(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 February 2024 (15.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/033901 A1

(51) International Patent Classification:
C12N 9/22 (2006.01) *C07K 14/32* (2006.01)
C12N 15/10 (2006.01) *C12N 15/11* (2006.01)
C12N 15/113 (2010.01)

(21) International Application Number:
PCT/IB2023/058160

(22) International Filing Date:
12 August 2023 (12.08.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/371,230 12 August 2022 (12.08.2022) US

(71) Applicant: **LIFEEDIT THERAPEUTICS, INC.** [US/US]; 300 Morris Street, Suite 300, Durham, North Carolina 27701 (US).

(72) Inventors: **CRAWLEY, Alexandra Briner**; 1220 Scott Pl., Cary, North Carolina 27511 (US). **COYLE, Michael**; 407 N. Elliot Road, Chapel Hill, North Carolina 27514 (US). **ARYA, Gunjan H.**; 8544 Rosiere Drive, Cary, North Carolina 27518 (US). **MOSE, Lisle**; 6509 Wynbrook Way, Raleigh, North Carolina 27612 (US).

(74) Agent: **BUCK, B. Logan** et al.; Womble Bond Dickinson (US) LLP, P.O. Box 7037, Atlanta, Georgia 30357-0037 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

(54) Title: RNA-GUIDED NUCLEASES AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

(57) Abstract: Compositions and methods for binding to a target sequence of interest are provided. The compositions find use in cleaving or modifying a target sequence of interest, visualization of a target sequence of interest, and modifying the expression of a sequence of interest. Compositions comprise RNA-guided nuclease (RGN) polypeptides, CRISPR RNAs, trans-activating CRISPR RNAs, guide RNAs, and nucleic acid molecules encoding the same. Vectors and host cells comprising the nucleic acid molecules are also provided. Further provided are RGN systems for binding a target sequence of interest, wherein the RGN system comprises an RNA-guided nuclease polypeptide and one or more guide RNAs.

WO 2024/033901 A1

RNA-GUIDED NUCLEASES AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority to US Provisional Patent Application Nos. 63/371,230, filed August 12, 2022, which is fully incorporated by reference herein.

REFERENCE TO A SEQUENCE LISTING SUBMITTED ELECTRONICALLY AS AN XML FILE

The instant application contains a Sequence Listing which has been submitted in xml format via USPTO Patent Center and is hereby incorporated by reference in its entirety. Said xml copy, created on August 11, 2023, is named L103438_1310WO_Seq_List.xml, and is 1.49 MB in size.

FIELD OF THE INVENTION

The present invention relates to the field of molecular biology and gene editing.

BACKGROUND OF THE INVENTION

Targeted genome editing or modification is rapidly becoming an important tool for basic and applied research. Initial methods involved engineering nucleases such as meganucleases, zinc finger fusion proteins or TALENs, requiring the generation of chimeric nucleases with engineered, programmable, sequence-specific DNA-binding domains specific for each particular target sequence. RNA-guided nucleases, such as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) proteins of the CRISPR-Cas bacterial system, allow for the targeting of specific sequences by complexing the nucleases with guide RNA that specifically hybridizes with a particular target sequence. Producing target-specific guide RNAs is less costly and more efficient than generating chimeric nucleases for each target sequence. Such RNA-guided nucleases can be used to edit genomes optionally through the introduction of a sequence-specific, double-stranded break that is repaired via error-prone non-homologous end-joining (NHEJ) to introduce a mutation at a specific genomic location. Alternatively, heterologous DNA may be introduced into the genomic site via homology-directed repair. RNA-guided nucleases (RGNs) can also be used for base editing when fused with a deaminase.

BRIEF SUMMARY OF THE INVENTION

Compositions and methods for binding a target sequence of interest in a target nucleic acid molecule are provided. The compositions find use in cleaving or modifying a target nucleic acid molecule of interest, detection of a target sequence of interest, and modifying the expression of a gene of interest comprising a target sequence. Compositions comprise RNA-guided nuclease (RGN) polypeptides, CRISPR RNAs (crRNAs), trans-activating CRISPR RNAs (tracrRNAs), guide RNAs (gRNAs) such as single guide RNAs

(sgRNAs), nucleic acid molecules encoding the same, compositions comprising the same, and vectors and host cells comprising the nucleic acid molecules. Also provided are RGN systems and ribonucleoprotein complexes for binding a target sequence of interest, wherein the RGN system and ribonucleoprotein complex comprises an RNA-guided nuclease polypeptide and one or more guide RNAs. Thus, methods disclosed herein are drawn to binding a target sequence of interest in a target nucleic acid molecule, and in some embodiments, cleaving or modifying the target nucleic acid molecule of interest. The target nucleic acid molecule of interest can be modified, for example, as a result of non-homologous end joining, homology-directed repair with an introduced donor sequence, or base editing.

In one aspect, the present disclosure provides a nucleic acid molecule comprising a polynucleotide encoding an RNA-guided nuclease (RGN) polypeptide, wherein the 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:1-20.

In some embodiments of the above aspect, the RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of the target sequence.

In some embodiments of the above aspect, the polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to the polynucleotide.

In some embodiments of the above aspect, the RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs:1-20. In some embodiments, the RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the RGN polypeptide is capable of cleaving the target nucleic acid molecule upon binding. In some embodiments, the RGN polypeptide is capable of generating a double-stranded break. In some embodiments, the RGN polypeptide is capable of generating a single-stranded break.

In some embodiments of the above aspect, the RGN polypeptide is nuclease inactive or is a nickase.

In some embodiments of the above aspect, the RGN polypeptide is operably fused to a base-editing polypeptide. In some embodiments, the base-editing polypeptide is a deaminase, such as a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90% or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the RGN polypeptide comprises one or more nuclear localization signals.

In some embodiments of the above aspect, the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

In some embodiments of the above aspect, the target sequence is located adjacent to a protospacer adjacent motif (PAM).

In another aspect, the present disclosure provides a vector comprising a nucleic acid molecule as described hereinabove.

In some embodiments of the above aspect, the vector further comprises at least one nucleotide sequence encoding the gRNA capable of hybridizing to the non- target strand of the target sequence.

In some embodiments of the above aspect, 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: 21; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1; b) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2; c) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3; d) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4; e) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5; a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6; g) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7; h) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or

100% sequence identity to SEQ ID NO: 28; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8; i) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9; j) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 10; k) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11; l) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12; m) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13; n) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14; o) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15; p) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16; q) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37; and ii) a tracrRNA having at least 90%, at least 95%, or 100%

sequence identity to SEQ ID NO: 58; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17; r) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18; s) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19; and t) a guide RNA comprising: i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41; and ii) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62; wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20.

In some embodiments of the above aspect, the gRNA is a single guide RNA. In some embodiments of the above aspect, the gRNA is a dual-guide RNA.

In another aspect, the present disclosure provides a cell comprising a nucleic acid molecule or vector described hereinabove. In some embodiments, the cell is a prokaryotic cell. In some embodiments, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the mammalian cell is a human cell. In some embodiments, the human cell is an immune cell. In some embodiments, the immune cell is a stem cell. In some embodiments, the stem cell is an induced pluripotent stem cell. In some embodiments, the eukaryotic cell is an insect or avian cell. In some embodiments, the eukaryotic cell is a fungal cell. In some embodiments, the eukaryotic cell is a plant cell.

In another aspect, the present disclosure provides a plant or a seed comprising a plant cell described hereinabove.

In another aspect, the present disclosure provides a method for making an RGN polypeptide comprising culturing a cell described hereinabove under conditions in which the RGN polypeptide is expressed.

In another aspect, the present disclosure provides a method for making an RGN polypeptide comprising introducing into a cell a heterologous nucleic acid molecule comprising a nucleotide sequence encoding an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20; and culturing the cell under conditions in which the RGN polypeptide is expressed.

In some embodiments of the above aspect, the RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non- target strand of the target sequence.

In some embodiments of the above aspect, the RGN polypeptide comprises an amino acid sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the method further comprises purifying the RGN polypeptide.

In some embodiments of the above aspect, the cell further expresses one or more guide RNAs capable of binding to the RGN polypeptide to form an RGN ribonucleoprotein complex. In some embodiments, the method further comprises purifying the RGN ribonucleoprotein complex.

In another aspect, the present disclosure provides an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non- target strand of the target sequence.

In some embodiments of the above aspect, the RGN polypeptide comprises an amino acid sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the RGN polypeptide is an isolated RGN polypeptide.

In some embodiments of the above aspect, the RGN polypeptide is capable of cleaving the target nucleic acid molecule upon binding. In some embodiments, cleavage by the RGN polypeptide generates a double-stranded break. In some embodiments, cleavage by the RGN polypeptide generates a single-stranded break.

In some embodiments of the above aspect, the RGN polypeptide is nuclease inactive or a nickase.

In some embodiments of the above aspect, the RGN polypeptide is operably fused to a base-editing polypeptide. In some embodiments, the base-editing polypeptide is a deaminase. In some embodiments, the deaminase is a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90%, at least 95%, or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the target sequence is located adjacent to a protospacer adjacent motif (PAM).

In some embodiments of the above aspect, the RGN polypeptide comprises one or more nuclear localization signals.

In another aspect, the present disclosure provides a ribonucleoprotein (RNP) complex comprising an RGN polypeptide described hereinabove and a guide RNA bound to the RGN polypeptide.

In another aspect, the present disclosure provides a nucleic acid molecule comprising a CRISPR RNA (crRNA) or a polynucleotide encoding a crRNA, wherein the crRNA comprises a spacer and a CRISPR repeat, wherein the CRISPR repeat comprises a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

In some embodiments of the above aspect, a guide RNA comprising the crRNA and a trans-activating CRISPR RNA (tracrRNA) hybridized to the CRISPR repeat of the crRNA is capable of hybridizing to the non-target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of the crRNA when the guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

In some embodiments of the above aspect, the polynucleotide encoding a crRNA is operably linked to a promoter heterologous to the polynucleotide.

In some embodiments of the above aspect, the CRISPR repeat comprises a nucleotide sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

In another aspect, the present disclosure provides a vector comprising a nucleic acid molecule comprising a polynucleotide encoding a crRNA described hereinabove.

In some embodiments of the above aspect, the vector further comprises a polynucleotide encoding the tracrRNA. In some embodiments, the tracrRNA is selected from the group consisting of: a) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21; b) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22; c) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23; d) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24; e) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; f) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26; g) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; h) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28; i) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29; j) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30; k) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52, wherein the

CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31; l) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32; m) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33; n) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34; o) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35; p) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36; q) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37; r) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39; s) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40; and t) a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41.

In some embodiments of the above aspect, the polynucleotide encoding the crRNA and the polynucleotide encoding the tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

In some embodiments of the above aspect, the polynucleotide encoding the crRNA and the polynucleotide encoding the tracrRNA are operably linked to separate promoters.

In some embodiments of the above aspect, the vector further comprises a polynucleotide encoding the RGN polypeptide. In some embodimetns, the RGN polypeptide is selected from the group consisting of: a) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42; b) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43; c) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040; d) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence

identity to SEQ ID NO: 24 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45; e) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042; f) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO:26 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043; g) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045; h) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49; i) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50; j) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 10, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51; k) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52; l) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53; m) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54; n) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55; o) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56; p) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16, wherein the

CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57; q) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58; r) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60; s) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61; and t) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20, wherein the CRISPR repeat has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62.

In another aspect, the present disclosure provides a nucleic acid molecule comprising a trans-activating CRISPR RNA (tracrRNA) or a polynucleotide encoding a tracrRNA comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

In some embodiments of the above aspect, a guide RNA comprising the tracrRNA and a crRNA comprising a spacer and a CRISPR repeat, wherein the tracrRNA hybridizes with the CRISPR repeat of the crRNA, is capable of hybridizing to the non- target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of the crRNA when the guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

In some embodiments of the above aspect, the polynucleotide encoding a tracrRNA is operably linked to a promoter heterologous to the polynucleotide.

In some embodiments of the above aspect, the tracrRNA comprises a nucleotide sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

In another aspect, the present disclosure provides a vector comprising the nucleic acid molecule comprising the polynucleotide encoding a tracrRNA described hereinabove.

In some embodiments of the above aspect, the vector further comprises a polynucleotide encoding the crRNA. In some embodiments, the crRNA comprises a CRISPR repeat selected from the group consisting of: a) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42; b) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22,

wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43; c) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040; d) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45; e) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042; f) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043; g) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045; h) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49; i) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50; j) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51; k) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52; l) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53; m) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54; n) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55; o) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56; p) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57; q) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58; r) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60; s) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40, wherein the tracrRNA has at least 90%, at least 95%, or 100%

sequence identity to SEQ ID NO: 61; and t) a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41, wherein the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62.

In some embodiments of the above aspect, the polynucleotide encoding the crRNA and the polynucleotide encoding the tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

In some embodiments of the above aspect, the polynucleotide encoding the crRNA and the polynucleotide encoding the tracrRNA are operably linked to separate promoters.

In some embodiments of the above aspect, the vector further comprises a polynucleotide encoding the RGN polypeptide. In some embodiments, the RGN polypeptide is selected from the group consisting of: a) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42; b) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43; c) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040; d) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45; e) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042; f) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043; g) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045; h) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID

NO: 28 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49; i) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50; j) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 10, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51; k) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52; l) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53; m) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54; n) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55; o) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56; p) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57; q) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58; r) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38, or 39 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60; s) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19, wherein the crRNA comprises a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61; and t) a RGN polypeptide having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20, wherein the crRNA comprises a CRISPR repeat having

at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and the tracrRNA has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62.

In another aspect, the present disclosure provides a cell comprising a nucleic acid molecule, a vector, a single guide RNA, or a double guide RNA described hereinabove.

In some embodiments of the above aspect, the cell is a prokaryotic cell.

In some embodiments of the above aspect, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the mammalian cell is a human cell. In some embodiments, the human cell is an immune cell. In some embodiments, the immune cell is a stem cell. In some embodiments, stem cell is an induced pluripotent stem cell. In some embodiments, the eukaryotic cell is an insect or avian cell. In some embodiments, the eukaryotic cell is a fungal cell. In some embodimetns, the eukaryotic cell is a plant cell.

In another aspect, the present disclosure provides a plant or a seed comprising a plant cell described hereinabove.

In another aspect, the present disclosure provides a system for binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, the system comprising: a) one or more guide RNAs capable of hybridizing to the non-target strand of the target 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: 1-20 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide; wherein the one or more guide RNAs are capable of forming a complex with the RGN polypeptide in order to direct the RGN polypeptide to bind to the target sequence.

In some embodiments of the above aspect, at least one of the nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding the RGN polypeptide is operably linked to a promoter heterologous to the nucleotide sequence.

In another aspect, the present disclosure provides a system for binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, the system comprising: a) one or more guide RNAs capable of hybridizing to the non- target strand of the target 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: 1-20; wherein the one or more guide RNAs are capable of forming a complex with the RGN polypeptide in order to direct the RGN polypeptide to bind to the target sequence.

In some embodiments of the above aspect, at least one of the nucleotides sequences encoding the one or more guide RNAs is operably linked to a promoter heterologous to the nucleotide sequence.

In some embodiments of the above aspect, the RGN polypeptide comprises an amino acid sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the RGN polypeptide and the one or more guide RNAs are not found complexed to one another in nature.

In some embodiments of the above aspect, the target sequence is a eukaryotic target sequence.

In some embodiments of the above aspect, the gRNA is a single guide RNA (sgRNA).

In some embodiments of the above aspect, the gRNA is a dual-guide RNA.

In some embodiments of the above aspect, the gRNA is selected from the group consisting of: a) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1; b) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2; c) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3; d) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4; e) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5; f) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6; g) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7; h) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8; i) a gRNA

comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9; j) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NOs:10; k) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11; l) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12; m) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13; n) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14; o) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15; p) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16; q) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17; r) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18; s) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61,

wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19; and t) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62, wherein the RGN polypeptide comprises an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20.

In some embodiments of the above aspect, the target sequence is located adjacent to a protospacer adjacent motif (PAM). In some embodiments, the target sequence is within a cell.

In some embodiments of the above aspect, wherein the one or more guide RNAs is capable of hybridizing to the non- target strand of the target sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target nucleic acid molecule.

In some embodiments of the above aspect, the cleavage generates a double-stranded break.

In some embodiments of the above aspect, the cleavage generates a single-stranded break.

In some embodiments of the above aspect, the RGN polypeptide is nuclease inactive or is a nickase.

In some embodiments of the above aspect, the RGN polypeptide is operably linked to a base-editing polypeptide. In some embodiments, the base-editing polypeptide is a deaminase. In some embodiments, the deaminase is a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90%, at least 95%, or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the RGN polypeptide comprises one or more nuclear localization signals.

In some embodiments of the above aspect, the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

In some embodiments of the above aspect, the nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding the RGN polypeptide are located on one vector.

In some embodiments of the above aspect, the system further comprises one or more donor polynucleotides.

In another aspect, the present disclosure provides a cell comprising a system described hereinabove.

In some embodiments of the above aspect, the cell is a prokaryotic cell.

In some embodiments of the above aspect, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the mammalian cell is a human cell. In some embodiments, the human cell is an immune cell. In some embodiments, the immune cell is a stem cell. In some embodiments, the stem cell is an induced pluripotent stem cell. In some embodiments, the eukaryotic cell is an insect or avian cell. In some embodiments, the eukaryotic cell is a fungal cell. In some embodiments, the eukaryotic cell is a plant cell.

In another aspect, the present invention provides a plant or a seed comprising a plant cell described hereinabove.

In another aspect, the present invention provides a pharmaceutical composition comprising a nucleic acid molecule, a vector, a cell, a RGN polypeptide, a RNP complex, or a system described hereinabove, and a pharmaceutically acceptable carrier.

In some embodiments of the above aspect, the pharmaceutically acceptable carrier is heterologous to the nucleic acid molecule, the vector, the cell, the RGN polypeptide, or the system.

In some embodiments of the above aspect, the pharmaceutically acceptable carrier is not naturally-occurring.

In some embodiments of the above aspect, the pharmaceutical composition is lipid-based. In some embodiments, the lipid-based pharmaceutical composition comprises liposomes or lipid nanoparticles (LNPs). In some embodiments, nucleic acid molecule, the vector, the cell, the RGN polypeptide, the RGN complex, or the system is encapsulated in, and/or non-covalently or covalently attached to the liposomes or the LNPs.

In another aspect, the present disclosure provides a method for binding a target sequence in a target nucleic acid molecule comprising delivering a system described hereinabove to the target sequence or a cell comprising the target sequence.

In some embodiments of the above aspect, the RGN polypeptide or the guide RNA further comprises a detectable label, thereby allowing for detection of the target sequence.

In some embodiments of the above aspect, the guide RNA or the RGN polypeptide further comprises an expression modulator, thereby modulating expression of a target gene comprising the target sequence.

In another aspect, the present disclosure provides a method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence comprising delivering a system described hereinabove to the target sequence or a cell comprising the target sequence, wherein cleavage or modification of the target nucleic acid molecule occurs.

In some embodiments of the above aspect, the modified target nucleic acid molecule comprises insertion of heterologous DNA into the target DNA sequence.

In some embodiments of the above aspect, the modified target nucleic acid molecule comprises deletion or mutation of at least one nucleotide from the target nucleic acid molecule.

In another aspect, the present disclosure provides a method for binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, wherein the method comprises: a) assembling an RNA-guided nuclease (RGN) ribonucleotide complex by combining: i) one or more guide RNAs capable of hybridizing to the non- target strand of the target sequence; and ii) an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs:1-20; under conditions suitable for formation of the RGN ribonucleotide complex; and b) contacting the target nucleic acid molecule or a cell comprising the target nucleic acid molecule with the assembled RGN ribonucleotide complex; wherein the one or more guide RNAs hybridize

to the non-target strand of the target sequence, thereby directing the RGN polypeptide to bind to the target sequence.

In some embodiments of the above aspect, the method is performed *in vitro*, *in vivo*, or *ex vivo*.

In some embodiments of the above aspect, the RGN polypeptide or the guide RNA further comprises a detectable label, thereby allowing for detection of the target sequence.

In some embodiments of the above aspect, the guide RNA or the RGN polypeptide further comprises an expression modulator, thereby allowing for the modulation of expression of a target gene comprising the target sequence.

In some embodiments of the above aspect, the RGN polypeptide further comprises a base-editing polypeptide, thereby allowing for the modification of the target nucleic acid molecule. In some embodiments, the base-editing polypeptide comprises a deaminase. In some embodiments, the deaminase is a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90%, at least 95%, or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the RGN polypeptide is capable of cleaving the target nucleic acid molecule, thereby allowing for the cleaving and/or modifying of the target nucleic acid molecule.

In another aspect, the present disclosure provides a method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence, wherein the target sequence comprises a target strand and a non-target strand, wherein the method comprises contacting the target nucleic acid molecule with: a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20; and b) one or more guide RNAs capable of targeting the RGN of (a) to the target sequence; wherein the one or more guide RNAs hybridize to the non-target strand of the target sequence, thereby directing the RGN polypeptide to bind to the target nucleic acid molecule and cleavage and/or modification of the target nucleic acid molecule occurs.

In some embodiments of the above aspect, cleavage by the RGN polypeptide generates a double-stranded break.

In some embodiments of the above aspect, cleavage by the RGN polypeptide generates a single-stranded break.

In some embodiments of the above aspect, the RGN polypeptide is nuclease inactive or a nickase and is operably fused to a base-editing polypeptide. In some embodiments, the base-editing polypeptide is a deaminase. In some embodiments, the deaminase is a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90%, at least 95%, or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the modified target nucleic acid molecule comprises insertion of heterologous DNA into the target nucleic acid molecule.

In some embodiments of the above aspect, the modified target nucleic acid molecule comprises deletion or mutation of at least one nucleotide from the target nucleic acid molecule.

In some embodiments of the above aspect, the target sequence is located adjacent to a protospacer adjacent motif (PAM).

In some embodiments of the above aspect, the target sequence is a eukaryotic target sequence.

In some embodiments of the above aspect, the gRNA is a single guide RNA (sgRNA).

In some embodiments of the above aspect, the gRNA is a dual-guide RNA.

In some embodiments of the above aspect, the RGN comprises an amino acid sequence having at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, a) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42; b) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43; c) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040; d) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45; e) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042; f) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043; g) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045; h) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49; i) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50; j) the RGN has

at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 10, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51; k) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52; l) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53; m) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54; n) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55; o) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56; p) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57; q) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58; r) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60; s) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61; and t) the RGN has at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20, the guide RNA comprises a crRNA repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62.

In some embodiments of the above aspect, the target sequence is within a cell. In some embodiments, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the mammalian cell is a human cell. In some embodiments, the human cell is an immune cell. In some embodiments, the immune cell is a stem cell. In some embodiments, the stem cell is

an induced pluripotent stem cell. In some embodiments, the eukaryotic cell is an insect or avian cell. In some embodiments, the cell is a prokaryotic cell. In some embodiments, the eukaryotic cell is a fungal cell. In some embodiments, the eukaryotic cell is a plant cell.

In some embodiments of the above aspect, the method further comprises culturing the cell under conditions in which the RGN polypeptide is expressed and cleaves and modifies the target nucleic acid molecule to produce a modified target nucleic acid molecule; and selecting a cell comprising the modified target nucleic acid molecule.

In another aspect, the present disclosure provides a cell comprising a modified target nucleic acid molecule described hereinabove.

In some embodiments of the above aspect, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the mammalian cell is a human cell. In some embodiments, the human cell is an immune cell. In some embodiments, the immune cell is a stem cell. In some embodiments, the stem cell is an induced pluripotent stem cell. In some embodiments, the eukaryotic cell is an insect or avian cell. In some embodiments, the cell is a prokaryotic cell. In some embodiments, the eukaryotic cell is a fungal cell. In some embodiments, the eukaryotic cell is a plant cell.

In another aspect, the present disclosure provides a plant or a seed comprising a plant cell described hereinabove.

In another aspect, the present disclosure provides a pharmaceutical composition comprising a cell described hereinabove and a pharmaceutically acceptable carrier.

In another aspect, the present disclosure provides a method for producing a genetically modified cell with a correction in a causal mutation for a genetically inherited disease, the method comprising introducing into the cell: a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding the RGN polypeptide, wherein the polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and b) a guide RNA (gRNA) or a polynucleotide encoding the gRNA, wherein the polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell; whereby the RGN and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

In some embodiments of the above aspect, the RGN is nuclease inactive or a nickase and is fused to a polypeptide which has base-editing activity. In some embodiments, the base-editing polypeptide is a deaminase. In some embodiments, the polypeptide with base-editing activity is a cytosine deaminase or an adenine deaminase. In some embodiments, the deaminase has at least 90%, at least 9%, or 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

In some embodiments of the above aspect, the genetically inherited disease is caused by a single nucleotide polymorphism.

In some embodiments of the above aspect, the genetically inherited disease is Hurler Syndrome.

In some embodiments of the above aspect, the gRNA further comprises a spacer that targets a region proximal to the causal single nucleotide polymorphism.

In another aspect, the present disclosure provides a method for producing a genetically modified cell with a deletion in a disease-causing expanded trinucleotide repeat, the method comprising introducing into the cell: a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding the RGN polypeptide, wherein the polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and b) a first guide RNA (gRNA) or a polynucleotide encoding the gRNA, wherein the polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein the gRNA comprises a spacer that targets the 5' flank of the expanded trinucleotide repeat; and c) a second guide RNA (gRNA) or a polynucleotide encoding the gRNA, wherein the polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein the second gRNA comprises a spacer that targets the 3'flank of the expanded trinucleotide repeat; whereby the RGN and the two gRNAs target to the expanded trinucleotide repeat and at least a portion of the expanded trinucleotide repeat is removed.

In some embodiments of the above aspect, the genetically inherited disease is Friedrich's Ataxia or Huntington's Disease.

In some embodiments of the above aspect, the first gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat. In some embodiments, the second gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat.

In some embodiments of the above aspect, the RGN polypeptide has at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the gRNA, the first gRNA, the second gRNA, or the first gRNA and the second gRNA is selected from a gRNA selected from the group consisting of: a) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1; b) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2; c) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3; d) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at

least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4; e) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5; f) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6; g) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7; h) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8; i) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9; j) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 10; k) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11; l) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12; m) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13; n) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90%, at least 95%, or

100% sequence identity to SEQ ID NO: 55, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14; o) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15; p) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16; q) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17; r) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58 or 60, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18; s) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19; and t) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20.

In some embodiments of the above aspect, the cell is an animal cell. In some embodiments, the animal cell is a mammalian cell. In some embodiments, the cell is derived from a dog, cat, mouse, rat, rabbit, horse, cow, pig, or human.

In another aspect, the present disclosure provides a method for producing a genetically modified mammalian hematopoietic progenitor cell having decreased BCL11A mRNA and protein expression, the method comprising introducing into an isolated human hematopoietic progenitor cell: a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding the RGN polypeptide, wherein the polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and b) a guide RNA (gRNA) or a polynucleotide encoding the gRNA, wherein the polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, whereby the RGN and gRNA are expressed in the cell and cleave at the BCL11A enhancer region, resulting in genetic modification of the human hematopoietic progenitor cell and reducing the mRNA and/or protein expression of BCL11A.

In some embodiments of the above aspect, the RGN polypeptide has at least 95% or 100% sequence identity to any one of SEQ ID NOs: 1-20.

In some embodiments of the above aspect, the gRNA is selected from the group consisting of: a) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 42, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 1; b) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 43, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 2; c) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 3; d) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 45, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 4; e) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 5; f) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 6; g) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 7; h) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 49, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 8; i) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 50, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 9; j) a gRNA comprising a

CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 51, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO:10; k) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 52, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 11; l) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 53, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 12; m) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 54, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 13; n) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 55, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 14; o) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 56, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 15; p) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 57, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 16; q) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 58, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 17; r) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 38 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 59 or 60, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 18; s) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 39 or 40 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 61, wherein the RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 19; and t) a gRNA comprising a CRISPR repeat having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 62, wherein the

RGN polypeptide has an amino acid sequence having at least 90%, at least 95%, or 100% sequence identity to SEQ ID NO: 20.

In some embodiments of the above aspect, the gRNA further comprises a spacer that targets a region within or proximal to the BCL11A enhancer region.

In another aspect, the present disclosure provides a method of treating a disease, disorder, or condition, the method comprising administering to a subject in need thereof a pharmaceutical composition described hereinabove.

In some embodiments of the above aspect, the disease, disorder, or condition is associated with a causal mutation and the pharmaceutical composition corrects the causal mutation.

In some embodiments of the above aspect, the subject is at risk of developing the disease, disorder, or condition.

In another aspect, the present disclosure provides a use of a nucleic acid molecule, a vector, a cell, an RGN polypeptide, an RNP complex, or a system described hereinabove for the treatment of a disease, disorder, or condition in a subject in need thereof.

In some embodiments of the above aspect, the disease, disorder, or condition is associated with a causal mutation and treating comprises correcting the causal mutation.

In some embodiments of the above aspect, the subject is at risk of developing the disease, disorder, or condition.

In another aspect, the present disclosure provides a use of a nucleic acid molecule, a vector, a cell, an RGN polypeptide, an RNP complex, or a system described hereinabove for the manufacture of a medicament useful for treating a disease, disorder, or condition.

In some embodiments of the above aspect, the disease is associated with a causal mutation and the medicament corrects the causal mutation.

In another aspect, the present disclosure provides a single guide RNA comprising a nucleic acid molecule comprising a crRNA described hereinabove and a nucleic acid molecule comprising a tracrRNA described hereinabove.

In another aspect, the present disclosure provides a dual guide RNA comprising a nucleic acid molecule comprising a crRNA described hereinabove and a nucleic acid molecule comprising a tracrRNA described hereinabove.

DETAILED DESCRIPTION

Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the inventions are not to be limited to the specific embodiments disclosed and that modifications and other embodiments are

intended to be included within the scope of the appended embodiments. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

I. Overview

RNA-guided nucleases (RGNs) allow for the targeted manipulation of specific site(s) within a genome and are useful in the context of gene targeting for therapeutic and research applications. In a variety of organisms, including mammals, RNA-guided nucleases have been used for genome engineering by stimulating non-homologous end joining and homologous recombination, for example. The compositions and methods described herein are useful for creating single- or double-stranded breaks in polynucleotides, modifying polynucleotides, detecting a particular site within a polynucleotide, or modifying the expression of a particular gene.

The RNA-guided nucleases disclosed herein can alter gene expression by modifying a target nucleic acid molecule comprising a target sequence. In specific embodiments, the RNA-guided nucleases are directed to a target sequence (e.g., target DNA sequence) by a guide RNA (gRNA) as part of a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) RNA-guided nuclease system. The RGNs are considered "RNA-guided" because guide RNAs form a complex with the RNA-guided nucleases to direct the RNA-guided nuclease to bind to a target sequence and in some embodiments, introduce a single-stranded or double-stranded break at the target sequence (e.g., target DNA sequence). After the target sequence has been cleaved, the break can be repaired such that the sequence of the target nucleic acid molecule is modified during the repair process. Thus, provided herein are methods for using the RNA-guided nucleases to modify a target nucleic acid molecule in a host cell. For example, RNA-guided nucleases can be used to modify a target sequence at a genomic locus of eukaryotic cells or prokaryotic cells.

II. RNA-guided nucleases

Provided herein are RNA-guided nucleases. The term RNA-guided nuclease (RGN) refers to a polypeptide that binds to a particular target sequence (e.g., target DNA sequence) in a sequence-specific manner and is directed to the target sequence by a guide RNA molecule that is complexed with the polypeptide and hybridizes with the target sequence. Although an RNA-guided nuclease can be capable of cleaving the target sequence upon binding, the term RNA-guided nuclease also encompasses nuclease-dead RNA-guided nucleases that are capable of binding to, but not cleaving, a target sequence. Cleavage of a target sequence by an RNA-guided nuclease can result in a single- or double-stranded break. RNA-guided nucleases only capable of cleaving a single strand of a double-stranded target nucleic acid molecule are referred to herein as nickases.

The RNA-guided nucleases disclosed herein include the LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207 and LPG10208 RNA-guided nucleases, the amino acid sequences of which are set forth, respectively, as SEQ ID

NOs: 1-20, and active fragments or variants thereof that retain the ability to bind to a target sequence in an RNA-guided sequence-specific manner. In some of these embodiments, the active fragment or variant of the LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, or LPG10208 RGN is capable of cleaving a single- or double-stranded target sequence. In some embodiments, an active variant of the LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, or LPG10208 RGN comprises an amino acid sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the amino acid sequence set forth as any one of SEQ ID NOs: 1-20.

