Tabel 4

AAVS-1 crD(R)NA

Lane	SEQ ID NO:
1	DNA Marker
2	No guide control
3	SEQ ID NO: 37
4	SEQ ID NO: 38
5	SEQ ID NO: 39
6	SEQ ID NO: 40
7	SEQ ID NO: 41
8	SEQ ID NO: 42
9	DNA Marker
10	DNA Marker
11	No guide control
12	SEQ ID NO: 1
13	SEQ ID NO: 43
14	SEQ ID NO: 44
15	SEQ ID NO: 45
16	SEQ ID NO: 46
17	SEQ ID NO: 47
18	SEQ ID NO: 48
19	SEQ ID NO: 49
20	DNA Marker
21	DNA Marker
22	No guide control
23	SEQ ID NO: 1
24	SEQ ID NO: 50
25	SEQ ID NO: 51
26	SEQ ID NO: 52
27	SEQ ID NO: 53
28	SEQ ID NO: 54
29	SEQ ID NO: 55
30	SEQ ID NO: 56
31	SEQ ID NO: 57
32	SEQ ID NO: 58
33	SEQ ID NO: 59
34	SEQ ID NO: 60
35	SEQ ID NO: 61
36	SEQ ID NO: 62
37	DNA Marker

Cleavage percentages are shown at the bottom of each lane. For crDNA or crD(R)NAs where no cleavage activity was observed (*e.g.*, **FIGURE 3**, 3; **FIGURE 3**, 5; **FIGURE 3**, 15;

FIGURE 3, 33; **FIGURE 3**, 34; **FIGURE** 3, 35) cleavage activity is expressed as n/d (indicating that cleavage activity was not detected).

The data presented in **FIGURE 3** demonstrate that the crD(R)NAs of the present disclosure facilitate Cas9 mediated site-specific cleavage of a target double-stranded DNA.

Example 4

crD(R)NA Activity against Multiple Targets

This example demonstrates the *in vitro* biochemical activity of crD(R)NAs comprising different spaces programmed to target specific sequences.

The sequences of the crDNA, crRNA and crD(R)NA (shown in Table 5) were provided to a commercial manufacturer for synthesis.

Table 5

crDNA, crRNA, and crD(R)NA sequences

Target	Guide RNA type	Sequences (RNA bases are bracketed)	SEQ ID NO
EMX1 target 1	crDNA	5' - GAGTCCGAGC AGAAGAAGAA GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 3
VEGFA target 1	crDNA	5' - GGGTGGGGGG AGTTTGCTCC GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 4
CD34 target 1	crDNA	5' - GTTTGTGTTT CCATAAACTG GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 5
CD34 target 2	crDNA	5' - TCTGTGATAA CCTCAGTTTA GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 6
STAT5a target 1	crDNA	5' - GGCCACTGTA GTCCTCCAGG GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 7
STAT5a target 2	crDNA	5' - GTCCCCCAGC CGGTCAGCCA GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 8
JAK1 target 1	crDNA	5' - GGCAGCCAGC ATGATGAGAC GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 9
JAK1 target 2	crDNA	5' - GAGGAGCTCC AAGAAGACTG GTCTCAGAGC TATGCTGTCC TG - 3'	SEQ ID NO: 10

EMX1 target 1	crRNA	5' - [G][A][G][U][C][C][G][A][G][C] [A][G][A][A][G][A][A][G][A][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 11
VEGFA target 1	crRNA	5' - [G][G][G][U][G][G][G][G][G][G] [A][G][U][U][U][G][C][U][C][C] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 12
CD34 target 1	crRNA	5' - [G][U][U][U][G][U][G][U][U][U] [C][C][A][U][A][A][A][C][U][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 13
CD34 target 2	crRNA	5' - [U][C][U][G][U][G][A][U][A][A] [C][C][U][C][A][G][U][U][U][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 14
STAT5a target 1	crRNA	5' - [G][G][C][C][A][C][U][G][U][A] [G][U][C][C][U][C][C][A][G][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 15
STAT5a target 2	crRNA	5' - [G][U][C][C][C][C][C][A][G][C] [C][G][G][U][C][A][G][C][C][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 16
JAK1 target 1	crRNA	5' - [G][G][C][A][G][C][C][A][G][C] [A][U][G][A][U][G][A][G][A][C] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 17
JAK1 target 2	crRNA	5' - [G][A][G][G][A][G][C][U][C][C] [A][A][G][A][A][G][A][C][U][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G] - 3'	SEQ ID NO: 18

EMX1 target 1	crD(R)NA	5' - GAGTCCGAGC AGAA[G][A][A][G][A][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 19
VEGFA target 1	crD(R)NA	5' - GGGTGGGGGG AGTT[U][G][C][U][C][C] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 20
CD34 target 1	crD(R)NA	5' - GTTTGTGTTT CCAT[A][A][A][C][U][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 21
CD34 target 2	crD(R)NA	5' - TCTGTGATAA CCTC[A][G][U][U][U][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 22
STAT5a target 1	crD(R)NA	5' - GGCCACTGTA GTCC[U][C][C][A][G][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 23
STAT5a target 2	crD(R)NA	5' - GTCCCCCAGC CGGT[C][A][G][C][C][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 24
JAK1 target 1	crD(R)NA	5' - GGCAGCCAGC ATGA[U][G][A][G][A][C] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 25
JAK1 target 2	crD(R)NA	5' - GAGGAGCTCC AAGA[A][G][A][C][U][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG - 3'	SEQ ID NO: 26

tracrRNA was constructed as described in Example 1.

Double stranded DNA targets were generated as described in Example 2 using the oligonucleotides shown in Table 3 corresponding to the appropriate target sequence.

crDNA/tracrRNA, crRNA/tracrRNA, and crD(R)NA/tracrRNA were hybridized and biochemical cleavage is carried out as described in Example 3.

FIGURE 4A and **FIGURE 4B** show the results for the biochemical cleavage of various spacers. **FIGURE 4A** shows biochemical cleavage percentages. Activity for EMX target 1 is shown in group 1: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for VEGFA target 1 is shown in group 2: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for CD34 target 1 is shown in group 3: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for CD34 target 2 is shown in group 4: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for STAT5a target 1 is shown in group 5: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for STAT5a target 2 is shown in group 6: where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for JAK1 target 1 is shown in group 7; where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. Activity for JAK1 target 2 is shown in group 8; where 'A' is a Cas9 only control, 'B' is the crDNA/tracrRNA/Cas9, 'C' is the crRNA/tracrRNA/Cas9, and 'D' is the crD(R)NA/tracrRNA/Cas9. For all Cas9 only samples (**FIGURE 4A**, 'A') and crDNA/tracrRNA/cas9 samples (**FIGURE 4B**, 'B'), no cleavage activity was detected (**FIGURE 4A**, 'n/d').

In **FIGURE 4B**, the percent cleavage is shown on the y-axis of the graph and the target is shown on the x-axis. Activity for EMX target 1 is shown in the bars of group 1. Activity for VEGFA target 1 is shown in the bars of group 2. Activity for CD34 target 1 is shown in the bars of group 3. Activity for CD34 target 2 is shown in the bars of group 4. Activity for STAT5a target 1 is shown in the bars of group 5. Activity for STAT5a target 2 is shown in the bars of group 6. Activity for JAK1 target 1 is shown in the bars of group 7. Activity for JAK1 target 2 is shown in the bars of group 8. 'C' and 'D' refer to the same reactions as in **FIGURE 4A**.

FIGURE 4 demonstrates that the Cas9 mediated biochemical cleavage of a double stranded DNA target using the crD(R)NA of the present disclosure is transferable across different target sequences.

Example 5

T7E1 Assay for Detection of Target Modifications in Eukaryotic Cells

This example illustrates the use of T7E1 assays to evaluate the percent cleavage *in vivo* of crD(R)NA relative to selected double-stranded DNA target sequences.

A. Cell Transfections Using Cas Polynucleotide Components

sgRNA and crD(R)NA/tracrRNAs comprising an AAVS-1 targeting sequence were transfected into HEK293 cells constitutively expressing SpyCas9-GFP fusion (HEK293-Cas9-GFP), using the Nucleofector® 96-well Shuttle System (Lonza, Allendale, NJ) and the following protocol. Equal molar amounts of guide RNA components were prepared in an annealing buffer (1.25mM HEPES, 0.625mM $MgCl_2$, 9.375mM KCl at pH 7.5), were incubated for 2 minutes at 95°C, were removed from thermocycler, allowed to equilibrate to room temperature, and dispensed in a 10µL final volume in triplicate in a 96-well plate. Culture medium was aspirated from HEK293-Cas9-GFP cells, and the cells were washed once with calcium and magnesium-free PBS then were trypsinized by the addition of TrypLE (Life Technologies, Grand Island, NY) followed by incubation at 37ºC for 3-5 minutes. Trypsinized cells were gently pipetted up and down to form a single cell suspension and added to DMEM complete culture medium composed of DMEM culture medium (Life Technologies, Grand Island, NY) containing 10% FBS (Fisher Scientific, Pittsburgh, PA) and supplemented with penicillin and streptomycin (Life Technologies, Grand Island, NY).

The cells were then pelleted by centrifugation for 3 minutes at 200 x g, the culture medium aspirated and cells were resuspended in PBS. The cells were counted using the Countess® II Automated Cell Counter (Life Technologies, Grand Island, NY). 2.2 x 10^7 cells were transferred to a 50ml tube and pelleted. The PBS was aspirated and the cells were resuspended in Nucleofector™ SF (Lonza, Allendale, NJ) solution to a density of 1 x 10^7 cells/mL. 20µL of the cell suspension were then added to individual wells containing 10uL of Cas polynucleotide components and the entire volume was transferred to the wells of a 96-well Nucleocuvette™ Plate (Lonza, Allendale, NJ). The plate was loaded onto the Nucleofector™ 96-well Shuttle™ (Lonza, Allendale, NJ) and cells were nucleofected using the 96-CM-130 Nucleofector™ program (Lonza, Allendale, NJ). Post-nucleofection, 70 µL DMEM complete culture medium was added to each well and 50µL of the cell suspension were transferred to a collagen coated 96-well cell culture plate containing 150µL pre-warmed DMEM complete

culture medium. The plate was then transferred to a tissue culture incubator and maintained at 37°C in 5% CO_2 for 48 hours.

