exploited to engineer RNA-programmable transcriptional silencing and activating devices or epigenetic modifiers.

Genome editing in mammalian cells has been limited, in part, by the size of Cas9 proteins. Cas9 from *Staphylococcus pyogenes* (SpyCas9), the enzyme most widely used to date, comprises approximately 4.2kb of DNA (W02013/176722) and a direct combination with cognate single guide RNAs (sgRNA) further increases the size. Adeno-associated viruses are among the vectors used for the delivery of Cas9 enzymes in gene therapy applications. However, AAV cargo size is restricted to about 4.5 kb. Because of the size constraints, delivering a Cas9 with its sgRNA and a potential DNA repair template can be an impediment to using the methods. Smaller Cas9 molecules have been characterized, but most of them suffer from a protospacer adjacent motif (PAM) sequence that is not as well defined as the one used by SpyCas9. For example, *Staphylococcus aureus* (SauCas9 uses an "NNGRR(T)" sequence, where R = A or G, and *Campylobacter jejuni* (Cja)Cas9 uses a "NNNACAC"/ "NNNRYAC" PAM (where Y = T or G), respectively. The PAM ambiguity increases the potential for undesirable activity of the enzyme at off-target sequences harbouring high or perfect sequence identity to the PAM. Specificity of these systems remains a concern, as targeting similar sites by accident ("off-targets") will increase the likelihood of adverse events.

Existing CRISPR-Cas 9 systems generally have one or more of the following disadvantages:

a) Their size is too large to be carried inside the genome of established therapeutically-suitable viral delivery systems like adeno associated viruses (AAVs).

b) Many of them are not substantially active in non-host environments, for example in eukaryotic cells, and in particular in mammalian cells.

c) Their nuclease can catalyze DNA strand cleavage when mismatches between spacer and protospacer sequences are present, leading to undesired off target effects that would for example make them unsuitable for gene therapeutic uses or other applications requiring high precision.

d) They may trigger an immune response that can limit their use for *in vivo* applications in mammals.

e) They require complex and/or long PAMs that restrict target selection for the DNA targeting segments.

f) They exhibit poor expression from plasmid or viral vectors.

Summary of the Disclosure

The present disclosure relates to improved Cas nucleases that at least can have increased specificity compared to certain wildtype enzymes, as well as having a relatively small size maintaining a small size.

In one example, provided herein is a M-SmallCas9 polypeptide selected from the group: M-

SauCas9_X according to SEQ ID NO. 1, M-SluCas9_X according to SEQ ID NO. 2, M-SpaCas9_X according to SEQ ID NO. 3, M-ShyCas9_X according to SEQ ID NO. 4, M-SmiCas9_X according to SEQ ID NO. 5, MGib11SpaCas9-3-E410A, according to SEQ ID NO. 8, MGib11Spa-1-M417L according to SEQ ID NO. 133, or any polypeptide sequence that is at least 95% identical to any of the foregoing polypeptides.

In one aspect, provided herein is a M-SmallCas9 polypeptide selected from the group: M-SauCas9-R420A, according to SEQ ID NO. 6; M-SluCas9-R414A, according to SEQ ID NO. 7; M-Gib11SpaCas9-3-E410A, according to SEQ ID NO. 8; and M-Gib11Spa-1-M417L according to SEQ ID NO. 133, or any polypeptide sequence that is at least 95% identical to any of the foregoing polypeptides.

In another aspect, provided herein is a composition comprising: (I) a M-SmallCas9 polypeptide disclosed herein; and (II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) that allow the generation of such one or more sgRNA(s) in situ, each sgRNA or DNA encoding an sgRNA comprising (a) an engineered DNA targeting segment that can hybridize to a target sequence in a polynucleotide locus, (b) a tracr mate sequence, and (c) a tracr RNA sequence, wherein the tracr mate sequence can hybridize to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation. In some examples, the engineered DNA targeting segment is directly adjacent to the PAM sequence on its 3' end or such PAM sequence is part of the DNA targeting sequence in its 3' portion.