In some specific embodiments, an active variant of the LPG10165 RGN comprises an amino acid sequence having at least 85% sequence identity to the amino acid sequence set forth in SEQ ID NO: 1 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10166 RGN comprises an amino acid sequence having at least 83% sequence identity to the amino acid sequence set forth in SEQ ID NO: 2 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10167 RGN comprises an amino acid sequence having at least 92% sequence identity to the amino acid sequence set forth in SEQ ID NO: 3 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10168 RGN comprises an amino acid sequence having at least 93% sequence identity to the amino acid sequence set forth in SEQ ID NO: 4 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10169 RGN comprises an amino acid sequence having at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO: 5 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10171 RGN comprises an amino acid sequence having at least 94% sequence identity to the amino acid sequence set forth in SEQ ID NO: 6 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10186 RGN comprises an amino acid sequence having at least 75% sequence identity to the amino acid sequence set forth in SEQ ID NO: 7 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10190 RGN comprises an amino acid sequence having at least 84% sequence identity to the amino acid sequence set forth in SEQ ID NO: 8 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10191 RGN comprises an amino acid sequence having at least 73% sequence identity to the amino acid sequence set forth in SEQ ID NO: 9 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10194 RGN comprises an amino acid sequence having at least 98% sequence identity to the amino acid sequence set forth in SEQ ID NO: 10 and retains RNA-guided sequence-specific binding activity. In some some

embodiments, an active variant of the LPG10195 RGN comprises an amino acid sequence having at least 88% sequence identity to the amino acid sequence set forth in SEQ ID NO: 11 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10196 RGN comprises an amino acid sequence having at least 75% sequence identity to the amino acid sequence set forth in SEQ ID NO: 12 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10197 RGN comprises an amino acid sequence having at least 92% sequence identity to the amino acid sequence set forth in SEQ ID NO: 13 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10198 RGN comprises an amino acid sequence having at least 97% sequence identity to the amino acid sequence set forth in SEQ ID NO: 14 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10200 RGN comprises an amino acid sequence having at least 98% sequence identity to the amino acid sequence set forth in SEQ ID NO: 15 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10203 RGN comprises an amino acid sequence having at least 78% sequence identity to the amino acid sequence set forth in SEQ ID NO: 16 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LEPG10204 RGN comprises an amino acid sequence having at least 81% sequence identity to the amino acid sequence set forth in SEQ ID NO: 17 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10205 RGN comprises an amino acid sequence having at least 81% sequence identity to the amino acid sequence set forth in SEQ ID NO: 18 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10207 RGN comprises an amino acid sequence having at least 88% sequence identity to the amino acid sequence set forth in SEQ ID NO: 19 and retains RNA-guided sequence-specific binding activity. In some some embodiments, an active variant of the LPG10208 RGN comprises an amino acid sequence having at least 81% sequence identity to the amino acid sequence set forth in SEQ ID NO: 20 and retains RNA-guided sequence-specific binding activity.

In certain embodiments, an active fragment of the LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, or LPG10208 RGN comprises at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, or more contiguous amino acid residues of the amino acid sequence set forth as any one of SEQ ID NOs: 1-20. RNA-guided nucleases provided herein can comprise at least one nuclease domain (*e.g.*, DNase, RNase domain) and at least one RNA recognition and/or RNA binding domain to interact with guide RNAs. Further domains that can be found in RNA-guided nucleases provided herein include, but are not limited to: DNA binding domains, helicase domains, protein-protein interaction domains, and dimerization domains. In specific embodiments, the RNA-guided nucleases provided herein can comprise at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or

more sequence identity to one or more of a DNA binding domain, helicase domain, protein-protein interaction domain, and dimerization domain.

A target sequence is bound by an RNA-guided nuclease provided herein. In those instances wherein the target sequence is double-stranded (e.g., double-stranded DNA), the non- target strand of the target sequence hybridizes with the guide RNA associated with the RNA-guided nuclease. The target strand and/or the non-target strand of the target sequence (e.g., target DNA sequence) can then be subsequently cleaved by the RNA-guided nuclease if the polypeptide possesses nuclease activity. The terms "cleave" or "cleavage" refer to the hydrolysis of at least one phosphodiester bond within the backbone of one or both strands of a double-stranded target sequence (e.g., target DNA sequence) that can result in either single-stranded or double-stranded breaks within the target sequence. The presently disclosed RGNs can cleave nucleotides within a polynucleotide, functioning as an endonuclease or can be an exonuclease, removing successive nucleotides from the end (the 5' and/or the 3' end) of a polynucleotide. In other embodiments, the disclosed RGNs can cleave nucleotides of a target polynucleotide within any position of a polynucleotide and thus function as both an endonuclease and exonuclease. The cleavage of a target polynucleotide by the presently disclosed RGNs can result in staggered breaks or blunt ends.

The presently disclosed RNA-guided nucleases can be wild-type sequences derived from bacterial or archaeal species. Alternatively, the RNA-guided nucleases can be variants or fragments of wild-type polypeptides. The wild-type RGN can be modified to alter nuclease activity or alter PAM specificity, for example. In some embodiments, the RNA-guided nuclease is not naturally-occurring.

In certain embodiments, the RNA-guided nuclease functions as a nickase, only cleaving a single strand of a double-stranded target sequence (e.g., target DNA sequence). Such RNA-guided nucleases have a single functioning nuclease domain. In particular embodiments, the nickase is capable of cleaving the target strand or the non-target strand of the double-stranded target sequence (e.g., target DNA sequence). In some of these embodiments, additional nuclease domains have been mutated such that the nuclease activity is reduced or eliminated. In embodiments wherein a nickase is used, in order to effect a double-stranded cleavage of a double-stranded target sequence (e.g., target DNA sequence), two nickases are needed, each of which nicks a single strand within the double-stranded target sequence.

In other embodiments, the RNA-guided nuclease lacks nuclease activity altogether and is referred to herein as nuclease-dead or nuclease inactive. Any method known in the art for introducing mutations into an amino acid sequence, such as PCR-mediated mutagenesis and site-directed mutagenesis, can be used for generating nickases or nuclease-dead RGNs. See, e.g., U.S. Publ. No. 2014/0068797 and U.S. Pat. No. 9,790,490; each of which is incorporated by reference in its entirety.

RNA-guided nucleases that lack nuclease activity can be used to deliver a fused polypeptide, polynucleotide, or small molecule payload to a particular genomic location. In some of these embodiments, the RGN polypeptide or guide RNA can be fused to a detectable label to allow for detection of a particular sequence. As a non-limiting example, a nuclease-dead RGN can be fused to a detectable label (*e.g.*,

fluorescent protein) and targeted to a particular sequence associated with a disease to allow for detection of the disease-associated sequence.

Alternatively, nuclease-dead RGNs can be targeted to particular genomic locations to alter the expression of a desired gene (i.e., target gene). In some embodiments, the binding of a nuclease-dead RNA-guided nuclease to a target sequence results in the reduction in expression of the target gene by interfering with the binding of RNA polymerase or transcription factors within the targeted genomic region. In other embodiments, the RGN (*e.g.*, a nuclease-dead RGN) or its complexed guide RNA further comprises an expression modulator that, upon binding to a target sequence within a target gene, serves to either repress or activate the expression of the target gene. In some of these embodiments, the expression modulator modulates the expression of the target gene through epigenetic mechanisms.

In other embodiments, the nuclease-dead RGNs or an RGN with nickase activity can be targeted to particular genomic locations to modify the sequence of a target polynucleotide through fusion to a base-editing polypeptide, for example a deaminase polypeptide or active variant or fragment thereof, that directly chemically modifies (e.g., deaminates) a nucleobase, resulting in conversion from one nucleobase to another. The base-editing polypeptide can be fused to the RGN at its N-terminal or C-terminal end. Additionally, the base-editing polypeptide may be fused to the RGN via a peptide linker. A non-limiting example of a deaminase polypeptide that is useful for such compositions and methods includes a cytosine deaminase or an adenine deaminase (such as the adenine deaminase base editor described in Gaudelli *et al.* (2017) *Nature* 551:464-471, U.S. Publ. Nos. 2017/0121693 and 2018/0073012, and International Publ. No. WO 2018/027078, or any of the deaminases disclosed in International Publ. Nos. WO 2020/139783 and WO 2022/056254, and International Appl. No. PCT/US2022/021271 filed March 22, 2022, each of which is herein incorporated by reference in its entirety). In one embodiment, the deaminase polypeptide that is useful for such compositions and methods is a cytosine deaminase or an adenine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 481-552. In one embodiment, the deaminase polypeptide that is useful for such compositions and methods is a cytosine deaminase or an adenine deaminase having a sequence that is at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater identity to any one of the amino acid sequences set forth as SEQ ID NOs: 481-552. In some embodiments, the deaminase polypeptide that is useful for such the presently disclosed compositions and methods is a deaminase disclosed in Table 17 of International Publ. No. WO 2020/139783, which is incorporated herein by reference in its entirety. Further, it is known in the art that certain fusion proteins between an RGN and a base-editing enzyme (e.g., cytosine deaminase) may also comprise at least one uracil stabilizing polypeptide that increases the mutation rate of a cytidine, deoxycytidine, or cytosine to a thymidine, deoxythymidine, or thymine in a nucleic acid molecule by a deaminase. Non-limiting examples of uracil stabilizing polypeptides include those disclosed in International Publ. No. WO 2021/217002, which is herein incorporated by reference in its entirety, including USP2 (SEQ ID NO: 564), and a uracil glycosylase inhibitor (UGI) domain (SEQ ID NO: 565), which may increase base editing efficiency. Therefore, a fusion

protein may comprise an RGN described herein or variant thereof, a deaminase, and optionally at least one uracil stabilizing polypeptide, such as UGI or USP2. In certain embodiments, the RGN that is fused to the base-editing polypeptide is a nickase that cleaves the DNA strand that is not acted upon by the base-editing polypeptide (e.g., deaminase).

RNA-guided nucleases that are fused to a polypeptide or domain can be separated or joined by a linker. The term "linker," as used herein, refers to a chemical group or a molecule linking two molecules or moieties, *e.g.*, a binding domain and a cleavage domain of a nuclease. In some embodiments, a linker joins a gRNA binding domain of an RNA guided nuclease and a base-editing polypeptide, such as a deaminase. In some embodiments, a linker joins a nuclease-dead RGN and a deaminase. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (*e.g.*, a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

The presently disclosed RNA-guided nucleases can comprise at least one nuclear localization signal (NLS) to enhance transport of the RGN to the nucleus of a cell. Nuclear localization signals are known in the art and generally comprise a stretch of basic amino acids (see, *e.g.*, Lange *et al.*, *J. Biol. Chem.* (2007) 282:5101-5105). In some embodiments, the RGN comprises 2, 3, 4, 5, 6 or more nuclear localization signals. The nuclear localization signal(s) can be a heterologous NLS. Non-limiting examples of nuclear localization signals useful for the presently disclosed RGNs are the nuclear localization signals of SV40 Large T-antigen, nucleoplasmin, and c-Myc (*see, e.g.*, Ray et al. (2015) *Bioconjug Chem* 26(6):1004-7). In particular embodiments, the RGN comprises the NLS sequence set forth as SEQ ID NO: 168 or 170. The RGN can comprise one or more NLS sequences at its N-terminus, C- terminus, or both the N-terminus and C-terminus. For example, the RGN can comprise two NLS sequences at the N-terminal region and four NLS sequences at the C-terminal region.

Other localization signal sequences known in the art that localize polypeptides to particular subcellular location(s) can also be used to target the RGNs, including, but not limited to, plastid localization sequences, mitochondrial localization sequences, and dual-targeting signal sequences that target to both the plastid and mitochondria (see, *e.g.*, Nassoury and Morse (2005) *Biochim Biophys Acta* 1743:5-19; Kunze and Berger (2015) *Front Physiol* dx.doi.org/10.3389/fphys.2015.00259; Herrmann and Neupert (2003) *IUBMB Life* 55:219-225; Soll (2002) *Curr Opin Plant Biol* 5:529-535; Carrie and Small (2013) *Biochim Biophys Acta* 1833:253-259; Carrie *et al.* (2009) *FEBS J* 276:1187-1195; Silva-Filho (2003) *Curr Opin Plant Biol* 6:589-595; Peeters and Small (2001) *Biochim Biophys Acta* 1541:54-63; Murcha *et al.* (2014) *J Exp Bot* 65:6301-6335; Mackenzie (2005) *Trends Cell Biol* 15:548-554; Glaser *et al.* (1998) *Plant Mol Biol* 38:311-338).

In certain embodiments, the presently disclosed RNA-guided nucleases comprise at least one cell-penetrating domain that facilitates cellular uptake of the RGN. Cell-penetrating domains are known in the art and generally comprise stretches of positively charged amino acid residues (*i.e.*, polycationic cell-penetrating domains), alternating polar amino acid residues and non-polar amino acid residues (*i.e.*, amphipathic cell-penetrating domains), or hydrophobic amino acid residues (*i.e.*, hydrophobic cell-penetrating domains) (*see, e.g.*, Milletti F. (2012) *Drug Discov Today* 17:850-860). A non-limiting example of a cell-penetrating domain is the trans-activating transcriptional activator (TAT) from the human immunodeficiency virus 1.

The nuclear localization signal, plastid localization signal, mitochondrial localization signal, dual-targeting localization signal, and/or cell-penetrating domain can be located at the amino-terminus (N-terminus), the carboxyl-terminus (C-terminus), or in an internal location of the RNA-guided nuclease.

The presently disclosed RGNs can be fused to an effector domain, such as a cleavage domain, a deaminase domain, or an expression modulator domain, either directly or indirectly via a linker peptide. Such a domain can be located at the N-terminus, the C-terminus, or an internal location of the RNA-guided nuclease. In some of these embodiments, the RGN component of the fusion protein is a nuclease-dead RGN or a nickase.

In some embodiments, the RGN fusion protein comprises a cleavage domain, which is any domain that is capable of cleaving a polynucleotide (*i.e.*, RNA, DNA, or RNA/DNA hybrid) and includes, but is not limited to, restriction endonucleases and homing endonucleases, such as Type IIS endonucleases (*e.g.*, *Fok*I) (*see, e.g.*, Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388; Linn *et al.* (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993).

In other embodiments, the RGN fusion protein comprises a deaminase domain that deaminates a nucleobase, resulting in conversion from one nucleobase to another, and includes, but is not limited to, a cytosine deaminase or an adenine deaminase (see, *e.g.*, Gaudelli *et al.* (2017) *Nature* 551:464-471, U.S. Publ. Nos. 2017/0121693 and 2018/0073012, and International Publ. No. WO 2018/027078, or any of the deaminases disclosed in International Publ. Nos. WO 2020/139783 and WO 2022/056254, and International Appl. No. PCT/US2022/021271 filed March 22, 2022, each of which is herein incorporated by reference in its entirety. In some embodiments, the effector domain of the RGN fusion protein can be an expression modulator domain, which is a domain that either serves to upregulate or downregulate transcription. The expression modulator domain can be an epigenetic modification domain, a transcriptional repressor domain or a transcriptional activation domain.

In some of these embodiments, the expression modulator of the RGN fusion protein comprises an epigenetic modification domain that covalently modifies DNA or histone proteins to alter histone structure and/or chromosomal structure without altering the DNA sequence, leading to changes in gene expression (*i.e.*, upregulation or downregulation). Non-limiting examples of epigenetic modifications include acetylation or methylation of lysine residues, arginine methylation, serine and threonine phosphorylation, and lysine ubiquitination and sumoylation of histone proteins, and methylation and hydroxymethylation of

cytosine residues in DNA. Non-limiting examples of epigenetic modification domains include histone acetyltransferase domains, histone deacetylase domains, histone methyltransferase domains, histone demethylase domains, DNA methyltransferase domains, and DNA demethylase domains.

In other embodiments, the expression modulator of the fusion protein comprises a transcriptional repressor domain, which interacts with transcriptional control elements and/or transcriptional regulatory proteins, such as RNA polymerases and transcription factors, to reduce or terminate transcription of at least one gene. Transcriptional repressor domains are known in the art and include, but are not limited to, Sp1-like repressors, IκB, and Krüppel associated box (KRAB) domains.

In yet other embodiments, the expression modulator of the fusion protein comprises a transcriptional activation domain, which interacts with transcriptional control elements and/or transcriptional regulatory proteins, such as RNA polymerases and transcription factors, to increase or activate transcription of at least one gene. Transcriptional activation domains are known in the art and include, but are not limited to, a herpes simplex virus VP16 activation domain and an NFAT activation domain.

The presently disclosed RGN polypeptides can comprise a detectable label or a purification tag. The detectable label or purification tag can be located at the N-terminus, the C-terminus, or an internal location of the RNA-guided nuclease, either directly or indirectly via a linker peptide. In some of these embodiments, the RGN component of the fusion protein is a nuclease-dead RGN. In other embodiments, the RGN component of the fusion protein is an RGN with nickase activity.

A detectable label is a molecule that can be visualized or otherwise observed. The detectable label may be fused to the RGN as a fusion protein (*e.g.*, fluorescent protein) or may be a small molecule conjugated to the RGN polypeptide that can be detected visually or by other means. Detectable labels that can be fused to the presently disclosed RGNs as a fusion protein include any detectable protein domain, including but not limited to, a fluorescent protein or a protein domain that can be detected with a specific antibody. Non-limiting examples of fluorescent proteins include green fluorescent proteins (*e.g.*, GFP, EGFP, ZsGreen1) and yellow fluorescent proteins (*e.g.*, YFP, EYFP, ZsYellow1). Non-limiting examples of small molecule detectable labels include radioactive labels, such as ^{3}H and ^{35}S.

RGN polypeptides can also comprise a purification tag, which is any molecule that can be utilized to isolate a protein or fused protein from a mixture (e.g., biological sample, culture medium). Non-limiting examples of purification tags include biotin, myc, maltose binding protein (MBP), glutathione-S-transferase (GST), and 3X FLAG tag.

III. Guide RNA

The present disclosure provides guide RNAs and polynucleotides encoding the same that target an associated RGN to a target sequence. The term "guide RNA" refers to a nucleotide sequence having sufficient complementarity with a target nucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of an associated RNA-guided nuclease to the target nucleotide sequence. More specifically, when the target nucleotide sequence is double-stranded as is the case with DNA, the target

nucleotide sequence is comprised of a target strand (which comprises the PAM sequence) and the non-target strand. In these embodiments, the guide RNA has sufficient complementarity with the non-target strand of a double-stranded target sequence (e.g., target DNA sequence) such that the guide RNA hybridizes with the non-target strand and directs sequence-specific binding of an associated RNA-guided nuclease (RGN) to the target sequence (e.g., target DNA sequence). Therefore, in some embodiments, a guide RNA includes a spacer that is identical to the sequence of the target strand except that uracil (U) replaces thymidine (T) in the guide RNA.

An RGN's respective guide RNA is one or more RNA molecules (generally, one or two), that can bind to the RGN and guide the RGN to bind to a particular target sequence, and in those embodiments wherein the RGN has nickase or nuclease activity, also cleave the target strand and/or the non-target strand. In general, a guide RNA comprises a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA), although some RGNs do not require a tracrRNA. Native guide RNAs that comprise both a crRNA and a tracrRNA generally comprise two separate RNA molecules that hybridize to each other through the repeat sequence of the crRNA and the anti-repeat sequence of the tracrRNA.

The present invention provides CRISPR RNAs (crRNAs) or polynucleotides encoding CRISPR RNAs that together with a tracrRNA, target an associated RGN to a target sequence. A crRNA comprises a spacer and a CRISPR repeat. The "spacer" has a nucleotide sequence that directly hybridizes with the non-target strand of a target sequence (e.g., target DNA sequence) of interest. The spacer is engineered to have full or partial complementarity with the non-target strand of a target sequence of interest. In some embodiments, the spacer can comprise from about 8 nucleotides to about 30 nucleotides, or more. For example, the spacer can be about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In some embodiments, the spacer is 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more nucleotides in length. In some embodiments, the spacer is about 10 to about 26 nucleotides in length, or about 12 to about 30 nucleotides in length. In some embodiments, the degree of complementarity between a spacer and the non-target strand of a target sequence (e.g., target DNA sequence), when optimally aligned using a suitable alignment algorithm, is between 50% and 99% or more, including but not limited to about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more. In some embodiments, the degree of complementarity between a spacer and the non-target strand of a target sequence (e.g., target DNA sequence), when optimally aligned using a suitable alignment algorithm, is 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more. In some embodiments, the spacer can be identical in sequence to the target strand of a target sequence. In some of those embodiments wherein the target sequence is a target DNA sequence, the spacer can be identical in sequence to the target strand of the target DNA sequence, with the exception of the

thymidines (Ts) in the target strand being replaced by uracils (Us) in the spacer. In particular embodiments, the spacer is free of secondary structure, which can be predicted using any suitable polynucleotide folding algorithm known in the art, including but not limited to mFold (see, *e.g.*, Zuker and Stiegler (1981) *Nucleic Acids Res*. 9:133-148) and RNAfold (see, *e.g.*, Gruber et al. (2008) *Cell* 106(1):23-24).

The presently disclosed crRNAs comprise a spacer capable of targeting a bound RGN polypeptide to a target DNA sequence, wherein the target strand of the target DNA sequence has the nucleotide sequence set forth in any one of SEQ ID NOs: 344-464, 573-641, 667-677, 684-747, 770-817, 826-1039, and 1046-1057.

Along with a spacer, crRNAs further comprise a CRISPR RNA repeat. The CRISPR RNA repeat comprises a nucleotide sequence that forms a structure, either on its own or in concert with a hybridized tracrRNA, that is recognized by the RGN molecule. In various embodiments, the CRISPR RNA repeat can comprise from about 8 nucleotides to about 30 nucleotides, or more. For example, the CRISPR repeat can be about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In particular embodiments, the CRISPR repeat is 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more nucleotides in length. In some embodiments, the degree of complementarity between a CRISPR repeat and its corresponding tracrRNA sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more. In particular embodiments, the degree of complementarity between a CRISPR repeat and its corresponding tracrRNA sequence, when optimally aligned using a suitable alignment algorithm, is 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more.

In particular embodiments, the CRISPR repeat comprises the nucleotide sequence of any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof that when comprised within a guide RNA, is capable of directing the sequence-specific binding of an associated RNA-guided nuclease provided herein to a target sequence of interest. In certain embodiments, an active CRISPR repeat variant of a wild-type sequence comprises a nucleotide sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045. In certain embodiments, an active CRISPR repeat fragment of a wild-type sequence comprises at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22 contiguous nucleotides of any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

In certain embodiments, the crRNA is not naturally-occurring. In some of these embodiments, the specific CRISPR repeat is not linked to the engineered spacer in nature and the CRISPR repeat is considered heterologous to the spacer. In certain embodiments, the spacer is an engineered sequence that is not naturally occurring.

Presently disclosed guide RNAs comprise a crRNA and a trans-activating CRISPR RNA (tracrRNA). A tracrRNA molecule comprises a nucleotide sequence comprising a region that has sufficient complementarity to hybridize to a CRISPR repeat of a crRNA, which is referred to herein as the anti-repeat. In some embodiments, the tracrRNA molecule further comprises a region with secondary structure (*e.g.*, stem-loop) or forms secondary structure upon hybridizing with its corresponding crRNA. In particular embodiments, the region of the tracrRNA that is fully or partially complementary to a CRISPR repeat is at the 5' end of the molecule and the 3' end of the tracrRNA comprises secondary structure. This region of secondary structure generally comprises several hairpin structures, including the nexus hairpin, which is found adjacent to the anti-repeat. The nexus forms the core of the interactions between the guide RNA and the RGN, and is at the intersection between the guide RNA, the RGN, and the target DNA. The nexus hairpin often has a conserved nucleotide sequence in the base of the hairpin stem, with the motif UNANNC (SEQ ID NO: 566) found in many nexus hairpins in tracrRNAs. In embodiments, guide RNAs or RGN systems of the disclosure use tracrRNAs that comprise non-canonical sequences in the base of the hairpin stem of their nexus hairpins, including UNANNG (SEQ ID NO: 567), CNANNC (SEQ ID NO: 568), CNANNU (SEQ ID NO: 569), UNANNU (SEQ ID NO: 570), CNANNG (SEQ ID NO: 571), and CNCNNU (SEQ ID NO: 572). There are often terminal hairpins at the 3' end of the tracrRNA that can vary in structure and number, but often comprise a GC-rich Rho-independent transcriptional terminator hairpin followed by a string of U's at the 3' end. See, for example, Briner et al. (2014) *Molecular Cell* 56:333-339, Briner and Barrangou (2016) *Cold Spring Harb Protoc*; doi: 10.1101/pdb.top090902, and U.S. Publication No. 2017/0275648, each of which is herein incorporated by reference in its entirety.

In various embodiments, the anti-repeat region of the tracrRNA that is fully or partially complementary to the CRISPR repeat comprises from about 8 nucleotides to about 30 nucleotides, or more. For example, the region of base pairing between the tracrRNA anti-repeat and the CRISPR repeat can be about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In particular embodiments, the region of base pairing between the tracrRNA anti-repeat and the CRISPR repeat is 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more nucleotides in length. In some embodiments, the degree of complementarity between a CRISPR repeat and its corresponding tracrRNA anti-repeat, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more. In particular embodiments, the degree of complementarity

between a CRISPR repeat and its corresponding tracrRNA anti-repeat, when optimally aligned using a suitable alignment algorithm, is 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more.

In various embodiments, the entire tracrRNA can comprise from about 60 nucleotides to more than about 210 nucleotides. For example, the tracrRNA can be about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, about 100, about 105, about 110, about 115, about 120, about 125, about 130, about 135, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 210, or more nucleotides in length. In particular embodiments, the tracrRNA is 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 190, 200, 210 or more nucleotides in length. In particular embodiments, the tracrRNA is about 57 to about 115 nucleotides in length, including about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70, about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about 92, about 93, about 94, about 95, about 96, about 97, about 98, about 99, about 100, about 101, about 102, about 103, about 104, about 105, about 106, about 107, about 108, about 109, about 110, about 111, about 112, about 113, about 114, and about 115 nucleotides in length. In particular embodiments, the tracrRNA is 59 to 115 nucleotides in length, including 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, and 115 nucleotides in length.

In particular embodiments, the tracrRNA comprises the nucleotide sequence of any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease provided herein to a target DNA sequence of interest. In certain embodiments, an active tracrRNA sequence variant comprises a nucleotide sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045. In certain embodiments, an active tracrRNA sequence fragment comprises at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more contiguous nucleotides of any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

Two polynucleotide sequences can be considered to be substantially complementary when the two sequences hybridize to each other under stringent conditions. Likewise, an RGN is considered to bind to a

particular target sequence within a sequence-specific manner if the guide RNA bound to the RGN binds to a target sequence under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which the two polynucleotide sequences will hybridize to each other to a detectably greater degree than to other sequences (*e.g.*, at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is at least about 30°C for short sequences (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long sequences (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

The Tm is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched sequence. For DNA-DNA hybrids, the Tm can be approximated from the equation of Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284: Tm = 81.5°C + 16.6 (log M) + 0.41 (%GC) - 0.61 (% form) - 500/L; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (Tm); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (Tm); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (Tm). Using the equation, hybridization and wash compositions, and desired Tm, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, New York); and Ausubel et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

The term "sequence specific" can also refer to the binding of a RGN polypeptide to a target sequence at a greater frequency than binding to a randomized background sequence.

The guide RNA can be a single guide RNA (sgRNA) or a dual-guide RNA. A single guide RNA comprises the crRNA and tracrRNA on a single molecule of RNA, whereas a dual-guide RNA comprises a crRNA and a tracrRNA present on two distinct RNA molecules, hybridized to one another through at least a portion of the CRISPR repeat of the crRNA and at least a portion of the tracrRNA (i.e., the antirepeat), which may be fully or partially complementary to the CRISPR repeat of the crRNA. In some of those embodiments wherein the guide RNA is a single guide RNA, the crRNA and tracrRNA are separated by a linker nucleotide sequence. In general, the linker nucleotide sequence is one that does not include complementary bases in order to avoid the formation of secondary structure within or comprising nucleotides of the linker nucleotide sequence. In some embodiments, the linker nucleotide sequence between the crRNA and tracrRNA is at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, or more nucleotides in length. In particular embodiments, the linker nucleotide sequence of a single guide RNA is at least 4 nucleotides in length. In certain embodiments, the linker nucleotide sequence is the nucleotide sequence set forth as SEQ ID NO: 84.

The single guide RNA or dual-guide RNA can be synthesized chemically or via *in vitro* transcription. Assays for determining sequence-specific binding between an RGN and a guide RNA are known in the art and include, but are not limited to, *in vitro* binding assays between an expressed RGN and the guide RNA, which can be tagged with a detectable label (e.g., biotin) and used in a pull-down detection assay in which the guide RNA:RGN complex is captured via the detectable label (e.g., with streptavidin beads). A control guide RNA with an unrelated sequence or structure to the guide RNA can be used as a negative control for non-specific binding of the RGN to RNA. In certain embodiments, the guide RNA has a backbone sequence that is any one of SEQ ID NOs: 63-83, 1040, 1041, 1042, 1043, 1044, or 1045.

In certain embodiments, the guide RNA can be introduced into a target cell, organelle, or embryo as an RNA molecule. The guide RNA can be transcribed *in vitro* or chemically synthesized. In other embodiments, a nucleotide sequence encoding the guide RNA is introduced into the cell, or embryo. In some of these embodiments, the nucleotide sequence encoding the guide RNA is operably linked to a promoter (*e.g.*, an RNA polymerase III promoter). The promoter can be a native promoter or heterologous to the guide RNA-encoding nucleotide sequence.

In various embodiments, the guide RNA can be introduced into a target cell, organelle, or embryo as a ribonucleoprotein complex, as described herein, wherein the guide RNA is bound to an RNA-guided nuclease polypeptide.

The guide RNA directs an associated RGN to a particular target nucleotide sequence of interest through hybridization of the guide RNA to the target sequence of interest. The target sequence can be bound (and in some embodiments, cleaved) by an RNA-guided nuclease *in vitro* or in a cell. A target sequence is within a target polynucleotide and can comprise DNA, RNA, or a combination of both and can be single-stranded or double-stranded. A target sequence can be genomic DNA (*i.e.*, chromosomal DNA), plasmid

DNA, or an RNA molecule (*e.g.*, messenger RNA, ribosomal RNA, transfer RNA, micro RNA, small interfering RNA). In those embodiments wherein the target sequence is a chromosomal sequence, the chromosomal sequence can be a nuclear, plastid, or mitochondrial chromosomal sequence. In the presently disclosed compositions and methods, the target sequence is within a target nucleic acid molecule that is double-stranded (e.g., a target DNA sequence. In embodiments, the target sequence is unique in the target genome.

The target sequence is adjacent to a protospacer adjacent motif (PAM) and the target strand of the target sequence is the strand that comprises the PAM. The PAM is immediately adjacent to the target sequence and often comprise Ns, which represent any nucleotide. In some embodiments, the protospacer adjacent motif comprises about 1 to about 10 Ns, including about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, or about 10 nucleotides. In particular embodiments, a PAM comprises 1 to 10 Ns, including 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 Ns. The PAM can be 5' or 3' of the target sequence on its target strand. The PAM of the presently disclosed RGNs is immediately 3' of the target sequence on its target strand. Generally, the PAM is a consensus sequence of about 3-4 nucleotides, but in particular embodiments it can be 2, 3, 4, 5, 6, 7, 8, 9, or more nucleotides in length. In various embodiments, the PAM sequence recognized by the presently disclosed RGNs comprises the consensus sequence set forth in any one of SEQ ID NOs: 127-147. In some embodiments of the above aspect, the crRNA is capable of binding to an RGN polypeptide capable of recognizing a full protospacer adjacent motif (PAM) having the nucleotide sequence set forth in any one of SEQ ID NOs: 127-147.

In particular embodiments, an RNA-guided nuclease having any one of SEQ ID NOs: 1-20 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as any one of SEQ ID NOs: 127-147. In some embodiments, the RGN binds to a guide RNA comprising a CRISPR repeat having a nucleotide sequence set forth in any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof, and a tracrRNA having a nucleotide sequence set forth in any one of SEQ ID NOs: 42-62, or nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof. The RGN systems are described further in Examples 1-3 and Tables 1 and 2 of the present specification.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 1 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 127, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 21 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 42 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 2 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as any one of SEQ ID NOs: 128-131, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID

NO: 22 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 43 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 3 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 132, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 23 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 4 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 132, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 24 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 45 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 5 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 133, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 6 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 134, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 26 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 7 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 135, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO: 1044 or nucleotides 27-95 of SEQ ID NO: 1045 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 8 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 136 or 137, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 28 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 49 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 9 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 138, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 29 or an

active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 50 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 10 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 139, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 30 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 51 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 11 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 140, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 31 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 52 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 12 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 141, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 32 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 53 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 13 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 142, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 33 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 54 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 14 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 143, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 34 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 55 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 15 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 144, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 35 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 56 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 16 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 145, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 36 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 57 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 17 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 146, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 37 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 58 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 18 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 146, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 38 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 59 or an active variant or fragment thereof. In some other embodiments, an RNA-guided nuclease having SEQ ID NO: 18 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 146, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 39 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 60 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 19 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 132, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 40 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO: 61 or an active variant or fragment thereof.

In some embodiments, an RNA-guided nuclease having SEQ ID NO: 20 or an active variant or fragment thereof binds a target nucleotide sequence adjacent to a PAM sequence set forth as SEQ ID NO: 146 or 147, when bound to a guide RNA comprising a CRISPR repeat sequence set forth as SEQ ID NO: 41 or an active variant or fragment thereof and a tracrRNA sequence set forth as SEQ ID NO:62 or an active variant or fragment thereof.It is well-known in the art that PAM sequence specificity for a given nuclease enzyme is affected by enzyme concentration (*see, e.g.*, Karvelis *et al.* (2015) *Genome Biol* 16:253), which may be modified by altering the promoter used to express the RGN, or the amount of ribonucleoprotein complex delivered to the cell, or embryo.

Upon recognizing its corresponding PAM sequence, the RGN can cleave one or both strands of a target DNA sequence at a specific cleavage site. As used herein, a cleavage site is made up of the two particular nucleotides within a target DNA sequence between which the strand of a target DNA locus is cleaved by an RGN. The cleavage site can comprise the 1st and 2nd, 2nd and 3rd, 3rd and 4th, 4th and 5th, 5th and 6th, 7th and 8th, or 8th and 9th nucleotides from the PAM in either the 5' or 3' direction. In some embodiments, the cleavage site may be over 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides from the PAM in either the 5’ or 3’ direction. As RGNs can cleave a target DNA sequence resulting in staggered ends, in some embodiments, the cleavage site is defined based on the distance of the two nucleotides from the PAM on the target strand of the target DNA sequence and for the non-target strand, the distance of the two nucleotides from the complement of the PAM.