B. Target Double-Stranded DNA Generation for T7E1 Assay

gDNA was isolated from HEK-293-SpyCas9 cells 48 hours after Cas polynucleotide component transfection using 50µL QuickExtract DNA Extraction solution (Epicentre, Madison, WI) per well followed by incubation at 37°C for 10 minutes, 65°C for 6 minutes and 95°C for 3 minutes to stop the reaction. gDNA was then diluted with 150µL water and samples were stored at -80°C.

DNA for T7E1 was generated by PCR amplification of a target double-stranded DNA sequence (*e.g.*, AAVS-1) from isolated gDNA. PCR reactions were set up using 8µL gDNA as template with KAPA HiFi Hot Start polymerase and containing 0.5U of polymerase, 1x reaction buffer, 0.4mM dNTPs and 300nM forward and reverse primers directed to the target double-stranded DNA (*e.g.*, AAVS-1, SEQ ID NOs: 75, 76 (Table 3)) in a total volume of 25uL. Target DNA was amplified using the following conditions: 95°C for 5 minutes, 4 cycles of 20 s at 98°C, 20 s at 70°C, minus 2°C/cycle, 30 s at 72°C, followed by 30 cycles of 15 s at 98°C, 20 s at 62°C, 20 s at 72°C, and a final extension at 72°C for 1 minute.

C. T7E1 Assay

PCR amplified target double-stranded DNA for T7E1 assays was denatured at 95°C for 10 minutes and then allowed to re-anneal by cooling to 25°C at -0.5°C/s in a thermal cycler. The re-annealed DNA was incubated with 0.5mL T7 Endonuclease I in 1x NEBuffer 2 buffer (New England Biolabs, Ipswich, MA) in a total volume of 15mL for 25 minutes at 37°C. T7E1 reactions were analyzed using the Fragment Analyzer™ system (Advanced Analytical Technologies, Inc., Ames, IA) and the DNF-910 double-stranded DNA Reagent Kit (Advanced Analytical Technologies, Inc., Ames, IA). The Fragment Analyzer™ system provides the concentration of each cleavage fragment and of the target double-stranded DNA that remains after cleavage.

Cleavage percentages of the target double-stranded DNA were calculated from the concentration of each cleavage fragment and the target double-stranded DNA, which remains after cleavage has taken place, using the following formula:

$$\%cleavage = \left(1 - \sqrt{1 - \frac{(frag1 + frag2)}{(frag1 + frag2 + parent)}}\right) \times N$$

EQUATION 1

In Equation 1, "frag1" and "frag2" concentrations correspond to the concentration of Cas cleavage fragments of the double-stranded DNA target and "parent" corresponds to the target double-stranded DNA that remains after cleavage has taken place.

FIGURE 5 shows the results of a T7E1 assay of gDNA prepped from cells transfected with crD(R)NAs at various concentrations. The average percent indels frequency detected was shown above each bar graph (calculated using Equation 1). The percent are the average of three samples, except for **FIGURE 5,** bar 4, in which activity was only detected in two samples and **FIGURE 5,** bar 5, in which activity was only detected in one sample. The concentration of either crD(R)NA/tracrRNA or sgRNA nucleofected into cells are shown in Table 6.

Table 6

Transfected Guide RNA Component Concentrations

#	SEQ ID NO.	pmol
1	SEQ ID NO: 43	500
2	SEQ ID NO: 43	750
3	SEQ ID NO: 43	1000
4	SEQ ID NO: 43	2000
5	SEQ ID NO: 43	3000
6	SEQ ID NO: 1	500

The T7E1 assay for detection of target modifications in eukaryotic cells provides data to demonstrate that the crD(R)NA/tracrRNA/Cas9 systems as described herein facilitate Cas-mediated site-specific *in vivo* cleavage of target double-stranded DNA.

Following the guidance describe herein, the T7E1 assay described in this example can be practiced by one of ordinary skill in the art to measure activity from cells modified with other CRISPR-Cas systems, including, but not limited to Cas9, Cas9-like, Cas1, Csn2, Cas4, Cpf1, C2c1, C2c2, C2c3, proteins encoded by Cas9 orthologs, Cas9-like synthetic proteins, Cas9 fusions, and variants and modifications thereof, combined with their cognate polynucleotide components modified as described herein to comprise a crD(R)NA.

Example 6

On/Off-Target crD(R)NA Cleavage Activity

This example illustrates the use of crD(R)NAs to evaluate the cleavage activity of a target at the intended target site (“on-target”) and predicted nearest neighbor (“off-target”) sites. Target sequences of on/off-target sites are shown in Table 7:

Table 7

On/Off-Target Site Sequences

Target	Site	Target Sequence	SEQ ID NO:
EMX-1	ON	5' - GAGTCCGAGC AGAAGAAGAA - 3'	SEQ ID NO: 27
EMX-1	OFF1	5' - GAGTTAGAGC AGAAGAAGAA - 3'	SEQ ID NO: 28
EMX-1	OFF2	5' - AGGTACTAGC AGAAGAAGAA - 3'	SEQ ID NO: 29
EMX-1	OFF3	5' - ACGTCTGAGC AGAAGAAGAA - 3'	SEQ ID NO: 30
EMX-1	OFF4	5' - AGGTGCTAGC AGAAGAAGAA - 3'	SEQ ID NO: 31
VEGFA-1	ON	5' - GGGTGGGGGG AGTTTGCTCC - 3'	SEQ ID NO: 32
VEGFA-1	OFF1	5' - GGATGGAGGG AGTTTGCTCC - 3'	SEQ ID NO: 33
VEGFA-1	OFF2	5' - GGGGAGGGGA AGTTTGCTCC - 3'	SEQ ID NO: 34
VEGFA-1	OFF3	5' - GGGAGGGTGG AGTTTGCTCC - 3'	SEQ ID NO: 35
VEGFA-1	OFF4	5' - CGGGGGAGGG AGTTTGCTCC - 3'	SEQ ID NO: 36

crRNA and crD(R)NA sequences were provided to a commercial manufacturer for synthesis. tracrRNA were constructed as described in Example 1.

Double stranded DNA targets were generated as described in Example 2 using the oligonucleotides shown in Table 8 corresponding to the appropriate target sequence.

Table 8

On/Off-Target DNA

Target	Site	Target Sequence
EMX-1	on	SEQ ID NOs.107, 108

EMX-1	OFF1	SEQ ID NOs. 111, 112
EMX-1	OFF2	SEQ ID NOs.113, 114
EMX-1	OFF3	SEQ ID NOs.115, 116
EMX-1	OFF4	SEQ ID NOs.117, 118
VEGFA-1	on	SEQ ID NOs.119, 120
VEGFA-1	OFF1	SEQ ID NOs.121, 122
VEGFA-1	OFF2	SEQ ID NOs.123, 124
VEGFA-1	OFF3	SEQ ID NOs.125, 126
VEGFA-1	OFF4	SEQ ID NOs.107, 108

crRNA/tracrRNA and crD(R)NA/tracrRNA were hybridized and biochemical cleavage was carried out as described in Example 3.

FIGURE 6 shows the comparison of biochemical activity of a crRNA/tracrRNA and crD(R)NA/tracrRNA at intended on-target sites and four computationally predicted off-target sites. Percent cleavage is shown on the y-axis and samples are shown on the x-axis. Table 9 lists the samples:

Table 9

crRNA and tracrRNA On/Off-target Activity

ID	Target Site	Guide RNA Component
1A	EMX-1 ON	crRNA
1B	EMX-1 ON	crD(R)NA
2A	EMX-1 OFF-1	crRNA
2B	EMX-1 OFF-1	crD(R)NA
3A	EMX-1 OFF-2	crRNA

3B	EMX-1 OFF-2	crD(R)NA
4A	EMX-1 OFF-3	crRNA
4B	EMX-1 OFF-3	crD(R)NA
5A	EMX-1 OFF-4	crRNA
5B	EMX-1 OFF-4	crD(R)NA
6A	VEGFA-1 ON	crRNA
6B	VEGFA-1 ON	crD(R)NA
7A	VEGFA-1 OFF-1	crRNA
7B	VEGFA-1 OFF-1	crD(R)NA
8A	VEGFA-1 OFF-2	crRNA
8B	VEGFA-1 OFF-2	crD(R)NA
9A	VEGFA-1 OFF-3	crRNA
9B	VEGFA-1 OFF-3	crD(R)NA
10A	VEGFA-1 OFF-4	crRNA
10B	VEGFA-1 OFF-4	crD(R)NA

The data presented in **FIGURE 7** show crD(R)NAs maintain high on-target activity when compared to crRNA. crD(R)NAs do not support off-target activity whereas the crRNAs have undesirable off-target activity.

Example 7

Deep Sequencing Analysis for Detection of Target Modifications in Eukaryotic Cells

This example illustrates the use of deep sequencing analysis to evaluate and compare the percent cleavage *in vivo* of selected sgD(R)NA/Cas9 protein complexes relative to selected double-stranded DNA target sequences.

A. Synthesis of sgD(R)NA

Six sgD(R)NA sequences targeting the human AAVS-1 locus and comprising different DNA/RNA compositions and phosphorothioate protected bonds were provided to a commercial manufacturer for synthesis. These sequences are shown in Table 10.

