655.
48. **Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C.** 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. **31:**686–688. http://dx.doi.org/10.1038/nbt.2650.
49. **Dickinson DJ, Ward JD, Reiner DJ, Goldstein B.** 2013. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat. Methods **10:**1028–1034. http://dx.doi.org/10.1038/nmeth.2641.
50. **Dianov GL, Hubscher U.** 2013. Mammalian base excision repair: the forgotten archangel. Nucleic Acids Res. **41:**3483–3490. http://dx.doi.org/10.1093/nar/gkt076.
51. **Wilson RC, Doudna JA.** 2013. Molecular mechanisms of RNA interference. Annu. Rev. Biophys. **42:**217–239. http://dx.doi.org/10.1146/annurev-biophys-083012-130404.
52. **Miller JC, Tan SY, Qiao GJ, Barlow KA, Wang JB, Xia DF, Meng XD, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ.** 2011. A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. **29:**143–148. http://dx.doi.org/10.1038/nbt.1755.
53. **Minczuk M, Papworth MA, Kolasinska P, Murphy MP, Klug A.** 2006. Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase. Proc. Natl. Acad. Sci. U. S. A. **103:**19689–19694. http://dx.doi.org/10.1073/pnas.0609502103.
54. **Li FY, Papworth M, Minczuk M, Rohde C, Zhang YY, Ragozin S, Jeltsch A.** 2007. Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. Nucleic Acids Res. **35:**100–112. http://dx.doi.org/10.1093/nar/gkl1035.
55. **Mali P, Esvelt KM, Church GM.** 2013. Cas9 as a versatile tool for engineering biology. Nat. Methods **10:**957–963. http://dx.doi.org/10.1038/nmeth.2649.
56. **Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS.** 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat. Protoc. **8:**2180–2196. http://dx.doi.org/10.1038/nprot.2013.132.
57. **Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F.** 2013. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science **343:**84–87. http://dx.doi.org/10.1126/science.1247005.
58. **Wang T, Wei JJ, Sabatini DM, Lander ES.** 2013. Genetic screens in human cells using the CRISPR/Cas9 system. Science **343:**80–84. http://dx.doi.org/10.1126/science.1246981.
59. **Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J.** 2013. RNA-programmed genome editing in human cells. eLife **2:**e00471. doi:http://dx.doi.org/10.7554/eLife.00471.
60. **Nakayama T, Fish MB, Fisher M, Oomen-Hajagos J, Thomsen GH, Grainger RM.** 2013. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. Genesis **51:**835–843. http://dx.doi.org/10.1002/dvg.22720.
61. **Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD.** 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. **31:**822–826. http://dx.doi.org/10.1038/nbt.2623.
62. **Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ.** 2013. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res. **23:**1233–1236. http://dx.doi.org/10.1038/cr.2013.123.
63. **Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJ, Severinov K.** 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. U. S. A. **108:**10098–10103. http://dx.doi.org/10.1073/pnas.1104144108.
64. **Wang Y, Yau YY, Perkins-Balding D, Thomson JG.** 2011. Recombinase technology: applications and possibilities. Plant Cell Rep. **30:**267–285. http://dx.doi.org/10.1007/s00299-010-0938-1.
65. **Groth AC, Calos MP.** 2004. Phage integrases: biology and applications. J. Mol. Biol. **335:**667–678. http://dx.doi.org/10.1016/j.jmb.2003.09.082.
66. **Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD.** 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. **11:**636–646. http://dx.doi.org/10.1038/nrg2842.
67. **Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I.** 1998. FokI dimerization is required for DNA cleavage. P. Natl. Acad. Sci. U. S. A. **95:**10570–10575. http://dx.doi.org/10.1073/pnas.95.18.10570.
68. **Maida Y, Kyo S, Lassmann T, Hayashizaki Y, Masutomi K.** 2013. Off-target effect of endogenous siRNA derived from RMRP in human cells. Int. J. Mol. Sci. **14:**9305–9318. http://dx.doi.org/10.3390/ijms14059305.
69. **Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS.** 2003. Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. **21:**635–637. http://dx.doi.org/10.1038/nbt831.
70. **Waaijers S, Portegijs V, Kerver J, Lemmens BB, Tijsterman M, van den Heuvel S, Boxem M.** 2013. CRISPR/Cas9-targeted mutagenesis in *Caenorhabditis elegans*. Genetics **195:**1187–1191. http://dx.doi.org/10.1534/genetics.113.156299.
71. **Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T.** 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res. **39:**9283–9293. http://dx.doi.org/10.1093/nar/gkr597.