IV. Nucleotides Encoding RNA-guided nucleases, CRISPR RNA, and/or tracrRNA

The present disclosure provides polynucleotides comprising the presently disclosed CRISPR RNAs, tracrRNAs, and/or gRNAs and polynucleotides comprising a nucleotide sequence encoding the presently disclosed RNA-guided nucleases, CRISPR RNAs, tracrRNAs, and/or gRNAs. Presently disclosed polynucleotides include those comprising or encoding a crRNA comprising a CRISPR repeat sequence having any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease to a target sequence of interest. Also disclosed are polynucleotides comprising or encoding a tracrRNA having any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease to a target sequence of interest. Polynucleotides are also provided that encode an RNA-guided nuclease having any one of the amino acid sequences set forth as SEQ ID NOs: 1-20, and active fragments or variants thereof that retain the ability to bind to a target sequence in an RNA-guided sequence-specific manner.

The use of the term "polynucleotide" or "nucleic acid molecule" is not intended to limit the present disclosure to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides (RNA) and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. These include peptide nucleic acids (PNAs), PNA-DNA chimers, locked nucleic acids (LNAs), and phosphothiorate linked sequences. The polynucleotides disclosed herein also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, DNA-RNA hybrids, triplex structures, stem-and-loop structures, and the like.

In some embodiments, the polynucleotide encoding a presently disclosed RGN is an mRNA (messenger RNA) molecule. An mRNA refers to any polynucleotide which encodes a polypeptide of interest and which is capable of being translated to produce the encoded polypeptide of interest *in vitro*, *in vivo*, *in situ*, or *ex vivo*. In embodiments, the basic components of an mRNA molecule include at least a coding region, a 5′UTR, a 3′UTR, a 5′ cap and a poly-A tail. In embodiments, an mRNA encoding an RGN useful in the presently disclosed methods and compositions can include one or more structural and/or chemical modifications or alterations which impart useful properties to the polynucleotide. For instance, a useful property of an mRNA includes the lack of a substantial induction of the innate immune response of a cell into which the mRNA is introduced. A "structural" feature or modification is one in which two or more linked nucleotides are inserted, deleted, duplicated, inverted or randomized in an mRNA without significant chemical modification to the nucleotides themselves. Because chemical bonds will necessarily be broken

and reformed to effect a structural modification, structural modifications are of a chemical nature and hence are chemical modifications. However, structural modifications will result in a different sequence of nucleotides. Chemical modifications to mRNA can involve inclusion of 5-methylcytosine, N1-methyl-pseudouridine, pseudouridine, 2-thiouridine, 4-thiouridine, 5-methoxyuridine, 2′Fluoroguanosine, 2′Fluorouridine, 5-bromouridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3(1-E-propenylamino)] uridine, α-thiocytidine, N6-methyladenosine, 5-methylcytidine, N4-acetylcytidine, 5-formylcytidine, or combinations thereof, in an mRNA.

The nucleic acid molecules encoding RGNs can be codon optimized for expression in an organism of interest. A "codon-optimized" coding sequence is a polynucleotide coding sequence having its frequency of codon usage designed to mimic the frequency of preferred codon usage or transcription conditions of a particular host cell. Expression in the particular host cell or organism is enhanced as a result of the alteration of one or more codons at the nucleic acid level such that the translated amino acid sequence is not changed. Nucleic acid molecules can be codon optimized, either wholly or in part. Codon tables and other references providing preference information for a wide range of organisms are available in the art (*see*, *e.g.*, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of plant-preferred codon usage). Methods are available in the art for synthesizing plant-preferred genes or mammalian (for example human) codon-optimized coding sequences. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference. Non-limiting examples of codon-optimized coding sequences for the presently disclosed RGNs are set forth as SEQ ID NOs: 148-167.

Polynucleotides encoding the RGNs, crRNAs, tracrRNAs, and/or gRNAs provided herein can be provided in expression cassettes for *in vitro* expression or expression in a cell, organelle, embryo, or organism of interest. The cassette will include 5' and 3' regulatory sequences operably linked to a polynucleotide encoding an RGN, crRNA, tracrRNAs, and/or gRNAs provided herein that allows for expression of the polynucleotide. The cassette may additionally contain at least one additional gene or genetic element to be cotransformed into the organism. Where additional genes or elements are included, the components are operably linked. The term "operably linked" is intended to mean a functional linkage between two or more elements. For example, an operable linkage between a promoter and a coding region of interest (*e.g.*, region coding for an RGN, crRNA, tracrRNAs, and/or gRNAs) is a functional link that allows for expression of the coding region of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by operably linked is intended that the coding regions are in the same reading frame. Alternatively, the additional gene(s) or element(s) can be provided on multiple expression cassettes. For example, the nucleotide sequence encoding a presently disclosed RGN can be present on one expression cassette, whereas the nucleotide sequence encoding a crRNA, tracrRNA, or guide RNA can be on a separate expression cassette. Such an expression cassette is provided with a plurality of restriction sites and/or recombination sites for insertion of the polynucleotides to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain a selectable marker gene.

The expression cassette will include in the 5'-3' direction of transcription, a transcriptional (and, in some embodiments, translational) initiation region (*i.e.*, a promoter), an RGN-, crRNA-, tracrRNA-and/or sgRNA- encoding polynucleotide of the invention, and a transcriptional (and in some embodiments, translational) termination region (*i.e.*, termination region) functional in the organism of interest. The promoters of the invention are capable of directing or driving expression of a coding sequence in a host cell. The regulatory regions (*e.g.*, promoters, transcriptional regulatory regions, and translational termination regions) may be endogenous or heterologous to the host cell or to each other. As used herein, "heterologous" in reference to a sequence is a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. As used herein, a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence.

Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acids Res.* 15:9627-9639.

Additional regulatory signals include, but are not limited to, transcriptional initiation start sites, operators, activators, enhancers, other regulatory elements, ribosomal binding sites, an initiation codon, termination signals, and the like. See, for example, U.S. Pat. Nos. 5,039,523 and 4,853,331; EPO 0480762A2; Sambrook et al. (1992) Molecular Cloning: A Laboratory Manual, ed. Maniatis et al. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), hereinafter "Sambrook 11"; Davis et al., eds. (1980) Advanced Bacterial Genetics (Cold Spring Harbor Laboratory Press), Cold Spring Harbor, N.Y., and the references cited therein.

In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, *e.g.*, transitions and transversions, may be involved.

A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, inducible, growth stage-specific, cell type-specific, tissue-preferred, tissue-specific, or other promoters for expression in the organism of interest. See, for example, promoters set forth in WO 99/43838 and in US Patent Nos: 8,575,425; 7,790,846; 8,147,856; 8,586832; 7,772,369; 7,534,939; 6,072,050; 5,659,026; 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611; herein incorporated by reference.

For expression in plants, constitutive promoters also include CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); and MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730).

Examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, the PPDK promoter and the pepcarboxylase promoter which are both inducible by light. Also useful are promoters which are chemically inducible, such as the In2-2 promoter which is safener induced (U.S. Pat. No. 5,364,780), the Axig1 promoter which is auxin induced and tapetum specific but also active in callus (PCT US01/22169), the steroid-responsive promoters (see, for example, the ERE promoter which is estrogen induced, and the glucocorticoid-inducible promoter in Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis *et al.* (1998) *Plant J.* 14(2):247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, and U.S. Pat. Nos. 5,814,618 and 5,789,156), herein incorporated by reference.

Tissue-specific or tissue-preferred promoters can be utilized to target expression of an expression construct within a particular tissue. In certain embodiments, the tissue-specific or tissue-preferred promoters are active in plant tissue. Examples of promoters under developmental control in plants include promoters that initiate transcription preferentially in certain tissues, such as leaves, roots, fruit, seeds, or flowers. A "tissue specific" promoter is a promoter that initiates transcription only in certain tissues. Unlike constitutive expression of genes, tissue-specific expression is the result of several interacting levels of gene regulation. As such, promoters from homologous or closely related plant species can be preferable to use to achieve efficient and reliable expression of transgenes in particular tissues. In some embodiments, the expression comprises a tissue-preferred promoter. A "tissue preferred" promoter is a promoter that initiates transcription preferentially, but not necessarily entirely or solely in certain tissues.

In some embodiments, the nucleic acid molecules encoding an RGN, crRNA, and/or tracrRNA comprise a cell type-specific promoter. A "cell type specific" promoter is a promoter that primarily drives expression in certain cell types in one or more organs. Some examples of plant cells in which cell type specific promoters functional in plants may be primarily active include, for example, BETL cells, vascular cells in roots, leaves, stalk cells, and stem cells. The nucleic acid molecules can also include cell type preferred promoters. A "cell type preferred" promoter is a promoter that primarily drives expression mostly, but not necessarily entirely or solely in certain cell types in one or more organs. Some examples of plant cells in which cell type preferred promoters functional in plants may be preferentially active include, for example, BETL cells, vascular cells in roots, leaves, stalk cells, and stem cells.

The nucleic acid sequences encoding the RGNs, crRNAs, tracrRNAs, and/or gRNAs can be operably linked to a promoter sequence that is recognized by a phage RNA polymerase for example, for *in vitro* mRNA synthesis. In such embodiments, the *in vitro*-transcribed RNA can be purified for use in the methods described herein. For example, the promoter sequence can be a T7, T3, or SP6 promoter sequence

or a variation of a T7, T3, or SP6 promoter sequence. In such embodiments, the expressed protein and/or RNAs can be purified for use in the methods of genome modification described herein.

In certain embodiments, the polynucleotide encoding the RGN, crRNA, tracrRNA, and/or gRNA also can be linked to a polyadenylation signal (*e.g.*, SV40 polyA signal and other signals functional in plants) and/or at least one transcriptional termination sequence. Additionally, the sequence encoding the RGN also can be linked to sequence(s) encoding at least one nuclear localization signal, at least one cell-penetrating domain, and/or at least one signal peptide capable of trafficking proteins to particular subcellular locations, as described elsewhere herein.

The polynucleotide encoding the RGN, crRNA, tracrRNA, and/or gRNA can be present in a vector or multiple vectors. A "vector" refers to a polynucleotide composition for transferring, delivering, or introducing a nucleic acid into a host cell. Suitable vectors include plasmid vectors, phagemids, cosmids, artificial/mini-chromosomes, transposons, and viral vectors (*e.g.*, lentiviral vectors, adeno-associated viral vectors, baculoviral vector). The vector can comprise additional expression control sequences (*e.g.*, enhancer sequences, Kozak sequences, polyadenylation sequences, transcriptional termination sequences), selectable marker sequences (*e.g.*, antibiotic resistance genes), origins of replication, and the like. Additional information can be found in "Current Protocols in Molecular Biology" Ausubel *et al.*, John Wiley & Sons, New York, 2003 or "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 3rd edition, 2001.

The vector can also comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D).

In some embodiments, the expression cassette or vector comprising the sequence encoding the RGN polypeptide can further comprise a sequence encoding a crRNA and/or a tracrRNA, or the crRNA and tracrRNA combined to create a gRNA. The sequence(s) encoding the crRNA and/or tracrRNA can be operably linked to at least one transcriptional control sequence for expression of the crRNA and/or tracrRNA in the organism or host cell of interest. For example, the polynucleotide encoding the crRNA and/or tracrRNA can be operably linked to a promoter sequence that is recognized by RNA polymerase III (Pol III). Examples of suitable Pol III promoters include, but are not limited to, mammalian U6, U3, H1, and 7SL RNA promoters and rice U6 and U3 promoters, such as the human U6 promoter set forth as SEQ ID NO: 173, as well as the promoters disclosed in U.S. Provisional Appl. No. 63/209,660, filed June 11, 2021, and PCT International Appl. No. PCT/US2022/032940, filed June 10, 2022, each of which is herein incorporated by reference in its entirety, including those set forth herein as SEQ ID NOs: 553-562.

As indicated, expression constructs comprising nucleotide sequences encoding the RGNs, crRNA, tracrRNA, and/or gRNA can be used to transform organisms of interest. Methods for transformation involve introducing a nucleotide construct into an organism of interest. By "introducing" is intended to introduce the

nucleotide construct to the host cell in such a manner that the construct gains access to the interior of the host cell. The methods of the invention do not require a particular method for introducing a nucleotide construct to a host organism, only that the nucleotide construct gains access to the interior of at least one cell of the host organism. The host cell can be a eukaryotic or prokaryotic cell. In particular embodiments, the eukaryotic host cell is a plant cell, a mammalian cell, an avian cell, or an insect cell. In some embodiments, the eukaryotic cell that comprises or expresses a presently disclosed RGN or that has been modified by a presently disclosed RGN is a human cell. In some embodiments, the eukaryotic cell that comprises or expresses a presently disclosed RGN or that has been modified by a presently disclosed RGN is a cell of hematopoietic origin, such as an immune cell (i.e., a cell of the innate or adaptive immune system) including but not limited to a B cell, a T cell, a natural killer (NK) cell, a pluripotent stem cell, an induced pluripotent stem cell, a chimeric antigen receptor T (CAR-T) cell, a monocyte, a macrophage, and a dendritic cell. In some embodiments, the eukaryotic cell that comprises or expresses a presently disclosed RGN or that has been modified by a presently disclosed RGN is an ocular cell, muscle cell (e.g., skeletal muscle cell), epithelial cell (e.g., lung epithelial cell), diseased cell (e.g., tumor cell).

Methods for introducing nucleotide constructs into plants and other host cells are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

The methods result in a transformed organism, such as a plant, including whole plants, as well as plant organs (*e.g.*, leaves, stems, roots, etc.), seeds, plant cells, propagules, embryos and progeny of the same. Plant cells can be differentiated or undifferentiated (*e.g.,* callus, suspension culture cells, protoplasts, leaf cells, root cells, phloem cells, pollen).

"Transgenic organisms" or "transformed organisms" or "stably transformed" organisms or cells or tissues refers to organisms that have incorporated or integrated a polynucleotide encoding an RGN, crRNA, and/or tracrRNA of the invention. It is recognized that other exogenous or endogenous nucleic acid sequences or DNA fragments may also be incorporated into the host cell. *Agrobacterium*-and biolistic-mediated transformation remain the two predominantly employed approaches for transformation of plant cells. However, transformation of a host cell may be performed by infection, transfection, microinjection, electroporation, microprojection, biolistics or particle bombardment, electroporation, silica/carbon fibers, ultrasound mediated, PEG mediated, calcium phosphate co-precipitation, polycation DMSO technique, DEAE dextran procedure, and viral mediated, liposome mediated and the like. Viral-mediated introduction of a polynucleotide encoding an RGN, crRNA, and/or tracrRNA includes retroviral, lentiviral, adenoviral, and adeno-associated viral mediated introduction and expression, as well as the use of Caulimoviruses, Geminiviruses, and RNA plant viruses.

Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of host cell (*e.g.*, monocot or dicot plant cell) targeted for transformation. Methods for transformation are known in the art and include those set forth in US Patent Nos: 8,575,425; 7,692,068; 8,802,934; 7,541,517; each of which is herein incorporated by reference.

See, also, Rakoczy-Trojanowska, M. (2002) *Cell Mol Biol Lett.* 7:849-858; Jones *et al.* (2005) *Plant Methods* 1:5; Rivera *et al.* (2012) *Physics of Life Reviews* 9:308-345; Bartlett *et al.* (2008) *Plant Methods* 4:1-12; Bates, G.W. (1999) *Methods in Molecular Biology* 111:359-366; Binns and Thomashow (1988) *Annual Reviews in Microbiology* 42:575-606; Christou, P. (1992) *The Plant Journal* 2:275-281; Christou, P. (1995) *Euphytica* 85:13-27; Tzfira *et al.* (2004) *TRENDS in Genetics* 20:375-383; Yao *et al.* (2006) *Journal of Experimental Botany* 57:3737-3746; Zupan and Zambryski (1995) *Plant Physiology* 107:1041-1047; Jones *et al.* (2005) *Plant Methods* 1:5;

Transformation may result in stable or transient incorporation of the nucleic acid into the cell. "Stable transformation" is intended to mean that the nucleotide construct introduced into a host cell integrates into the genome of the host cell and is capable of being inherited by the progeny thereof. "Transient transformation" is intended to mean that a polynucleotide is introduced into the host cell and does not integrate into the genome of the host cell.

Methods for transformation of chloroplasts are known in the art. See, for example, Svab et al. (1990) *Proc. Nail. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation can be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase. Such a system has been reported in McBride et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:7301-7305.

The cells that have been transformed may be grown into a transgenic organism, such as a plant, in accordance with conventional ways. See, for example, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides transformed seed (also referred to as "transgenic seed") having a nucleotide construct of the invention, for example, an expression cassette of the invention, stably incorporated into their genome.

Alternatively, cells that have been transformed may be introduced into an organism. These cells could have originated from the organism, wherein the cells are transformed in an *ex vivo* approach.

The sequences provided herein may be used for transformation of any plant species, including, but not limited to, monocots and dicots. Examples of plants of interest include, but are not limited to, corn (maize), sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, and oilseed rape, Brassica sp., alfalfa, rye, millet, safflower, peanuts, sweet potato, cassava, coffee, coconut, pineapple, citrus trees, cocoa, tea, banana, avocado, fig, guava, mango, olive, papaya, cashew, macadamia, almond, oats, vegetables, ornamentals, and conifers.

Vegetables include, but are not limited to, tomatoes, lettuce, green beans, lima beans, peas, and members of the genus Curcumis such as cucumber, cantaloupe, and musk melon. Ornamentals include, but are not limited to, azalea, hydrangea, hibiscus, roses, tulips, daffodils, petunias, carnation, poinsettia, and chrysanthemum. In specific embodiments, plants of the present invention are crop plants (for example, maize, sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, oilseed rape, etc.).

As used herein, the term plant includes plant cells, plant protoplasts, plant cell tissue cultures from which plants can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, and the like. Grain is intended to mean the mature seed produced by commercial growers for purposes other than growing or reproducing the species. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced polynucleotides. Further provided is a processed plant product or byproduct that retains the sequences disclosed herein, including for example, soymeal.

The polynucleotides encoding the RGNs, crRNAs, and/or tracrRNAs or comprising the crRNAs and/or tracrRNAs can also be used to transform any prokaryotic species, including but not limited to, archaea and bacteria (*e.g.*, *Bacillus* sp., *Klebsiella* sp. *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., Agrobacterium, *Lactobacillus* sp.).

The polynucleotides encoding the RGNs, crRNAs, and/or tracrRNAs or comprising the crRNAs and/or tracrRNAs can be used to transform any eukaryotic species, including but not limited to animals (*e.g.*, mammals, insects, fish, birds, and reptiles), fungi, amoeba, algae, and yeast.

Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids in mammalian, insect, or avian cells or target tissues. Such methods can be used to administer nucleic acids encoding components of an RGN system to cells in culture, or in a host organism. Non-viral vector delivery systems include DNA plasmids, RNA (e.g. a transcript of a vector described herein), naked nucleic acid, and nucleic acid complexed with a delivery vehicle, such as a liposome. Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. For a review of gene therapy procedures, see Anderson, Science 256: 808- 813 (1992); Nabel & Feigner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10): 1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada et al., in Current Topics in Microbiology and Immunology, Doerfler and Bohm (eds) (1995); and Yu et al., Gene Therapy 1:13-26 (1994).

Methods of non-viral delivery of nucleic acids include lipofection, nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid: nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat.

Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam ™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Feigner, WO 91/17424; WO 91/16024. Delivery can be to cells (e.g. in vitro or ex vivo administration) or target tissues (e.g. in vivo administration). The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (see, e.g., Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291- 297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787).

The use of RNA or DNA viral based systems for the delivery of nucleic acids takes advantage of highly evolved processes for targeting a virus to specific cells in the body and trafficking the viral payload to the nucleus. Viral vectors can be administered directly to patients (in vivo) or they can be used to treat cells in vitro, and the modified cells may optionally be administered to patients (ex vivo). Conventional viral based systems could include retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, often resulting in long term expression of the inserted transgene. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues.

The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and typically produce high viral titers. Selection of a retroviral gene transfer system would therefore depend on the target tissue. Retroviral vectors are comprised of cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof (see, e.g., Buchscher et al., J. Viral. 66:2731-2739 (1992); Johann et al., J. Viral. 66:1635-1640 (1992); Sommnerfelt et al., J. Viral. 176:58-59 (1990); Wilson et al., J. Viral. 63:2374-2378 (1989); Miller et al., J. Viral. 65:2220-2224 (1991); PCT/US94/05700).

In applications where transient expression is preferred, adenoviral based systems may be used. Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. With such vectors, high titer and levels of expression have been obtained. This vector can be produced in large quantities in a relatively simple system. Adeno-associated virus ("AAV") vectors may also be used to transduce cells with target nucleic acids, e.g., in the in vitro production of nucleic acids and peptides, and for in vivo and ex vivo gene therapy procedures (see, e.g., West et al., Virology 160:38-47

(1987); U.S. Pat. No. 4,797,368; WO 93/24641; Katin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). Construction of recombinant AAV vectors is described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); and Samulski et al., *1.* Viral. 63:03822-3828 (1989). Packaging cells are typically used to form virus particles that are capable of infecting a host cell. Such cells include 293 cells, which package adenovirus, and ψJ2 cells or PA317 cells, which package retrovirus.

Viral vectors used in gene therapy are usually generated by producing a cell line that packages a nucleic acid vector into a viral particle. The vectors typically contain the minimal viral sequences required for packaging and subsequent integration into a host, other viral sequences being replaced by an expression cassette for the polynucleotide(s) to be expressed. The missing viral functions are typically supplied in trans by the packaging cell line. For example, AAV vectors used in gene therapy typically only possess ITR sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely rep and cap, but lacking ITR sequences.

The cell line may also be infected with adenovirus as a helper. The helper virus promotes replication of the AAV vector and expression of AAV genes from the helper plasmid. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV. Additional methods for the delivery of nucleic acids to cells are known to those skilled in the art. See, for example, US20030087817, incorporated herein by reference.

In some embodiments, a host cell is transiently or non-transiently transfected with one or more vectors described herein. In some embodiments, a cell is transfected as it naturally occurs in a subject. In some embodiments, a cell that is transfected is taken from a subject. In some embodiments, the cell is derived from cells taken from a subject, such as a cell line. In some embodiments, the cell line may be mammalian, insect, or avian cells. A wide variety of cell lines for tissue culture are known in the art. Examples of cell lines include, but are not limited to, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLaS3, Huhl, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panel, PC-3, TFl, CTLL-2, CIR, Rat6, CVI, RPTE, AlO, T24, 182, A375, ARH-77, Calul, SW480, SW620, SKOV3, SK-UT, CaCo2, P388Dl, SEM-K2, WEHI- 231, HB56, TIB55, lurkat, *145.01,* LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4. COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-Ll, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-I cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10Tl/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-Kl, CHO-K2, CHO-T, CHO Dhfr-/-, COR-L23, COR-L23/CPR, COR-L235010, CORL23/ R23, COS-7, COV-434, CML Tl, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293,

HeLa, Hepalclc7, HL-60, HMEC, HT-29, lurkat, *lY* cells, K562 cells, Ku812, KCL22, KGl, KYOl, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-l0A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCKII, MDCKII, MOR/ 0.2R, MONO-MAC 6, MTD-lA, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN/OPCT cell lines, Peer, PNT-lA/ PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THPl cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassas, Va.)).

In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the components of an RGN system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of an RGN system, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In some embodiments, cells transiently or non-transiently transfected with one or more vectors described herein, or cell lines derived from such cells are used in assessing one or more test compounds.

In some embodiments, one or more vectors described herein are used to produce a non-human transgenic animal or transgenic plant. In some embodiments, the transgenic animal is a mammal, such as a mouse, rat, hamster, rabbit, cow, or pig. In some embodiments, the transgenic animal is a bird, such as a chicken or a duck. In some embodiments, the transgenic animal is an insect, such as a mosquito or a tick.

V. Variants and Fragments of Polypeptides and Polynucleotides

The present disclosure provides active variants and fragments of a naturally-occurring (*i.e.*, wild-type) RNA-guided nuclease, the amino acid sequence of which is set forth as any one of SEQ ID NOs: 1-20, as well as active variants and fragments of naturally-occurring CRISPR repeats, such as any one of the sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and active variant and fragments of naturally-occurring tracrRNAs, such as any one of the sequences set forth as SEQ ID NOs: 42-62, , nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045and polynucleotides encoding the same.

While the activity of a variant or fragment may be altered compared to the polynucleotide or polypeptide of interest, the variant and fragment should retain the functionality of the polynucleotide or polypeptide of interest. For example, a variant or fragment may have increased activity, decreased activity, different spectrum of activity or any other alteration in activity when compared to the polynucleotide or polypeptide of interest.

Fragments and variants of naturally-occurring RGN polypeptides, such as those disclosed herein, will retain sequence-specific, RNA-guided DNA-binding activity. In particular embodiments, fragments and

variants of naturally-occurring RGN polypeptides, such as those disclosed herein, will retain nuclease activity (single-stranded or double-stranded).

Fragments and variants of naturally-occurring CRISPR repeats, such as those disclosed herein, will retain the ability, when part of a guide RNA (comprising a tracrRNA), to bind to and guide an RNA-guided nuclease (complexed with the guide RNA) to a target sequence (e.g., target DNA sequence) in a sequence-specific manner.

Fragments and variants of naturally-occurring tracrRNAs, such as those disclosed herein, will retain the ability, when part of a guide RNA (comprising a CRISPR RNA), to guide an RNA-guided nuclease (complexed with the guide RNA) to a target sequence (e.g., target DNA sequence) in a sequence-specific manner.

The term “fragment” refers to a portion of a polynucleotide or polypeptide sequence of the invention. "Fragments" or "biologically active portions" include polynucleotides comprising a sufficient number of contiguous nucleotides to retain the biological activity (*i.e.*, binding to and directing an RGN in a sequence-specific manner to a target nucleotide sequence when comprised within a guide RNA). "Fragments" or "biologically active portions" include polypeptides comprising a sufficient number of contiguous amino acid residues to retain the biological activity (*i.e.*, binding to a target sequence in a sequence-specific manner when complexed with a guide RNA). Fragments of the RGN proteins include those that are shorter than the full-length sequences due to the use of an alternate downstream start site. A biologically active portion of an RGN protein can be a polypeptide that comprises, for example, 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, or more contiguous amino acid residues of any one of SEQ ID NOs: 1-20. Such biologically active portions can be prepared by recombinant techniques and evaluated for sequence-specific, RNA-guided DNA-binding activity. A biologically active fragment of a CRISPR repeat sequence can comprise at least 8 contiguous amino acids of any one of SEQ ID NOs:21-41, or nucleotides 1-14 of SEQ ID NO: 1040, nucleotides 1-17 of SEQ ID NO: 1041 or 1042, nucleotides 1-19 of SEQ ID NO: 1043, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045. A biologically active portion of a CRISPR repeat sequence can be a polynucleotide that comprises, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 21 contiguous nucleotides of any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045. A biologically active portion of a tracrRNA can be a polynucleotide that comprises, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more contiguous nucleotides of any one of SEQ ID NOs:42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

In general, "variants" is intended to mean substantially similar sequences. For polynucleotides, a variant comprises a deletion and/or addition of one or more nucleotides at one or more internal sites within the native polynucleotide and/or a substitution of one or more nucleotides at one or more sites in the native

polynucleotide. As used herein, a "native" or "wild type" polynucleotide or polypeptide comprises a naturally occurring nucleotide sequence or amino acid sequence, respectively. For polynucleotides, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the native amino acid sequence of the gene of interest. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant polynucleotides also include synthetically derived polynucleotides, such as those generated, for example, by using site-directed mutagenesis but which still encode the polypeptide or the polynucleotide of interest. Generally, variants of a particular polynucleotide disclosed herein will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

Variants of a particular polynucleotide disclosed herein (*i.e.*, the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides disclosed herein is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity.

In particular embodiments, the presently disclosed polynucleotides encode an RNA-guided nuclease polypeptide comprising an amino acid sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater identity to any one of the amino acid sequences set forth as SEQ ID NOs: 1-20.

A biologically active variant of an RGN polypeptide of the invention may differ by as few as about 1-15 amino acid residues, as few as about 1-10, such as about 6-10, as few as 5, as few as 4, as few as 3, as few as 2, or as few as 1 amino acid residue. In specific embodiments, the polypeptides can comprise an N-terminal or a C-terminal truncation, which can comprise at least a deletion of 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350 amino acids or more from either the N or C terminus of the polypeptide.

In certain embodiments, the presently disclosed polynucleotides comprise or encode a CRISPR repeat comprising a nucleotide sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater identity to any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

The presently disclosed polynucleotides can comprise or encode a tracrRNA comprising a nucleotide sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater identity to any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

Biologically active variants of a CRISPR repeat or tracrRNA of the invention may differ by as few as about 1-15 nucleotides, as few as about 1-10, such as about 6-10, as few as 5, as few as 4, as few as 3, as few as 2, or as few as 1 nucleotide. In specific embodiments, the polynucleotides can comprise a 5' or 3' truncation, which can comprise at least a deletion of 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 95, 100, 105, 110 nucleotides or more from either the 5' or 3' end of the polynucleotide.

It is recognized that modifications may be made to the RGN polypeptides, CRISPR repeats, and tracrRNAs provided herein creating variant proteins and polynucleotides. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques. Alternatively, native, as yet unknown or as yet unidentified polynucleotides and/or polypeptides structurally and/or functionally related to the sequences disclosed herein may also be identified that fall within the scope of the present invention. Conservative amino acid substitutions may be made in nonconserved regions that do not alter the function of the RGN proteins. Alternatively, modifications may be made that improve the activity of the RGN.

Variant polynucleotides and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different RGN proteins disclosed herein (e.g., SEQ ID NOs: 1-20) is manipulated to create a new RGN protein possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the RGN sequences provided herein and other known RGN genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased K_m in the case of an enzyme. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Crameri *et al.* (1997) *Nature Biotech.* 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al. (1997) Proc. Natl. Acad. Sci. USA* 94:4504-4509; Crameri *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458. A "shuffled" nucleic acid is a nucleic acid produced by a shuffling procedure such as any shuffling procedure set forth herein. Shuffled nucleic acids are produced by recombining (physically or virtually) two or more nucleic acids (or character strings), for example in an artificial, and optionally recursive, fashion. Generally, one or more screening steps are used in shuffling processes to identify nucleic acids of interest; this screening step can be performed before or after any recombination step. In some (but not all) shuffling embodiments, it is desirable to perform multiple rounds of recombination prior to selection to increase the diversity of the pool

to be screened. The overall process of recombination and selection are optionally repeated recursively. Depending on context, shuffling can refer to an overall process of recombination and selection, or, alternately, can simply refer to the recombinational portions of the overall process.

As used herein, "sequence identity" or "identity" in the context of two polynucleotides or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (*e.g.*, charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well known to those of skill in the art. Typically, this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 using the following parameters: % identity and % similarity for a nucleotide sequence using GAP Weight of 50 and Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix; or any equivalent program thereof. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

Two sequences are "optimally aligned" when they are aligned for similarity scoring using a defined amino acid substitution matrix (*e.g.*, BLOSUM62), gap existence penalty and gap extension penalty so as to arrive at the highest score possible for that pair of sequences. Amino acid substitution matrices and their use

in quantifying the similarity between two sequences are well-known in the art and described, *e.g.*, in Dayhoff et al. (1978) "A model of evolutionary change in proteins." In "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 3 (ed. M. O. Dayhoff), pp. 345-352. Natl. Biomed. Res. Found., Washington, D.C. and Henikoff et al. (1992) Proc. Natl. Acad. Sci. USA 89:10915-10919. The BLOSUM62 matrix is often used as a default scoring substitution matrix in sequence alignment protocols. The gap existence penalty is imposed for the introduction of a single amino acid gap in one of the aligned sequences, and the gap extension penalty is imposed for each additional empty amino acid position inserted into an already opened gap. The alignment is defined by the amino acid positions of each sequence at which the alignment begins and ends, and optionally by the insertion of a gap or multiple gaps in one or both sequences, so as to arrive at the highest possible score. While optimal alignment and scoring can be accomplished manually, the process is facilitated by the use of a computer-implemented alignment algorithm, *e.g.*, gapped BLAST 2.0, described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402, and made available to the public at the National Center for Biotechnology Information Website (www.ncbi.nlm.nih.gov). Optimal alignments, including multiple alignments, can be prepared using, *e.g.*, PSI-BLAST, available through www.ncbi.nlm.nih.gov and described by Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402.

With respect to an amino acid sequence that is optimally aligned with a reference sequence, an amino acid residue "corresponds to" the position in the reference sequence with which the residue is paired in the alignment. The "position" is denoted by a number that sequentially identifies each amino acid in the reference sequence based on its position relative to the N-terminus. Owing to deletions, insertion, truncations, fusions, etc., that must be taken into account when determining an optimal alignment, in general the amino acid residue number in a test sequence as determined by simply counting from the N-terminal will not necessarily be the same as the number of its corresponding position in the reference sequence. For example, in a case where there is a deletion in an aligned test sequence, there will be no amino acid that corresponds to a position in the reference sequence at the site of deletion. Where there is an insertion in an aligned reference sequence, that insertion will not correspond to any amino acid position in the reference sequence. In the case of truncations or fusions there can be stretches of amino acids in either the reference or aligned sequence that do not correspond to any amino acid in the corresponding sequence.

VI. Antibodies

Antibodies to the RGN polypeptides or ribonucleoproteins comprising the RGN polypeptides of the present invention, including those having any one of the amino acid sequences set forth as SEQ ID NOs: 1-20 or active variants or fragments thereof, are also encompassed. Methods for producing antibodies are well known in the art (see, for example, Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; and U.S. Pat. No. 4,196,265). These antibodies can be used in kits for the detection and isolation of RGN polypeptides or ribonucleoproteins. Thus, this disclosure provides kits comprising antibodies that specifically bind to the polypeptides or ribonucleoproteins described

herein, including, for example, polypeptides having any one of the amino acid sequences set forth as SEQ ID NOs: 1-20.