Table 10

sgD(R)NA Sequences

Name	Sequence (RNA bases are bracketed, phosphorothioate bonds are shown with an *)	SEQ ID NO:

sgD(R)NA -01	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U][G][C][U] - 3'	SEQ ID NO: 127
sgD(R)NA -02	5' - G*G*GGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U][G][C][U] - 3'	SEQ ID NO: 128
sgD(R)NA -03	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A]GC TGCT[G][A][A][A]AGC AUAGC[A][A][G][U][U] [A][A][A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] [C][G][A][G][U][C][G][G][U][G] [C][U] - 3'	SEQ ID NO: 129
sgD(R)NA -04	5' - G*G*GGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A][A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] [C][G][A][G][U][C][G][G][U][G] [C][U] - 3'	SEQ ID NO: 130
sgD(R)NA -05	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A][A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C][U][U][G][A][A] [A][A][A][G][U][G][G]CAC CG[A][G][U]CGGTG [C][U] - 3'	SEQ ID NO: 131

sgD(R)NA -06	5' - G*G*GGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A][A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C][U][U][G][A][A] [A][A][A][G][U][G][G]CAC CG[A][G][U]CGGTG [C][U] - 3'	SEQ ID NO: 132

B. Formation of RNP Complexes of sgD(R)NA/Cas9 protein

Cas9 protein was expressed from a bacterial expression vector in *E. coli* (BL21 (DE3)) and purified using affinity ion exchange and size exclusion chromatography according to methods described in Jinek *et al*. (Science; 337(6096):816-21(2012)). The coding sequence for *Streptococcus pyogenes* Cas9 included two nuclear localization sequences (NLS) at the C-terminus. Ribonucleoprotein (RNP) complexes were assembled, in triplicate, at two concentrations, 20pmol Cas9:60pmols sgD(R)NA and 200pmols Cas9:600pmols sgD(R)NA. The sgD(R)NA components were mixed in equimolar amounts in an annealing buffer (1.25mM HEPES, 0.625mM $MgCl_2$, 9.375mM KCl at pH7.5) to the desired concentration (60pmols or 600pmols) in a final volume of 5µL, incubated for 2 minutes at 95°C, removed from the thermocycler and allowed to equilibrate to room temperature. Cas9 protein was diluted to an appropriate concentration in binding buffer (20mM HEPES, 100mM KCl, 5mM $MgCl_2$, 1mM DTT, and 5% glycerol at pH 7.4) to a final volume of 5µL and mixed with the 5µL of heat-denatured crD(R)NAs followed by incubation at 37°C for 30 minutes.

C. Cell Transfections Using sgD(R)NA/Cas9 Protein RNPs

RNP complexes were transfected into K562 cells (ATCC, Manassas, VA), using the Nucleofector® 96-well Shuttle System (Lonza, Allendale, NJ) and the following protocol. RNP complexes were dispensed in a 10µL final volume into individual wells of a 96-well plate. K562 cells suspended in media were transferred from a culture flask to a 50mL conical tube. Cells were pelleted by centrifugation for 3 minutes at 200 x g, the culture medium aspirated, and the cells were washed once with calcium and magnesium-free PBS. K562 cells were then pelleted by centrifugation for 3 minutes at 200 x g, the PBS aspirated and cell pellet were resuspended in 10mL of calcium and magnesium-free PBS.

The cells were counted using the Countess® II Automated Cell Counter (Life Technologies, Grand Island, NY). 2.2 x 10^7 cells were transferred to a 50ml tube and pelleted. The PBS was aspirated and the cells were resuspended in Nucleofector™ SF (Lonza, Allendale,

NJ) solution to a density of 1 x 10^7 cells/mL. 20µL of the cell suspension were added to individual wells containing 10µL of RNP complexes and the entire volume was transferred to the wells of a 96-well Nucleocuvette™ Plate (Lonza, Allendale, NJ). The plate was loaded onto the Nucleofector™ 96-well Shuttle™ (Lonza, Allendale, NJ) and cells were nucleofected using the 96-FF-120 Nucleofector™ program (Lonza, Allendale, NJ). Post-nucleofection, 70µL Iscove's Modified Dulbecco's Media (IMDM; Life Technologies, Grand Island, NY), supplemented with 10% FBS (Fisher Scientific, Pittsburgh, PA), penicillin and streptomycin (Life Technologies, Grand Island, NY) was added to each well and 50µL of the cell suspension were transferred to a 96-well cell culture plate containing 150µL pre-warmed IMDM complete culture medium. The plate was then transferred to a tissue culture incubator and maintained at 37°C in 5% CO_2 for 48 hours.

D. Target Double-stranded DNA Generation for Deep Sequencing

gDNA was isolated from K562 cells 48 hours after RNP transfection using 50µL QuickExtract DNA Extraction solution (Epicentre, Madison, WI) per well followed by incubation at 37°C for 10 minutes, 65°C for 6 minutes and 95°C for 3 minutes to stop the reaction. The isolated gDNAs were diluted with 50µL water and samples stored at -80°C.

Using the isolated gDNA, a first PCR was performed using Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA) at 1x concentration, primers at 0.5µM each (SEQ ID NOs: 93, 94), 3.75µL of gDNA in a final volume of 10uL and amplified 98°C for 1 minute, 35 cycles of 10s at 98°C, 20s at 60°C, 30s at 72°C, and a final extension at 72°C for 2 min. PCR reaction were diluted 1:100 in water.

A "barcoding" PCR was set up using unique primers for each sample to facilitate multiplex sequencing. The samples and corresponding primer pairs are shown in Table 11.

Table 11

Barcoding Primers

ID	Sample	SEQ ID NO:
BARCODING PRIMER set-1	sgD(R)NA-01 60pmol rep1	SEQ ID NO: 95, 101
BARCODING PRIMER set-2	sgD(R)NA-02 60pmol rep1	SEQ ID NO: 95, 102
BARCODING PRIMER set-3	sgD(R)NA-03 60pmol rep1	SEQ ID NO: 95, 103
BARCODING PRIMER set-4	sgD(R)NA-04 60pmol rep1	SEQ ID NO: 95, 104
BARCODING PRIMER set-5	sgD(R)NA-05 60pmol rep1	SEQ ID NO: 95, 105
BARCODING PRIMER set-6	sgD(R)NA-06 60pmol rep2	SEQ ID NO: 95, 106
BARCODING PRIMER set-7	sgD(R)NA-01 60pmol rep2	SEQ ID NO: 96, 101
BARCODING PRIMER set-8	sgD(R)NA-02 60pmol rep2	SEQ ID NO: 96, 102
BARCODING PRIMER set-9	sgD(R)NA-03 60pmol rep2	SEQ ID NO: 96, 103
BARCODING PRIMER set-10	sgD(R)NA-04 60pmol rep2	SEQ ID NO: 96, 104
BARCODING PRIMER set-11	sgD(R)NA-05 60pmol rep2	SEQ ID NO: 96, 105
BARCODING PRIMER set-12	sgD(R)NA-06 60pmol rep2	SEQ ID NO: 96, 106
BARCODING PRIMER set-13	sgD(R)NA-01 60pmol rep3	SEQ ID NO: 97, 101
BARCODING PRIMER set-14	sgD(R)NA-02 60pmol rep3	SEQ ID NO: 97, 102
BARCODING PRIMER set-15	sgD(R)NA-03 60pmol rep3	SEQ ID NO: 97, 103
BARCODING PRIMER set-16	sgD(R)NA-04 60pmol rep3	SEQ ID NO: 97, 104
BARCODING PRIMER set-17	sgD(R)NA-05 60pmol rep3	SEQ ID NO: 97, 105
BARCODING PRIMER set-18	sgD(R)NA-06 60pmol rep3	SEQ ID NO: 97, 106

BARCODING PRIMER set-19	sgD(R)NA-01 600pmol rep1	SEQ ID NO: 98, 101
BARCODING PRIMER set-20	sgD(R)NA-02 600pmol rep1	SEQ ID NO: 98, 102
BARCODING PRIMER set-21	sgD(R)NA-03 600pmol rep1	SEQ ID NO: 98, 103
BARCODING PRIMER set-22	sgD(R)NA-04 600pmol rep1	SEQ ID NO: 98, 104
BARCODING PRIMER set-23	sgD(R)NA-05 600pmol rep1	SEQ ID NO: 98, 105
BARCODING PRIMER set-24	sgD(R)NA-06 600pmol rep2	SEQ ID NO: 98, 106
BARCODING PRIMER set-25	sgD(R)NA-01 600pmol rep2	SEQ ID NO: 99, 101
BARCODING PRIMER set-26	sgD(R)NA-02 600pmol rep2	SEQ ID NO: 99, 102
BARCODING PRIMER set-27	sgD(R)NA-03 600pmol rep2	SEQ ID NO: 99, 103
BARCODING PRIMER set-28	sgD(R)NA-04 600pmol rep2	SEQ ID NO: 99, 104
BARCODING PRIMER set-29	sgD(R)NA-05 600pmol rep2	SEQ ID NO: 99, 105
BARCODING PRIMER set-30	sgD(R)NA-06 600pmol rep2	SEQ ID NO: 99, 106
BARCODING PRIMER set-31	sgD(R)NA-01 600pmol rep3	SEQ ID NO: 100, 101
BARCODING PRIMER set-32	sgD(R)NA-02 600pmol rep3	SEQ ID NO: 100, 102
BARCODING PRIMER set-33	sgD(R)NA-03 600pmol rep3	SEQ ID NO: 100, 103
BARCODING PRIMER set-34	sgD(R)NA-04 600pmol rep3	SEQ ID NO: 100, 104
BARCODING PRIMER set-35	sgD(R)NA-05 600pmol rep3	SEQ ID NO: 100, 105
BARCODING PRIMER set-36	sgD(R)NA-06 600pmol rep3	SEQ ID NO: 100, 106

The barcoding PCR was performed using Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA) at 1x concentration, primers at 0.5μM each, 1μL of 1:100 diluted first PCR, in a final volume of 10μL and amplified 98°C for 1 minutes, 12 cycles of 10s at 98°C, 20s at 60°C, 30s at 72°C, and a final extension at 72°C for 2 min.

E. SPRIselect Clean-up

PCR reactions were pooled into a single microfuge tube for SPRIselect (Beckman Coulter, Pasadena, CA) bead-based clean-up of amplicons for sequencing.

To the pooled amplicons, 0.9x volumes of SPRIselect beads were added, and mixed and incubated at room temperature (RT) for 10 minutes. The microfuge tube was placed on a magnetic tube stand (Beckman Coulter, Pasadena, CA) until solution had cleared. Supernatant was removed and discarded, and the residual beads were washed with 1 volume of 85% ethanol, and incubated at RT for 30 seconds. After incubation, ethanol was aspirated and beads are air dried at RT for 10 min. The microfuge tube was then removed from the magnetic stand and 0.25x volumes of Qiagen EB buffer (Qiagen, Venlo, Limburg) was added to the beads, mixed vigorously, and incubated for 2 minutes at room temperature. The microfuge tube was returned to the magnet, incubated until solution had cleared, and supernatant containing the purified amplicons was dispensed into a clean microfuge tube. The purified amplicon library was quantified using the Nanodrop™ 2000 system (Thermo Scientific, Wilmington, DE) and library-quality analyzed using the Fragment Analyzer™ system (Advanced Analytical Technologies, Inc., Ames, IA) and the DNF-910 double-stranded DNA Reagent Kit (Advanced Analytical Technologies, Inc. Ames, IA).