In another aspect, provided herein is a method of targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell *ex vivo* or *in vitro*, the method comprising (I) introducing a heterologous M-SmallCas9 polypeptide disclosed herein or a nucleic acid encoding a M-SmallCas9 disclosed herein into the cell or into the in vitro environment; and (II) introducing one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) encoding such one or more sgRNA(s) in the cell *ex vivo* or the *in vitro* environment, each sgRNA or DNA encoding the sgRNA comprising: (a) an engineered DNA targeting segment comprising an RNA and capable of hybridizing to a target sequence in a polynucleotide locus, (b) a tracr mate sequence comprised of RNA, and (c) a tracr RNA sequence comprised of RNA, wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation; and (III) creating one or more nicks or cuts or base edits in the target DNA, wherein the M-SmallCas9 polypeptide is directed to the target DNA by the sgRNA in its processed or unprocessed form.

In one example, provided herein is a method of targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell or in vitro, the method comprising (I) introducing a heterologous M-SmallCas9 polypeptide disclosed herein or a nucleic acid encoding a M-SmallCas9 disclosed herein into the cell or into the in vitro environment; and (II) introducing one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) encoding such one or more sgRNA(s) in the cell or the in vitro environment, each sgRNA or DNA encoding the sgRNA comprising: (a) an engineered DNA targeting segment comprising an RNA and capable of hybridizing to a target sequence in a polynucleotide locus,

(b) a tracr mate sequence comprised of RNA, and (c) a tracr RNA sequence comprised of RNA, wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation; and (III) creating one or more nicks or cuts or base edits in the target DNA, wherein the M-SmallCas9 polypeptide is directed to the target DNA by the sgRNA in its processed or unprocessed form.

In a further aspect, provided herein is of a composition comprising

(I) a M-SmallCas9 polypeptide as described herein or a nucleic acid encoding the same; and/or

(II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA in situ, each comprising:

(a) an engineered DNA targeting segment comprised of RNA and capable of hybridizing to such target sequence in a polynucleotide locus,

(b) a tracr mate sequence comprised of RNA, and

(c) a tracr RNA sequence comprised of RNA,

wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation;

for targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell *ex vivo* or *in vitro*.

In one example, provided herein is use of a composition comprising (I) a M-SmallCas9 polypeptide disclosed herein or a nucleic acid encoding the same; and/or (II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA in situ, each comprising: (a) an engineered DNA targeting segment comprised of RNA and capable of hybridizing to such target sequence in a polynucleotide locus, (b) a tracr mate sequence comprised of RNA, and (c) a tracr RNA sequence comprised of RNA, wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation; for targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell or in vitro.

In another aspect, provided herein is a cell comprising (I) a M-SmallCas9 polypeptide as disclosed herein, or a nucleic acid encoding a M-SmallCas9 polypeptide disclosed herein; and (II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA in situ, each comprising: (a) an engineered DNA targeting segment that can hybridizing to a target sequence in a polynucleotide locus, (b) a tracr mate sequence, and (c) a tracr RNA sequence, wherein the tracr mate sequence that can hybridize to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation.

In yet another aspect, provided herein is a kit comprising (I) a nucleic acid sequence encoding a

M-SmallCas9 polypeptide as disclosed herein, wherein the nucleic acid sequence encoding the M-SmallCas9 is operably linked to a promoter; and (II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA in situ, each sgRNA comprising: (a) an engineered DNA targeting segment that can hybridize to a target sequence in a polynucleotide locus, (b) a tracr mate sequence, and (c) a tracr RNA sequence, wherein the tracr mate sequence can hybridize to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation.

The entire disclosure of each patent document and scientific article referred to herein, and those patent documents and scientific articles cited thereby, is expressly incorporated by reference herein for all purposes.

Additional features and advantages of the disclosure are more particularly described below.

Brief Description of the Drawings

Figures 1A and 1B are bar graphs depicting the results of experiments testing the relative cleavage activity (y axis) of Cas nucleases as a function a) position of nucleic acid within the target sequence as shown on the x-axis and b) the respective nucleic acid residue in that position, which is visualized by differently shaded boxes that are stacked. The absence of, *e.g.*, the box for "G" in a column means that no cleavage activity was seen in that position when the residue "G" was present. Figure 1A shows these data for SluCas9 according to SEQ ID NO. 9; Figure 1B shows these data for MSluCas9R414A according to SEQ ID NO. 7.