VII. RGN Systems and Ribonucleoprotein Complexes for Binding a Target Sequence of Interest and Methods of Making the Same

The present disclosure provides a system for binding a target sequence (e.g., target DNA sequence) of interest, wherein the system comprises at least one RNA-guided nuclease or a nucleotide sequence encoding the same and one or more guide RNAs capable of forming a complex with the RGN polypeptide (ribonucleoprotein complex). The guide RNA hybridizes to the non-target strand of a target sequence of interest and also forms a complex with the RGN polypeptide, thereby directing the RGN polypeptide to bind to the target DNA sequence. In some of these embodiments, the RGN comprises any one of the amino acid sequences set forth as SEQ ID NOs: 1-20, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat having any one of the nucleotide sequences set forth as SEQ ID NOs:21-41, or nucleotides 1-14 of SEQ ID NO: 1040, nucleotides 1-17 of SEQ ID NO: 1041 or 1042, nucleotides 1-19 of SEQ ID NO: 1043, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA having any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA. In particular embodiments, the system comprises an RNA-guided nuclease that is heterologous to the guide RNA, wherein the RGN and guide RNA are not found complexed to one another (i.e., bound to one another) in nature.

The system for binding a target sequence of interest provided herein can be a ribonucleoprotein complex, which is at least one molecule of an RNA bound to at least one protein. The ribonucleoprotein complexes provided herein comprise at least one guide RNA as the RNA component and an RNA-guided nuclease as the protein component. Such ribonucleoprotein complexes can be purified from a cell or organism that naturally expresses an RGN polypeptide and has been engineered to express a particular guide RNA that is specific for a target sequence of interest. Alternatively, the ribonucleoprotein complex can be purified from a cell or organism that has been transformed with polynucleotides that encode an RGN polypeptide and a guide RNA (or a polynucleotide that comprises a guide RNA) and cultured under conditions to allow for the expression of the RGN polypeptide and guide RNA. Thus, methods are provided for making an RGN polypeptide or an RGN ribonucleoprotein complex. Such methods comprise culturing a cell comprising a nucleotide sequence encoding an RGN polypeptide, and in some embodiments a nucleotide sequence encoding or comprising a guide RNA, under conditions in which the RGN polypeptide (and in some embodiments, the guide RNA) is expressed. The RGN polypeptide or RGN ribonucleoprotein can then be purified from a lysate of the cultured cells. In embodiments, the nucleotide sequence encoding an RGN polypeptide includes a mRNA (messenger RNA). In some embodiments, methods for assembling

an RNP complex comprise combining one or more of the presently disclosed guide RNAs and one or more of the presently disclosed RGN polypeptides under conditions suitable for formation of the RNP complex.

Methods for purifying an RGN polypeptide or RGN ribonucleoprotein complex from a lysate of a biological sample are known in the art (*e.g.*, size exclusion and/or affinity chromatography, 2D-PAGE, HPLC, reversed-phase chromatography, immunoprecipitation). In particular methods, the RGN polypeptide is recombinantly produced and comprises a purification tag to aid in its purification, including but not limited to, glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein, thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG (e.g., 3X FLAG tag), HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, 10xHis, biotin carboxyl carrier protein (BCCP), and calmodulin. Generally, the tagged RGN polypeptide or RGN ribonucleoprotein complex is purified using immobilized metal affinity chromatography. It will be appreciated that other similar methods known in the art may be used, including other forms of chromatography or for example immunoprecipitation, either alone or in combination.

An "isolated" or "purified" polypeptide, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polypeptide as found in its naturally occurring environment. Thus, an isolated or purified polypeptide is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, optimally culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals. Similarly, an “isolated” polynucleotide or nucleic acid molecule is removed from its naturally occurring environment. An isolated polynucleotide is substantially free of chemical precursors or other chemicals when chemically synthesized or has been removed from a genomic locus via the breaking of phosphodiester bonds. An isolated polynucleotide can be part of a vector, a composition of matter or can be contained within a cell so long as the cell is not the original environment of the polynucleotide.

Particular methods provided herein for binding and/or cleaving a target nucleic acid molecule comprising a target sequence of interest involve the use of an *in vitro* assembled RGN ribonucleoprotein complex. *In vitro* assembly of an RGN ribonucleoprotein complex can be performed using any method known in the art in which an RGN polypeptide is contacted with a guide RNA under conditions to allow for binding of the RGN polypeptide to the guide RNA. As used herein, "contact", contacting", "contacted," refer to placing the components of a desired reaction together under conditions suitable for carrying out the desired reaction. The RGN polypeptide can be purified from a biological sample, cell lysate, or culture medium, produced via *in vitro* translation, or chemically synthesized. The guide RNA can be purified from a biological sample, cell lysate, or culture medium, transcribed *in vitro*, or chemically synthesized. The

RGN polypeptide and guide RNA can be brought into contact in solution (*e.g.*, buffered saline solution) to allow for *in vitro* assembly of the RGN ribonucleoprotein complex.

VIII. Methods of Binding, Cleaving, or Modifying a Target Nucleic Acid Molecule

The present disclosure provides methods for binding, cleaving, and/or modifying a target nucleic acid molecule (e.g., target DNA) of interest comprising a target sequence. The methods include delivering a system comprising at least one guide RNA or a polynucleotide encoding the same, and at least one RGN polypeptide or a polynucleotide encoding the same to the target sequence or a cell, organelle, or embryo comprising the target sequence. In some of these embodiments, the RGN comprises any one of the amino acid sequences set forth as SEQ ID NOs:1-20, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat comprising any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA comprising any one of the nucleotide sequences set forth as SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA.

The RGN of the system may be a nuclease dead RGN, have nickase activity, or may be a fusion polypeptide. In some embodiments, the fusion polypeptide comprises a base-editing polypeptide, for example a cytosine deaminase or an adenine deaminase. In other embodiments, the RGN fusion protein comprises a reverse transcriptase. In other embodiments, the RGN fusion protein comprises a polypeptide that recruits members of a functional nucleic acid repair complex, such as a member of the nucleotide excision repair (NER) or transcription coupled-nucleotide excision repair (TC-NER) pathway (Wei *et al.*, 2015, *PNAS USA* 112(27):E3495-504 ; Troelstra *et al.*, 1992, *Cell* 71:939-953; Marnef *et al.*, 2017, *J Mol Biol* 429(9):1277-1288), as described in U.S. Provisional Application No. 63/332,486, which was filed on April 19, 2022, and is incorporated by reference in its entirety. In some embodiments, the RGN fusion protein comprises CSB (van den Boom *et al.*, 2004, *J Cell Biol* 166(1):27-36; van Gool *et al.*, 1997, *EMBO J* 16(19):5955-65; an example of which is set forth as SEQ ID NO: 563), which is a member of the TC-NER (nucleotide excision repair) pathway and functions in the recruitment of other members. In further embodiments, the RGN fusion protein comprises an active domain of CSB, such as the acidic domain of CSB which comprises amino acid residues 356-394 of SEQ ID NO: 563 (Teng *et al.*, 2018, *Nat Commun* 9(1):4115).

In particular embodiments, the RGN and/or guide RNA is heterologous to the cell, organelle, or embryo to which the RGN and/or guide RNA (or polynucleotide(s) encoding at least one of the RGN and guide RNA) are introduced.

In those embodiments wherein the method comprises delivering a polynucleotide encoding a guide RNA and/or an RGN polypeptide, the cell or embryo can then be cultured under conditions in which the guide RNA and/or RGN polypeptide are expressed. In various embodiments, the method comprises contacting a target nucleic acid molecule with an RGN ribonucleoprotein complex. The RGN ribonucleoprotein complex may comprise an RGN that is nuclease dead or has nickase activity. In some embodiments, the RGN of the ribonucleoprotein complex is a fusion polypeptide comprising a base-editing polypeptide. In certain embodiments, the method comprises introducing into a cell, organelle, or embryo comprising a target nucleic acid molecule an RGN ribonucleoprotein complex. The RGN ribonucleoprotein complex can be one that has been purified from a biological sample, recombinantly produced and subsequently purified, or *in vitro*-assembled as described herein. In those embodiments wherein the RGN ribonucleoprotein complex that is contacted with the target nucleic acid molecule or a cell organelle, or embryo has been assembled *in vitro*, the method can further comprise the *in vitro* assembly of the complex prior to contact with the target nucleic acid molecule, cell, organelle, or embryo.

A purified or *in vitro* assembled RGN ribonucleoprotein complex can be introduced into a cell, organelle, or embryo using any method known in the art, including, but not limited to electroporation. Alternatively, an RGN polypeptide and/or polynucleotide encoding or comprising the guide RNA can be introduced into a cell, organelle, or embryo using any method known in the art (*e.g.*, electroporation).

Upon delivery to or contact with the target nucleic acid molecule or cell, organelle, or embryo comprising the target nucleic acid molecule, the guide RNA directs the RGN to bind to the target sequence within the target nucleic acid molecule in a sequence-specific manner. In those embodiments wherein the RGN has nuclease activity, the RGN polypeptide cleaves the target sequence of interest upon binding. The target DNA sequence can subsequently be modified via endogenous repair mechanisms, such as non-homologous end joining, or homology-directed repair with a provided donor polynucleotide.

Methods to measure binding of an RGN polypeptide to a target sequence are known in the art and include chromatin immunoprecipitation assays, gel mobility shift assays, DNA pull-down assays, reporter assays, microplate capture and detection assays. Likewise, methods to measure cleavage or modification of a target nucleic acid molecule comprising a target sequence are known in the art and include *in vitro* or *in vivo* cleavage assays wherein cleavage is confirmed using PCR, sequencing, or gel electrophoresis, with or without the attachment of an appropriate label (*e.g.*, radioisotope, fluorescent substance) to the target sequence to facilitate detection of degradation products. Alternatively, the nicking triggered exponential amplification reaction (NTEXPAR) assay can be used (see, *e.g.*, Zhang et al. (2016) *Chem. Sci.* 7:4951-4957). *In vivo* cleavage can be evaluated using the Surveyor assay (Guschin et al. (2010) *Methods Mol Biol* 649:247-256).

In some embodiments, the methods involve the use of a single type of RGN complexed with more than one guide RNA. The more than one guide RNA can target different regions of a single gene or can target multiple genes.

In those embodiments wherein a donor polynucleotide is not provided, a double-stranded break introduced by an RGN polypeptide can be repaired by a non-homologous end-joining (NHEJ) repair process. Due to the error-prone nature of NHEJ, repair of the double-stranded break can result in a modification to the target sequence. As used herein, a "modification" in reference to a nucleic acid molecule refers to a change in the nucleotide sequence of the nucleic acid molecule, which can be a deletion, insertion, or substitution of one or more nucleotides, or a combination thereof. Modification of the target nucleic acid molecule comprising a target sequence can result in the expression of an altered protein product or inactivation of a coding sequence.

In those embodiments wherein a donor polynucleotide is present, the donor sequence in the donor polynucleotide can be integrated into or exchanged with the target nucleotide sequence during the course of repair of the introduced double-stranded break, resulting in the introduction of the exogenous donor sequence. A donor polynucleotide thus comprises a donor sequence that is desired to be introduced into a target sequence of interest. In some embodiments, the donor sequence alters the original target nucleotide sequence such that the newly integrated donor sequence will not be recognized and cleaved by the RGN. Integration of the donor sequence can be enhanced by the inclusion within the donor polynucleotide of flanking sequences, referred to herein as "homology arms" that have substantial sequence identity with the sequences flanking the target nucleotide sequence, allowing for a homology-directed repair process. In some embodiments, homology arms have a length of at least 50 base pairs, at least 100 base pairs, and up to 2000 base pairs or more, and have at least 90%, at least 95%, or more, sequence homology to their corresponding sequence within the target nucleotide sequence.

In those embodiments wherein the RGN polypeptide introduces double-stranded staggered breaks, the donor polynucleotide can comprise a donor sequence flanked by compatible overhangs, allowing for direct ligation of the donor sequence to the cleaved target nucleotide sequence comprising overhangs by a non-homologous repair process during repair of the double-stranded break.

In those embodiments wherein the method involves the use of an RGN that is a nickase (*i.e.*, is only able to cleave a single strand of a double-stranded polynucleotide), the method can comprise introducing two RGN nickases that target identical or overlapping target sequences and cleave different strands of the polynucleotide. For example, an RGN nickase that only cleaves the positive (+) strand of a double-stranded polynucleotide can be introduced along with a second RGN nickase that only cleaves the negative (-) strand of a double-stranded polynucleotide.

In various embodiments, a method is provided for binding a target nucleotide sequence and detecting the target sequence, wherein the method comprises introducing into a cell, organelle, or embryo at least one guide RNA or a polynucleotide encoding the same, and at least one RGN polypeptide or a polynucleotide encoding the same, expressing the guide RNA and/or RGN polypeptide (if coding sequences are introduced), wherein the RGN polypeptide is a nuclease-dead RGN and further comprises a detectable label, and the method further comprises detecting the detectable label. The detectable label may be fused to

the RGN as a fusion protein (*e.g.*, fluorescent protein) or may be a small molecule conjugated to or incorporated within the RGN polypeptide that can be detected visually or by other means.

Also provided herein are methods for modulating the expression of a target gene of interest comprising a target sequence or a gene under the regulation of a target sequence. The methods comprise introducing into a cell, organelle, or embryo at least one guide RNA or a polynucleotide encoding the same, and at least one RGN polypeptide or a polynucleotide encoding the same, expressing the guide RNA and/or RGN polypeptide (if coding sequences are introduced), wherein the RGN polypeptide is a nuclease-dead RGN. In some of these embodiments, the nuclease-dead RGN is a fusion protein comprising an expression modulator domain (*i.e.*, epigenetic modification domain, transcriptional activation domain or a transcriptional repressor domain) as described herein.

The present disclosure also provides methods for binding and/or modifying a target nucleic acid molecule of interest comprising a target sequence. The methods include delivering a system comprising at least one guide RNA or a polynucleotide encoding the same, and at least one fusion polypeptide comprising an RGN of the invention and a base-editing polypeptide, for example a cytosine deaminase or an adenine deaminase, or a polynucleotide encoding the fusion polypeptide, to the target sequence or a cell, organelle, or embryo comprising the target sequence.

In some embodiments wherein a fusion polypeptide comprising an RGN and a base-editing polypeptide is utilized, the binding of the fusion protein to a target sequence results in the modification of nucleotide(s) adjacent to the target sequence. The nucleobase adjacent to the target sequence that is modified by the deaminase may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 base pairs from the 5' or 3' end of the target sequence.

One of ordinary skill in the art will appreciate that any of the presently disclosed methods can be used to target a single target sequence or multiple target sequences. Thus, methods comprise the use of a single RGN polypeptide in combination with multiple, distinct guide RNAs, which can target multiple, distinct sequences within a single gene and/or multiple genes. Also encompassed herein are methods wherein multiple, distinct guide RNAs are introduced in combination with multiple, distinct RGN polypeptides. These guide RNAs and guide RNA/RGN polypeptide systems can target multiple, distinct sequences within a single gene and/or multiple genes.

In one aspect, the invention provides kits containing any one or more of the elements disclosed in the above methods and compositions, including crRNAs, tracrRNAs, guide RNAs, RGNs, and/or polynucleotides encoding the same, cells, and RGN systems. In some embodiments, the kit comprises a vector system and instructions for using the kit. In some embodiments, the vector system comprises (a) a first regulatory element operably linked to a DNA sequence encoding a guide RNA and one or more insertion sites for inserting a guide sequence upstream of the encoded guide RNA, wherein when expressed, the guide RNA directs sequence-specific binding of an RGN complex to a target sequence in a eukaryotic cell, wherein the RGN complex comprises an RGN enzyme complexed with the guide RNA polynucleotide; and/or (b) a second regulatory element operably linked to an enzyme coding sequence encoding said RGN

enzyme comprising a nuclear localization sequence. In some embodiments, the kit further comprises a homologous recombination template polynucleotide. Elements may be provided individually or in combinations, and may be provided in any suitable container, such as a vial, a bottle, or a tube.

In some embodiments, the kit includes instructions in one or more languages. In some embodiments, a kit comprises one or more reagents for use in a process utilizing one or more of the elements described herein. Reagents may be provided in any suitable container. For example, a kit may provide one or more reaction or storage buffers. Reagents may be provided in a form that is usable in a particular assay, or in a form that requires addition of one or more other components before use (e.g. in concentrate or lyophilized form). A buffer can be any buffer, including but not limited to a sodium carbonate buffer, a sodium bicarbonate buffer, a borate buffer, a Tris buffer, a MOPS buffer, a HEPES buffer, and combinations thereof. In some embodiments, the buffer is alkaline. In some embodiments, the buffer has a pH from about 7 to about 10.

In one aspect, the invention provides methods for using one or more elements of an RGN system. The RGN system of the invention provides an effective means for modifying a target polynucleotide. The RGN system of the invention has a wide variety of utility including modifying (e.g., deleting, inserting, translocating, inactivating, activating, base editing) a target polynucleotide in a multiplicity of cell types. As such the RGN system of the invention has a broad spectrum of applications in, e.g., gene therapy, drug screening, disease diagnosis, and prognosis. An exemplary RGN system, or RGN complex, comprises an RGN enzyme complexed with a guide sequence capable of binding to a target sequence.

IX. Target Polynucleotides

In one aspect, the invention provides for methods of modifying a target polynucleotide comprising a target sequence or modifying the expression of a target polynucleotide in a eukaryotic cell, which may be in vivo, ex vivo or in vitro. In some embodiments, the method comprises sampling a cell or population of cells from a human or non-human animal or plant (including microalgae) and modifying the cell or cells. Culturing may occur at any stage ex vivo. The cell or cells may even be re-introduced into the non-human animal or plant (including micro-algae).

Using natural variability, plant breeders combine most useful genes for desirable qualities, such as yield, quality, uniformity, hardiness, and resistance against pests. These desirable qualities also include growth, day length preferences, temperature requirements, initiation date of floral or reproductive development, fatty acid content, insect resistance, disease resistance, nematode resistance, fungal resistance, herbicide resistance, tolerance to various environmental factors including drought, heat, wet, cold, wind, and adverse soil conditions including high salinity The sources of these useful genes include native or foreign varieties, heirloom varieties, wild plant relatives, and induced mutations, e.g., treating plant material with mutagenic agents. Using the present invention, plant breeders are provided with a new tool to induce mutations. Accordingly, one skilled in the art can analyze the genome for sources of useful genes, and in varieties having desired characteristics or traits employ the present invention to induce the rise of useful

genes, with more precision than previous mutagenic agents and hence accelerate and improve plant breeding programs.

The target polynucleotide of an RGN system can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA). Without wishing to be bound by theory, the target strand of the target sequence should be adjacent to a PAM (protospacer adjacent motif); that is, a short sequence recognized by the RGN system. The precise sequence and length requirements for the PAM differ depending on the RGN used, but PAMs are typically 2-7 base pair sequences adjacent to the protospacer (that is, the target sequence).

The target polynucleotide of an RGN system may include a number of disease-associated genes and polynucleotides as well as signaling biochemical pathway-associated genes and polynucleotides. Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease associated gene or polynucleotide. A "disease-associated" gene or polynucleotide refers to any gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated gene also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease (e.g., a causal mutation). The transcribed or translated products may be known or unknown, and further may be at a normal or abnormal level. In some embodiments, the disease may be an animal disease. In some embodiments, the disease may be an avian disease. In other embodiments, the disease may be a mammalian disease. In further embodiments, the disease may be a human disease. Examples of disease-associated genes and polynucleotides in humans are available from McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), available on the World Wide Web.

Although RGN systems are particularly useful for their relative ease in targeting to genomic sequences of interest, there still remains an issue of what the RGN can do to address a causal mutation. One approach is to produce a fusion protein between an RGN (e.g., an inactive or nickase variant of the RGN) and a base-editing enzyme or the active domain of a base editing enzyme, such as a cytosine deaminase or an adenine deaminase base editor (U.S. Patent No. 9,840,699, herein incorporated by reference). In some embodiments, the methods comprise contacting a DNA molecule comprising a target sequence with (a) a fusion protein comprising an RGN of the invention or a nickase variant thereof and a base-editing polypeptide such as a deaminase; and (b) a gRNA targeting the fusion protein of (a) to the target sequence;

wherein the DNA molecule is contacted with the fusion protein and the gRNA in an amount effective and under conditions suitable for the deamination of a nucleobase. In some embodiments, the target DNA sequence comprises a sequence associated with a disease or disorder, and wherein the deamination of the nucleobase results in a sequence that is not associated with a disease or disorder. In some embodiments, the target DNA sequence resides in an allele of a crop plant, wherein the particular allele of the trait of interest results in a plant of lesser agronomic value. The deamination of the nucleobase results in an allele that improves the trait and increases the agronomic value of the plant.

In some embodiments, the target DNA sequence comprises a T→C or A→G point mutation associated with a disease or disorder, and wherein the deamination of the mutant C or G base results in a sequence that is not associated with a disease or disorder. In some embodiments, the deamination corrects a point mutation in the sequence associated with the disease or disorder.

In some embodiments, the sequence associated with the disease or disorder encodes a protein, and wherein the deamination introduces a stop codon into the sequence associated with the disease or disorder, resulting in a truncation of the encoded protein. In some embodiments, the contacting is performed *in vivo* in a subject susceptible to having, having, or diagnosed with the disease or disorder. In some embodiments, the disease or disorder is a disease associated with a point mutation, or a single-base mutation, in the genome. In some embodiments, the disease is a genetic disease, a cancer, a metabolic disease, or a lysosomal storage disease.

X. Pharmaceutical Compositions and Methods of Treatment

Pharmaceutical compositions comprising the presently disclosed RGN polypeptides and active variants and fragments thereof, as well as polynucleotides encoding the same, the presently disclosed crRNAs and active varants and fragments thereof or polynucleotides encoding the same, the presently disclosed tracrRNAs and active variants and fragments thereof or polynucleotides encoding the same, the presently disclosed gRNAs or polynucleotides encoding the same, the presently disclosed systems, or cells comprising any of the RGN polypeptides or RGN-encoding polynucleotides, gRNA or gRNA-encoding polynucleotides, or the RGN systems, and a pharmaceutically acceptable carrier are provided.

A pharmaceutical composition is a composition that is employed to prevent, reduce in intensity, cure or otherwise treat a target condition or disease that comprises an active ingredient (i.e., RGN polypeptides, RGN-encoding polynucleotides, gRNA, gRNA-encoding polynucleotides, RGN systems, or cells comprising any one of these) and a pharmaceutically acceptable carrier.

As used herein, a "pharmaceutically acceptable carrier" refers to a material that does not cause significant irritation to an organism and does not abrogate the activity and properties of the active ingredient (i.e., RGN polypeptides, RGN-encoding polynucleotides, gRNA, gRNA-encoding polynucleotides, RGN systems, or cells comprising any one of these). Carriers must be of sufficiently high purity and of sufficiently low toxicity to render them suitable for administration to a subject being treated. The carrier can be inert, or it can possess pharmaceutical benefits. In some embodiments, a pharmaceutically acceptable

carrier comprises one or more compatible solid or liquid filler, diluents or encapsulating substances which are suitable for administration to a human or other vertebrate animal. In some embodiments, the pharmaceutically acceptable carrier is not naturally-occurring. In some embodiments, the pharmaceutically acceptable carrier and the active ingredient are not found together in nature.

Pharmaceutical compositions used in the presently disclosed methods can be formulated with suitable carriers, excipients, and other agents that provide suitable transfer, delivery, tolerance, and the like. A multitude of appropriate formulations are known to those skilled in the art. See, *e.g.*, Remington, The Science and Practice of Pharmacy (21st ed. 2005). Suitable formulations include, for example, powders, pastes, ointments, jellies, waxes, oils, lipids, lipid (cationic or anionic) containing vesicles (such as LIPOFECTIN vesicles), lipid nanoparticles, DNA conjugates, anhydrous absorption pastes, oil-in-water and water-in-oil emulsions, emulsions carbowax (polyethylene glycols of various molecular weights), semi-solid gels, and semi-solid mixtures containing carbowax. Pharmaceutical compositions for oral or parenteral use may be prepared into dosage forms in a unit dose suited to fit a dose of the active ingredients. Such dosage forms in a unit dose include, for example, tablets, pills, capsules, injections (ampoules), suppositories, etc.

The disclosure provides for pharmaceutical compositions comprising lipid-based formulations including an active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such). In some embodiments, the lipid-based formulations include liposomes. In some embodiments, the lipid-based formulations include lipid nanoparticles (LNPs). In some embodiments, an active ingredient is encapsulated in the lipid particle and/or disposed on the surface of the lipid particle. In some embodiments, an active ingredient is covalently attached to the lipid particle. In some embodiments, an active ingredient is non-covalently associated with the lipid particle. A covalent attachment includes the sharing of electrons in a chemical bond. Non-covalent interactions include dispersed electromagnetic interactions such as hydrogen bonds, ionic bonds, van der Waals interactions, and hydrophobic bonds.

In some embodiments, an active ingredient is encapsulated in the lipid particle. The term "encapsulate" means to enclose, surround or encase. As it relates to the formulation of the compounds of the disclosure, encapsulation may be substantial, complete or partial. The term "substantially encapsulated" or "substantial encapsulation" means that greater than 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.9%, or greater of the pharmaceutical composition or active ingredient of the disclosure may be enclosed, surrounded, or encased within a delivery agent (e.g., liposome or LNP). The term "partially encapsulated" or "partial encapsulation" means that less than 50%, 40%, 30%, 20%, 10%, or less of the pharmaceutical composition or active ingredient of the disclosure may be enclosed, surrounded, or encased within the delivery agent. Encapsulation may be determined by measuring the escape or the activity of the pharmaceutical composition or active ingredient of the disclosure using fluorescence and/or electron microscopy. For example, at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.9%, or greater of the pharmaceutical composition or active ingredient of the disclosure is encapsulated in a delivery agent (e.g., liposome or LNP).

Liposomes are spherical vesicle structures composed of a uni- or multilamellar lipid bilayer surrounding internal aqueous compartments and a relatively impermeable outer lipophilic phospholipid bilayer. Liposomes have gained considerable attention as drug delivery carriers because they are biocompatible, nontoxic, can deliver both hydrophilic and lipophilic drug molecules, protect their cargo from degradation by plasma enzymes, and transport their load across biological membranes and the blood brain barrier (BBB) (see, e.g., Spuch and Navarro (2011) Journal of drug delivery 2011).

Liposomes can be made from several different types of lipids; however, phospholipids are most commonly used to generate liposomes as drug carriers. Although liposome formation is spontaneous when a lipid film is mixed with an aqueous solution, it can also be expedited by applying force in the form of shaking by using a homogenizer, sonicator, or an extrusion apparatus (see, e.g., Spuch and Navarro (2011) Journal of drug delivery 2011).

A conventional liposome formulation is mainly comprised of natural phospholipids and phospholipids such as 1,2-distearoryl-*sn*-glycero-3-phosphatidyl choline (DSPC), sphingomyelin, egg phosphatidylcholines, and monosialoganglioside. In some embodiments, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) increases stability of a liposome.

Additives may be added to liposomes in order to modify their structure and properties. In some embodiments, cholesterol and/or sphingomyelin may be added to a liposomal mixture to help stabilize the liposomal structure and to prevent leakage of the liposomal inner cargo. In some embodiments, addition of cholesterol to a conventional liposome formulation reduces rapid release of the encapsulated active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) into the plasma. In some embodiments, liposomes are prepared from hydrogenated egg phosphatidylcholine or egg phosphatidylcholine, cholesterol, and dicetyl phosphate. In some embodiments, mean liposome vesicle size is adjusted to about 50 or 100 nm.

In some embodiments, Trojan Horse liposomes (also known as Molecular Trojan Horses or PEGylated immunoliposomes) may be used in pharmaceutical compositions for delivery of an active ingredient across the BBB (described on World Wide Web at cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long). Without being bound by any theory, it is believed that neutral lipid particles with specific antibodies conjugated to the surface allows crossing of the BBB via endocytosis. In some embodiments, pharmaceutical compositions comprising Trojan Horse liposomes may be used to deliver an active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) to the brain via an intravascular injection.

In some embodiments, liposomes include stable nucleic-acid-lipid particles (SNALP) (see, e.g., Morrissey et al. (2005) Nature Biotechnology 23(8):1002-1007; Zimmerman et al. (2006) Nature 441: 111-114). SNALPs include a mixture of cationic and fusogenic lipids and coated with polyethylene glycol (PEG) that allow cellular uptake and endosomal release of an active ingredient cargo. In some embodiments, a SNALP is a class of LNP and includes an ionizable lipid that is cationic at low pH (e.g., DLinDMA), a neutral helper lipid, cholesterol, and a diffusible polyethylene glycol (PEG)-lipid. In some embodiments, a

SNALP formulation includes the following lipids: 3-*N*-(-methoxy poly(ethylene glycol)2000) carbamoyl-1,2-dimyrestyloxy-propylamine (PEG-cDMA); 1,2-dilinoleyloxy-*N*,*N*-dimethyl-3-aminopropane (DLinDMA); 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC); and cholesterol. In some embodiments, a SNALP includes synthetic cholesterol, dipalmitoylphosphatidylcholine (DOPC), PEG-cDMA, and DLinDMA (see, e.g., Geisbert et al. (2010) Lancet 375:1896-1905). In some embodiments, a SNALP includes synthetic cholesterol, DSPC, PEG-cDMA, and DLinDMA (see, e.g., Judge et al. (2009) J. Clin. Invest. 119:661-673). In some embodiments, SNALP liposomes are about 80-100 nm in size. SNALPs have been used as effective delivery molecules to highly vascularized HepG2-derived liver tumors (see, e.g., Li et al. (2012) Gene Therapy 19:775- 780).

Without being bound by any one theory, during formulation of SNALPs, the ionizable lipid serves to condense lipid with an active ingredient (e.g., a nucleic acid molecule) during particle formation. When positively charged under increasingly acidic endosomal conditions, the ionizable lipid may mediate the fusion of a SNALP with the endosomal membrane, enabling release of the active ingredient into the cytoplasm. The PEG-lipid may stabilize the particle and reduce aggregation during formulation, and subsequently may provide a neutral hydrophilic exterior that improves pharmacokinetic properties. In some embodiments, SNALP liposomes are prepared by formulating DLinDMA and PEG-cDMA with DSPC, cholesterol and an active ingredient using a 25:1 lipid: active ingredient ratio and a 48:40:10:2 molar ratio of cholesterol: DLinDMA: DSPC: PEG-cDMA.

In some embodiments, a pharmaceutical composition of the disclosure includes LNPs. In some embodiments, lipids may be formulated with an active ingredient of the present disclosure to form LNPs. An LNP comprises a plurality of lipid molecules physically associated with each other by intermolecular forces. In some embodiments, LNPs include liposomes. In some embodiments, LNPs differ from liposomes in not having a continuous lipid bilayer. In some embodiments, LNPs comprise solid particles having a mixture of solid and liquid lipids. In some embodiments, LNPs include dendrimer lipid nanoparticles (DLNPs), SNALPs, and lipid-like nanoparticles (LLNPs). In general, a "nanoparticle" refers to any particle having a diameter of less than 1000 nanometers (nm). In some embodiments, nanoparticles have a diameter of 500 nm or less. In some embodiments, nanoparticles have a diameter ranging between 25 nm and 200 nm, or 100 nm or less. In some embodiments, nanoparticles have a diameter ranging between 35 nm and 60 nm. In some embodiments, an LNP includes a lipid particle between about 1 and about 100 nm in size.

LNPs include four components: ionizable cationic lipids, fusogenic zwitterionic phospholipids, cholesterol, and PEGylated (PEG) lipids. In some embodiments, the ionizable cationic lipid component complexes a negatively charged polynucleotide and enhances endosomal escape). In some embodiments, the phospholipid component functions in modifying lipid bilayer structure. In some embodiments, the cholesterol component helps to stabilize an LNP. In some embodiments, the PEG lipid component decreases LNP aggregation and non-specific uptake.

Ionizable cationic lipids useful in LNPs include: 1,2-dilineoyl-3-dimethylammonium-propane (DLinDAP); DLinDMA; l,2-dilinoleyloxy-keto-*N*,*N*-dimethyl-3-aminopropane (DLinK-DMA); 1,2-

dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLinKC2-DMA); 5A2-SC8 (Zhou et al. (2016) Proc. Natl Acad. Sci. USA 113:520–525); C12-200 (Love et al. (2010) Proc. Natl Acad. Sci. USA 107:1864–1869); 246C10 (Kim et al. (2021) Sci Adv 7(9): eabf4398); cKK-E12 (Fenton et al. (2016) Advanced Materials 28(15):2939-2943); 1,2-distearyloxy-*N,N*-dimethyl-3-aminopropane (DSDMA); 1,2-dioleyloxy-*N,N* -dimethyl-3-aminopropane (DODMA); 1,2-dilinolenyloxy-*N,N* -dimethyl-3-aminopropane (DLenDMA); and dilinoleylmethyl-4-dimethylaminobutyrate (Dlin-MC3-DMA; Jayaraman et al. (2012) Angew Chem Int Ed Engl. 51(34): 8529–8533). Cationic lipids are further described in International Publication Nos. WO2012040184, WO2011153120, WO2011149733, WO2011090965, WO2011043913, WO2011022460, WO2012061259, WO2012054365, WO2012044638, WO2010080724, WO201021865 and WO2008103276, US Patent Nos. 7,893,302 and 7,404,969 and US Patent Publication No. US20100036115, each of which is herein incorporated by reference in their entirety.