F. Deep Sequencing Set-up

The amplicon library was normalized to a 4 nmolar concentration as calculated from Nanodrop values and size of the amplicons. The library were analyzed on MiSeq Sequencer (Illumina, San Diego, CA) with MiSeq Reagent Kit v2 (Illumina, San Diego, CA) for 300 cycles with two 151-cycle paired-end run plus two eight-cycle index reads.

G. Deep Sequencing Data Analysis

The identity of products in the sequencing data were determined based on the index barcode sequences adapted onto the amplicons in the barcoding round of PCR. A computational script was used to process the MiSeq data by executing the following tasks:

- Reads were aligned to the human genome (build GRCh38/38) using Bowtie (http://bowtie-bio.sourceforge.net/index.shtml) software.
- Aligned reads were compared to the expected wild-type AAVS-1 locus sequence, reads not aligning to any part of the AAVS-1 locus were discarded.
- Reads matching wild-type AAVS-1 sequence were tallied.
- Reads with indels (insertion or the deletion of bases) were categorized by indel type and tallied.
- Total indel reads were divided by the sum of wild-type reads and indel reads give the percent indels detected.

FIGURE 7 shows the results of an analysis of the AAVS-1 target locus from human K562 cells nucleofected with sgD(R)NA/Cas9 targeting a region in the AAVS-1 locus. The x-axis shows the SEQ ID NO. For the sgD(R)NA used, the y-axis shows the percent indel detected from MiSeq data. Series A shows the average percent indels detected for three independent replicates for a given sgD(R)NA at 20pmols Cas9:120pmols sgD(R)NA, and Series B shows the average percent indels detected for three independent replicates for a given sgD(R)NA at 100pmols Cas9:600pmols sgD(R)NA. Standard deviation of the average percent of the three replicates is represented by vertical black lines. The numbers below the bars correspond to the SEQ ID NO. of the sgD(R)NA used in the transfection, sequences of the sgD(R)NA are provided in Table 10. This data shows the ability of various types of sgD(R)NA to induce modifications at a target region in human cells in a sequence specific and dose dependent manner.

The methods described herein were practiced by one of ordinary skill in the art to demonstrate *in viv*o activity of a sgD(R)NA/Cas9 through analysis of deep sequencing.

Example 8

Screening of Multiple crD(R)NAs Comprising DNA Target-Binding Sequences

This example illustrates the use of crD(R)NAs of the present disclosure to modify targets present in human genomic DNA and measure the level of cleavage activity at those sites. Target sites can first be selected from genomic DNA and then crD(R)NAs can then be designed to target those selected sequences. Measurements can then be carried out to determine the level of target cleavage that has taken place. Not all of the following steps are required for every screening nor must the order of the steps be as presented, and the screening can be coupled to other experiments, or form part of a larger experiment.

A. Select a DNA Target Region from Genomic DNA

Identify all PAM sequences (*e.g.*, 'NGG') within the selected genomic region.

Identify and select one or more 20 nucleotide sequence long sequences (target DNA sequence) that are 5' adjacent to PAM sequences.

Selection criteria can include but are not limited to: homology to other regions in the genome; percent G-C content; melting temperature; presences of homopolymer within the spacer; and other criteria known to one skilled in the art.

Append an appropriate crD(R)NA sequence to the 3' end of the identified target DNA sequence. A crD(R)NA construct is typically synthesized by a commercial manufacturer and the cognate tracrRNA is produced as described in Example 1 by *in vitro* transcription.

A crD(R)NA as described herein can be used with cognate tracrRNA to complete a crD(R)NA/tracrRNA system for use with a cognate Cas protein.

B. Determination of Cleavage Percentages and Specificity

In vitro cleavage percentages and specificity associated with a crD(R)NA/tracrRNA system are compared, for example, using the Cas cleavage assays of Example 3, as follows:

(a) If only a single target DNA sequence is identified or selected, the cleavage percentage and specificity for the DNA target region can be determined. If so desired, cleavage percentage and/or specificity can be altered in further experiments using methods of the present disclosure including but not limited to modifying the crD(R)NA, introducing effector proteins/effector protein-binding sequences or ligand/ligand binding moieties.

(b) The percentage cleavage data and site-specificity data obtained from the cleavage assays can be compared between different DNAs comprising the target binding sequence to identify the target DNA sequences having the best cleavage percentage and highest specificity. Cleavage percentage data and specificity data provide criteria on which to base choices for a variety of applications. For example, in some situations the activity of the crD(R)NA may be the most important factor. In other situations, the specificity of the cleavage site may be relatively more important than the cleavage percentage. If so desired, cleavage percentage and/or specificity are altered in further experiments using methods of the present disclosure including but not limited to modifying the crD(R)NA, introducing effector proteins/effector protein-binding sequences or ligand/ligand binding moieties.

Optionally, or instead of, the *in vitro* analysis, *in vivo* cleavage percentages and specificity associated with a crD(R)NA system are compared, for example, using the T7E1 assay described in Example 5, as follows:

(a) If only a target DNA sequence is identified the cleavage percentage and specificity for the DNA target region can be determined. If so desired, cleavage percentage and/or specificity are altered in further experiments using methods of the present disclosure including but not limited to modifying the crD(R)NA, introducing effector proteins/effector protein-binding sequences or ligand/ligand binding moieties.

(b) The percentage cleavage data and site-specificity data obtained from the cleavage assays can be compared between different target DNAs to identify a crD(R)NA sequence that results in the highest percentage cleavage of target DNA and the highest specificity for the target DNA. Cleavage percentage data and specificity data provide criteria on which to base choices for a variety of applications. For example, certain embodiments may rely on the activity of a crD(R)NA and may be the most important factor. In certain embodiments, the specificity of the cleavage site may be relatively more important than the cleavage percentage. In certain embodiments, cleavage percentage and/or specificity can be altered using methods of the present

disclosure including but not limited to modifying the RNA, introducing effector proteins/effector protein-binding sequences or ligand/ligand binding moieties.

Following the guidance of the present specification and examples, the screening described in this example can be practiced by one of ordinary skill in the art with other Class II CRISPR Cas proteins, including, but not limited to Cas9, Cas9-like, Cas, Cas3, Csn2, Cas4, proteins encoded by Cas9 orthologs, Cas9-like synthetic proteins, Cas9 fusions, Cpf1, Cpf1-like, C2c1, C2c2, C2c3, and variants and modifications thereof, combined with their cognate polynucleotide components modified as described herein to comprise a crD(R)NA.

Example 9

crD(R)NA:tracrRNA and sgD(R)NA Mediated Nicking

This example illustrates the method through which a crD(R)NA:tracrRNA complex or sgD(R)NA of the present disclosure might be used to induced nicks in a double stranded DNA (dsDNA) plasmid target in conjunction with *S. pyogenes* Cas9 containing a D10A mutation (Cas9-D10A) rendering the RuvC nuclease lobe inactive. Not all of the following steps are required, nor must the order of the steps be as presented.

The *S. pyogenes* Cas9 has two active nuclease domains, the RuvC and the HNH domains. A mutation of the aspartic acid at the 10th amino acid position of the *S. pyogenes* Cas9, converting it to an alanine, reduces the nuclease capability of the RuvC domain. The HNH domain remains active but the Cas9-D10A site-directed polypeptide can only cause nicks in the phosphodiester backbone of the DNA target strand complementary to the spacer sequence.

Examples of suitable vectors, media, culture conditions, etc. are described. Modifications of these components and conditions will be understood by one of ordinary skill in the art in view of the teachings of the present specification.

Guide reagents were generated according to Example 1 of the present specification.

The dsDNA target was generated as described in Example 2 using SEQ ID NOs 133 and 134. The amplified fragment was then cloned into suitable LIC compatible vector. One such suitable vector is the commercially available pET His6 LIC cloning vector (Addgene, Cambridge, MA). The plasmid was transformed into bacterial strain for plasmid expression, using commercially availble XL1-Blu bacterial cells (Agilent, Santa Clara, CA).

Bacterial cells containing the LIC vectors were grown in LB media supplemented with 100ug/mL ampicillin (Sigma-Aldrich, St. Louis, MO) for 18 hours at 37°C. Cells were centrifuged at 5,000 rpm for 15 minutes, after which the plasmid was extracted using Qiagen Plasmid Kit (Qiagen, Venlo, Netherlands).

Biochemical cleavage of purified plasmid was performed as detailed in Example 3 of the present specification, with the modification that DNA target was replaced with the purified plasmid at a final concentration of 1nM in the reaction. crD(R)NA were hybridized with tracrRNA (SEQ ID NO: 2) in the manner described in Example 3.

Biochemical reactions were analyzed by running on a 1% agarose gel stained with SYBR gold (Life Technologies, Grand Island, NY). Nicking efficiency was calculated based upon the disappearances of supercoiled plasmid form and the appearance of the nicked-open circular form of the plasmid (nicked plasmid), which was distinguishable by the shift in the migration rate of the plasmid on the gel.

Percentages of the nicked plasmid were calculated from the intensities of stained bands on the gel containing the nicked plasmid and the supercoiled plasmid. Intensities were measured using area under the curve values as calculated by FIJI (ImageJ; an open source Java image processing program). Percentages of nicking were calculated by dividing the staining intensity of the nicked plasmid by the sum of both the staining intensities of the nicked plasmid species and the supercoiled plasmid species.

SEQ ID NOs for the crD(R)NA and sgD(R)NA used in this experiment are shown in Table 12.

