

Claims

1. A nucleic acid molecule comprising a polynucleotide encoding an RNA-guided nuclease (RGN) polypeptide, wherein said polynucleotide comprises a nucleotide sequence encoding an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123;

wherein said RGN polypeptide is capable of binding a target DNA sequence in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence, and

wherein said polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to said polynucleotide.

2. The nucleic acid molecule of claim 1, wherein said RGN polypeptide comprises the amino acid sequence of any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123.

3. The nucleic acid molecule of claim 1 or 2, wherein said RGN polypeptide is nuclease inactive or is a nickase.

4. The nucleic acid molecule of any one of claims 1-3, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

5. The nucleic acid molecule of any one of claims 1-4, wherein the RGN polypeptide comprises one or more nuclear localization signals.

6. The nucleic acid molecule of any one of claims 1-5, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

7. The nucleic acid molecule of any one of claims 1-6, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

8. A vector comprising the nucleic acid molecule of any one of claims 1-7.

9. The vector of claim 8, further comprising at least one nucleotide sequence encoding said gRNA capable of hybridizing to said target DNA sequence.

10. The vector of claim 9, wherein the guide RNA is selected from the group consisting of:
- a) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 84; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 85;wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 83;
 - b) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 84; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 78;wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 83;
 - c) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 64; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 65;wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 63 and 570-579;
 - d) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 71; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 72;wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 70;
 - e) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 2; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 3;wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;
 - f) a guide RNA comprising:
 - i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 9; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 10;

- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;
- g) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 16; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 17;
- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;
- h) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 23; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 24;
- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 22;
- i) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 30; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 31;
- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 29;
- j) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 37; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 38;
- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 36;
- k) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 44; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45;
- wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 43;
- l) a guide RNA comprising:
- i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 51; and
 - ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 50;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 57; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 56;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 77; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 78;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 76;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 90; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 91;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 89;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 97; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 98;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 96;

q) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 104; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 105;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 103;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 112;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 110;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 119;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 117; and

t) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 124; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 123.

11. A cell comprising the nucleic acid molecule of any one of claims 1-7 or the vector of any one of claims 8-10.

12. An isolated RNA-guided nuclease (RGN) polypeptide, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123; and

wherein said RGN polypeptide is capable of binding a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence.

13. The isolated RGN polypeptide of claim 12, wherein said RGN polypeptide comprises the amino acid sequence set forth in any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123.

14. The isolated RGN polypeptide of claim 12 or 13, wherein said RGN polypeptide is nuclease inactive or a nickase.

15. The isolated RGN polypeptide of any one of claims 12-14, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

16. The isolated RGN polypeptide of any one of claims 12-15, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

17. The isolated RGN polypeptide of any one of claims 12-16, wherein the RGN polypeptide comprises one or more nuclear localization signals.

18. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

a) one or more guide RNAs capable of hybridizing to said target DNA sequence or one or more polynucleotides comprising one or more nucleotide sequences encoding the one or more guide RNAs (gRNAs); and

b) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123, or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide;

wherein at least one of said nucleotide sequences encoding the one or more guide RNAs and said nucleotide sequence encoding the RGN polypeptide is operably linked to a promoter heterologous to said nucleotide sequence;

wherein the one or more guide RNAs are capable of hybridizing to the target DNA sequence, and

wherein the one or more guide RNAs are capable of forming a complex with the RGN polypeptide in order to direct said RGN polypeptide to bind to said target DNA sequence of the DNA molecule.

19. The system of claim 18, wherein at least one of said nucleotides sequences encoding the one or more guide RNAs is operably linked to a promoter heterologous to said nucleotide sequence.

20. The system of claim 18 or 19, wherein said RGN polypeptide comprises the amino acid sequence set forth in any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123.

21. The system of any one of claims 18-20, wherein said target DNA sequence is a eukaryotic target DNA sequence.

22. The system of any one of claims 18-21, wherein said gRNA is selected from the group consisting of:

- a) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 84 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 85, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 83;
- b) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 84 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 78, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 83;
- c) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 64 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 65, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 63 and 570-579;
- d) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 71 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 72, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 70;
- e) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 2 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 3, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;
- f) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 9 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 10, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;
- g) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 16 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 17, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;
- h) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 24, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 22;
- i) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 31, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 29;

j) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 38, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 36;

k) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 44 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 43;

l) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 51 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 50;

m) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 57 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 56;

n) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 77 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 78, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 76;

o) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 90 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 91, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 89;

p) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 97 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 98, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 96;

q) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 104 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 105, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 103;

r) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 112, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 110;

s) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 119, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 117; and

t) a gRNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 124 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 123.

23. The system of any one of claims 18-22, wherein said RGN polypeptide is nuclease inactive or is a nickase.

24. The system of any one of claims 18-23, wherein the RGN polypeptide is operably linked to a base-editing polypeptide.

25. The system of any one of claims 18-24, wherein nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding an RGN polypeptide are located on one vector.

26. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

a) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence set forth as any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123, or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide; and

b) one or more guide RNAs comprising (i) a spacer sequence that is complementary to said target DNA sequence and (ii) a CRISPR repeat sequence recognized by the RGN; or one or more polynucleotides comprising one or more nucleotide sequences encoding the one or more guide RNAs.

27. A complex comprising (i) an isolated RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence set forth as any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123 bound to (ii) a non-natural guide RNA (gRNA).

28. The complex of claim 27, wherein the non-natural gRNA is selected from: (a) a gRNA comprising one or more synthetic ribonucleotide analogues; (b) a single guide RNA; (c) a

gRNA comprising one or more PNA, PNA-DNA chimera, a locked nucleic acid (LNA), and/or a phosphothiorate linkage.

29. A method for cleaving and/or modifying a target DNA sequence of a DNA molecule, comprising contacting the DNA molecule with:

a) an RNA-guided nuclease (RGN) polypeptide, wherein said RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 83, 63, 570-579, 70, 1, 8, 15, 22, 29, 36, 43, 50, 56, 76, 89, 96, 103, 110, 117, and 123; and

b) one or more guide RNAs capable of targeting the RGN of (a) to the target DNA sequence;

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said RGN polypeptide to bind to said target DNA sequence and cleavage and/or modification of said target DNA sequence occurs.

30. The method of claim 29, wherein cleavage by said RGN polypeptide generates a double-stranded break or a single-stranded break.

31. The method of claim 29 or 30, wherein said modified target DNA sequence comprises:

insertion of heterologous DNA into the target DNA sequence;

deletion of at least one nucleotide from the target DNA sequence; or

mutation of at least one nucleotide in the target DNA sequence.