Zwitterionic phospholipids useful for LNPs include DSPC, DOPE, and DOPC.

PEG lipids useful for LNPs include: 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol (PEG-DMG); (3-o-[2''-(methoxypolyethyleneglycol 2000) succinoyl]-l,2-dimyristoyl-sn-glycol (PEG-S-DMG); R-3-[(ω-methoxy-poly(ethylene glycol)2000) carbamoyl]-1,2-dimyristyloxlpropyl-3-amine (PEG-C-DOMG); and C16 PEG-ceramide. In some embodiments, an LNP includes 50: 10: 38.5: 1.5 molar ratio of DLinKC2-DMA or C12-200: DSPC: cholesterol: PEG-DMG (see, e.g., Basha et al. (2011) Molecular Therapy 19(12):2186-2200). In some embodiments, an LNP includes 26.5: 20: 52: 1.5 ionizable lipid: DOPE: cholesterol: PEG lipid (see, e.g., Han et al. (2022) Sci Adv 8(3): eabj6901; Kim et al. (2021) Sci Adv 7(9): eabf4398). PEG lipids are further described in WO2012099755. In some embodiments, the ratio of PEG in the LNP formulations may be increased or decreased and/or the carbon chain length of the PEG lipid may be modified from C14 to C18 to alter the pharmacokinetics and/or biodistribution of the LNP formulations.

In some embodiments, the charge of an LNP is taken into consideration. Cationic lipids may combine with negatively charged lipids to induce non-bilayer structures that facilitate intracellular delivery. Because charged LNPs are rapidly cleared from circulation following intravenous injection, ionizable cationic lipids with pKa values below 7 were developed (see, e.g., Basha et al. (2011) Molecular Therapy 19(12):2186-2200). Negatively charged polymers such as polynucleotides may be loaded into LNPs at low pH values (e.g., pH 4) where the ionizable lipids display a positive charge. However, at physiological pH values, the LNPs exhibit a low surface charge compatible with longer circulation times.

Preparation of LNPs and encapsulation of an active ingredient are described in e.g., Basha et al. (2011) Molecular Therapy 19(12):1286-2200; Han et al. (2022) Sci Adv 8(3): eabj6901; Kim et al. (2021) Sci Adv 7(9): eabf4398; Finn et al. (2018) Cell Reports 22:2227–2235; Wei et al. (2020) Nature Communications 11:3232; WO2011127255; and WO2008103276. Lipids are commercially available (e.g., from Tekmira Pharmaceuticals, Vancouver, Canada; Avanti Polar Lipids, Inc., Alabaster, AL) or may be synthesized (e.g., Kim et al. (2021) Sci Adv 7(9): eabf4398). Synthesis of cationic lipids are also described in International Publication Nos. WO2012040184, WO2011153120, WO2011149733, WO2011090965,

WO2011043913, WO2011022460, WO2012061259, WO2012054365, WO2012044638, WO2010080724 and WO201021865. Cholesterol is commercially available (e.g., from Sigma-Aldrich, St Louis, MO).

In some embodiments, encapsulation may be performed by dissolving lipid mixtures comprising cationic lipid (e.g., Dlin-DMA): phospholipid (e.g., DSPC, DOPE): cholesterol: PEG-lipid (e.g., at 40: 10: 40: 10 molar ratio) in ethanol. An active ingredient (e.g., a polynucleotide comprising or encoding a guide RNA or RGN of the disclosure) may be dissolved in an acidic buffer (e.g., citrate, acetate), pH 3 or 4. In some embodiments, the lipid solution and active ingredient solution may be mixed using a microfluidics system (Chen et al. (2012) J. Amer. Chem. Soc. 134:6948-6951; e.g., NanoAssemblr from Precision Nanosystems) or by dropwise addition of the lipid solution to the active ingredient solution. Removal of ethanol and neutralization of formulation buffer may be performed by dialysis for, e.g., 16 hours or overnight, against phosphate-buffered saline (PBS) using dialysis cassettes (e.g., 3500 molecular weight cut-off cassettes from Life Technologies). Dynamic light scattering may be used to assess LNP size, polydispersity index (PDI), and zeta potential. Encapsulation efficiency of an active ingredient such as RNA may be determined by assays such as Quant-it™ Ribogreen Assay (Thermo Fisher). In some embodiments wherein the encapsulated active ingredient is a polynucleotide, the polynucleotide may be extracted from the eluted nanoparticles and quantified at 260 nm. LNP pKa may be assessed using a 2-(*p*-toluidino)-6-napthalene sulfonic acid (TNS) assay (Zhang et al. (2011) Langmuir 27(5):1907–1914). In some embodiments, a final lipid: active ingredient weight ratio includes 12: 1, 11: 1, 10: 1, 9: 1, 8: 1, 7: 1, 6: 1, and 5: 1.

In some embodiments where a pharmaceutical composition comprises a ribonucleoprotein (RNP) complex (i.e. RGN and guide RNA) encapsulated in an LNP, inclusion of an additional permanent cationic lipid (e.g., 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP)) allows formation of LNPs comprising RNP by mixing an ethanol solution of lipids with a solution of RNP at physiological pH (e.g., PBS buffer; Wei et al. (2020) Nature Communications 11:3232). In some embodiments, the permanent cationic lipid is included at 10 to 20 mole % of total lipids in LNPs.

In some embodiments, the LNP formulations described herein may additionally comprise a permeability enhancer molecule. Non-limiting permeability enhancer molecules are described in US2005/0222064.

In some embodiments, the LNP compositions are biodegradable, in that they do not accumulate to cytotoxic levels *in vivo* at a therapeutically effective dose. LNP formulations may be improved by replacing the cationic lipid with a biodegradable cationic lipid which is known as a rapidly eliminated lipid nanoparticle (reLNP). In some embodiments, the rapid metabolism of the rapidly eliminated lipids can improve the tolerability and therapeutic index of LNPs by an order of magnitude from a 1 mg/kg dose to a 10 mg/kg dose in rat. Inclusion of an enzymatically degraded ester linkage can improve the degradation and metabolism profile of the cationic component, while still maintaining the activity of the reLNP formulation. The ester linkage can be internally located within the lipid chain or it may be terminally

located at the terminal end of the lipid chain. The internal ester linkage may replace any carbon in the lipid chain.

In some embodiments, the LNP compositions do not cause an innate immune response that leads to substantial adverse effects at a therapeutic dose level. In some embodiments, the LNP compositions provided herein do not cause toxicity at a therapeutic dose level.

In some embodiments, the active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) is formulated as a solid lipid nanoparticle. A solid lipid nanoparticle (SLN) may be spherical with an average diameter between 10 to 1000 nm. SLN possess a solid lipid core matrix that can solubilize lipophilic molecules and may be stabilized with surfactants and/or emulsifiers. In a further embodiment, the lipid nanoparticle may be a self-assembly lipid-polymer nanoparticle (see, e.g., Zhang et al. (2008) ACS Nano 2(8):1696-1702).

In some embodiments, a lipid-based formulation including an active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) can be formulated for controlled release and/or targeted delivery. As used herein, "controlled release" refers to a pharmaceutical composition or compound release profile that conforms to a particular pattern of release to effect a therapeutic outcome.

In some embodiments, a lipid-based formulation including an active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) includes at least one controlled release coating. Controlled release coatings include: OPADRY® (Colorcon Inc., Harleysville, PA); polyvinylpyrrolidone/vinyl acetate copolymer; polyvinylpyrrolidone; hydroxypropyl methylcellulose; hydroxypropyl cellulose; hydroxyethyl cellulose; EUDRAGIT RL® (Evonik, Essen, Germany); EUDRAGIT RS® (Evonik, Essen, Germany); and cellulose derivatives such as ethylcellulose aqueous dispersions (AQUACOAT® and SURELEASE®, Colorcon Inc., Harleysville, PA). In some embodiments, the controlled release and/or targeted delivery formulation may comprise at least one degradable polyester which may contain polycationic side chains. Degradable polyesters include poly(serine ester), poly(L-lactide-*co*-L-lysine), poly(4-hydroxy-L-proline ester), and combinations thereof. In some embodiments, the degradable polyesters may include a PEG conjugation to form a PEGylated polymer.

In some embodiments, LNP formulations may be prepared such that they passively or actively are directed to different cell types *in vivo*, including hepatocytes, immune cells, tumor cells, endothelial cells, antigen presenting cells, and leukocytes (Akinc et al. (2010) Mol Ther. 18: 1357-1364; Song et al. (2005) Nat Biotechnol. 23:709-717; Judge et al. (2009) J Clin Invest. 119:661-673; Kaufmann et al. (2010) Microvasc Res 80:286-293; Santel et al. (2006) Gene Ther 13:1222-1234; Santel et al. (2006) Gene Ther 13:1360-1370; Gutbier et al. (2010) Pulm Pharmacol. Ther. 23:334-344; Basha et al. (2011) Mol. Ther. 19:2186-2200; Fenske and Cullis (2008) Expert Opin Drug Deliv. 5:25-44; Peer et al. (2008) Science 319:627-630; Peer and Lieberman (2011) Gene Ther. 18:1127-1133; all of which are incorporated herein by reference in their entirety). One example of passive targeting of formulations to liver cells includes the DLin-DMA, DLin-KC2-DMA and MC3 -based lipid nanoparticle formulations which have been shown to

bind to apolipoprotein E and promote binding and uptake of these formulations into hepatocytes *in vivo* (Akinc et al. (2010) Mol Ther. 18: 1357-1364).

LNP formulations can also be selectively targeted through expression of different ligands on their surface such as folate, transferrin, N-acetylgalactosamine (GalNAc), and antibody targeted approaches (Kolhatkar et al. (2011) Curr Drug Discov Technol. 8:197-206; Musacchio and Torchilin (2011) Front Biosci. 16:1388-1412; Yu et al. (2010) Mol Membr Biol. 27:286-298; Patil et al. (2008) Crit Rev Ther Drug Carrier Syst. 25:1- 61; Benoit et al. (2011) Biomacromolecules. 12:2708-2714; Zhao et al. (2008) Expert Opin Drug Deliv. 5:309-319; Akinc et al. (2010) Mol Ther. 18: 1357-1364; Srinivasan et al. (2012) Methods Mol Biol. 820:105-116; Ben-Arie et al. (2012) Methods Mol Biol. 757:497- 507; Peer, D (2010) J of controlled release 148(1):63-68; Peer et al. (2007) Proc Natl Acad Sci USA. 104:4095-4100; Kim et al. (2011) Methods Mol Biol. 721:339-353; Subramanya et al. (2010) Mol Ther. 18:2028-2037; Song et al. (2005) Nat Biotechnol. 23:709-717; Peer et al. (2008) Science 319:627-630; Peer and Lieberman (2011) Gene Ther. 18:1127-1133; all of which are incorporated herein by reference in their entirety).

In some embodiments, an active ingredient (i.e., guide RNAs and/or RGNs, or polynucleotides comprising or encoding such) may be encapsulated into an LNP and the LNP may then be encapsulated into a polymer, polymer matrix, hydrogel and/or surgical sealant described herein and/or known in the art. In some embodiments, the polymer, hydrogel or surgical sealant includes: poly(lactic-*co*-glycolic acid (PLGA); ethylene vinyl acetate (EVAc); poloxamer; GELSITE® (Nanotherapeutics, Inc. Alachua, FL); HYLENEX® (Halozyme Therapeutics, San Diego CA); surgical sealants such as fibrinogen polymers (Ethicon Inc., Cornelia, GA) and TISSELL® (Baxter International, Inc Deerfield, IL); PEG-based sealants; and COSEAL® (Baxter International, Inc Deerfield, IL).

LNPs and LNP formulations are further described in, e.g., U.S. Pat. Nos. 7,982,027; 7,799,565; 8,058,069; 8,283,333; 7,901,708; 7,745,651; 7,803,397; 8,101,741; 8,188,263; 7,915,399; 8,236,943 and 7,838,658; European Pat. Nos. 1766035; 1519714; 1781593; and 1664316.

In some embodiments wherein cells comprising or modified with the presently disclosed RGN, gRNAs, RGN systems or polynucleotides encoding the same are administered to a subject, the cells are administered as a suspension with a pharmaceutically acceptable carrier. One of skill in the art will recognize that a pharmaceutically acceptable carrier to be used in a cell composition will not include buffers, compounds, cryopreservation agents, preservatives, or other agents in amounts that substantially interfere with the viability of the cells to be delivered to the subject. A formulation comprising cells can include, for example, osmotic buffers that permit cell membrane integrity to be maintained, and optionally, nutrients to maintain cell viability or enhance engraftment upon administration. Such formulations and suspensions are known to those of skill in the art and/or can be adapted for use with the cells described herein using routine experimentation.

A cell composition can also be emulsified or presented as a liposome composition, provided that the emulsification procedure does not adversely affect cell viability. The cells and any other active ingredient

can be mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient, and in amounts suitable for use in the therapeutic methods described herein.

Additional agents included in a cell composition can include pharmaceutically acceptable salts of the components therein. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) that are formed with inorganic acids, such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, tartaric, mandelic and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases, such as, for example, sodium, potassium, ammonium, calcium or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine and the like.

Physiologically tolerable and pharmaceutically acceptable carriers are well known in the art. Exemplary liquid carriers are sterile aqueous solutions that contain no materials in addition to the active ingredients and water, or contain a buffer such as sodium phosphate at physiological pH value, physiological saline or both, such as phosphate-buffered saline. Still further, aqueous carriers can contain more than one buffer salt, as well as salts such as sodium and potassium chlorides, dextrose, polyethylene glycol and other solutes. Liquid compositions can also contain liquid phases in addition to and to the exclusion of water. Exemplary of such additional liquid phases are glycerin, vegetable oils such as cottonseed oil, and water-oil emulsions. The amount of an active compound used in the cell compositions that is effective in the treatment of a particular disorder or condition can depend on the nature of the disorder or condition, and can be determined by standard clinical techniques.

The presently disclosed RGN polypeptides, guide RNAs, RGN systems or polynucleotides encoding the same can be formulated with pharmaceutically acceptable excipients such as carriers, solvents, stabilizers, adjuvants, diluents, etc., depending upon the particular mode of administration and dosage form. In some embodiments, these pharmaceutical compositions are formulated to achieve a physiologically compatible pH, and range from a pH of about 3 to a pH of about 11, about pH 3 to about pH 7, depending on the formulation and route of administration. In some embodiments, the pH can be adjusted to a range from about pH 5.0 to about pH 8. In some embodiments, the compositions can comprise a therapeutically effective amount of at least one compound as described herein, together with one or more pharmaceutically acceptable excipients. In some embodiments, the compositions comprise a combination of the compounds described herein, or include a second active ingredient useful in the treatment or prevention of bacterial growth (for example and without limitation, anti-bacterial or anti-microbial agents), or include a combination of reagents of the present disclosure.

Suitable excipients include, for example, carrier molecules that include large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Other exemplary excipients can include antioxidants (for example and without limitation, ascorbic acid), chelating agents (for example and without limitation, EDTA), carbohydrates (for example and without limitation, dextrin, hydroxyalkylcellulose, and

hydroxyalkylmethylcellulose), stearic acid, liquids (for example and without limitation, oils, water, saline, glycerol and ethanol), wetting or emulsifying agents, pH buffering substances, and the like.

In some embodiments, the formulations are provided in unit-dose or multi-dose containers, for example sealed ampules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring the addition of the sterile liquid carrier, for example, saline, water-for-injection, a semi-liquid foam, or gel, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets of the kind previously described. In some embodiments, the active ingredient is dissolved in a buffered liquid solution that is frozen in a unit-dose or multi-dose container and later thawed for injection or kept/stabilized under refrigeration until use.

The therapeutic agent(s) may be contained in controlled release systems. In order to prolong the effect of a drug, it often is desirable to slow the absorption of the drug from subcutaneous, intrathecal, or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material with poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle. In some embodiments, the use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term sustained release implants are well-known to those of ordinary skill in the art.

Methods of treating a disease in a subject in need thereof are provided herein. The methods comprise administering to a subject in need thereof an effective amount of a presently disclosed RGN polypeptide or active variant or fragment thereof or a polynucleotide encoding the same, a presently disclosed gRNA or a polynucleotide encoding the same, a presently disclosed RGN system, or a cell modified by or comprising any one of these compositions.

In some embodiments, the treatment comprises *in vivo* gene editing by administering a presently disclosed RGN polypeptide, gRNA, or RGN system or polynucleotide(s) encoding the same. In some embodiments, the treatment comprises *ex vivo* gene editing wherein cells are genetically modified *ex vivo* with a presently disclosed RGN polypeptide, gRNA, or RGN system or polynucleotide(s) encoding the same and then the modified cells are administered to a subject. In some embodiments, the genetically modified cells originate from the subject that is then administered the modified cells, and the transplanted cells are referred to herein as autologous. In some embodiments, the genetically modified cells originate from a different subject (i.e., donor) within the same species as the subject that is administered the modified cells (i.e., recipient), and the transplanted cells are referred to herein as allogeneic. In some examples described herein, the cells can be expanded in culture prior to administration to a subject in need thereof.

In some embodiments, the disease to be treated with the presently disclosed compositions is one that can be treated with immunotherapy, such as with a chimeric antigen receptor (CAR) T cell. Such diseases include but are not limited to cancer. In some embodiments, the disease to be treated with the presently disclosed compositions is associated with a causal mutation. As used herein, a “causal mutation” refers to a

particular nucleotide, nucleotides, or nucleotide sequence in the genome that contributes to the severity or presence of a disease or disorder in a subject. The correction of the causal mutation leads to the improvement of at least one symptom resulting from a disease or disorder. In some embodiments, the causal mutation is adjacent to a PAM site recognized by an RGN disclosed herein. The causal mutation can be corrected with a presently disclosed RGN or a fusion polypeptide comprising a presently disclosed RGN and a base-editing polypeptide (i.e., a base editor). Non-limiting examples of diseases associated with a causal mutation include cystic fibrosis, Hurler syndrome, Friedreich's Ataxia, Huntington's Disease, and sickle cell disease. Additional non-limiting examples of disease-associated genes and mutations are set forth in Table 6 and more are available from McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), available on the World Wide Web.

In some embodiments, a method of treating a disease in a subject in need thereof comprises creating an induced pluripotent stem cell (iPSC) or isolating a mesenchymal stem cell from the subject, contacting the iPSC or mesenchymal stem cell with any one of the RGN polypeptides, systems, compositions comprising the same, or pharmaceutical compositions disclosed herein in order to genetically modify a target nucleic acid molecule within the cell, differentiating the modified iPSC or the modified mesenchymal stem cell into a genetically-modified mature cell or precursor thereof, and administering the genetically-modified mature cell or precursor thereof into the subject. In some embodiments, the iPSC or the mesenchymal stem cell is an autologous or an allogeneic cell. In some embodiments, the iPSC or the mesenchymal stem cell is derived from a donor that is a perfect human leukocyte antigen (HLA) match for the subject. In some embodiments, the subject is administered a myeloablative therapy prior to administration of the modified cells.

Any method known in the art for creating patient specific iPS cells can be used, including but not limited to that described in Takahashi and Yamanaka, Cell 126(4):663-76, 2006. For example, the creating step can comprise: a) isolating a somatic cell, such as a skin cell or fibroblast, from the subject; and b) introducing a set of pluripotency-associated genes into the somatic cell in order to induce the cell to become a pluripotent stem cell. The set of pluripotency-associated genes can be one or more of the genes selected from the group consisting of OCT4, SOX1, SOX2, SOX3, SOX15, SOX18, NANOG, KLF1, KLF2, KLF4, KLF5, c-MYC, n-MYC, REM2, TERT and LIN28. Mesenchymal stem cells can be isolated according to any method known in the art, such as from a patient's bone marrow or peripheral blood. For example, marrow aspirate can be collected into a syringe with heparin. Cells can be washed and centrifuged on a Percoll. The cells can be cultured in Dulbecco's modified Eagle's medium (DMEM) (low glucose) containing 10% fetal bovine serum (FBS) (Pittinger M F, Mackay A M, Beck S C et al., Science 1999; 284:143-147).

Genetically modified cells of the disclosure administered to a subject include autologous and allogeneic cells. Allogeneic cells refer to cells that are from a donor or donors (i.e., an individual or individuals from which the genetically modified cells are derived). Autologous cells refer to cells that are

from the subject undergoing treatment (i.e., the recipient of the genetically modified cells). Due to the risk of transplant rejection, an effort is made to optimize the degree of major histocompatibility complex (MHC)/human leukocyte antigen (HLA) matching between donor tissue and recipient. HLA are found on the surface of cells and help the body in identifying self versus non-self, so that the body can attack foreign entities such as bacteria and viruses. HLA typing of donor tissue and the recipient concerns determining the genotype of six HLA antigens or alleles between a donor(s) and recipient to assess the degree to which the six HLA match. HLA alleles usually refer to two each at the loci HLA-A, HLA-B and HLA-DR, or one each at the loci HLA-A, HLA-B and HLA-C and one each at the loci HAL-DRB1, HLA-DQB1 and HLA-DPB1 (see e.g., Kawase et al., 2007, Blood 110:2235-2241). In some embodiments, 4 of 6 HLA matching between donor(s) and recipient are sufficient for administration to the recipient of cells derived from a donor. In some embodiments, 5 of 6 HLA matching between donor(s) and recipient are sufficient for administration to the recipient of cells derived from a donor(s). In some embodiments, 6 of 6 HLA are matched between donor(s) and recipient for administration to the recipient of cells derived from the donor(s). In general, a 4/6, 5/6, or a 6/6 HLA match is the standard of clinical care. When all 6 HLA match between donor(s) and recipient, the match is referred to as being a perfect match.

As used herein, "treatment" or "treating," or "palliating" or "ameliorating" are used interchangeably. These terms refer to an approach for obtaining beneficial or desired results including but not limited to a therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant any therapeutically relevant improvement in or effect on one or more diseases, conditions, or symptoms under treatment. For prophylactic benefit, the compositions may be administered to a subject at risk of developing a particular disease, condition, or symptom, or to a subject reporting one or more of the physiological symptoms of a disease, even though the disease, condition, or symptom may not have yet been manifested. In some embodiments, treatment may be administered after one or more symptoms have developed and/or after a disease has been diagnosed. In particular embodiments, treatment may be administered in the absence of symptoms, e.g., to prevent or delay onset of a symptom or inhibit onset or progression of a disease. For example, treatment may be administered to a susceptible individual prior to the onset of symptoms (e.g., in light of a history of symptoms and/or in light of genetic or other susceptibility factors). Treatment may also be continued after symptoms have resolved, for example, to prevent or delay their prevention or recurrence.

The term "effective amount" or "therapeutically effective amount" refers to the amount of an agent that is sufficient to effect beneficial or desired results. The therapeutically effective amount may vary depending upon one or more of: the subject and disease condition being treated, the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art. The specific dose may vary depending on one or more of: the particular agent chosen, the dosing regimen to be followed, whether it is administered in combination with other compounds, timing of administration, and the delivery system in which it is carried.

The term "administering" refers to the placement of an active ingredient into a subject, by a method or route that results in at least partial localization of the introduced active ingredient at a desired site, such as

a site of injury or repair, such that a desired effect(s) is produced. In some embodiments, the disclosure provides methods comprising delivering any of the RGN polypeptides, nucleic acid molecules, ribonucleoprotein complexes, vectors, pharmaceutical compositions and/or gRNAs described herein. In some embodiments, the disclosure further provides cells produced by such methods, and organisms (such as animals or plants) comprising or produced from such cells. In some embodiments, a RGN polypeptide and/or nucleic acid molecules as described herein in combination with (and optionally complexed with) a guide sequence is delivered to a cell.

In those embodiments wherein cells are administered, the cells can be administered by any appropriate route that results in delivery to a desired location in the subject where at least a portion of the implanted cells or components of the cells remain viable. The period of viability of the cells after administration to a subject can be as short as a few hours, e.g., twenty-four hours, to a few days, to as long as several years, or even the lifetime of the patient, i.e., long-term engraftment. For example, in some aspects described herein, an effective amount of photoreceptor cells or retinal progenitor cells is administered via a systemic route of administration, such as an intraperitoneal or intravenous route.

In some embodiments, the administering comprises administering by viral delivery. Viral vectors comprising a nucleic acid encoding the RGN polypeptides, ribonucleoprotein complexes, or vectors disclosed herein may be administered directly to patients (i.e., *in vivo*) or they may be used to treat cells *in vitro*, and the modified cells may optionally be administered to patients (i.e., *ex vivo*). Conventional viral based systems may include, without limitation, retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, often resulting in long term expression of the inserted transgene. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and typically produce high viral titers. In applications where transient expression is preferred, adenoviral based systems may be used. Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division.

In some embodiments, the administering comprises administering by other non-viral delivery of nucleic acids. Exemplary non-viral delivery methods, without limitation, include RNP complexes, lipofection, nucleofection, microinjection, biolistics, virosomes, liposomes, LNPs, immunoliposomes, polycation or lipidmucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat. Nos.5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Feigner, WO1991/17424; WO 1991/16024. Delivery can be to cells (e.g., in vitro or ex vivo administration) or target tissues (e.g., in vivo administration). In some embodiments, administration of a pharmaceutical composition of the disclosure includes daily intravenous injections of about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 mg/kg/day, or more of an active ingredient in a pharmaceutical composition comprising a liposome or LNP. In some embodiments, administration of a pharmaceutical composition comprising a liposome or LNP includes doses

of about 0.01 to 1 mg per kg of body weight. In some embodiments, administration of a pharmaceutical composition comprising a liposome or LNP includes doses of about 1 to 10 mg per kg of body weight.

Suitable routes of administering the pharmaceutical compositions described herein include, without limitation: topical, subcutaneous, transdermal, intradermal, intralesional, intraarticular, intraperitoneal, intravesical, transmucosal, gingival, intradental, intracochlear, transtympanic, intraorgan, epidural, intrathecal, intramuscular, intravenous, intravascular, intraosseus, periocular, intratumoral, intracerebral, and intracerebroventricular administration.

In embodiments, the pharmaceutical composition described herein is administered to a subject by injection, inhalation (e.g., of an aerosol), by means of a catheter, by means of a suppository, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including a membrane, such as a sialastic membrane, or a fiber. In embodiments, the pharmaceutical composition is formulated for delivery to a subject, e.g., for gene editing.

In embodiments, the pharmaceutical composition is formulated in accordance with routine procedures as a composition adapted for intravenous or subcutaneous administration to a subject, e.g., a human. In embodiments, pharmaceutical composition for administration by injection are solutions in sterile isotonic aqueous buffer. Where necessary, the pharmaceutical can also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the pharmaceutical is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the pharmaceutical composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients can be mixed prior to administration.

In embodiments, the pharmaceutical composition can be contained within a lipid particle or vesicle, such as a liposome or microcrystal, which is also suitable for parenteral administration.

Although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions which are suitable for administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to animals or organisms of all sorts.

As used herein, the term "subject" refers to any individual for whom diagnosis, treatment or therapy is desired. In some embodiments, the subject is an animal. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human being.

The efficacy of a treatment can be determined by the skilled clinician. However, a treatment is considered an "effective treatment," if any one or all of the signs or symptoms of a disease or disorder are altered in a beneficial manner (e.g., decreased by at least 10%), or other clinically accepted symptoms or markers of disease are improved or ameliorated. Efficacy can also be measured by failure of an individual to worsen as assessed by hospitalization or need for medical interventions (e.g., progression of the disease is

halted or at least slowed). Methods of measuring these indicators are known to those of skill in the art. Treatment includes: (1) inhibiting the disease, e.g., arresting, or slowing the progression of symptoms; or (2) relieving the disease, e.g., causing regression of symptoms; and (3) preventing or reducing the likelihood of the development of symptoms.

A. Modifying causal mutations using base-editing

An example of a genetically inherited disease which could be corrected using an approach that relies on an RGN-base editor fusion protein of the invention is Hurler Syndrome. Hurler Syndrome, also known as MPS-1, is the result of a deficiency of α-L-iduronidase (IDUA) resulting in a lysosomal storage disease characterized at the molecular level by the accumulation of dermatan sulfate and heparan sulfate in lysosomes. This disease is generally an inherited genetic disorder caused by mutations in the IDUA gene encoding α-L-iduronidase. Common IDUA mutations are W402X and Q70X, both nonsense mutations resulting in premature termination of translation. Such mutations are well addressed by precise genome editing (PGE) approaches, since reversion of a single nucleotide, for example by a base-editing approach, would restore the wild-type coding sequence and result in protein expression controlled by the endogenous regulatory mechanisms of the genetic locus. Additionally, since heterozygotes are known to be asymptomatic, a PGE therapy that targets one of these mutations would be useful to a large proportion of patients with this disease, as only one of the mutated alleles needs to be corrected (Bunge et al. (1994) Hum. Mol. Genet. 3(6): 861-866, herein incorporated by reference).

Current treatments for Hurler Syndrome include enzyme replacement therapy and bone marrow transplants (Vellodi et al. (1997) Arch. Dis. Child. 76(2): 92-99; Peters et al. (1998) Blood 91(7): 2601-2608, herein incorporated by reference). While enzyme replacement therapy has had a dramatic effect on the survival and quality of life of Hurler Syndrome patients, this approach requires costly and time-consuming weekly infusions. Additional approaches include the delivery of the IDUA gene on an expression vector or the insertion of the gene into a highly expressed locus such as that of serum albumin (U.S. Patent No. 9,956,247, herein incorporated by reference). However, these approaches do not restore the original IDUA locus to the correct coding sequence. A genome-editing strategy would have a number of advantages, most notably that regulation of gene expression would be controlled by the natural mechanisms present in healthy individuals. Additionally, using base editing does not necessitate causing a double-stranded DNA breaks, which could lead to large chromosomal rearrangements, cell death, or oncogenecity by the disruption of tumor suppression mechanisms. A general strategy may be directed toward using RGN-base editor fusion proteins of the invention, for example those comprising LPG10165, LPG10167, LPG10168, LPG10171, LPG10186, LPG10190, LPG10194, LPG10195, LPG10200, LPG10203, or LPG10207 to target and correct certain disease-causing mutations in the human genome. It will be appreciated that similar approaches to target diseases that can be corrected by base-editing may also be pursued. It will be further appreciated that similar approaches to target disease-causing mutations in other species, particularly common household pets or livestock, can also be deployed using the RGNs of the

invention. Common household pets and livestock include dogs, cats, horses, pigs, cows, sheep, chickens, donkeys, snakes, ferrets, and fish including salmon and shrimp.

B. Modifying causal mutations by targeted deletion

RGNs of the invention could also be useful in human therapeutic approaches where the causal mutation is more complicated. For example, some diseases such as Friedreich's Ataxia and Huntington's Disease are the result of a significant increase in repeats of a three nucleotide motif (i.e., an expanded trinucleotide repeat") at a particular region of a gene, which affects the ability of the expressed protein to function or to be expressed. Friedreich's Ataxia (FRDA) is an autosomal recessive disease resulting in progressive degeneration of nervous tissue in the spinal cord. Reduced levels of the frataxin (FXN) protein in the mitochondria cause oxidative damages and iron deficiencies at the cellular level. The reduced FXN expression has been linked to a GAA triplet expansion within the intron 1 of the somatic and germline FXN gene. In FRDA patients, the GAA repeat frequently consists of more than 70, sometimes even more than 1000 (most commonly 600-900) triplets, whereas unaffected individuals have about 40 repeats or less (Pandolfo et al. (2012) Handbook of Clinical Neurology 103: 275-294; Campuzano et al. (1996) Science 271: 1423-1427; Pandolfo (2002) Adv. Exp. Med. Biol. 516: 99-118; all herein incorporated by reference).

The expansion of the trinucleotide repeat sequence causing Friedreich's Ataxia (FRDA) occurs in a defined genetic locus within the FXN gene, referred to as the FRDA instability region. RNA guided nucleases (RGNs) may be used for excising the expanded trinucleotide repeat in FRDA patient cells. This approach requires 1) an RGN and guide RNA sequence that can be programmed to target the allele in the human genome; and 2) a delivery approach for the RGN and guide sequence. Many nucleases used for genome editing, such as the commonly used Cas9 nuclease from *S. pyogenes* (SpCas9), are too large to be packaged into adeno-associated viral (AAV) vectors, especially when considering the length of the SpCas9 gene and the guide RNA in addition to other genetic elements required for functional expression cassettes. This makes an approach using SpCas9 more difficult.

Certain RNA guided nucleases of the invention, for example, LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, and LPG10208, are well suited for packaging into an AAV vector along with a guide RNA. The present invention encompasses a strategy using RGNs of the invention in which an expanded trinucleotide repeat is removed. Such a strategy is applicable to other diseases and disorders which have a similar genetic basis, such as Huntington's Disease. Similar strategies using RGNs of the invention may also be applicable to similar diseases and disorders in non-human animals of agronomic or economic importance, including dogs, cats, horses, pigs, cows, sheep, chickens, donkeys, snakes, ferrets, and fish including salmon and shrimp.

C. Modifying causal mutations by targeted mutagenesis

RGNs of the invention could also be to introduce disruptive mutations that may result in a beneficial effect. Genetic defects in the genes encoding hemoglobin, particularly the beta globin chain (the HBB gene), can be responsible for a number of diseases known as hemoglobinopathies, including sickle cell anemia and thalassemias.