Table 12

Nicking crD(R)NA and sgD(R)NA

Sample ID	Description	SEQ ID NO:
A	crD(R)NA	SEQ ID NO: 38
B	crD(R)NA w/ 18nt spacer	SEQ ID NO: 135
C	crD(R)NA	SEQ ID NO: 41
D	crD(R)NA w/ 17nt spacer	SEQ ID NO: 136
E	crD(R)NA	SEQ ID NO: 43
F	crD(R)NA w/ 18nt spacer	SEQ ID NO: 137
H	sgD(R)NA	SEQ ID NO: 127
I	sgRNA control	SEQ ID NO: 1
H	target plasmid only	-

FIGURE 8 shows the results of the biochemical nicking activity of a crD(R)NA or sgD(R)NA with a Cas9-D10A protein against a plasmid target. Nicking percentages are shown on the y-axis. crD(R)NA and sgD(R)NA samples are shown on the x-axis and correspond to the

sample IDs shown in Table 12. The data show the ability of crD(R)NA and sgD(R)NA to support nicking activity of the Cas9-D10A protein against a target plasmid. The data also show that truncation of the spacer sequence from the 5' end of the spacer (SEQ ID NOs: 135, 136, and 137) is capable of nicking activity.

Following the guidance of the present specification and the examples herein, the design and validation of the nicking activity of crD(R)NA:tracrRNA and sgD(R)NA can be practiced by one of ordinary skill in the art.

Example 10

Identification and Screening of CRISPR RNA and Trans-activating CRISPR RNA

This example illustrates the method through which CRISPR RNAs (crRNAs) and trans-activating CRISPR RNAs (tracrRNAs) of a CRISPR-Cas Type II system may be identified. The method presented here is adapted from Chylinski, et. al., (RNA Biol;10(5):726-37 (2013)). Not all of the following steps are required for screening nor must the order of the steps be as presented.

A. Identify a Bacterial Species Containing a CRISPR-Cas9 Type-II System

Using the Basic Local Alignment Search Tool (BLAST, blast.ncbi.nlm.nih.gov/Blast.cgi), a search of various species' genomes is conducted to identify Cas9 or Cas9-like proteins. Type II CRISPR-Cas9 systems exhibit a high diversity in sequence across bacterial species, however Cas9 orthologs exhibit conserved domain architecture of central HNH endonuclease domain and a split RuvC/RNase H domain. Primary BLAST results are filtered for identified domains; incomplete or truncated sequences are discarded and Cas9 orthologs identified.

When a Cas9 ortholog is identified in a species, sequences adjacent to the Cas9 ortholog's coding sequence are probed for other Cas proteins and an associated repeat-spacer array in order to identify all sequences belonging to the CRISPR-Cas locus. This may be done by alignment to other CRISPR-Cas Type-II loci already known in the public domain, with the knowledge that closely related species exhibit similar CRISPR-Cas9 locus architecture (*i.e.*, Cas protein composition, size, orientation, location of array, location of tracrRNA, etc.).

B. Identification of Putative crRNA and tracrRNA

Within the locus, the crRNAs are readily identifiable by the nature of their repeat sequences interspaced by fragments of foreign DNA and make up the repeat-spacer array. If the repeat sequence is from a known species, it is identified in and retrieved from the CRISPRdb database (crispr.u-psud.fr/crispr/). If the repeat sequence is not known to be associated with a species, repeat sequences are predicted using CRISPRfinder software (crispr.u-psud.fr/Server/) using the sequence identified as a CRISPR-Cas Type-II locus for the species as described above.

Once the sequence of the repeat sequence is identified for the species, the tracrRNA is identified by its sequence complementarity to the repeat sequence in the repeat-spacer array (tracr anti-repeat sequence). *In silico* predictive screening is used to extract the anti-repeat sequence to identify the associated tracrRNA. Putative anti-repeats are screened, for example, as follows.

The identified repeat sequence for a given species is used to probe the CRISPR-Cas9 locus for the anti-repeat sequence (*e.g.*, using the BLASTp algorithm or the like). The search is typically restricted to intronic regions of the CRISPR-Cas9 locus.

An identified anti-repeat region is validated for complementarity to the identified repeat sequence.

A putative anti-repeat region is probed both 5' and 3' of the putative anti-repeat for a Rho-independent transcriptional terminator (TransTerm HP, transterm.cbcb.umd.edu/).

Thus, the identified sequence comprising the anti-repeat element and the Rho-independent transcriptional terminator is determined to be the putative tracrRNA of the given species.

C. Preparation of RNA-Seq Library

The putative crRNA and tracrRNA that were identified *in silico* are further validated using RNA sequencing (RNAseq).

Cells from species from which the putative crRNA and tracrRNA were identified are procured from a commercial repository (*e.g.*, ATCC, Manassas, VA; DSMZ, Braunschweig, Germany).

Cells are grown to mid-log phase and total RNA prepped using Trizol reagent (Sigma-Aldrich, St. Louis, MO) and treated with DNaseI (Fermentas, Vilnius, Lithuania).

10ug of the total RNA is treated with Ribo-Zero rRNA Removal Kit (Illumina, San Diego, CA) and the remaining RNA purified using RNA Clean and Concentrators (Zymo Research, Irvine, CA).

A library is then prepared using TruSeq Small RNA Library Preparation Kit (Illumina, San Diego, CA) following the manufacturer's instructions, which results in the presence of adapter sequences associated with the cDNA.

The resulting cDNA library is sequenced using MiSeq Sequencer (Illumina, San Diego, CA).

D. Processing of Sequencing Data

Sequencing reads of the cDNA library can be processed using the following method.

Adapter sequences are removed using cutadapt 1.1 (pypi.python.org/pypi/cutadapt/1.1) and 15 nt are trimmed from the 3'end of the read to improve read quality.

Reads are aligned back to each respective species' genome (from which the putative tracrRNA was identified) with a mismatch allowance of 2 nucleotides.

Read coverage is calculated using BedTools (bedtools.readthedocs.org/en/latest/).

Integrative Genomics Viewer (IGV, www.broadinstitute.org/igv/) is used to map the starting (5') and ending (3') position of reads. Total reads retrieved for the putative tracrRNA are calculated from the SAM file of alignments.

The RNA-seq data is used to validate that a putative crRNA and tracrRNA element is actively transcribed *in vivo*. Confirmed hits from the composite of the *in silico* and RNA-seq screens are validated for functional ability of the identified crRNA and tracrRNA sequences to support Cas9 mediated cleavage of a double-stranded DNA target using methods outline herein (*see* Examples 1, 2, and 3).

Following the guidance of the present specification and the examples herein, the identification of novel crRNA and tracrRNA sequences can be practiced by one of ordinary skill in the art.

Example 11

Design of crD(R)NA and sgD(R)NA

This example illustrates the method through which crD(R)NA and sgD(R)NA are designed from crRNA and tracrRNA, respectively. Not all of the following steps are required for screening nor must the order of the steps be as presented.

Identification of the crRNA and tracrRNA guide sequences for a given species are performed as described in Example 10.

Identified crRNA and tracrRNA sequences are reverse-transcribed *in silico* to DNA. Upper stem, lower stem and bulge elements are identified from the sequences of the crRNA and tracrRNA. RNA bases are introduced into the DNA sequence of the crDNA and tracrDNA sequences creating crD(R)NA and sgD(R)NA, respectively. The placement, number and distribution of RNA bases within the crDNA and tracrRNA can be chosen using either computational or experimental screening methods. A collection of crD(R)NAs are designed with ribonucleotides placed in a number of different locations within the molecule. Preferably, deoxyriboucleotides within the lower stem are substituted for ribonucleotides in some crD(R)NA sequences. Ribonucleotides are substituted at the 3' end of the spacer sequence in some crD(R)NA sequences. Additional crD(R)NA and sgD(R)NA sequences are designed, for example, as follows.

Repositories of 3-dimensional protein structures (*e.g.*, RCSB PDB; rcsb.org) in the public domain are searched to identify Cas endonuclease structures. The repository is searched for high

resolution coordinate files of Cas endonucleases bound to their cognate crRNA and tracrRNA. Structural neighbors, defined by sequence or tertiary structural similarities to the Cas endonuclease of interest are used if there is no solved structure for the Cas endonuclease of interest. Deposited coordinate files are downloaded. Using visualization software, such as PyMOL (PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC), the coordinates are analyzed to identify ribose-specific interactions between the Cas endonuclease protein and the nucleotides of the crRNA and tracrRNA. Positions where the protein makes direct or indirect contact (*i.e.*, through a water or metal intermediate) with the nucleotides of the crRNA and tracrRNA are used to identify favored positions within the guide sequences for replacing deoxyribonucleotides with ribonucleotides or other nucleotide variants.

crRNA and tracrRNA sequences are conserved when compared with Cas9 proteins from related species. Alignment of a guide sequence with the other known guide sequences from similar species provides additional information on conserved bases that would confer a preference for ribonucleotides. Multiple sequence alignments of crRNA or tracrRNA are performed using the web-based software MUSCLE (ebi.ac.uk/Tools/mas/muscle/). Alignments are then assessed for conserved nucleotide sequence positions along the backbone.

Nucleic acid secondary structure prediction software (e.g RNAfold; rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) is used to analyze the folding of the guide backbone. Regions where RNA specific torsion angles would be favored are used to inform placement of ribonucleotide locations in both the crDNA and/or tracrDNA.

Combinations of secondary structure, protein-nucleic acid interaction, and sequence conservation are used to inform the positioning of ribonucleotides within crD(R)NA, tracrD(R)NA and sgD(R)NA sequence. Multiple designs of crD(R)NA and tracrD(R)NA are tested with the understanding that different configurations may support different desired properties (*i.e.*, activity, specificity, stability, etc.). The crD(R)NA and tracrD(R)NA can be joined into a single molecule by a linker to form a sgD(R)NA. The combining of the crD(R)NA and tracrD(R)NA may be accompanied by a reduction in the total number of nucleotides at the 3' end of the crD(R)NA and 5' end of the tracrD(R)NA that together would form the upper stem. SEQ ID NOs 138-142, 147-150, 154-157, and 161-164 show designs for crD(R)NAs and tracrD(R)NAs. SEQ ID NOs 143-146, 151-153, 158-160, and 165-167 show designs for sgD(R)NAs. Table 13 gives the identity of sequences.