Figure 2 is a bar graph depicting the results of experiments detecting the total specificity values of four different Cas enzymes: SluCas9 (SEQ ID NO: 9), M-SluCas9R414A (SEQ ID NO. 7),

CLAIMS

1. A M-SmallCas9 polypeptide selected from the group M-SauCas9-R420A, according to SEQ ID NO. 6; M-SluCas9-R414A, according to SEQ ID NO. 7; M-Gib11SpaCas9-3-E410A, according to SEQ ID NO. 8; and M-Gib11Spa-1-M417L according to SEQ ID NO. 133 or any polypeptide sequence that is at least 95% identical to any of the above.

2. A composition comprising
    (I) a M-SmallCas9 polypeptide according to claim 1; and
    (II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) that allow the generation of such one or more sgRNA(s) *in situ*, each sgRNA or DNA encoding an sgRNA comprising:
        (a) an engineered DNA targeting segment that can hybridize to a target sequence in a polynucleotide locus,
        (b) a tracr mate sequence, and
        (c) a tracr RNA sequence,

    wherein the tracr mate sequence can hybridize to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation.

3. The composition according to claim 2, wherein the engineered DNA targeting segment is directly adjacent to the PAM sequence on its 3' end or such PAM sequence is part of the DNA targeting sequence in its 3' portion.

4. A method of targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell *ex vivo* or *in vitro*, the method comprising
    (I) introducing a heterologous M-SmallCas9 polypeptide according to claim 1 or a nucleic acid encoding a M-SmallCas9 of claim 1 into the cell *ex vivo* or into the *in vitro* environment; and
    (II) introducing one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) encoding such one or more sgRNA(s) in the cell *ex vivo* or the *in vitro* environment, each sgRNA or DNA encoding the sgRNA comprising:
        (a) an engineered DNA targeting segment comprising an RNA and capable of hybridizing to a target sequence in a polynucleotide locus,

(b) a tracr mate sequence comprised of RNA, and

(c) a tracr RNA sequence comprised of RNA,

wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (b) are arranged in a 5' to 3' orientation; and

(III) creating one or more nicks or cuts or base edits in the target DNA, wherein the M-SmallCas9 polypeptide is directed to the target DNA by the sgRNA in its processed or unprocessed form.

5. Use of a composition comprising

(I) a M-SmallCas9 polypeptide according to claim 1 or a nucleic acid encoding the same; and/or

(II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA in situ, each comprising:

(a) an engineered DNA targeting segment comprised of RNA and capable of hybridizing to such target sequence in a polynucleotide locus,

(b) a tracr mate sequence comprised of RNA, and

(c) a tracr RNA sequence comprised of RNA,

wherein the tracr mate sequence hybridizes to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5' to 3' orientation;

for targeting, editing, modifying, or manipulating a target DNA at one or more locations in a cell *ex vivo* or in vitro.

6. A cell comprising

(I) a M-SmallCas9 polypeptide according to claim 1, or a nucleic acid encoding a M-SmallCas9 polypeptide according to claim 1; and

(II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA *in situ*, each comprising:

(a) an engineered DNA targeting segment that can hybridizing to a target sequence in a polynucleotide locus,

(b) a tracr mate sequence, and

(c) a tracr RNA sequence,

wherein the tracr mate sequence that can hybridize to the tracr sequence, and wherein (a), (b), and

(c) are arranged in a 5’ to 3’ orientation.

7. A kit comprising

(I) a nucleic acid sequence encoding a M-SmallCas9 polypeptide according to claim 1, wherein the nucleic acid sequence encoding the M-SmallCas9 is operably linked to a promoter; and

(II) one or more single heterologous guide RNA(s) (sgRNA) or DNA(s) suitable for the generation of such one or more sgRNA *in situ*, each sgRNA comprising:

(a) an engineered DNA targeting segment that can hybridize to a target sequence in a polynucleotide locus,

(b) a tracr mate sequence, and

(c) a tracr RNA sequence,

wherein the tracr mate sequence can hybridize to the tracr sequence, and wherein (a), (b), and (c) are arranged in a 5’ to 3’ orientation.