In adult humans, hemoglobin is a heterotetramer comprising two alpha (α)-like globin chains and two beta (β)-like globin chains and 4 heme groups. In adults the α2β2 tetramer is referred to as Hemoglobin A (HbA) or adult hemoglobin. Typically, the alpha and beta globin chains are synthesized in an approximate 1:1 ratio and this ratio seems to be critical in terms of hemoglobin and red blood cell (RBC) stabilization. In a developing fetus, a different form of hemoglobin, fetal hemoglobin (HbF), is produced which has a higher binding affinity for oxygen than Hemoglobin A such that oxygen can be delivered to the baby's system via the mother's blood stream. Fetal hemoglobin also contains two α globin chains, but in place of the adult β-globin chains, it has two fetal gamma (γ)-globin chains (i.e., fetal hemoglobin is α2γ2). The regulation of the switch from production of gamma- to beta-globin is quite complex, and primarily involves a down-regulation of gamma globin transcription with a simultaneous up-regulation of beta globin transcription. At approximately 30 weeks of gestation, the synthesis of gamma globin in the fetus starts to drop while the production of beta globin increases. By approximately 10 months of age, the newborn's hemoglobin is nearly all α2β2 although some HbF persists into adulthood (approximately 1-3% of total hemoglobin). In the majority of patients with hemoglobinopathies, the genes encoding gamma globin remain present, but expression is relatively low due to normal gene repression occurring around parturition as described above.

Sickle cell disease is caused by a V6E mutation in the β globin gene (HBB) (a GAG to GTG at the DNA level), where the resultant hemoglobin is referred to as "hemoglobinS" or "HbS." Under lower oxygen conditions, HbS molecules aggregate and form fibrous precipitates. These aggregates cause the abnormality or 'sickling' of the RBCs, resulting in a loss of flexibility of the cells. The sickling RBCs are no longer able to squeeze into the capillary beds and can result in vaso-occlusive crisis in sickle cell patients. In addition, sickled RBCs are more fragile than normal RBCs, and tend towards hemolysis, eventually leading to anemia in the patient.

Treatment and management of sickle cell patients is a life-long proposition involving antibiotic treatment, pain management and transfusions during acute episodes. One approach is the use of hydroxyurea, which exerts its effects in part by increasing the production of gamma globin. Long term side effects of chronic hydroxyurea therapy are still unknown, however, and treatment gives unwanted side effects and can have variable efficacy from patient to patient. Despite an increase in the efficacy of sickle cell treatments, the life expectancy of patients is still only in the mid to late 50's and the associated morbidities of the disease have a profound impact on a patient's quality of life.

Thalassemias (alpha thalassemias and beta thalassemia) are also diseases relating to hemoglobin and typically involve a reduced expression of globin chains. This can occur through mutations in the regulatory regions of the genes or from a mutation in a globin coding sequence that results in reduced expression or reduced levels or functional globin protein. Treatment of thalassemias usually involves blood transfusions

and iron chelation therapy. Bone marrow transplants are also being used for treatment of people with severe thalassemias if an appropriate donor can be identified, but this procedure can have significant risks.

One approach that has been proposed for the treatment of both sickle cell disease (SCD) and beta thalassemias is to increase the expression of gamma globin so that HbF functionally replaces the aberrant adult hemoglobin. As mentioned above, treatment of SCD patients with hydroxyurea is thought to be successful in part due to its effect on increasing gamma globin expression (DeSimone (1982) Proc Nat'l Acad Sci USA 79(14):4428-31; Ley, et al., (1982) N. Engl. J. Medicine, 307: 1469-1475; Ley, et al., (1983) Blood 62: 370-380; Constantoulakis et al., (1988) Blood 72(6):1961-1967, all herein incorporated by reference). Increasing the expression of HbF involves identification of genes whose products play a role in the regulation of gamma globin expression. One such gene is BCL11A. BCL11A encodes a zinc finger protein that expressed in adult erythroid precursor cells, and down-regulation of its expression leads to an increase in gamma globin expression (Sankaran et at (2008) Science 322: 1839, herein incorporated by reference). Use of an inhibitory RNA targeted to the BCL11A gene has been proposed (e.g., U.S. Patent Publication 2011/0182867, herein incorporated by reference) but this technology has several potential drawbacks, including that complete knock down may not be achieved, delivery of such RNAs may be problematic, and the RNAs must be present continuously, requiring multiple treatments for life.

RGNs of the invention, for example LPG10165, LPG10166, LPG10167, LPG10168, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, or LPG10208 may be used to target the BCL11A enhancer region to disrupt expression of BCL11A, thereby increasing gamma globin expression. This targeted disruption can be achieved by non-homologous end joining (NHEJ), whereby an RGN of the invention targets to a particular sequence within the BCL11A enhancer region, makes a double-stranded break, and the cell's machinery repairs the break, typically simultaneously introducing deleterious mutations. Similar to what is described for other disease targets, RGNs of the invention may have advantages over other known RGNs due to their relatively small size, which enables packaging expression cassettes for the RGN and its guide RNA into a single AAV vector for *in vivo* delivery. Similar strategies using RGNs of the invention may also be applicable to similar diseases and disorders in both humans and in non-human animals of agronomic or economic importance.

XI. Cells Comprising a Polynucleotide Genetic Modification

Provided herein are cells and organisms comprising a target nucleic acid molecule that has been modified using a process mediated by an RGN, crRNA, and/or tracrRNA as described herein. In some of these embodiments, the RGN comprises any one of the amino acid sequences set forth as SEQ ID NOs: 1-20, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat comprising any one of the nucleotide sequences set forth as SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA comprising any one of

the nucleotide sequences set forth as SEQ ID NOs:42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA.

The modified cells can be eukaryotic (*e.g.*, mammalian, plant, insect, avian cell) or prokaryotic. Also provided are organelles and embryos comprising at least one nucleotide sequence that has been modified by a process utilizing an RGN, crRNA, and/or tracrRNA as described herein. The genetically modified cells, organisms, organelles, and embryos can be heterozygous or homozygous for the modified nucleotide sequence.

The chromosomal modification of the cell, organism, organelle, or embryo can result in altered expression (up-regulation or down-regulation), inactivation, or the expression of an altered protein product or an integrated sequence. In those embodiments wherein the chromosomal modification results in either the inactivation of a gene or the expression of a non-functional protein product, the genetically modified cell, organism, organelle, or embryo is referred to as a "knock out". The knock out phenotype can be the result of a deletion mutation (*i.e.*, deletion of at least one nucleotide), an insertion mutation (*i.e.*, insertion of at least one nucleotide), or a nonsense mutation (*i.e.*, substitution of at least one nucleotide such that a stop codon is introduced).

Alternatively, the chromosomal modification of a cell, organism, organelle, or embryo can produce a "knock in", which results from the chromosomal integration of a nucleotide sequence that encodes a protein. In some of these embodiments, the coding sequence is integrated into the chromosome such that the chromosomal sequence encoding the wild-type protein is inactivated, but the exogenously introduced protein is expressed.

In other embodiments, the chromosomal modification results in the production of a variant protein product. The expressed variant protein product can have at least one amino acid substitution and/or the addition or deletion of at least one amino acid. The variant protein product encoded by the altered chromosomal sequence can exhibit modified characteristics or activities when compared to the wild-type protein, including but not limited to altered enzymatic activity or substrate specificity.

In yet other embodiments, the chromosomal modification can result in an altered expression pattern of a protein. As a non-limiting example, chromosomal alterations in the regulatory regions controlling the expression of a protein product can result in the overexpression or downregulation of the protein product or an altered tissue or temporal expression pattern.

The cells that have been modified can be grown into an organism, such as a plant, in accordance with conventional ways. See, for example, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same modified strain or different strains, and the resulting hybrid having the genetic modification. The present invention provides genetically modified seed. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention,

provided that these parts comprise the genetic modification. Further provided is a processed plant product or byproduct that retains the genetic modification, including for example, soymeal.

The methods provided herein may be used for modification of any plant species, including, but not limited to, monocots and dicots. Examples of plants of interest include, but are not limited to, corn (maize), sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, and oilseed rape, Brassica sp., alfalfa, rye, millet, safflower, peanuts, sweet potato, cassava, coffee, coconut, pineapple, citrus trees, cocoa, tea, banana, avocado, fig, guava, mango, olive, papaya, cashew, macadamia, almond, oats, vegetables, ornamentals, and conifers.

Vegetables include, but are not limited to, tomatoes, lettuce, green beans, lima beans, peas, and members of the genus Curcumis such as cucumber, cantaloupe, and musk melon. Ornamentals include, but are not limited to, azalea, hydrangea, hibiscus, roses, tulips, daffodils, petunias, carnation, poinsettia, and chrysanthemum. In specific embodiments, plants of the present invention are crop plants (for example, maize, sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, oilseed rape, etc.).

The methods provided herein can also be used to genetically modify any prokaryotic species, including but not limited to, archaea and bacteria (*e.g.*, *Bacillus* sp., *Klebsiella* sp. *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., Agrobacterium, *Lactobacillus* sp.).

The methods provided herein can be used to genetically modify any eukaryotic species or cells therefrom, including but not limited to animals (*e.g.*, mammals, insects, fish, birds, and reptiles), fungi, amoeba, algae, and yeast. In some embodiments, the cell that is modified by the presently disclosed methods include cells of hematopoietic origin, such as cells of the immune system including but not limited to B cells, T cells, natural killer (NK) cells, pluripotent stem cells, induced pluripotent stem cells, chimeric antigen receptor T (CAR-T) cells, monocytes, macrophages, and dendritic cells.

Cells that have been modified may be introduced into an organism. These cells could have originated from the same organism (e.g., person) in the case of autologous cellular transplants, wherein the cells are modified in an *ex vivo* approach. Alternatively, the cells originated from another organism within the same species (e.g., another person) in the case of allogeneic cellular transplants.

The article "a" and "an" are used herein to refer to one or more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, "a polypeptide" means one or more polypeptides.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this disclosure pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended embodiments.

Non-limiting embodiments include:

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:1-20.

2. The nucleic acid molecule of embodiment 1, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of the target sequence.

3. The nucleic acid molecule of embodiment 1 or 2, wherein said polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to said polynucleotide.

4. The nucleic acid molecule of any one of embodiments 1-3, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs:1-20.

5. The nucleic acid molecule of any one of embodiments 1-3, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

6. The nucleic acid molecule of any one of embodiments 1-5, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule upon binding.

7. The nucleic acid molecule of embodiment 6, wherein said RGN polypeptide is capable of generating a double-stranded break.

8. The nucleic acid molecule of embodiment 6, wherein said RGN polypeptide is capable of generating a single-stranded break.

9. The nucleic acid molecule of any one of embodiments 1-5, wherein said RGN polypeptide is nuclease inactive or is a nickase.

10. The nucleic acid molecule of any one of embodiments 1-9, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

11. The nucleic acid molecule of embodiment 10, wherein the base-editing polypeptide is a deaminase.

12. The nucleic acid molecule of embodiment 11, wherein the deaminase is a cytosine deaminase or an adenine deaminase.

13. The nucleic acid molecule of embodiment 11, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

14. The nucleic acid molecule of embodiment 11, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

15. The nucleic acid molecule of any one of embodiments 1-14, wherein the RGN polypeptide comprises one or more nuclear localization signals.

16. The nucleic acid molecule of any one of embodiments 1-15, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

17. The nucleic acid molecule of any one of embodiments 1-16, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

18. A vector comprising the nucleic acid molecule of any one of embodiments 1-17.

19. The vector of embodiment 18, further comprising at least one nucleotide sequence encoding said gRNA capable of hybridizing to the non-target strand of said target sequence.

20. The vector of embodiment 19, 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: 21; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 22; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) 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: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 24; 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: 4;

e) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 26; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 28; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 29; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) 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: 51;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10;

k) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 31; 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: 11;

l) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 32; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 33; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 34; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 35; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 36; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) 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: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 17;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 38 or 39; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 40; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; 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: 41; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

21. The vector of embodiment 19, 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 95% sequence identity to SEQ ID NO: 21; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 22; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 23; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 24; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 26; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO: 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 7;

h) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 28; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 8;

i) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 29; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 9;

j) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 30; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 10;

k) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 31; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

l) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 32; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 33; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 34; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 35; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 36; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16;

q) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 37; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 17;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 38 or 39; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO:59 or 60;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 18;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 40; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 19; and

t) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 41; and

ii) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62;

wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 20.

22. The vector of embodiment 19, 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 100% sequence identity to SEQ ID NO: 21; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 22; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 23; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 24; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 26; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 28; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 29; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 30; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 10;

k) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 31; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 32; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 33; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 34; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 35; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 36; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 37; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 38 or 39; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 40; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 41; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

23. The vector of any one of embodiments 19-22, where said gRNA is a single guide RNA.

24. The vector of any one of embodiments 19-22, wherein said gRNA is a dual-guide RNA.

25. A cell comprising the nucleic acid molecule of any one of embodiments 1-17 or the vector of any one of embodiments 18-24.

26. The cell of embodiment 25, wherein the cell is a prokaryotic cell.

27. The cell of embodiment 25, wherein the cell is a eukaryotic cell.

28. The cell of embodiment 27, wherein the eukaryotic cell is a mammalian cell.

29. The cell of embodiment 28, wherein the mammalian cell is a human cell.

30. The cell of embodiment 29, wherein the human cell is an immune cell.

31. The cell of embodiment 30, wherein the immune cell is a stem cell.

32. The cell of embodiment 31, wherein the stem cell is an induced pluripotent stem cell.

33. The cell of embodiment 27, wherein the eukaryotic cell is an insect or avian cell.

34. The cell of embodiment 27, wherein the eukaryotic cell is a fungal cell.

35. The cell of embodiment 27, wherein the eukaryotic cell is a plant cell.

36. A plant comprising the cell of embodiment 35.

37. A seed comprising the cell of embodiment 35.

38. A method for making an RGN polypeptide comprising culturing the cell of any one of embodiments 25-35 under conditions in which the RGN polypeptide is expressed.

39. A method for making an RGN polypeptide comprising introducing into a cell a heterologous nucleic acid molecule comprising a nucleotide sequence encoding an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20; and culturing said cell under conditions in which the RGN polypeptide is expressed.

40. The method of embodiment 39, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of said target sequence.

41. The method of embodiment 39 or 40, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

42. The method of embodiment 39 or 40, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

43. The method of any one of embodiments 38-42, further comprising purifying said RGN polypeptide.

44. The method of any one of embodiments 38-42, wherein said cell further expresses one or more guide RNAs capable of binding to said RGN polypeptide to form an RGN ribonucleoprotein complex.

45. The method of embodiment 42, further comprising purifying said RGN ribonucleoprotein complex.

46. An 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: 1-20.

47. The RGN polypeptide of embodiment 46, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of said target sequence.

48. The RGN polypeptide of embodiment 46 or 47, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

49. The RGN polypeptide of embodiment 46 or 47, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

50. The RGN polypeptide of any one of embodiments 46-49, wherein said RGN polypeptide is an isolated RGN polypeptide.

51. The RGN polypeptide of any one of embodiments 46-50, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule upon binding.

52. The RGN polypeptide of embodiment 51, wherein cleavage by said RGN polypeptide generates a double-stranded break.

53. The RGN polypeptide of embodiment 51, wherein cleavage by said RGN polypeptide generates a single-stranded break.

54. The RGN polypeptide of any one of embodiments 46-50, wherein said RGN polypeptide is nuclease inactive or a nickase.

55. The RGN polypeptide of any one of embodiments 46-54, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

56. The RGN polypeptide of embodiment 55, wherein the base-editing polypeptide is a deaminase.

57. The RGN polypeptide of embodiment 56, wherein the deaminase is a cytosine deaminase or an adenine deaminase.

58. The RGN polypeptide of embodiment 56, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

59. The RGN polypeptide of embodiment 56, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

60. The RGN polypeptide of any one of embodiments 46-59, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

61. The RGN polypeptide of any one of embodiments 46-60, wherein the RGN polypeptide comprises one or more nuclear localization signals.

62. A ribonucleoprotein (RNP) complex comprising the RGN polypeptide of any one of embodiments 46-61 and a guide RNA bound to the RGN polypeptide.

63. A nucleic acid molecule comprising a CRISPR RNA (crRNA) or a polynucleotide encoding a crRNA, wherein said crRNA comprises a spacer and a CRISPR repeat, wherein said CRISPR repeat comprises a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

64. The nucleic acid molecule of embodiment 63, wherein a guide RNA comprising:

a) said crRNA; and

b) a trans-activating CRISPR RNA (tracrRNA) hybridized to said CRISPR repeat of said crRNA;

is capable of hybridizing to the non-target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

65. The nucleic acid molecule of embodiment 63 or 64, wherein said polynucleotide encoding a crRNA is operably linked to a promoter heterologous to said polynucleotide.

66. The nucleic acid molecule of any one of embodiments 63-65, wherein said CRISPR repeat comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

67. The nucleic acid molecule of any one of embodiments 63-65, wherein said CRISPR repeat comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

68. A vector comprising the nucleic acid molecule comprising said polynucleotide encoding said crRNA of any one of embodiments 63-67.

69. The vector of embodiment 68, wherein said vector further comprises a polynucleotide encoding said tracrRNA.

70. The vector of embodiment 69, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 21;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 22;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 23;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 24;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 1-19 of SEQ ID NO :1043, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 26;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 28;

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 29;

j) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 30;

k) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 31;

l) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 32;

m) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 33;

n) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 34;

o) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 35;

p) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 36;

q) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 37;

r) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 38 or 39;

s) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 40; and

t) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 41.

71. The vector of embodiment 69, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 21;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 22;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 23;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 24;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 26;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 28;

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 29;

j) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 30;

k) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 31;

l) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 32;

m) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 33;

n) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 34;

o) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 35;

p) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 36;

q) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 37;

r) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 59 or 60, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 38 or 39;

s) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 40; and

t) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 41.

72. The vector of embodiment 69, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 21;

b) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 22;

c) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 23;

d) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 24;

e) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042;

f) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 26;

g) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045;

h) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 28;

i) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 29;

j) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 30;

k) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 31;

l) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 32;

m) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 33;

n) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 34;

o) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 35;

p) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 36;

q) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 37;

r) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 38 or 39;

s) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 40; and

t) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 41.

73. The vector of any one of embodiments 69-72, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

74. The vector of any one of embodiments 69-72, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to separate promoters.

75. The vector of any one of embodiments 69-74, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide.

76. The vector of embodiment 75, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 1, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 2, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 3, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 4, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 5, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 6, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO:26 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 7, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 8, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 9, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 10, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 11, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 12, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 13, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 14, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 15, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 16, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 17, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 18, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 38 or 39 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 19, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 20, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

77. The vector of embodiment 75, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 1, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 2, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 3, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 4, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 5, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 6, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO:26 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 7, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 8, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 9, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 10, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 11, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 12, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 13, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 14, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 15, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 16, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 17, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 18, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 38 or 39 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 19, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 20, wherein said CRISPR repeat has at least 95% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 62.

78. The vector of embodiment 75, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 1, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 2, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 3, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 4, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 5, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 6, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO:26 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 7, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 8, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 9, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 10, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 11, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 12, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 13, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 14, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 15, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 16, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 17, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 18, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 38 or 39 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 19, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 20, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

79. A nucleic acid molecule comprising a trans-activating CRISPR RNA (tracrRNA) or a polynucleotide encoding a tracrRNA comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

80. The nucleic acid molecule of embodiment 79, wherein a guide RNA comprising:

a) said tracrRNA; and

b) a crRNA comprising a spacer and a CRISPR repeat, wherein said tracrRNA hybridizes with said CRISPR repeat of said crRNA;

is capable of hybridizing to the non-target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

81. The nucleic acid molecule of embodiment 79 or 80, wherein said polynucleotide encoding a tracrRNA is operably linked to a promoter heterologous to said polynucleotide.

82. The nucleic acid molecule of any one of embodiments 79-81, wherein said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

83. The nucleic acid molecule of any one of embodiments 79-81, wherein said tracrRNA comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

84. A vector comprising the nucleic acid molecule comprising said polynucleotide encoding said tracrRNA of any one of embodiments 79-83.

85. The vector of embodiment 84, wherein said vector further comprises a polynucleotide encoding said crRNA.

86. The vector of embodiment 85, wherein said crRNA comprises a CRISPR repeat selected from the group consisting of:

a) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

87. The vector of embodiment 85, wherein said crRNA comprises a CRISPR repeat selected from the group consisting of:

a) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 21, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 42;

b) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 22, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 43;

c) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 23, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 24, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 45;

e) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 26, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 28, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 49;

i) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 29, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 50;

j) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 30, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 51;

k) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 31, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 52;

l) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 32, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 53;

m) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 33, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 54;

n) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 34, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 55;

o) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 35, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 56;

p) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 36, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 57;

q) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 37, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 58;

r) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 38 or 39, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 59 or 60;

s) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 40, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 61; and

t) a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 41, wherein said tracrRNA has at least 95% sequence identity to SEQ ID NO: 62.

88. The vector of embodiment 85, wherein said crRNA comprises a CRISPR repeat selected from the group consisting of:

a) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

89. The vector of any one of embodiments 85-88, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

90. The vector of any one of embodiments 85-88, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to separate promoters.

91. The vector of any one of embodiments 85-90, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide.

92. The vector of embodiment 91, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 1, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 2, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 3, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 4, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 5, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 6, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 7, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 8, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 9, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 10, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 11, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 12, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 13, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 14, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 15, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 16, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 17, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 18, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38, or 39 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 19, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 20, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

93. The vector of embodiment 91, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 1, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 2, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 3, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 4, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 5, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 6, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 26 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 7, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 8, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 9, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 10, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 11, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 12, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 13, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 14, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 15, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 16, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 17, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 18, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 38, or 39 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 19, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 95% sequence identity to SEQ ID NO: 20, wherein said crRNA comprises a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 95% sequence identity to SEQ ID NO: 62.

94. The vector of embodiment 91, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 1, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 2, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 3, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and said

tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 4, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 5, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 6, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 7, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 8, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 9, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 10, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 11, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 12, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 13, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 14, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 15, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 16, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 17, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 18, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38, or 39 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 19, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 20, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

95. A cell comprising the nucleic acid molecule of any one of embodiments 63-67 and 79-83, the vector of any one of embodiments 68-78 and 84-94, the single guide RNA of embodiment 254, or the double guide RNA of embodiment 255.

96. The cell of embodiment 95, wherein the cell is a prokaryotic cell.

97. The cell of embodiment 95, wherein the cell is a eukaryotic cell.

98. The cell of embodiment 97, wherein the eukaryotic cell is a mammalian cell.

99. The cell of embodiment 98, wherein the mammalian cell is a human cell.

100. The cell of embodiment 99, wherein the human cell is an immune cell.

101. The cell of embodiment 100, wherein the immune cell is a stem cell.

102. The cell of embodiment 101, wherein the stem cell is an induced pluripotent stem cell.

103. The cell of embodiment 97, wherein the eukaryotic cell is an insect or avian cell.

104. The cell of embodiment 97, wherein the eukaryotic cell is a fungal cell.

105. The cell of embodiment 97, wherein the eukaryotic cell is a plant cell.

106. A plant comprising the cell of embodiment 105.

107. A seed comprising the cell of embodiment 105.

108. A system for binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, said system comprising:

a) one or more guide RNAs capable of hybridizing to the non-target strand of said target 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: 1-20 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide;

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 sequence.

109. The system of embodiment 108, 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.

110. A system for binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, said system comprising:

a) one or more guide RNAs capable of hybridizing to the non-target strand of said target 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: 1-20;

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 sequence.

111. The system of any one of embodiments 108-110, 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.

112. The system of any one of embodiments 108-111, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

113. The system of any one of embodiments 108-111, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

114. The system of any one of embodiments 108-113, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

115. The system of any one of embodiments 108-114, wherein said target sequence is a eukaryotic target sequence.

116. The system of any one of embodiments 108-115, wherein said gRNA is a single guide RNA (sgRNA).

117. The system of any one of embodiments 108-115, wherein said gRNA is a dual-guide RNA.

118. The system of any one of embodiments 108-117, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 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: 4;

e) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NOs:10;

k) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 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: 11;

l) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 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: 17;

r) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

119. The system of any one of embodiments 108-117, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NOs:10;

k) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 20.

120. The system of any one of embodiments 108-117, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of

SEQ ID NO: 1040, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NOs:10;

k) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

121. The system of any one of embodiments 108-120, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

122. The system of any one of embodiments 108-120, wherein the target sequence is within a cell.

123. The system of any one of embodiments 108-122, wherein the one or more guide RNAs is capable of hybridizing to the non-target strand of said target sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target nucleic acid molecule.

124. The system of embodiment 123, wherein the cleavage generates a double-stranded break.

125. The system of embodiment 123, wherein the cleavage generates a single-stranded break.

126. The system of any one of embodiments 108-122, wherein said RGN polypeptide is nuclease inactive or is a nickase.

127. The system of any one of embodiments 108-126, wherein the RGN polypeptide is operably linked to a base-editing polypeptide.

128. The system of embodiment 127, wherein the base-editing polypeptide is a deaminase.

129. The system of embodiment 128, wherein the deaminase is a cytosine deaminase or an adenine deaminase.

130. The system of embodiment 128, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

131. The system of embodiment 128, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

132. The system of any one of embodiments 108-131, wherein the RGN polypeptide comprises one or more nuclear localization signals.

133. The system of any one of embodiments 108-132, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

134. The system of any one of embodiments 108-133, wherein the nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding the RGN polypeptide are located on one vector.

135. The system of any one of embodiments 108-134, wherein said system further comprises one or more donor polynucleotides.

136. A cell comprising the system of any one of embodiments 108-135.

137. The cell of embodiment 136, wherein the cell is a prokaryotic cell.

138. The cell of embodiment 136, wherein the cell is a eukaryotic cell.

139. The cell of embodiment 138, wherein the eukaryotic cell is a mammalian cell.

140. The cell of embodiment 139, wherein the mammalian cell is a human cell.

141. The cell of embodiment 140, wherein the human cell is an immune cell.

142. The cell of embodiment 141, wherein the immune cell is a stem cell.

143. The cell of embodiment 142, wherein the stem cell is an induced pluripotent stem cell.

144. The cell of embodiment 138, wherein the eukaryotic cell is an insect or avian cell.

145. The cell of embodiment 138, wherein the eukaryotic cell is a fungal cell.

146. The cell of embodiment 138, wherein the eukaryotic cell is a plant cell.

147. A plant comprising the cell of embodiment 146.

148. A seed comprising the cell of embodiment 146.

149. A pharmaceutical composition comprising the nucleic acid molecule of any one of embodiments 1-17, 63-67, and 79-83, the vector of any one of embodiments 18-24, 68-78, and 84-94, the cell of any one of embodiments 27-32, 97-102, and 138-143, the RGN polypeptide of any one of embodiments 46-61, the RNP complex of embodiment 62, or the system of any one of embodiments 108-135, and a pharmaceutically acceptable carrier.

150. The pharmaceutical composition of embodiment 149, wherein the pharmaceutically acceptable carrier is heterologous to said nucleic acid molecule, said vector, said cell, said RGN polypeptide, or said system.

151. The pharmaceutical composition of embodiment 149 or 150, wherein the pharmaceutically acceptable carrier is not naturally-occurring.

152. The pharmaceutical composition of any one of embodiments 149-151, wherein the pharmaceutical composition is lipid-based.

153. The pharmaceutical composition of embodiment 152, wherein the lipid-based pharmaceutical composition comprises liposomes or lipid nanoparticles (LNPs).

154. The pharmaceutical composition of embodiment 153, wherein the nucleic acid molecule, the vector, the cell, the RGN polypeptide, the RNP complex, or the system is encapsulated in, and/or non-covalently or covalently attached to the liposomes or the LNPs.

155. A method for binding a target sequence in a target nucleic acid molecule comprising delivering a system according to any one of embodiments 108-135, to said target sequence or a cell comprising the target sequence.

156. The method of embodiment 155, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target sequence.

157. The method of embodiment 155, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby modulating expression of a target gene comprising said target sequence.

158. A method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence, said method comprising delivering a system according to any one of embodiments 108-135 to said target sequence or a cell comprising the target sequence, wherein cleavage or modification of said target nucleic acid molecule occurs.

159. The method of embodiment 158, wherein said modified target nucleic acid molecule comprises insertion of heterologous DNA into the target nucleic acid molecule.

160. The method of embodiment 158, wherein said modified target nucleic acid molecule comprises deletion of at least one nucleotide from the target nucleic acid molecule.

161. The method of embodiment 158, wherein said modified target nucleic acid molecule comprises mutation of at least one nucleotide in the target nucleic acid molecule.

162. A method for binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, said method comprising:

a) assembling an RNA-guided nuclease (RGN) ribonucleotide complex by combining:

i) one or more guide RNAs capable of hybridizing to the non-target strand of the target sequence; and

ii) an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs:1-20;

under conditions suitable for formation of the RGN ribonucleotide complex; and

b) contacting said target nucleic acid molecule or a cell comprising said target nucleic acid molecule with the assembled RGN ribonucleotide complex;

wherein the one or more guide RNAs hybridize to the non-target strand of the target sequence, thereby directing said RGN polypeptide to bind to said target sequence.

163. The method of embodiment 162, wherein said method is performed *in vitro*, *in vivo*, or *ex vivo*.

164. The method of embodiment 162 or 163, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target sequence.

165. The method of embodiment 162 or 163, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby allowing for the modulation of expression of a target gene comprising said target sequence.

166. The method of embodiment 162 or 163, wherein said RGN polypeptide further comprises a base-editing polypeptide, thereby allowing for the modification of said target nucleic acid molecule.

167. The method of embodiment 166, wherein said base-editing polypeptide comprises a deaminase.

168. The method of embodiment 167, wherein said deaminase is a cytosine deaminase or an adenine deaminase.

169. The method of embodiment 167, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

170. The method of embodiment 167, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

171. The method of embodiment 162 or 163, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule, thereby allowing for the cleaving and/or modifying of said target nucleic acid molecule.

172. A method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence, wherein said target sequence comprises a target strand and a non-target strand, said method comprising contacting the target nucleic acid 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: 1-20; and

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

wherein the one or more guide RNAs hybridize to the non-target strand of the target sequence, thereby directing said RGN polypeptide to bind to said target nucleic acid molecule and cleavage and/or modification of said target nucleic acid molecule occurs.

173. The method of embodiment 172, wherein cleavage by said RGN polypeptide generates a double-stranded break.

174. The method of embodiment 172, wherein cleavage by said RGN polypeptide generates a single-stranded break.

175. The method of embodiment 172, wherein said RGN polypeptide is nuclease inactive or a nickase and is operably fused to a base-editing polypeptide.

176. The method of embodiment 175, wherein the base-editing polypeptide is a deaminase.

177. The method of embodiment 175, wherein the deaminase is a cytosine deaminase or an adenine deaminase.

178. The method of embodiment 176, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

179. The method of embodiment 176, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

180. The method of embodiment 172, wherein said modified target nucleic acid molecule comprises insertion of heterologous DNA into the target nucleic acid molecule.

181. The method of embodiment 172, wherein said modified target nucleic acid molecule comprises deletion of at least one nucleotide from the target nucleic acid molecule.

182. The method of embodiment 172, wherein said modified target nucleic acid molecule comprises mutation of at least one nucleotide in the target nucleic acid molecule.

183. The method of any one of embodiments 162-182, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

184. The method of any one of embodiments 162-183, wherein said target sequence is a eukaryotic target sequence.

185. The method of any one of embodiments 162-184, wherein said gRNA is a single guide RNA (sgRNA).

186. The method of any one of embodiments 162-184, wherein said gRNA is a dual-guide RNA.

187. The method of any one of embodiments 162-186, wherein said RGN comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

188. The method of any one of embodiments 162-186, wherein said RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

189. The method of any one of embodiments 162-186, wherein:

a) said RGN has at least 90% sequence identity to SEQ ID NO: 1, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42;

b) said RGN has at least 90% sequence identity to SEQ ID NO: 2, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43;

c) said RGN has at least 90% sequence identity to SEQ ID NO: 3, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) said RGN has at least 90% sequence identity to SEQ ID NO: 4, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45;

e) said RGN has at least 90% sequence identity to SEQ ID NO: 5, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) said RGN has at least 90% sequence identity to SEQ ID NO: 6, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) said RGN has at least 90% sequence identity to SEQ ID NO: 7, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) said RGN has at least 90% sequence identity to SEQ ID NO: 8, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49;

i) said RGN has at least 90% sequence identity to SEQ ID NO: 9, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50;

j) said RGN has at least 90% sequence identity to SEQ ID NO: 10, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51;

k) said RGN has at least 90% sequence identity to SEQ ID NO: 11, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52;

l) said RGN has at least 90% sequence identity to SEQ ID NO: 12, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53;

m) said RGN has at least 90% sequence identity to SEQ ID NO: 13, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54;

n) said RGN has at least 90% sequence identity to SEQ ID NO: 14, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55;

o) said RGN has at least 90% sequence identity to SEQ ID NO: 15, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56;

p) said RGN has at least 90% sequence identity to SEQ ID NO: 16, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57;

q) said RGN has at least 90% sequence identity to SEQ ID NO: 17, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58;

r) said RGN has at least 90% sequence identity to SEQ ID NO: 18, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) said RGN has at least 90% sequence identity to SEQ ID NO: 19, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61; and

t) said RGN has at least 90% sequence identity to SEQ ID NO: 20, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62.