Table 13

crD(R)NA, tracrD(R)NA, and sgD(R)NA

ID	Genus/Species	Guide Description
SEQ ID NO: 138	*Staphylococcus aureus*	crD(R)NA
SEQ ID NO: 139	*Staphylococcus aureus*	crD(R)NA
SEQ ID NO: 140	*Staphylococcus aureus*	crD(R)NA
SEQ ID NO: 141	*Staphylococcus aureus*	crD(R)NA
SEQ ID NO: 142	*Staphylococcus aureus*	tracrRNA
SEQ ID NO: 143	*Staphylococcus aureus*	sgD(R)NA
SEQ ID NO: 144	*Staphylococcus aureus*	sgD(R)NA
SEQ ID NO: 145	*Staphylococcus aureus*	sgD(R)NA
SEQ ID NO: 146	*Staphylococcus aureus*	sgD(R)NA
SEQ ID NO: 147	*Streptococcus thermophilus* CRISPR-I	crD(R)NA
SEQ ID NO: 148	*Streptococcus thermophilus* CRISPR-I	crD(R)NA
SEQ ID NO: 149	*Streptococcus thermophilus* CRISPR-I	crD(R)NA
SEQ ID NO: 150	*Streptococcus thermophilus* CRISPR-I	tracrRNA
SEQ ID NO: 151	*Streptococcus thermophilus* CRISPR-I	sgD(R)NA
SEQ ID NO: 152	*Streptococcus thermophilus* CRISPR-I	sgD(R)NA
SEQ ID NO: 153	*Streptococcus thermophilus* CRISPR-I	sgD(R)NA

SEQ ID NO: 154	*Neisseria meningitidis*	crD(R)NA
SEQ ID NO: 155	*Neisseria meningitidis*	crD(R)NA
SEQ ID NO: 156	*Neisseria meningitidis*	crD(R)NA
SEQ ID NO: 157	*Neisseria meningitidis*	tracrRNA
SEQ ID NO: 158	*Neisseria meningitidis*	sgD(R)NA
SEQ ID NO: 159	*Neisseria meningitidis*	sgD(R)NA
SEQ ID NO: 160	*Neisseria meningitidis*	sgD(R)NA
SEQ ID NO: 161	*Streptococcus pasteurianus*	crD(R)NA
SEQ ID NO: 162	*Streptococcus pasteurianus*	crD(R)NA
SEQ ID NO: 163	*Streptococcus pasteurianus*	crD(R)NA
SEQ ID NO: 164	*Streptococcus pasteurianus*	tracrRNA
SEQ ID NO: 165	*Streptococcus pasteurianus*	sgD(R)NA
SEQ ID NO: 166	*Streptococcus pasteurianus*	sgD(R)NA
SEQ ID NO: 167	*Streptococcus pasteurianus*	sgD(R)NA

Sequences are provided to a commercial manufacturer (*e.g.*, Integrated DNA Technologies, Coralville, IA) for synthesis.

crD(R)NA, tracrD(R)NA, and sgD(R)NA are tested experimentally to determine the activity of different sequences to support Cas9 mediated cleavage of a double-stranded DNA target using methods set forth herein (*see* Examples 1, 2, and 3).

Following the guidance of the present specification and the examples herein, the design and validation of novel crD(R)NA, tracrD(R)NA, and sgD(R)NA sequences can be practiced by one of ordinary skill in the art.

Example 12

Design of Type V Cpf1 crD(R)NA and sgD(R)NA Elements and Use with Cpf1 to Modify DNA

Tables 14 and 15 below provide exemplary dual guide crD(R)NAs and sgD(R)NAs for use with Type V CRISPR systems. The reference to exemplary figures and SEQ ID NOs is not intended to be limiting in anyway and it is understood by one of skill in the art that, based on the disclosure in Tables 14, 15, and the associated SEQ ID Nos and exemplary figures, dual guide crD(R)NAs and sgD(R)NAs for use with Type V CRISPR systems can be designed to target any desired sequence within a target nucleic acid.

Table 14

Description of Type V crD(R)NA 5' and 3' Elements and Combinations Used to Form Dual Guide crD(R)NAs and to Direct Cpf1 Activity to DNA Sequence of Interest

Description of Sequence	Exemplary Figure	SEQ ID NO:
Type V Cpf1 crRNA 5' element	12B, 13D, 13E, 13H	SEQ ID NO: 168
Type V Cpf1 crD(R)NA 5' element	12C, 13B, 13C, 13F, 13G	SEQ ID NO: 169
Phosphorothioate-protected Type V Cpf1 crRNA 5' element	12B, 13D, 13E, 13H	SEQ ID NO: 170
Phosphorothioate-protected Type V Cpf1 crD(R)NA 5' element	12C, 13B, 13C, 13F, 13G	SEQ ID NO: 171
Type V Cpf1 crRNA 3' element with 25 nucleotide RNA targeting region	12D	SEQ ID NO: 172
Type V Cpf1 crRNA 3' element with 20 nucleotide RNA targeting region	12D	SEQ ID NO: 173
Phosphorothioate-protected Type V Cpf1 crRNA 3' element with 25 nucleotide RNA targeting region	12D	SEQ ID NO: 174
Phosphorothioate-protected Type V Cpf1 crRNA 3' element with 20 nucleotide RNA targeting region	12D	SEQ ID NO: 175
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA targeting region	12F,13E ,13F	SEQ ID NO: 176

Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO: 177
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO:178
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO: 179
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide RNA targeting region	12E, 13C, 13D	SEQ ID NO: 180
Phosphorothioate-protected Type V Cpf1 crD(R)NA 3' element with 25 nucleotide RNA targeting region	12E, 13C, 13D	SEQ ID NO: 181
Type V Cpf1 crD(R)NA 3' element with 20 nucleotide RNA targeting region	12E, 13C, 13D	SEQ ID NO: 182
Phosphorothioate-protected Type V Cpf1 crD(R)NA 3' element with 20 nucleotide RNA targeting region	12E, 13C, 13D	SEQ ID NO: 183
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA targeting region	12G, 13G, 13H	SEQ ID NO: 184
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO: 185
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO: 186
Type V Cpf1 crD(R)NA 3' element with 25 nucleotide DNA/RNA targeting region	12H, 12I	SEQ ID NO: 187
Dual guide Type V Cpf1 crRNA containing 3' and 5' elements	13A	SEQ ID NO: 168; SEQ ID NO: 172
Dual guide Type V Cpf1 crRNA containing phosphorothioate protected 3' and 5' elements	13A	SEQ ID NO: 170; SEQ ID NO: 173

Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13B	SEQ ID NO: 169; SEQ ID NO: 172
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13C	SEQ ID NO: 169; SEQ ID NO: 180
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13D	SEQ ID NO: 168; SEQ ID NO: 180
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13E	SEQ ID NO: 168; SEQ ID NO: 176
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13F	SEQ ID NO: 169; SEQ ID NO: 176
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13G	SEQ ID NO: 169; SEQ ID NO: 184
Dual guide Type V Cpf1 cr(D)RNA containing 3' and 5' elements	13H	SEQ ID NO: 168; SEQ ID NO: 184

Table 15

Description of Type V sgD(R)NA Designs

Description of Sequence	Exemplary Figure	SEQ ID NO:

Type V Cpf1 sgD(R)NA with 25 nucleotide RNA targeting region	10A	SEQ ID NO: 188
Type V Cpf1 sgD(R)NA with 25 nucleotide RNA targeting region	10B	SEQ ID NO: 189
Type V Cpf1 sgD(R)NA with 25 nucleotide RNA targeting region	10C	SEQ ID NO: 190
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA targeting region	11D	SEQ ID NO: 191
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA targeting region	11B	SEQ ID NO: 192
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 193
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 194
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 195
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA targeting region	11A	SEQ ID NO: 196
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 197
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 198
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 199
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA targeting region	11C	SEQ ID NO: 200
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 201
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 202
Type V Cpf1 sgD(R)NA with 25 nucleotide DNA/RNA targeting region	11E	SEQ ID NO: 203

A. Design of Type V Cpf1 crD(R)NA and sgD(R)NA Elements

Cpf1 orthologs are identified using sequence analysis programs such as PSI-BLAST, PHI-BLAST and HMMer. Once a Cpf1 ortholog is identified, nearby sequences are searched to identify the associated CRISPR array. crRNA sequences are identified as repeat sequences located within the CRISPR array as described in Zetsche *et al.* (Cell;163(3):759-71(2015)). Type V crRNA sequences contain a stem loop within the repeat sequence, located 5' to the targeting region sequence. The stem loop comprises a 5' element and a 3' element. The sequences of both the 5' element, the 3' element, and the loop of the crRNA are identified. The sequence of these crRNA elements are reverse-transcribed *in silico* to DNA. 5' elements are designed containing mixtures of ribonucleotides and deoxyribonucleotides. Examples of 5' elements are shown in **FIGURE 12**, **FIGURE 13** and Table 14. 3' elements are designed containing mixtures of ribonucleotides and deoxyribonucleotides. Examples of 3' elements are shown in **FIGURE 12**, **FIGURE 13** and Table 14. Targeting region sequences are selected to be adjacent to PAM sequences in the the DNA of interest and are appended to the 3' end of 3' crRNA elements. Targeting region sequences are designed containing DNA, DNA and RNA, or RNA nucleotides. By combining crD(R)NA 3' elements and crD(R)NA 5' elements together (Table 14, **FIGURE 12, FIGURE 13**) to form dual guide TypeV crD(R)NAs, Cpf1 is directed to cut target nucleic acid sequences in the target nucleic acid of interest. A collection of crD(R)NAs for testing are designed with ribonucleotides placed in a number of different locations within the crD(R)NA sequences. Preferably, deoxyriboucleotides within the 3'stem and 5'stem are substituted for ribonucleotides in some crD(R)NA sequences. Ribonucleotides are substituted at the 5' end of the targeting region sequence in some crD(R)NA sequences.

Using combinations of targeting region, 3' elements, and 5' elements connected by a loop sequence, different versions of sgD(R)NA are designed. The placement, number, and distribution of RNA bases within the sgD(R)NA can be chosen using either computational or experimental screening methods. A collection of sgD(R)NAs are designed with ribonucleotides placed in a number of different locations within the sgD(R)NAs. Preferably, deoxyriboucleotides within the 3'stem and 5'stem are substituted for ribonucleotides in some sgD(R)NA sequences. Ribonucleotides are substituted at the 5' end of the targeting region sequence in some sgD(R)NA sequences. Examples of designed sgD(R)NAs are listed in Table 15, and shown in **FIGURES 10A-C** and **FIGURES 11A-E**.