190. The method of any one of embodiments 162-186, wherein:

a) said RGN has at least 95% sequence identity to SEQ ID NO: 1, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42;

b) said RGN has at least 95% sequence identity to SEQ ID NO: 2, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43;

c) said RGN has at least 95% sequence identity to SEQ ID NO: 3, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) said RGN has at least 95% sequence identity to SEQ ID NO: 4, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45;

e) said RGN has at least 95% sequence identity to SEQ ID NO: 5, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) said RGN has at least 95% sequence identity to SEQ ID NO: 6, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) said RGN has at least 95% sequence identity to SEQ ID NO: 7, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO:

1044 or 1045, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) said RGN has at least 95% sequence identity to SEQ ID NO: 8, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49;

i) said RGN has at least 95% sequence identity to SEQ ID NO: 9, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50;

j) said RGN has at least 95% sequence identity to SEQ ID NO: 10, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51;

k) said RGN has at least 95% sequence identity to SEQ ID NO: 11, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52;

l) said RGN has at least 95% sequence identity to SEQ ID NO: 12, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53;

m) said RGN has at least 95% sequence identity to SEQ ID NO: 13, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54;

n) said RGN has at least 95% sequence identity to SEQ ID NO: 14, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55;

o) said RGN has at least 95% sequence identity to SEQ ID NO: 15, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56;

p) said RGN has at least 95% sequence identity to SEQ ID NO: 16, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57;

q) said RGN has at least 95% sequence identity to SEQ ID NO: 17, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58;

r) said RGN has at least 95% sequence identity to SEQ ID NO: 18, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 59 or 60;

s) said RGN has at least 95% sequence identity to SEQ ID NO: 19, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61; and

t) said RGN has at least 95% sequence identity to SEQ ID NO: 20, said guide RNA comprises a crRNA repeat having at least 95% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62.

191. The method of any one of embodiments 162-186, wherein:

a) said RGN has at least 100% sequence identity to SEQ ID NO: 1, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42;

b) said RGN has at least 100% sequence identity to SEQ ID NO: 2, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43;

c) said RGN has at least 100% sequence identity to SEQ ID NO: 3, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) said RGN has at least 100% sequence identity to SEQ ID NO: 4, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45;

e) said RGN has at least 100% sequence identity to SEQ ID NO: 5, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) said RGN has at least 100% sequence identity to SEQ ID NO: 6, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) said RGN has at least 100% sequence identity to SEQ ID NO: 7, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) said RGN has at least 100% sequence identity to SEQ ID NO: 8, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49;

i) said RGN has at least 100% sequence identity to SEQ ID NO: 9, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50;

j) said RGN has at least 100% sequence identity to SEQ ID NO: 10, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51;

k) said RGN has at least 100% sequence identity to SEQ ID NO: 11, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52;

l) said RGN has at least 100% sequence identity to SEQ ID NO: 12, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53;

m) said RGN has at least 100% sequence identity to SEQ ID NO: 13, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54;

n) said RGN has at least 100% sequence identity to SEQ ID NO: 14, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55;

o) said RGN has at least 100% sequence identity to SEQ ID NO: 15, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56;

p) said RGN has at least 100% sequence identity to SEQ ID NO: 16, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57;

q) said RGN has at least 100% sequence identity to SEQ ID NO: 17, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58;

r) said RGN has at least 100% sequence identity to SEQ ID NO: 18, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) said RGN has at least 100% sequence identity to SEQ ID NO: 19, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61; and

t) said RGN has at least 100% sequence identity to SEQ ID NO: 20, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62.

192. The method of any one of embodiments 155-191, wherein the target sequence is within a cell.

193. The method of embodiment 192, wherein the cell is a eukaryotic cell.

194. The method of embodiment 193, wherein the eukaryotic cell is a mammalian cell.

195. The method of embodiment 194, wherein said mammalian cell is a human cell.

196. The method of embodiment 195, wherein said human cell is an immune cell.

197. The method of embodiment 196, wherein said immune cell is a stem cell.

198. The method of embodiment 197, wherein said stem cell is an induced pluripotent stem cell.

199. The method of embodiment 193, wherein the eukaryotic cell is an insect or avian cell.

200. The method of embodiment 192, wherein the cell is a prokaryotic cell.

201. The method of embodiment 193, wherein the eukaryotic cell is a fungal cell.

202. The method of embodiment 193, wherein the eukaryotic cell is a plant cell.

203. The method of any one of embodiments 192-202, further comprising culturing the cell under conditions in which the RGN polypeptide is expressed and cleaves and modifies the target nucleic acid molecule to produce a modified target nucleic acid molecule; and selecting a cell comprising said modified target nucleic acid molecule.

204. A cell comprising a modified target nucleic acid molecule according to the method of embodiment 203.

205. The cell of embodiment 204, wherein the cell is a eukaryotic cell.

206. The cell of embodiment 205, wherein the eukaryotic cell is a mammalian cell.

207. The cell of embodiment 206, wherein said mammalian cell is a human cell.

208. The cell of embodiment 207, wherein said human cell is an immune cell.

209. The cell of embodiment 208, wherein said immune cell is a stem cell.

210. The cell of embodiment 209, wherein said stem cell is an induced pluripotent stem cell.

211. The cell of embodiment 205, wherein the eukaryotic cell is an insect or avian cell.

212. The cell of embodiment 204, wherein the cell is a prokaryotic cell.

213. The cell of embodiment 205, wherein the eukaryotic cell is a fungal cell.

214. The cell of embodiment 205, wherein the eukaryotic cell is a plant cell.

215. A plant comprising the cell of embodiment 214.

216. A seed comprising the cell of embodiment 214.

217. A pharmaceutical composition comprising the cell of any one of embodiments 205-210 and a pharmaceutically acceptable carrier.

218. A method for producing a genetically modified cell with a correction in a causal mutation for a genetically inherited disease, the method comprising introducing into the cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell

whereby the RGN and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

219. The method of embodiment 218, wherein the RGN is nuclease inactive or a nickase and is fused to a polypeptide which has base-editing activity.

220. The method of embodiment 219, wherein the base-editing polypeptide is a deaminase.

221. The method of embodiment 220, wherein the polypeptide with base-editing activity is a cytosine deaminase or an adenine deaminase.

222. The method of embodiment 218, wherein the deaminase has at least 90% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

223. The method of embodiment 218, wherein the deaminase has 100% sequence identity to an amino acid sequence of any one of SEQ ID NOs: 481-552.

224. The method of any one of embodiments 218-223, wherein the genetically inherited disease is caused by a single nucleotide polymorphism.

225. The method of any one of embodiments 218-223, wherein the genetically inherited disease is Hurler Syndrome.

226. The method of embodiment 224, wherein the gRNA further comprises a spacer that targets a region proximal to the causal single nucleotide polymorphism.

227. A method for producing a genetically modified cell with a deletion in a disease-causing expanded trinucleotide repeat, the method comprising introducing into the cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a first guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein the gRNA comprises a spacer that targets the 5' flank of the expanded trinucleotide repeat; and

c) a second guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein said second gRNA comprises a spacer that targets the 3'flank of the expanded trinucleotide repeat;

whereby the RGN and the two gRNAs target to the expanded trinucleotide repeat and at least a portion of the expanded trinucleotide repeat is removed.

228. The method of embodiment 227, wherein the genetically inherited disease is Friedrich's Ataxia or Huntington's Disease.

229. The method of embodiment 227, wherein the first gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat.

230. The method of embodiment 229, wherein the second gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat.

231. The method of any one of embodiments 218-230, wherein said RGN polypeptide has at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

232. The method of any one of embodiments 218-230, wherein said RGN polypeptide has 100% sequence identity to any one of SEQ ID NOs: 1-20.

233. The method of any one of embodiments 218-230, wherein said gRNA, said first gRNA, said second gRNA, or said first gRNA and said second gRNA is selected from a gRNA selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10;

k) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

234. The method of any one of embodiments 218-230, wherein said gRNA, said first gRNA, said second gRNA, or said first gRNA and said second gRNA is selected from a gRNA selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 10;

k) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 20.

235. The method of any one of embodiments 218-230, wherein said gRNA, said first gRNA, said second gRNA, or said first gRNA and said second gRNA is selected from a gRNA selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 10;

k) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

236. The method of any one of embodiments 218-235, wherein the cell is an animal cell.

237. The method of embodiment 236, wherein the animal cell is a mammalian cell.

238. The method of embodiment 236, wherein the cell is derived from a dog, cat, mouse, rat, rabbit, horse, cow, pig, or human.

239. A method for producing a genetically modified mammalian hematopoietic progenitor cell having decreased BCL11A mRNA and protein expression, the method comprising introducing into an isolated human hematopoietic progenitor cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell,

whereby the RGN and gRNA are expressed in the cell and cleave at the BCL11A enhancer region, resulting in genetic modification of the human hematopoietic progenitor cell and reducing the mRNA and/or protein expression of BCL11A.

240. The method of embodiment 239, wherein said RGN polypeptide has at least 95% sequence identity to any one of SEQ ID NOs: 1-20.

241. The method of embodiment 239, wherein said RGN polypeptide has 100% sequence identity to any one of SEQ ID NOs: 1-20.

242. The method of embodiment 239, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID

NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO:10;

k) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 39 or 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

243. The method of embodiment 239, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO:10;

k) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 38 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 39 or 40 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 95% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 95% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 20.

244. The method of embodiment 239, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO:10;

k) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 39 or 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

245. The method of any one of embodiments 239-244, wherein the gRNA further comprises a spacer that targets a region within or proximal to the BCL11A enhancer region.

246. A method of treating a disease, disorder, or condition, said method comprising administering to a subject in need thereof a pharmaceutical composition of any one of embodiments 149-154 and 217.

247. The method of embodiment 246, wherein said disease, disorder, or condition is associated with a causal mutation and said pharmaceutical composition corrects said causal mutation.

248. The method of embodiment 246 or 247, wherein said subject is at risk of developing said disease, disorder, or condition.

249. Use of the nucleic acid molecule of any one of embodiments 1-17, 63-67, and 79-83, the vector of any one of embodiments 18-24, 68-78, and 84-94, the cell of any one of embodiments 27-32, 97-102, 138-143, and 205-210, the RGN polypeptide of any one of embodiments 46-61, the RNP complex of embodiment 62, or the system of any one of embodiments 108-135 for the treatment of a disease, disorder, or condition in a subject in need thereof.

250. The use of embodiment 249, wherein said disease, disorder, or condition is associated with a causal mutation and said treating comprises correcting said causal mutation.

251. The use of embodiment 249 or 250, wherein said subject is at risk of developing said disease, disorder, or condition.

252. Use of the nucleic acid molecule of any one of embodiments 1-17, 63-67, and 79-83, the vector of any one of embodiments 18-24, 68-78, and 84-94, the cell of any one of embodiments 27-32, 97-102, 138-143, and 205-210, the RGN polypeptide of any one of embodiments 46-61, the RNP complex of embodiment 62, or the system of any one of embodiments 108-135 for the manufacture of a medicament useful for treating a disease, disorder, or condition.

253. The use of embodiment 252, wherein said disease is associated with a causal mutation and said medicament corrects said causal mutation.

254. A single guide RNA comprising the nucleic acid molecule comprising said crRNA of any one of embodiments 63-67 and the nucleic acid molecule comprising said tracrRNA of any one of embodiments 79-83.

255. A dual guide RNA comprising the nucleic acid molecule comprising said crRNA of any one of embodiments 63-67 and the nucleic acid molecule comprising said tracrRNA of any one of embodiments 79-83.

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

Example 1. Identification of RNA-guided nuclease

Nineteen distinct CRISPR-associated RNA-guided nucleases (RGNs) were identified and are described in Table 1 below. Table 1 provides the name of each RGN, its amino acid sequence, the source

from which it was derived, and processed crRNA and tracrRNA sequences (see Example 2 for methods of identification). Table 1 further provides a generic single guide RNA (sgRNA) sequence, where the poly-N indicates the location of the spacer which determines the nucleic acid target sequence of the sgRNA. Two different crRNA and tracrRNA sequences were predicted for LPG10205. For RGN systems LPG10165, LPG10190, LPG10195, LPG10196, LPG10197, LPG10203, LPG10204, LPG10205, and LPG10208, the conserved sequence in the base of the hairpin stem of the tracrRNA is UNANNC, which is set forth as SEQ ID NO: 566. For RGN system LPG10166, the conserved sequence in the base of the hairpin stem of the tracrRNA is CNANNU, which is set forth as SEQ ID NO: 569. For RGN systems LPG10167, LPG10168, LPG10171, LPG10198, and LPG10207, the conserved sequence in the base of the hairpin stem of the tracrRNA is UNANNG, which is set forth as SEQ ID NO: 567. For RGN systems LPG10169 and LPG10200, the conserved sequence in the base of the hairpin stem of the tracrRNA is CNANNC, which is set forth as SEQ ID NO: 568. For RGN system LPG10186, the conserved sequence in the base of the hairpin stem of the tracrRNA is UNANNU, which is set forth as SEQ ID NO: 570. For RGN system LPG10191, the conserved sequence in the base of the hairpin stem of the tracrRNA is CNANNG, which is set forth as SEQ ID NO: 571. For RGN system LPG10194, the conserved sequence in the base of the hairpin stem of the tracrRNA is CNCNNU, which is set forth as SEQ ID NO: 572.

Table 1: Summary of SEQ IDs and CRISPR associated systems

RGN ID	SEQ ID NO.	Source	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA backbone* (SEQ ID NO)
LPG10165	1	*Bacillus cytotoxicus*	21	42	63
LPG10166	2	*Nitrobacter winogradskyi*	22	43	64
LPG10167	3	*Bacillus toyonensis*	23	44	65
LPG10168	4	*Bacillus toyonensis*	24	45	66
LPG10169	5	*Staphylococcus_A sciuri*	25	46	67
LPG10171	6	*B. mosaicus*	26	47	68
LPG10186	7	Contaminated	27	48	69
LPG10190	8	Not Determined	28	49	70
LPG10191	9	Not Determined	29	50	71
LPG10194	10	Not Determined	30	51	72
LPG10195	11	Not Determined	31	52	73
LPG10196	12	Not Determined	32	53	74
LPG10197	13	Not Determined	33	54	75
LPG10198	14	Not Determined	34	55	76
LPG10200	15	Not Determined	35	56	77
LPG10203	16	Not Determined	36	57	78

RGN ID	SEQ ID NO.	Source	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA backbone* (SEQ ID NO)
LPG10204	17	Not Determined	37	58	79
LPG10205	18	Not Determined	38, 39	59, 60	80, 81
LPG10207	19	Not Determined	40	61	82
LPG10208	20	Not Determined	41	62	83

*sgRNA backbone: crRNA repeat-linker-tracrRNA

LPG10167 can also use the LPG10167_altv1 backbone sequence set forth as SEQ ID NO: 1040. LPG10169 can also use the APG01604_altv1 backbone sequence or LPG10169_altv1 backbone sequence set forth as SEQ ID NO: 1041 or 1042, respectively. LPG10171 can also use the LPG10171_altv1 backbone sequence set forth as SEQ ID NO: 1043. LPG10186 can also use the LPG10186_altv1 backbone sequence or LPG10186_altv2 backbone sequence set forth as SEQ ID NO: 1044 or 1045, respectively.

Example 2: Guide RNA identification and sgRNA construction

Guide RNAs were determined through identification of tracrRNAs and crRNAs encoded in the genomic CRISPR loci. An internal database of validated and predicted tracrRNAs was used to determine a putative RNA. When no homology-based tracrRNA could be identified, the anti-repeat portion of the tracrRNA was identified using permissive BLASTn parameters using the consensus CRISPR repeat following the protocol determined in Briner, Cold Spring Harb Protoc 7: pbd. prot086785 (2016). Manual curation of RNAs was performed using secondary structure prediction by RNAfold, an RNA folding software. sgRNA cassettes were prepared by DNA synthesis and were generally designed as follows (5'->3'): 20-30 bp spacer, operably linked at its 3' end to the processed repeat portion of the crRNA, operably linked to a 4 bp noncomplementary linker (AAAG; SEQ ID NO: 84), operably linked at its 3' end to the processed tracrRNA. Other 4 bp noncomplementary linkers may also be used.

For *in vitro* assays, sgRNAs were synthesized by *in vitro* transcription of the sgRNA cassettes with a GeneArt™ Precision gRNA Synthesis Kit (ThermoFisher). Processed crRNA and tracrRNA sequences for each of the RGN polypeptides are identified and are set forth in Table 1. See Example 3 for the sgRNAs constructed for PAM libraries 1 and 2.

Example 3: Determination of PAM requirements for each RGN

PAM requirements for each RGN were determined using an in vitro translation PAM determination assay essentially adapted from Karvelis et al. (2015) *Genome Biol* 253. Briefly, two plasmid libraries (L1 and L2) were generated in a pUC18 backbone (ampR), with each containing a distinct 30bp protospacer (target) sequence flanked by 8 random nucleotides (i.e., the PAM region). The target sequence and flanking PAM region of library 1 and library 2 for each RGN are set forth in Table 2.

Briefly, the proteins were translated in vitro using the PURExpress In Vitro Protein Synthesis kit (New England Biolabs) starting from a T7 promoter driven bacterial expression plasmid. In the presence of the appropriate sgRNA, plasmids containing a PAM that is recognizable by the RGN will be cleaved. Adapters were ligated to cleaved DNA using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), and after cleanup, the product was PCR amplified using primers complementary to the ligated adapter and plasmid backbone sequence such that the enriched PAM sequences were contained in the amplicon. Deep sequencing (75bp single end reads) was performed on a MiSeq (Illumina) by a service provider (MoGene, St. Louis, MO). Typically, 500,000 reads were obtained per amplicon. PAM regions were extracted, counted, and normalized to total reads for each sample. PAMs that lead to plasmid cleavage were identified by being enriched when compared to controls (i.e., the starting frequencies for the PAM in the libraries). To represent PAM requirements for a novel RGN, the depletion ratios (frequency in sample/frequency in control) for all sequences in the region in question were converted to enrichment values with a -log base 2 transformation. Sufficient PAMs were defined as those with enrichment values >3.5. PAMs above this threshold in both libraries were collected and used to generate web logos, which for example can be generated using a web-based service on the internet known as "weblogo". PAM sequences were identified and reported when there was a consistent pattern in the top enriched PAMs. A consensus PAM (having an enrichment factor (EF) >3.5) for each RGN is provided in Table 2. The PAM orientation is also indicated in Table 2.

Table 2: PAM or PAM-like determination

RGN ID	sgRNA L1 (SEQ ID NO.)	sgRNA L2 (SEQ ID NO.)	PAM (SEQ ID NO.)	PAM orientation
LPG10165	85	106	127	5'-target-PAM-3'
LPG10166	86	107	128, 129, 130, 131	5'-target-PAM-3'
LPG10167	87	108	132	5'-target-PAM-3'
LPG10168	88	109	132	5'-target-PAM-3'
LPG10169	89	110	133	5'-target-PAM-3'
LPG10171	90	111	134	5'-target-PAM-3'
LPG10186	91	112	135	5'-target-PAM-3'
LPG10190	92	113	136, 137	5'-target-PAM-3'
LPG10191	93	114	138	5'-target-PAM-3'
LPG10194	94	115	139	5'-target-PAM-3'
LPG10195	95	116	140	5'-target-PAM-3'
LPG10196	96	117	141	5'-target-PAM-3'
LPG10197	97	118	142	5'-target-PAM-3'
LPG10198	98	119	143	5'-target-PAM-3'
LPG10200	99	120	144	5'-target-PAM-3'
LPG10203	100	121	145	5'-target-PAM-3'
LPG10204	101	122	146, 147	5'-target-PAM-3'
LPG10205	102, 103	123, 124	146, 147	5'-target-PAM-3'
LPG10207	104	125	132	5'-target-PAM-3'

LPG10208	105	126	146, 147	5'-target-PAM-3'

Example 4: Demonstration of gene editing activity in mammalian cells

RGN expression cassettes were produced and introduced into vectors for mammalian expression. Each RGN was codon-optimized for human expression (SEQ ID NOs: 148-167), and operably fused at the 5' end to an SV40 nuclear localization sequence (NLS; SEQ ID NO: 168) and to 3xFLAG tags (SEQ ID NO: 169), and operably fused at the 3' end to nucleoplasmin NLS sequences (SEQ ID NO: 170). Two copies of the NLS sequence were used, operably fused in tandem. Each expression cassette was under control of a cytomegalovirus (CMV) promoter (SEQ ID NO: 171). It is known in the art that the CMV transcription enhancer (SEQ ID NO: 172) may also be included in constructs comprising the CMV promoter. Guide RNA expression constructs encoding a single gRNA each under the control of a human RNA polymerase III U6 promoter (SEQ ID NO: 173) were produced and introduced into the pTwist High Copy Amp vector. Sequences for the target sequences for each guide are in Table 3.

Several of the constructs described above were introduced into mammalian cells. One day prior to transfection, $1x10^5$ HEK293T cells (Sigma) were plated in 24-well dishes in Dulbecco's modified Eagle medium (DMEM) plus 10% (vol/vol) fetal bovine serum (Gibco) and 1% Penicillin-Streptomycin (Gibco). The next day when the cells were at 50-60% confluency, 500 ng of an RGN expression plasmid plus 500 ng of a single gRNA expression plasmid were co-transfected using 1.5 µL of Lipofectamine 3000 (Thermo Scientific) per well, following the manufacturer's instructions. After 48 hours of growth, total genomic DNA was harvested using a genomic DNA isolation kit (Machery-Nagel) according to the manufacturer's instructions.

The total genomic DNA was then analyzed to determine the rate of editing for each RGN for each genomic target. First, oligonucleotides were produced to be used for PCR amplification and subsequent analysis of the amplified genomic target site. Oligonucleotide sequences used are listed in Table 4.

All PCR reactions were performed using 10 µL of 2X Master Mix Phusion High-Fidelity DNA polymerase (Thermo Scientific) in a 20 µL reaction including 0.5 µM of each primer. Large genomic regions encompassing each target gene were first amplified using PCR#1 primers, using a program of: 98°C., 1 min; 30 cycles of [98°C., 10 sec; 62°C., 15 sec; 72°C., 5 min]; 72°C., 5 min; 12°C., forever. One microliter of this PCR reaction was then further amplified using primers specific for each guide (PCR#2 primers), using a program of: 98°C., 1 min; 35 cycles of [98°C., 10 sec; 67°C., 15 sec; 72°C., 30 sec]; 72°C., 5 min; 12°C., forever. Primers for PCR#2 include Nextera Read 1 and Read 2 Transposase Adapter overhang sequences for Illumina sequencing.

A number of different genes in the human genome were targeted for RNA-guided cleavage. These loci are included in Table 3 below, along with the reference to the SEQ ID NO of the sgRNA. The indel percentage, which is an indication of RGN activity, is also shown.

Table 3: Target and sgRNA sequences for guide RNAs used to test gene editing activity in mammalian cells

Editing System	Gene Name	Name	sgRNA (SEQ ID NO.)	Target Sequence (SEQ ID NO.)
LPG10165	TRAC	SGN003657	174	344
LPG10165	TRAC	SGN003656	175	345
LPG10165	TRAC	SGN003655	176	346
LPG10165	TRAC	SGN003654	177	347
LPG10165	B2M	SGN003653	178	348
LPG10165	B2M	SGN003652	179	349
LPG10165	B2M	SGN003651	180	350
LPG10165	B2M	SGN003650	181	351
LPG10166	TRAC	SGN003665	182	352
LPG10166	TRAC	SGN003664	183	353
LPG10166	TRAC	SGN003663	184	354
LPG10166	TRAC	SGN003662	185	355
LPG10166	TRAC	SGN003661	186	356
LPG10166	B2M	SGN003660	187	357
LPG10166	B2M	SGN003659	188	358
LPG10166	B2M	SGN003658	189	359
LPG10167	TRAC	SGN003673	190	360
LPG10167	TRAC	SGN003672	191	361
LPG10167	TRAC	SGN003671	192	362
LPG10167	TRAC	SGN003670	193	363
LPG10167	B2M	SGN003669	194	364
LPG10167	B2M	SGN003668	195	365
LPG10167	B2M	SGN003667	196	366
LPG10167	B2M	SGN003666	197	367
LPG10169	B2M	SGN001987	821	826
LPG10169	B2M	SGN001988	822	827
LPG10169	B2M	SGN001990	823	828
LPG10169	TRAC	SGN001991	824	829
LPG10169	TRAC	SGN001992	825	830
LPG10171	TRAC	SGN003687	198	345
LPG10171	TRAC	SGN003686	199	368
LPG10171	TRAC	SGN003685	200	369
LPG10171	TRAC	SGN003684	201	370
LPG10171	B2M	SGN003683	202	371
LPG10171	B2M	SGN003682	203	372
LPG10171	B2M	SGN003681	204	373

Editing System	Gene Name	Name	sgRNA (SEQ ID NO.)	Target Sequence (SEQ ID NO.)
LPG10171	B2M	SGN003680	205	374
LPG10186	TRAC	SGN003695	206	375
LPG10186	TRAC	SGN003694	207	376
LPG10186	TRAC	SGN003693	208	377
LPG10186	TRAC	SGN003692	209	356
LPG10186	B2M	SGN003691	210	349
LPG10186	B2M	SGN003690	211	378
LPG10186	B2M	SGN003689	212	379
LPG10186	B2M	SGN003688	213	380
LPG10190	TRAC	SGN004196	214	381
LPG10190	TRAC	SGN004195	215	382
LPG10190	B2M	SGN004194	216	383
LPG10190	B2M	SGN004193	217	384
LPG10190	TRAC	SGN004192	218	385
LPG10190	TRAC	SGN004191	219	361
LPG10190	TRAC	SGN004190	220	360
LPG10190	TRAC	SGN004189	221	386
LPG10190	B2M	SGN004188	222	365
LPG10190	B2M	SGN004187	223	387
LPG10190	B2M	SGN004186	224	366
LPG10190	B2M	SGN004185	225	364
LPG10191	TRAC	SGN004205	226	388
LPG10191	TRAC	SGN004204	227	389
LPG10191	TRAC	SGN004203	228	390
LPG10191	TRAC	SGN004202	229	391
LPG10191	Dnmt3a	SGN004201	230	392
LPG10191	Dnmt3a	SGN004200	231	393
LPG10191	Dnmt3a	SGN004199	232	394
LPG10191	Dnmt3a	SGN004198	233	395
LPG10191	Dnmt3a	SGN004197	234	396
LPG10194	B2M	SGN005313	235	397
LPG10194	B2M	SGN005312	236	398
LPG10194	B2M	SGN005311	237	399
LPG10194	B2M	SGN005310	238	400
LPG10194	B2M	SGN005309	239	372
LPG10194	TRAC	SGN005308	240	401
LPG10194	TRAC	SGN005307	241	402
LPG10194	TRAC	SGN005306	242	403
LPG10194	TRAC	SGN005305	243	404

Editing System	Gene Name	Name	sgRNA (SEQ ID NO.)	Target Sequence (SEQ ID NO.)
LPG10194	TRAC	SGN005304	244	405
LPG10195	B2M	SGN005323	245	406
LPG10195	B2M	SGN005322	246	372
LPG10195	B2M	SGN005321	247	407
LPG10195	B2M	SGN005320	248	408
LPG10195	B2M	SGN005319	249	409
LPG10195	TRAC	SGN005318	250	410
LPG10195	TRAC	SGN005317	251	411
LPG10195	TRAC	SGN005316	252	361
LPG10195	TRAC	SGN005315	253	360
LPG10195	TRAC	SGN005314	254	356
LPG10196	B2M	SGN005333	255	412
LPG10196	B2M	SGN005332	256	413
LPG10196	B2M	SGN005331	257	414
LPG10196	B2M	SGN005330	258	415
LPG10196	B2M	SGN005329	259	416
LPG10196	TRAC	SGN005328	260	417
LPG10196	TRAC	SGN005327	261	418
LPG10196	TRAC	SGN005326	262	419
LPG10196	TRAC	SGN005325	263	420
LPG10196	TRAC	SGN005324	264	421
LPG10197	B2M	SGN005343	265	422
LPG10197	B2M	SGN005342	266	423
LPG10197	B2M	SGN005341	267	424
LPG10197	B2M	SGN005340	268	425
LPG10197	B2M	SGN005339	269	426
LPG10197	TRAC	SGN005338	270	391
LPG10197	TRAC	SGN005337	271	427
LPG10197	TRAC	SGN005336	272	428
LPG10197	TRAC	SGN005335	273	429
LPG10197	TRAC	SGN005334	274	430
LPG10198	B2M	SGN005353	275	431
LPG10198	B2M	SGN005352	276	432
LPG10198	B2M	SGN005351	277	433
LPG10198	B2M	SGN005350	278	434
LPG10198	B2M	SGN005349	279	435
LPG10198	TRAC	SGN005348	280	436
LPG10198	TRAC	SGN005347	281	437
LPG10198	TRAC	SGN005346	282	438

Editing System	Gene Name	Name	sgRNA (SEQ ID NO.)	Target Sequence (SEQ ID NO.)
LPG10198	TRAC	SGN005345	283	439
LPG10198	TRAC	SGN005344	284	440
LPG10200	B2M	SGN005363	285	441
LPG10200	B2M	SGN005362	286	442
LPG10200	B2M	SGN005361	287	426
LPG10200	B2M	SGN005360	288	443
LPG10200	B2M	SGN005359	289	379
LPG10200	TRAC	SGN005358	290	444
LPG10200	TRAC	SGN005357	291	445
LPG10200	TRAC	SGN005356	292	446
LPG10200	TRAC	SGN005355	293	447
LPG10200	TRAC	SGN005354	294	360
LPG10203	B2M	SGN005372	295	448
LPG10203	B2M	SGN005371	296	449
LPG10203	B2M	SGN005370	297	450
LPG10203	B2M	SGN005369	298	451
LPG10203	B2M	SGN005368	299	452
LPG10203	TRAC	SGN005367	300	345
LPG10203	TRAC	SGN005366	301	346
LPG10203	TRAC	SGN005365	302	347
LPG10203	TRAC	SGN005364	303	453
LPG10204	B2M	SGN005382	304	454
LPG10204	B2M	SGN005381	305	455
LPG10204	B2M	SGN005380	306	456
LPG10204	TRAC	SGN005379	307	457
LPG10204	TRAC	SGN005378	308	458
LPG10204	TRAC	SGN005377	309	459
LPG10204	TRAC	SGN005376	310	460
LPG10204	TRAC	SGN005375	311	461
LPG10204	B2M	SGN005374	312	372
LPG10204	B2M	SGN005373	313	462
LPG10205	B2M	SGN005392	314	463
LPG10205	B2M	SGN005391	315	455
LPG10205	B2M	SGN005390	316	456
LPG10205	TRAC	SGN005389	317	457
LPG10205	TRAC	SGN005388	318	458
LPG10205	TRAC	SGN005387	319	459
LPG10205	TRAC	SGN005386	320	460
LPG10205	TRAC	SGN005385	321	461

Editing System	Gene Name	Name	sgRNA (SEQ ID NO.)	Target Sequence (SEQ ID NO.)
LPG10205	B2M	SGN005384	322	372
LPG10205	B2M	SGN005383	323	462
LPG10207	B2M	SGN005412	324	463
LPG10207	B2M	SGN005411	325	364
LPG10207	B2M	SGN005410	326	365
LPG10207	B2M	SGN005409	327	366
LPG10207	B2M	SGN005408	328	379
LPG10207	TRAC	SGN005407	329	447
LPG10207	TRAC	SGN005406	330	464
LPG10207	TRAC	SGN005405	331	360
LPG10207	TRAC	SGN005404	332	363
LPG10207	TRAC	SGN005403	333	361
LPG10205	B2M	SGN005393	334	462
LPG10205	B2M	SGN005394	335	372
LPG10205	TRAC	SGN005395	336	461
LPG10205	TRAC	SGN005396	337	460
LPG10205	TRAC	SGN005397	338	459
LPG10205	TRAC	SGN005398	339	458
LPG10205	TRAC	SGN005399	340	457
LPG10205	B2M	SGN005400	341	456
LPG10205	B2M	SGN005401	342	455
LPG10205	B2M	SGN005402	343	454

Table 4: Oligonucleotides for detection of gene editing activity in mammalian cells

Primer Set	Forward Primer (SEQ ID NO)	Reverse Primer (SEQ ID NO)	Targets amplified by primer (SEQ ID NOs.)
1	465	466	349-351, 357-359, 364-367, 371-374, 378-380, 383-384, 387, 406, 407-409, 414-416, 426, 431-435, 441-443, 448-452, 454-456, 462-463
2	467	468	387-400, 412-413, 422-425, 463
3	469	470	392
4	471	472	396
5	473	474	395
6	475	476	393-394
7	477	478	344-347, 352-356, 360-363, 368-370, 375-377, 381-382, 385-391, 401-405, 410-411, 417-421, 427-430, 436-440, 444-447, 453, 457-461
8	479	480	464

Purified genomic DNA was subjected to PCR#1 and PCR#2 as above. Following the second PCR amplification, DNA was cleaned using a PCR cleanup kit (Zymo) according to the manufacturer's instructions and eluted in water. 200-500ng of purified PCR#2 product was combined with 2 μL of 10X NEB Buffer 2 and water in a 20 μL reaction and annealed to form heteroduplex DNA using a program of: 95°C., 5 min; 95-85°C, cooled at a rate of 2°C / sec; 85-25°C, cooled at a rate of 0.1°C / sec.; 12°C., forever. Following annealing 5 μL of DNA was removed as a no enzyme control, and 1 μL of T7 Endonuclease I (NEB) was added and the reaction incubated at 37°C for 1 hr. After incubation 5x FlashGel loading dye (Lonza) was added and 5 μL of each reaction and controls were analyzed by a 2.2% agarose FlashGel (Lonza) using gel electrophoresis. Following visualization of the gel, the percentage of non-homologous end joining (NHEJ) was determined using the following equation: %NHEJ events = 100 x [1-(1-fraction cleaved)(½)], where (fraction cleaved) is defined as: (density of digested products)/(density of digested products + undigested parental band).