In the following, sgD(R)NA sequences are used, but it is understood that pairs of 3' and 5' crD(R)NA elements (examples of which are shown in Table 14) can be used in place of the sgD(R)NA.

B. Digestion of Nucleic Acid Sequences with Cpf1 and sgD(R)NA

Cpf1 sgD(R)NA can be used together with Cpf1 to target and cut nucleic acid sequences. Target nucleic acid is either RNA, genomic DNA, plasmid DNA, or amplified DNA. Amplified target DNA can be prepared as described in Example 2. sgD(R)NA sequences are synthesized containing spacer sequences targeting sequences of interest in the target DNA. Cleavage assays are carried out as described in Zetsche *et al.* (2015) and analyzed using methods described in Example 3. In summary, target nucleic acid is incubated with Cpf1 and the sgD(R)NA sequence or sequences in an appropriate buffer chosen to support Cpf1 activity. Nucleic acid is analyzed to determine whether digestion has taken place as described in Example 3. Two or more Cpf1/sgD(R)NA complexes can be used to cut sections of DNA from a target DNA. The section of DNA has overhanging ends and can be ligated to complementary sequence adaptors or vectors after it has been separated from the parent DNA.

C. Genome Editing with Cpf1 sgD(R)NA Ribonucleoprotein Complexes

An *E. coli* expression vector is constructed by synthesizing a codon-optimized open-reading frame encoding Cpf1 and cloning the open-reading frame into an expression plasmid (*e.g.*, pET27b). The coding sequence can include an affinity tag for purification of the protein, and a NLS sequence at the C-terminus to drive nuclear localization in eukaryotic cells. Cpf1 protein can be expressed in *E. coli* from the expression vector and purified using a combination of affinity, ion exchange and size exclusion chromatography. The purified protein is concentrated to 10 mg/ml and combined with the sgD(R)NA to make a ribonucleoprotein complex. 200pmol of Cpf1 is combined in separate reaction tubes with 50pmol, 100pmol, 200pmol, 400pmol, 600pmol, 800pmol, 1000 pmol of sgD(R)NA and a reaction buffer. Cpf1-sgD(R)NA complexes are electroporated in replicate into HEK293 cells according to the methods described in Example 7. Cells are grown at 37°C and genomic DNA is harvested from each reaction after 4, 8, 16, 24, 48, and 72 hours. Genomic DNA is analyzed using PCR and Illumina sequencing to determine that the genome has been edited according to the methods described in Example 7.

D. Genome Editing using Cpf1 Expression Vectors and sgD(R)NA in Eukaryotic Cells

A mammalian expression vector can be constructed by synthesizing a codon-optimized open-reading frame encoding Cpf1 and cloning the open-reading frame into a suitable mammalian expression plasmid (*e.g.*, pcDNA3.1). The coding sequence can include a HA affinity tag for purification or detection of the protein, and a NLS sequence at the C-terminus to drive nuclear localization in eukaryotic cells. The coding sequence can be operably linked to the CMV promoter in the plasmid. Cpf1-expressing plasmids are combined in separate reaction tubes with 50pmol, 100pmol, 200pmol, 400pmol, 600pmol, 800pmol, 1000 pmol of sgD(R)NA

and a reaction buffer. Reaction mixtures are electroporated in replicate into HEK293 cells according to methods described in Example 7. Cells are grown at 37°C and genomic DNA is harvested from each reaction after 4, 8, 16, 24, 48, and 72 hours. Genomic DNA is analyzed using PCR and Illumina sequencing to determine that the genome has been edited according to the methods described in

Example 7.

Example 13

In planta Modification of Maize Embryos

This example illustrates the method by which single guide D(R)NA can be used to modify maize embryos. The method presented here is adapted from Svitashev, *et. al.* (Plant Physiol; 169(2):931–945 (2015)). Not all of the following steps are required for screening nor must the order of the steps be as presented.

This example illustrates the use of single guide D(R)NAs to guide a Cas endonucleases to cleave chromosomal DNA in maize embyos. Six single guide D(R)NAs (sgD(R)NAs) were designed targeting a region near the *liguleless 1* gene and the fertility gene *Ms45* (Table 16), and were delivered into a maize line containing a pre-integrated constitutively expressing *S. pyogenes* Cas9 gene. The maize *liguleless 1* and *Ms45* genomic loci were examined by deep sequencing for the presence of mutations induced by sgD(R)NAs/Cas9 mediated cleavage.

Table 16

Maize *Liguleless 1* and *Ms45* Targeting sgD(R)NA

Locus	Location	Sequence (RNA bases are bracketed, phosphorothioate bonds are shown with an *)	SEQ ID NO:
liguleless 1	Chr. 2: 28.45cM	5' - T*A*CGCGTACG CGTA[C][G][U][G][U][G] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U][G][C][U] - 3'	204

liguleless 1	Chr. 2: 28.45cM	5' - T*A*CGCGTACG CGTA[C][G][U][G][U][G] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][A][U][A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C][U][U] [G][A][A][A][A][A][G][U][G][G] [C][A][C][C][G][A][G][U][C][G] [G][U][G][C][U] - 3'	205
liguleless 1	Chr. 2: 28.45cM	5' - T*A*CGCGTACG CGTA[C][G][U][G][U][G] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][A][U][A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C][U][U] [G][A][A][A][A][A][G][U][G][G] CACCG[A][G][U]CG GTG[C][U] - 3'	206
Ms45	Chr. 9: 119.15 cM	5' - G*G*CCGAGGTC GACT[A][C][C][G][G][C] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U][G][C][U] - 3'	224
Ms45	Chr. 9: 119.15 cM	5' - G*G*CCGAGGTC GACT[A][C][C][G][G][C] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][A][U][A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C][U][U] [G][A][A][A][A][A][G][U][G][G] [C][A][C][C][G][A][G][U][C][G] [G][U][G][C][U] - 3'	225

Ms45	Chr. 9: 119.15 cM	5' - G*G*CCGAGGTC GACT[A][C][C][G][G][C] [G][U][U][U][U][A][G][A]GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][A][U][A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C][U][U] [G][A][A][A][A][A][G][U][G][G] CACCG[A][G][U]CG GTG[C][U] - 3'	226

A pre-integrated constitutively expressing *S. pyogenes* Cas9 maize line was generated as described in Svitashev *et al*. (2015).

sgD(R)NAs desgins were provided to a commercial manufacturer for synthesis (Eurofins Scientific, Huntsville, AL).

sgRNAs (SEQ ID NOS: 207 and 227) were constructed as described in Example 1.

Biolistic-mediated transformation of immature maize embryos (IMEs) derived from the constitutively expressing *S. pyogenes* Cas9 line with the sgD(R)NAs was carried-out as described in Svitashev *et. al.* (2015). Briefly, 100 ng of each sgD(R)NA was delivered to 60-90 IMEs in the presence of cell-division stimulating genes, *ZmODP2* (US Publ. No. 20050257289) and *ZmWUS*2 (US Pat. No. 7,256,322), as described in Ananiev *et. al.* (Chromosoma;118(2):157-77 (2009)). Since particle gun transformation can be highly variable, a visual selectable marker DNA expression cassette, MoPAT-DsRED, was also co-delivered with the cell-division promoting genes as described in Svitashev *et. al.* (2015). Embryos transformed with 100 ng of T7 transcribed single guide RNA (sgRNA) targeting the same region for cleavage (SEQ ID NOS: 207 and 227) served as a positive control and embryos transformed with only the *ZmODP2*, Zm*WUS2* and Mo-PAT-DsRED expression cassettes served as a negative control. After 3 days, the 20-30 most uniformly transformed embryos from each treatment were selected based on DsRED fluorescence, pooled and total genomic DNA was extracted. The region surrounding the intended target site was PCR amplified with Phusion® HighFidelity PCR Master Mix (M0531L, New England Biolabs, Ipswich, MA) adding on the sequences necessary for amplicon-specific barcodes and Illumnia sequencing using "tailed" primers through two rounds of PCR. The primers used in the primary PCR reaction are shown in Table 17 and the primers used in the secondary PCR reaction were SEQ ID NO: 214 and 215.

Table 17

PCR Primer Sequences

ID	Sample	Primers
BARCODING PRIMER set-37	SEQ ID NO. 204	SEQ ID NOs: 208, 209
BARCODING PRIMER set-38	SEQ ID NO. 205	SEQ ID NOs: 208, 210
BARCODING PRIMER set-39	SEQ ID NO. 206	SEQ ID NOs: 208, 211
BARCODING PRIMER set-40	SEQ ID NO. 207	SEQ ID NOs: 208, 212
BARCODING PRIMER set-41	No guide RNA (negative control)	SEQ ID NOs: 208, 213
BARCODING PRIMER set-42	SEQ ID NO. 224	SEQ ID NOs: 228, 229
BARCODING PRIMER set-43	SEQ ID NO. 225	SEQ ID NOs: 228, 230
BARCODING PRIMER set-44	SEQ ID NO. 226	SEQ ID NOs: 228, 231
BARCODING PRIMER set-45	SEQ ID NO. 227	SEQ ID NOs: 228, 232
BARCODING PRIMER set-46	No guide RNA (negative control)	SEQ ID NOs: 228, 233

The resulting PCR amplifications were purified with a Qiagen PCR purification spin column, concentration measured with a Hoechst dye-based fluorometric assay, combined in an equimolar ratio, and single read 100 nucleotide-length deep sequencing was performed on the Illumina MiSeq Personal Sequencer with a 25% (v/v) spike of PhiX control v3 (Illumina, FC-110-3001) to off-set sequence bias. Only those reads with a ≥1 nucleotide indel arising within the 10 nucleotide window centered over the expected site of cleavage and not found in a similar level in the negative control were classified as mutant. Mutant reads with the same mutation were counted and collapsed into a single read and visually confirmed as having a mutation arising within the expected site of cleavage. The total numbers of visually confirmed mutations were then used to calculate the percent mutant reads based on the total number of reads of an appropriate length containing a perfect match to the barcode and forward primer.

As shown in Table 18, mutations were recovered in all treatments indicating that sgD(R)NAs may be used to guide Cas endonucleases to cleave maize cellular chromosomal DNA. Furthermore, certain sgD(R)NA designs (SEQ ID NOS. 205 and 226) exhibited mutation frequencies near that of the T7 transcribed sgRNA (SEQ ID NOS. 207 and 227). Examples of the mutations recovered with the sgD(R)NAs are shown in **FIGURE 14A** (corresponding to SEQ

ID NOs: 217-223, wherein SEQ ID NO: 216 is the reference maize sequence comprising the *liguleless 1* target locus) and **FIGURE 14B** (corresponding to SEQ ID NOS: 235-254, wherein SEQ ID NO: 234 is the reference maize sequence comprising the *Ms45* target locus).

Table 18

Mutant Reads at maize *liguleless 1* and *Ms45* Target Loci Produced by sgD(R)NA/Cas Endonuclease System Compared to the sgRNA/Cas Endonuclease System

Treatment	Total Number of Reads	Number of Mutant Reads
Ligueless 1 No Guide RNA (Negative Control)	2,849,145	0
SEQ ID NO. 207	3,155,695	552
SEQ ID NO. 204	2,816,705	5
SEQ ID NO. 205	3,053,967	192
SEQ ID NO. 206	2,979,282	9
Ms45 No Guide RNA (Negative Control)	1,248,142	16
SEQ ID NO. XX4	1,194,050	8,784
SEQ ID NO. XX1	1,192,758	190
SEQ ID NO. XX2	1,206,632	114
SEQ ID NO. XX3	1,192,110	878

Although the foregoing disclosure provides description and examples of specific embodiments of the present invention, it is not intended to be limiting in any way and it is within the knowledge of one of skill in the art to modify the examples disclosed in order to adapt a particular method, composition or step to achieve the desired result within the scope of the present invention. All such modifications are intended to be within the scope of the present invention.

The Claims defining the invention are as follows:

1. A set of two Class 2 CRISPR polynucleotides comprising:

(i) a first Class 2 CRISPR polynucleotide wherein the first Class 2 CRISPR polynucleotide comprises a targeting region comprising deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) and an activating region adjacent to said targeting region; and,

(ii) a second Class 2 CRISPR polynucleotide wherein the second Class 2 CRISPR polynucleotide comprises an activating region wherein the activating region comprises DNA and a sequence that is complementary to a sequence in said activating region of the first Class 2 CRISPR polynucleotide, wherein said activating region of the first Class 2 CRISPR polynucleotide and said activating region of the second Class 2 CRISPR polynucleotide are capable of hybridizing to each other to form an activating duplex region, wherein said activating duplex region comprises a stem and a bulge, and wherein said activating duplex region is capable of binding with a Cas9.

2. The set of two Class 2 CRISPR polynucleotides of claim 1, wherein said targeting region of the first Class 2 CRISPR polynucleotide comprises a mixture of DNA and RNA.

3. The set of two Class 2 CRISPR polynucleotides of claim 1, wherein said activating duplex region comprises a compound selected from the group consisting of phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, alkylphosphonates, 5'-alkylene phosphonates, chiral phosphonates, phosphinates, phosphoramidates, 3'-amino phosphoramidate, amino alkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, and boranophosphates.

4. The set of two Class 2 CRISPR polynucleotides of claim 1, wherein said activating duplex region comprises a lower stem, a bulge, and an upper stem.

5. A Class 2 CRISPR system consisting essentially of:

(i) a set of two Class 2 CRISPR polynucleotides comprising (a) a first Class 2 CRISPR polynucleotide wherein the first Class 2 CRISPR polynucleotide comprises a targeting region comprising deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) and an activating region adjacent to said targeting region; and (b) a second Class 2 CRISPR polynucleotide wherein the second Class 2 CRISPR polynucleotide comprises an activating region wherein the activating region comprises DNA and a sequence that is complementary to a sequence in said activating region of the first Class 2 CRISPR polynucleotide, wherein said activating region of the first Class 2 CRISPR polynucleotide and said activating region of the second Class 2 CRISPR polynucleotide are capable of hybridizing to each other to form an activating duplex region, wherein said activating duplex region comprises a stem and a bulge, and wherein said activating duplex region is capable of binding with a Cas9; and

(ii) a Cas9.

6. The Class 2 CRISPR system of claim 5, wherein said activating duplex region comprises a lower stem, a bulge, and an upper stem.

7. The Class 2 CRISPR system of claim 5, wherein said targeting region of the first Class 2 CRISPR polynucleotide comprises a mixture of DNA and RNA.

8. The Class 2 CRISPR system of claim 5, wherein said activating duplex region comprises a mixture of DNA and RNA.

9. The Class 2 CRISPR system of claim 5, wherein said activating duplex region comprises a compound selected from the group consisting of phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, alkylphosphonates, 5'-alkylene phosphonates, chiral phosphonates, phosphinates, phosphoramidates, 3'-amino phosphoramidate, amino alkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, and boranophosphates.

10. The Class 2 CRISPR system of claim 5, further comprising a donor polynucleotide.

11. A method of modifying a target nucleic acid molecule in a non-human organism, an isolated cell, or *in vitro*, wherein said method comprises:

introducing into a cell

(i) a set of two Class 2 CRISPR polynucleotides comprising (a) a first Class 2 CRISPR polynucleotide wherein the first Class 2 CRISPR polynucleotide comprises a targeting region configured to hybridize to a target sequence within the target nucleic acid and an activating region adjacent to said targeting region; and (b) a second Class 2 CRISPR polynucleotide wherein the second Class 2 CRISPR polynucleotide comprises an activating region wherein the activating region comprises DNA and a sequence that is complementary to a sequence in said activating region of the first Class 2 CRISPR polynucleotide, wherein said activating region of the first Class 2 CRISPR polynucleotide and said activating region of the second Class 2 CRISPR polynucleotide are capable of hybridizing to each other to form an activating duplex region, and wherein said activating duplex region comprises a stem and a bulge; and

(ii) a Cas9, wherein the Cas9 binds with said activating duplex region of the two Class 2 CRISPR polynucleotides, wherein said targeting region of the first Class 2 CRISPR polynucleotide hybridizes to said target sequence, and wherein said target nucleic acid is cleaved.

12. The method of claim 11, wherein the Cas9 is encoded by an expression vector comprising a coding sequence for the Cas9.

13. The method of claim 11, wherein the set of two Class 2 CRISPR polynucleotides and the Cas9 form a nucleoprotein complex prior to introduction into the cell.

14. The method of claim 11, wherein the Cas9 comprises a nuclear localization signal (NLS).

15. The method of claim 11, wherein the set of two Class 2 CRISPR polynucleotides and the Cas9 are introduced into the cell by lipofection, electroporation, nucleofection, microinjection,

biolistics, liposomes, immunoliposomes, polycation, lipid:nucleic acid conjugates, or combinations thereof.

16. The method of claim 11, wherein said activating duplex region comprises a compound selected from the group consisting of phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, alkylphosphonates, 5'-alkylene phosphonates, chiral phosphonates, phosphinates, phosphoramidates, 3'-amino phosphoramidate, amino alkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, and boranophosphates.

17. The method of claim 11, wherein said activating duplex region comprises a lower stem, a bulge, and an upper stem.

18. The method of claim 11, wherein the cell is selected from the group consisting of a bacterial cell, an archaeal cell, a plant cell, an algal cell, a fungal cell, an invertebrate cell, a vertebrate cell, a mammalian cell, and a human cell.

19. The method of claim 11, wherein the targeting region comprises RNA.

20. The method of claim 11, further comprising introducing a donor polynucleotide into the cell.

21. A method of modulating transcription of at least one gene within a target nucleic acid molecule in a non-human organism, an isolated cell, or *in vitro*, wherein said method comprises:

introducing into a cell

(i) a set of two Class 2 CRISPR polynucleotides comprising (a) a first Class 2 CRISPR polynucleotide wherein the first Class 2 CRISPR polynucleotide comprises a targeting region configured to hybridize to a target sequence within the open reading frame of the at least one gene or a target sequence within the promoter sequence of the at least one gene and an activating region adjacent to said targeting region; and (b) a second Class 2 CRISPR polynucleotide wherein the second Class 2 CRISPR polynucleotide comprises an activating

region wherein the activating region comprises DNA and a sequence that is complementary to a sequence in said activating region of the first Class 2 CRISPR polynucleotide, wherein said activating region of the first Class 2 CRISPR polynucleotide and said activating region of the second Class 2 CRISPR polynucleotide are capable of hybridizing to each other to form an activating duplex region, and wherein said activating duplex region comprises a stem and a bulge; and

(ii) a Cas9, wherein the Cas9 has no nuclease activity, wherein the Cas9 binds with said activating duplex region of the two Class 2 CRISPR polynucleotides, wherein said targeting region of the first Class 2 CRISPR polynucleotide hybridizes to said target sequence, and wherein the transcription of at least one gene within the target nucleic acid molecule is modulated.

22. The method of claim 21, wherein the Cas9 is encoded by an expression vector comprising a coding sequence for the Cas9.

23. The method of claim 21, wherein the set of two Class 2 CRISPR polynucleotides and the Cas9 form a nucleoprotein complex prior to introduction into the cell.

24. The method of claim 21, wherein the Cas9 comprises a nuclear localization signal (NLS).

25. The method of claim 21, wherein the set of two Class 2 CRISPR polynucleotides and the Cas9 are introduced into the cell by lipofection, electroporation, nucleofection, microinjection, biolistics, liposomes, immunoliposomes, polycation, lipid:nucleic acid conjugates, or combinations thereof.

26. The method of claim 21, wherein said activating duplex region comprises a compound selected from the group consisting of phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, alkylphosphonates, 5'-alkylene phosphonates, chiral phosphonates, phosphinates, phosphoramidates, 3'-amino

phosphoramidate, amino alkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, and boranophosphates.

27. The method of claim 21, wherein said activating duplex region comprises a lower stem, a bulge, and an upper stem.

28. The method of claim 21, wherein the cell is selected from the group consisting of a bacterial cell, an archaeal cell, a plant cell, an algal cell, a fungal cell, an invertebrate cell, a vertebrate cell, a mammalian cell, and a human cell.

29. The method of claim 21, wherein said targeting region comprises RNA.