Additionally, products from PCR#2 containing Illumina overhang sequences underwent library preparation following the Illumina 16S Metagenomic Sequencing Library protocol. Deep sequencing was performed on an Illumina Mi-Seq platform by a service provider (MOGene). Typically, 200,000 of 250 bp paired-end reads (2 x 100,000 reads) are generated per amplicon. The reads were analyzed using CRISPResso (Pinello, et al. 2016 Nature Biotech, 34:695-697) to calculate the rates of editing. Output alignments were hand-curated to confirm insertion and deletion sites as well as identify microhomology sites at the recombination sites. The rates of editing are shown in Table 5. All experiments were performed in human cells. A selected example of mammalian genome editing is shown in Table 5.

Table 5: Editing Rates for RGNs in Mammalian Cells

Name	Guide Name	Percent Indels
LPG10166	SGN003661	0.24
LPG10166	SGN003662	0
LPG10166	SGN003663	0
LPG10166	SGN003664	0
LPG10166	SGN003665	0.13
LPG10167	SGN003666	0.14
LPG10167	SGN003667	1.16
LPG10167	SGN003668	0.79
LPG10167	SGN003669	0.34
LPG10167	SGN003670	0
LPG10167	SGN003671	0
LPG10167	SGN003672	0.21
LPG10167	SGN003673	0.16
LPG10168	SGN003666	0.16
LPG10168	SGN003668	0.61

Name	Guide Name	Percent Indels
LPG10168	SGN003669	3.68
LPG10168	SGN003670	0
LPG10168	SGN003671	0
LPG10168	SGN003672	0
LPG10168	SGN003673	0.31
LPG10169	SGN001987	40.09
LPG10169	SGN001988	13.03
LPG10169	SGN001990	3.9
LPG10169	SGN001991	2.53
LPG10169	SGN001992	0.56
LPG10171	SGN003680	0.22
LPG10171	SGN003681	0.96
LPG10171	SGN003682	15.16
LPG10171	SGN003683	12.13
LPG10171	SGN003684	0.05
LPG10171	SGN003685	0
LPG10171	SGN003686	0
LPG10171	SGN003687	0.25
LPG10186	SGN003688	0
LPG10186	SGN003689	8.93
LPG10186	SGN003690	0
LPG10186	SGN003691	0.89
LPG10186	SGN003692	0.69
LPG10186	SGN003693	0
LPG10186	SGN003694	0
LPG10186	SGN003695	0.06

Example 5: Testing different delivery formats

To determine if the RGNs are capable of delivery in different formats, mRNA and RNP nucleofection delivery was tested with primary T-cells. Purified CD3+ T-cells or peripheral blood mononuclear cells (PBMCs) were thawed, activated using CD3/CD28 beads (ThermoFisher) for 3 days, then nucleofected using the Lonza 4D-Nucleofector X unit and Nucleocuvette strips. The P3 Primary Cell kit was used for both mRNA and RNP delivery. Cells were transfected using the EO-115 and EH-115 programs for mRNA and RNP delivery respectively. Cells were cultured in CTS OpTimizer T cell expansion medium (ThermoFisher) containing IL-2, IL-7, and IL-15 (Miltenyi Biotec) for 4 days post nucleofection before being harvested using a Nucleospin Tissue genomic DNA isolation kit (Machery Nagel).

Amplicons surrounding the editing sites were generated by PCR using primers identified in Table 4 and subjected to NGS sequencing using the Illumina Nextera platform using 2x250bp paired end sequencing following the method in Example 4.

Example 6: Identification of Disease Targets

A database of clinical variants was obtained from NCBI ClinVar database, which is available through the world wide web at the NCBI ClinVar website. Pathogenic Single Nucleotide Polymorphisms (SNPs) were identified from this list. Using the genomic locus information, CRISPR targets in the region overlapping and surrounding each SNP were identified. A selection of SNPs that can be corrected using base editing in combination with the RGNs of the invention to target the causal mutation ("Casl Mut.") is listed in Table 6. In Table 6, only one alias of each disease is listed. The "RS#" corresponds to the RS accession number through the SNP database at the NCBI website. The AlleleID corresponds to a causal allele accession number, and the Chromosome Accession number also provides accession reference information found through the NCBI website. Table 6 also provides genomic target sequence information suitable for the RGN listed for each disease. The target sequence information also provides a sample protospacer sequence for the production of the necessary sgRNA, along with the presently disclosed crRNA repeat and tracrRNA sequences for the corresponding RGN.

Table 6: Disease Targets for RGNs

Disease	RS# (dbSNP)	RGN	Casl. Mut	Allele ID	Gene	Target Sequence (SEQ ID NO)
Stargardt disease	1800553	LPG10166	G>A	22927	ABCA4	573
Stargardt disease	1800553	LPG10167,LPG10168,LPG10207	G>A	22927	ABCA4	574
Stargardt disease	1800553	LPG10171	G>A	22927	ABCA4	575
Stargardt disease	1800553	LPG10186	G>A	22927	ABCA4	576
Stargardt disease	1800553	LPG10190	G>A	22927	ABCA4	577
Stargardt disease	1800553	LPG10194	G>A	22927	ABCA4	578
Stargardt disease	1800553	LPG10195	G>A	22927	ABCA4	579
Stargardt disease	1800553	LPG10200	G>A	22927	ABCA4	580
Stargardt disease	1800553	LPG10203	G>A	22927	ABCA4	581
Stargardt disease	1800553	LPG10204,LPG10205,LEG206	G>A	22927	ABCA4	582
Phenylketonuria	5030857	LPG10191	C>T	98638	PAH	583
Phenylketonuria	5030858	LPG10166	C>T	15616	PAH	584
Phenylketonuria	5030858	LPG10167,LPG10168,LPG10207	C>T	15616	PAH	585
Phenylketonuria	5030858	LPG10171	C>T	15616	PAH	586
Phenylketonuria	5030858	LPG10186	C>T	15616	PAH	587
Phenylketonuria	5030858	LPG10190	C>T	15616	PAH	588
Phenylketonuria	5030858	LPG10195	C>T	15616	PAH	589
Phenylketonuria	5030858	LPG10197	C>T	15616	PAH	590
Phenylketonuria	5030858	LPG10198	C>T	15616	PAH	591

Disease	RS# (dbSNP)	RGN	Casl. Mut	Allele ID	Gene	Target Sequence (SEQ ID NO)
Phenylketonuria	5030858	LPG10200	C>T	15616	PAH	592
Hyperphenylalaninemia, non-pku	5030860	LPG10203	A>G	15632	PAH	593
Alpha-1-antitrypsin deficiency	28929474	LPG10166	G>A	33006	SERPINA1	594
Alpha-1-antitrypsin deficiency	28929474	LPG10167,LPG10168,LPG10207	G>A	33006	SERPINA1	595
Alpha-1-antitrypsin deficiency	28929474	LPG10171	G>A	33006	SERPINA1	596
Alpha-1-antitrypsin deficiency	28929474	LPG10186	G>A	33006	SERPINA1	597
Alpha-1-antitrypsin deficiency	28929474	LPG10190	G>A	33006	SERPINA1	598
Alpha-1-antitrypsin deficiency	28929474	LPG10194	G>A	33006	SERPINA1	599
Alpha-1-antitrypsin deficiency	28929474	LPG10196	G>A	33006	SERPINA1	600
Alpha-1-antitrypsin deficiency	28929474	LPG10200	G>A	33006	SERPINA1	601
Alpha-1-antitrypsin deficiency	28929474	LPG10203	G>A	33006	SERPINA1	602
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10166	G>A	22745	PMM2	603
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10169	G>A	22745	PMM2	604
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10171	G>A	22745	PMM2	605
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10186	G>A	22745	PMM2	606

Disease	RS# (dbSNP)	RGN	Casl. Mut	Allele ID	Gene	Target Sequence (SEQ ID NO)
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10195	G>A	22745	PMM2	607
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10197	G>A	22745	PMM2	608
Carbohydrate-deficient glycoprotein syndrome type I	28936415	LPG10200	G>A	22745	PMM2	609
Familial Mediterranean fever	28940579	LPG10191	T>C	17579	MEFV	610
Carnitine palmitoyltransferase II deficiency	74315294	LPG10171	C>T	23992	CPT2	611
Carnitine palmitoyltransferase II deficiency	74315294	LPG10186	C>T	23992	CPT2	612
Carnitine palmitoyltransferase II deficiency	74315294	LPG10190	C>T	23992	CPT2	613
Carnitine palmitoyltransferase II deficiency	74315294	LPG10200	C>T	23992	CPT2	614
Carnitine palmitoyltransferase II deficiency	74315294	LPG10203	C>T	23992	CPT2	615
Carnitine palmitoyltransferase II deficiency	74315294	LPG10204,LPG10205,LEG206	C>T	23992	CPT2	616
Deafness	80338945	LEG165	T>C	32055	GJB2	617
Deafness	80338945	LEG165	T>C	32055	GJB2	618

Disease	RS# (dbSNP)	RGN	Casl. Mut	Allele ID	Gene	Target Sequence (SEQ ID NO)
Deafness	80338945	LPG10166	T>C	32055	GJB2	619
Deafness	80338945	LPG10167,LPG10168,LPG10207	T>C	32055	GJB2	620
Deafness	80338945	LPG10169	T>C	32055	GJB2	621
Deafness	80338945	LPG10171	T>C	32055	GJB2	622
Deafness	104894396	LPG10186	G>A	32041	GJB2	623
Deafness	104894396	LPG10190	G>A	32041	GJB2	624
Deafness	104894396	LPG10194	G>A	32041	GJB2	625
Deafness	104894396	LPG10195	G>A	32041	GJB2	626
Deafness	104894396	LPG10196	G>A	32041	GJB2	627
Deafness	104894396	LPG10200	G>A	32041	GJB2	628
Deafness	104894396	LPG10204,LPG10205,LEG206	G>A	32041	GJB2	629
Familial hypercholesterolemia	121908026	LPG10190	C>T	18725	LDLR	630
Familial hypercholesterolemia	121908026	LPG10203	C>T	18725	LDLR	631
Familial hypercholesterolemia	137929307	LPG10186	G>A	171217	LDLR	632
Familial hypercholesterolemia	137929307	LPG10194	G>A	171217	LDLR	633
Familial hypercholesterolemia	137929307	LPG10195	G>A	171217	LDLR	634
Familial hypercholesterolemia	137929307	LPG10196	G>A	171217	LDLR	635
Familial hypercholesterolemia	137929307	LPG10200	G>A	171217	LDLR	636
Noonan syndrome	267607048	LEG165	A>G	21860	SHOC2	637
Noonan syndrome	267607048	LPG10169	A>G	21860	SHOC2	638
Noonan syndrome	267607048	LPG10195	A>G	21860	SHOC2	639

Disease	RS# (dbSNP)	RGN	Casl. Mut	Allele ID	Gene	Target Sequence (SEQ ID NO)
Noonan syndrome	267607048	LPG10197	A>G	21860	SHOC2	640
Noonan syndrome	267607048	LPG10200	A>G	21860	SHOC2	641

Example 7: Targeting mutations responsible for Hurler Syndrome

The following describes a potential treatment for Hurler Syndrome, also referred to as MPS-1, using an RNA directed base editing system that corrects a mutation responsible for Hurler syndrome in a large proportion of patients with the disease. This approach utilizes a base editing fusion protein that is RNA guided and that can be packaged into a single AAV vector for delivery to a wide range of tissue types. Depending on the exact regulatory elements and base editor domain used, it may also be possible to engineer a single vector that encodes for both the base editing fusion protein and a single guide RNA to target the diseased locus.

Example 7.1: Identifying RGN with ideal PAM

The genetic disease MPS-1 is a lysosomal storage disease characterized at the molecular level by the accumulation of dermatan sulfate and heparan sulfate in lysosomes. This disease is generally an inherited genetic disorder caused by mutations in the IDUA gene (NCBI Reference sequence NG_008103.1), which encodes α-L-iduronidase. The disease is a result of a deficiency of α-L-iduronidase. The most common IDUA mutations found in studies of individuals of Northern European background are W402X and Q70X, both nonsense mutations resulting in premature termination of translation (Bunge et al. (1994), Hum. Mol. Genet, 3(6): 861-866, herein incorporated by reference). Reversion of a single nucleotide would restore the wild-type coding sequence and result in protein expression controlled by the endogenous regulatory mechanisms of the genetic locus.

The W402X mutation of the human Idua gene accounts for a high proportion of MPS-1H cases. Base editors can target a narrow sequence window relative to the binding site of the protospacer component of the guide RNA and thus the presence of a PAM sequence a specific distance from the target locus is essential for the success of the strategy. Given the constraints that the target mutation must be on the exposed non-target strand (NTS) during the interaction of the base editing protein and that the footprint of the RGN domain will block access to the region near the PAM, an accessible locus is thought to be 10-30 bp from the PAM. To avoid editing and mutagenesis of other nearby adenosine bases in this window, different linkers are screened. The ideal window is 12-16 bp from the PAM.

PAM sequences compatible with LPG10165, LPG10167, LPG10168, LPG10207, LPG10171, LPG10186, LPG10190, LPG10194, LPG10195, LPG10200, or LPG10203 are readily apparent at the genetic locus. These nucleases have a PAM sequence of 5' nnnnGRRA-3' (SEQ ID NO: 127), 5'-nnnnCC-3' (SEQ ID NO: 132), 5'- nnnnGnnA-3' (SEQ ID NO: 134)), 5'- nVG-3' (SEQ ID NO: 135), 5'- nnnnCCD-3' (SEQ ID NO: 136), 5'- VGGNR-3' (SEQ ID NO: 139), 5'- NNRA-3' (SEQ ID NO: 140), 5'- nGR-3' (SEQ ID NO: 144) and 5'-nnnnGGG-3' (SEQ ID NO: 145) respectively.

Cas9 from *S. pyogenes* (SpyCas9) requires a PAM sequence of NGG (SEQ ID NO: 642), which is present near the W402X locus, but the size of SpyCas9 prevents packaging into a single AAV vector, and thus forgoes the aforementioned advantages of this approach. While a dual delivery strategy may be employed (for example, Ryu et al, (2018), Nat. Biotechnol., 36(6): 536-539, herein incorporated by reference), it would add significant manufacturing complexity and cost. Additionally, dual viral vector delivery significantly decreases the efficiency of gene correction, since a successful edit in a given cell requires infection with both vectors and assembly of the fusion protein in the cell.

A commonly used Cas9 ortholog from *S. aureus* (SauCas9) is considerably smaller in size relative to SpyCas9 but has a more complex PAM requirement – NGRRT (SEQ ID NO: 643). This sequence is not within a range expected to be useful for base editing of the causative locus.

Example 7.2: RGN fusion constructs and sgRNA sequences

A DNA sequence encoding a fusion protein with the following domains is produced using standard molecular biology techniques: 1) an RGN domain with mutations that inactivate the DNA cleavage activity ("dead" or "nickase"); 2) an adenosine deaminase useful for base editing. All constructs described in the table below comprise a fusion protein with the base editing active domain, in this example LPG50148 (SEQ ID NO: 644), operably fused to the N-terminal end of a nickase version of RGN LPG10165, LPG10167, LPG10168, LPG10171, LPG10186, LPG10190, LPG10194, LPG10195, LPG10200, LPG10203, or LPG10207 (set forth as SEQ ID NOs: 645-655, respectively). Other adenosine deaminases useful for base editing DNA may also be used (see for example PCT Application Publication Nos. WO 2020/139783 and WO 2022/056254, each of which is herein incorporated by reference in its entirety). It is known in the art that a fusion protein could also be made with the base-editing enzyme at the C-terminal end of the RGN. Additionally, the RGN and the base editor of the fusion protein are typically separated by a linker amino sequence. It is known in the art that lengths of standard linkers range from 15-30 amino acids.

Table 7: Constructs for RNA-targeted base editing

SEQ ID NO.	Construct	RGN	Dead (D)or Nickase (N)	Base editor
656	SV40-LPG50148-Linker-nLPG10165 -Linker-Nuc	LPG10165	N	LPG50148
657	SV40-LPG50148-Linker-nLPG10167 -Linker-Nuc	LPG10167	N	LPG50148
658	SV40-LPG50148-Linker-nLPG10168 -Linker-Nuc	LPG10168	N	LPG50148
659	SV40-LPG50148-Linker-nLPG10171 -Linker-Nuc	LPG10171	N	LPG50148
660	SV40-LPG50148-Linker-nLPG10186 -Linker-Nuc	LPG10186	N	LPG50148
661	SV40-LPG50148-Linker-nLPG10190 -Linker-Nuc	LPG10190	N	LPG50148

SEQ ID NO.	Construct	RGN	Dead (D)or Nickase (N)	Base editor
662	SV40-LPG50148-Linker-nLPG10194 -Linker-Nuc	LPG10194	N	LPG50148
663	SV40-LPG50148-Linker-nLPG10195 -Linker-Nuc	LPG10195	N	LPG50148
664	SV40-LPG50148-Linker-nLPG10200 -Linker-Nuc	LPG10200	N	LPG50148
665	SV40-LPG50148-Linker-nLPG10203 -Linker-Nuc	LPG10203	N	LPG50148
666	SV40-LPG50148-Linker-nLPG10207 -Linker-Nuc	LPG10207	N	LPG50148

The accessible editing sites of an RGN are determined by the PAM sequence. When combining an RGN with a base editing domain, the target residue for editing must reside on the non-target strand (NTS), since the NTS is single stranded while the RGN is associated with the locus. Evaluating a number of nucleases and corresponding guide RNAs enables the selection of the most appropriate gene editing tool for this particular locus. Several potential PAM sequences that can be targeted by the constructs described above in the human *Idua* gene are in the proximity of the mutant nucleotide responsible for the W402X mutation. A sequence encoding a guide RNA transcript containing 1) a "spacer" that is complementary to the non-coding DNA strand at the disease locus; and 2) RNA sequence required for association of the guide RNA with the RGN is also produced. Non-limiting exemplary spacers would be complementary to the non-coding DNA strand of the target sequences set forth as SEQ ID NOs: 667-677 for the RGNs LPG10165, LPG10167, LPG10168, LPG10207, LPG10171, LPG10186, LPG10190, LPG10194, LPG10195, LPG10200, LPG10203. Non-limiting examples of sgRNA molecules include those comprising such spacers and the crRNA repeat and tracrRNA sequences disclosed herein for the corresponding RGN, and similar sgRNAs that may be devised by one of skill in the art, can be evaluated for their efficiency in directing the base editors above to the locus of interest.

Example 7.3: Assay for activity in cells from Hurler disease patients

To verify the genotype strategy and evaluate the constructs described above, fibroblasts from Hurler disease patients are used. A vector is designed containing appropriate promoters upstream of the fusion protein coding sequence and the sgRNA encoding sequence for expression of these in human cells, similar to those vectors described in Example 4. It is recognized that promoters and other DNA elements (for example enhancers, or terminators) which either are known for high levels of expression in human cells or may specifically express well in fibroblast cells may also be used. The vector is transfected into the fibroblasts using standard techniques, for example transfection similar to what is described in Example 4. Alternatively, electroporation may be used. The cells are cultured for 1 - 3 days. Genomic DNA (gDNA) is isolated using standard techniques. The editing

efficiency is determined by performing a qPCR genotyping assay and/or next generation sequencing on the purified gDNA, as described further below.

Taqman™ qPCR analysis utilizes probes specific for the wild-type and mutant allele. These probes bear fluorophores which are resolved by their spectral excitation and/or emission properties using a qPCR instrument. A genotyping kit containing PCR primers and probes can be obtained commercially (i.e., Thermo Fisher Taqman™ SNP genotyping assayID C_27862753_10 for SNP ID rs121965019) or designed. An example of a designed primer and probe set is shown in Table 8.

Table 8: RT-PCR primers and probes

Description	Sequence	SEQ ID NO.
Forward Amplification Primer	5'-GACTCCTTCACCAAG-3'	678
Reverse Amplification Primer	5'-GTAGATCAGCACCG-3'	679
Wild Type Probe	5'-CTCT**G**GGCCGAAGT-3'	680
W402X Probe	5'-CTCT**A**GGCCGAAGT-3'	681

Following the editing experiment, the gDNA is subjected to qPCR analysis using standard methods and the primers and probes described above. Expected results are shown in Table 9. This *in vitro* system can be used to expediently evaluate constructs and choose one with high editing efficiency for further studies. The systems will be evaluated in comparison with cells with and without the W402X mutation, and preferably with some that are heterozygous for this mutation. The Ct values will be compared to either a reference gene or the total amplification of the locus using a dye such as Sybr green.

Table 9: Expected qPCR results

Genotype	Transfected with base editor	Expected PCR result
$Idua^{WT/WT}$	No	Homozygous WT
$Idua^{WT/W402X}$	No	Heterozygous: 50% WT, 50% W402X
$Idua^{W402X/W402X}$	No	Homozygous W402X
$Idua^{W402X/W402X}$	Yes	Variable

The tissues can also be analyzed by next generation sequencing. Primer binding sites such as the ones shown below (Table 10), or other suitable primer binding sites that can be identified by a person of skill in the art, can be used. Following PCR amplification, products containing Illumina Nextera XT overhang sequences undergo library preparation following the Illumina 16S Metagenomic Sequencing Library protocol. Deep sequencing is performed on an Illumina Mi-Seq platform. Typically, 200,000 of 250 bp paired-end reads (2 x 100,000 reads) are generated per amplicon. The reads are analyzed using CRISPResso (Pinello et al., 2016) to calculate the rates of editing. Output alignments are hand-curated to confirm insertion and deletion sites as well as identify microhomology sites at the recombination sites.

Table 10: NGS primer binding sites

Direction	Sequence	SEQ ID NO.
Forward	5'-ACTTCCTCCAGCC-3'	682
Reverse	5'-GAACCCCGGCTTA-3'	683

Western blotting of cell lysate of transfected cells and control cells using an anti-IDUA antibody is performed to verify expression of the full-length protein and an enzyme activity assay on the cell lysate using substrate 4-methylumbelliferyl a-L-iduronide verifies that the enzyme is catalytically active (Hopwood et al., Clin.Chim. ACta (1979), 92(2): 257-265, incorporated by reference herein). These experiments are performed in comparison with the original $Idua^{W402X/W402X}$ cell line (without transfection), the $Idua^{W402X/W402X}$ cell line transfected with the base editing construct and a random guide sequence, and a cell line expressing wild-type IDUA.

Example 7.4: Disease treatment validation in a murine model

To verify the efficacy of this therapeutic approach, a mouse model with a nonsense mutation in the analogous amino acid is used. The mouse strain bears a W392X mutation in its Idua gene (Gene ID: 15932) which corresponds to the homologous mutation in Hurler syndrome patients (Bunge et al., (1994), Hum. Mol. Genet. 3(6): 861-866, incorporated by reference herein). This locus comprises a distinct nucleotide sequence relative to that in humans, which lacks the PAM sequence necessary for correction with the base editors described in the previous examples, and thus necessitates design of a distinct fusion protein to perform the nucleotide correction. Amelioration of the disease in this animal can validate the therapeutic approach of correcting the mutation in tissues accessible by a gene delivery vector.

Mice homozygous for this mutation display a number of phenotypic characteristics similar to Hurler syndrome patients. A base editing-RGN fusion protein as described above (Table 7) along with an RNA guide sequence are incorporated into an expression vector that allows protein expression and RNA transcription in mice. A study design is shown below in Table 11. The study includes groups that are treated with a high dose of the expression vector comprising the base-editing fusion protein and RNA guide sequence, a low dose of same expression vector, control which is the model mouse treated with an expression vector that does not comprise the base editing fusion protein or the guide RNA, and a second control which is a wild type mouse treated with the same empty vector.

Table 11: Genome editing experiment in murine model

Group	Mouse strain	N	Treatment
1	Idua-W392X[1]	≥ 5	Low dose of vector
2	Idua-W392X	≥ 5	High dose of vector
3	Idua-W392X	≥ 5	Vehicle
4	129/Sv (WT)	5	Vehicle

Endpoints to evaluate include body weight, urine GAG excretion, serum IDUA enzymatic activity, IDUA activity in tissues of interest, tissue pathology, genotyping of tissues of interest to verify correction of the SNP, and behavioral and neurological evaluation. Since some endpoints are terminal, additional groups may be added for evaluation of, for example, tissue pathology and tissue IDUA activities before the end of the study. Additional examples of endpoints can be found in published papers establishing Hurler syndrome animal models (Shull et al. (1994), Proc. Natl. Acad. Sci. U.S.A., 91(26): 12937-12941; Wang et al. (2010), Mol. Genet. Metab., 99(1): 62-71; Hartung et al. (2004), Mol. Ther., 9(6): 866-875; Liu et al. (2005), Mol. Ther., 11(1): 35-47; Clarke et al. (1997), Hum. Mol. Genet. 6(4): 503-511; all herein incorporated by reference).

One possible delivery vector utilizes the adeno associated virus (AAV). A vector is produced to include a base editor-dRGN or base editor-nRGN fusion protein coding sequence (for example, those described in Table 7) preceded by a CMV enhancer (SEQ ID NO: 172) and promoter (SEQ ID NO: 171), or other suitable enhancer and promoter combination), optionally a Kozak sequence, and operably fused at the 3' end to a terminator sequence and a poly-adenylation sequence such as the minimal sequence described in Levitt, N.; Briggs, D.; Gil, A.; Proudfoot, N. J. Definition of an Efficient Synthetic Poly(A) Site. Genes Dev. 1989, 3 (7), 1019–1025. The vector may further comprise an expression cassette encoding for a single guide RNA operably linked at its 5' end to a human U6 promoter (SEQ ID NO: 173) or another promoter suitable for production of small non-coding RNAs, and further comprising inverted terminal repeat (ITR) sequences necessary and well-known in the art for packaging into the AAV capsid. Production and viral packaging is performed by standard methods, such as those described in U.S. Patent No. 9,587,250, herein incorporated by reference.

Other possible viral vectors include adenovirus and lentivirus vectors, which are commonly used and would contain similar elements, with different packaging capabilities and requirements. Non-viral delivery methods also be used, such as mRNA and sgRNA encapsulated by lipid nanoparticles (Cullis, P. R. and Allen, T. M. (2013), Adv. Drug Deliv. Rev. 65(1): 36-48; Finn et al. (2018), Cell Rep. 22(9): 2227-2235, both incorporated by reference) hydrodynamic injection of plasmid DNA (Suda T and Liu D,)2007) Mol. Ther. 15(12): 2063-2069, herein incorporated by reference), or ribonucleoprotein complexes of sgRNA and associated with gold nanoparticles (Lee, K.; Conboy, M.; Park, H. M.; Jiang, F.; Kim, H. J.; Dewitt, M. A.; Mackley, V. A.; Chang, K.; Rao, A.; Skinner, C.; et al. Nanoparticle Delivery of Cas9 Ribonucleoprotein and Donor DNA in Vivo Induces Homology-Directed DNA Repair. Nat. Biomed. Eng. 2017, 1 (11), 889–90).

Example 7.5: Disease correction in a murine model with a humanized locus

To evaluate the efficacy of an identical base editor construct as would be used for human therapy, a mouse model in which the nucleotides near W392 are altered to match the sequence in humans around W402 is needed. This can be accomplished by a variety of techniques, including use of an RGN and an HDR template to cut and replace the locus in mouse embryos.

Due to the high degree of amino acid conservation, most nucleotides in the mouse locus can be altered to those of the human sequence with silent mutations as shown in Table 12. The only base changes resulting in altered coding sequence in the resulting engineered mouse genome occur after the introduced stop codon.

Table 12: Nucleotide mutations to generate a humanized mouse locus

	Human (W402X)		Mouse (W392X)		Humanized Mouse	
Feature	**Nucleotide (SEQ ID NO: 818)**	**Encoded AA**	**Nucleotide (SEQ ID NO: 819)**	**Encoded AA**	**Nucleotide (SEQ ID NO: 820)**	**Encoded AA**
Protospacer	G	E	A	G	G	G
	G	E	G	E	G	E
	A		A		A	
	G		A		G	
	C	Q	C	Q	C	Q
	A		A		A	
	G		A		G	
	C	L	C	L	C	L
	T		T		T	
	C		C		C	
	T	STOP	T	STOP	T	STOP
	A		A		A	
	G		G		G	
	G	A	G	A	G	A
	C		C		C	
	C		A		C	
	G	E	G	E	G	E
	A		A		A	
	A		G		A	
	G	V	G	V	G	V
	T		T		T	
	G		C		G	
	T	S	T	S	T	S
	C		C		C	
	G		A		G	
	C	Q	A	K	C	Q

	Human (W402X)		Mouse (W392X)		Humanized Mouse	
PAM, non-critical	A		A		A	
	G		G		G	
	G	A	G	A	G	A
PAM, critical	C		C		C	
	C		T		C	

Upon engineering of this mouse strain, similar experiments will be performed as described in Example 7.4.

Example 8: Targeting mutations responsible for Friedreich Ataxia

The expansion of the trinucleotide repeat sequence causing Friedreich's Ataxia (FRDA) occurs in a defined genetic locus within the FXN gene, referred to as the FRDA instability region. RNA guided nucleases (RGNs) may be used for excising the instability region in FRDA patient cells. This approach requires 1) an RGN and guide RNA sequence that can be programmed to target the allele in the human genome; and 2) a delivery approach for the RGN and guide sequence. Many nucleases used for genome editing, such as the commonly used Cas9 nuclease from *S. pyogenes* (SpCas9), are too large to be packaged into adeno-associated viral (AAV) vectors, especially when considering the length of the SpCas9 gene and the guide RNA in addition to other genetic elements required for functional expression cassettes. This makes a viable approach using SpCas9 unlikely.

RNA guided nucleases of the invention, for example RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208, are uniquely well suited for the excision of the FRDA instability region. RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 have a PAM requirement that is in the vicinity of the FRDA instability region.

Table 13 shows the location of genomic target sequences suitable for targeting RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 to the 5' and 3' flanks of the FRDA instability region. Once at the locus, the RGN would excise the FA instability region. Excision of the region can be verified with Illumina sequencing of the locus. Non-limiting exemplary spacers would be complementary to the non-coding DNA strand of the target sequences set forth in Table 13 and non-

limiting examples of sgRNA molecules include those comprising such spacers and the crRNA repeat and tracrRNA sequences disclosed herein for the corresponding RGN.

Table 13: Genomic target sequences for RGN systems

RGN	Location relative to FRDA instability region	Genome target sequence (SEQ ID NO.)
LPG10165	5'	684
LPG10165	5'	685
LPG10165	3'	686
LPG10165	3'	687
LPG10166	5'	688
LPG10166	5'	689
LPG10166	3'	690
LPG10166	3'	691
LPG10167	5'	692
LPG10167	5'	693
LPG10167	3'	694
LPG10167	3'	695
LPG10168	5'	692
LPG10168	5'	693
LPG10168	3'	694
LPG10168	3'	695
LPG10207	5'	692
LPG10207	5'	693
LPG10207	3'	694
LPG10207	3'	695
LPG10169	5'	696
LPG10169	5'	697
LPG10169	3'	698
LPG10169	3'	699
LPG10171	5'	700
LPG10171	5'	701
LPG10171	3'	702
LPG10171	3'	703
LPG10186	5'	704
LPG10186	5'	705
LPG10186	3'	706
LPG10186	3'	707
LPG10190	5'	708
LPG10190	5'	709
LPG10190	3'	710
LPG10190	3'	711
LPG10191	5'	712
LPG10191	5'	713
LPG10191	3'	714
LPG10191	3'	715
LPG10194	5'	716
LPG10194	5'	717
LPG10194	3'	718
LPG10194	3'	719
LPG10195	5'	720

RGN	Location relative to FRDA instability region	Genome target sequence (SEQ ID NO.)
LPG10195	5'	721
LPG10195	3'	722
LPG10195	3'	723
LPG10196	5'	724
LPG10196	5'	725
LPG10196	3'	726
LPG10196	3'	727
LPG10197	5'	728
LPG10197	5'	729
LPG10197	3'	730
LPG10197	3'	731
LPG10198	5'	732
LPG10198	5'	733
LPG10198	3'	734
LPG10198	3'	735
LPG10200	5'	736
LPG10200	5'	737
LPG10200	3'	738
LPG10200	3'	739
LPG10203	5'	740
LPG10203	5'	741
LPG10203	3'	742
LPG10203	3'	743
LPG10204	5'	744
LPG10204	5'	745
LPG10204	3'	746
LPG10204	3'	747
LPG10205	5'	744
LPG10205	5'	745
LPG10205	3'	746
LPG10205	3'	747
LPG10208	5'	744
LPG10208	5'	745
LPG10208	3'	746
LPG10208	3'	747

Example 9: Targeting mutations responsible for Sickle Cell Diseases

Targeting sequences within the BCL11A enhancer region (SEQ ID NO: 748) may provide a mechanism for increasing fetal hemoglobulin (HbF) to either cure or alleviate the symptoms of sickle cell diseases. For example, genome wide association studies have identified a set of genetic variations at BCL11A that are associated with increased HbF levels. These variations are a collection of SNPs found in non-coding regions of BCL11A that function as a stage-specific, lineage-restricted enhancer region. Further investigation revealed that this BCL11A enhancer is required in erythroid cells for BCL11A expression (Bauer et al, (2013) Science 343:253-257, incorporated by reference herein). The enhancer region was found within intron 2 of the BCL11A gene, and three areas of DNaseI

hypersensitivity (often indicative of a chromatin state that is associated with regulatory potential) in intron 2 were identified. These three areas were identified as "+62", "+58" and "+55" in accordance with the distance in kilobases from the transcription start site of BCL11A. These enhancer regions are roughly 350 (+55); 550 (+58); and 350 (+62) nucleotides in length (Bauer et al., 2013).

Example 9.1: Identifying preferred RGN systems

Here we describe a potential treatment for beta-hemoglobinopathies using an RGN system that disrupts BCL11A binding to its binding site within the HBB locus, which is the gene responsible for making beta-globin in adult hemoglobin. This approach uses NHEJ which is more efficient in mammalian cells. In addition, this approach uses a nuclease of sufficiently small size that can be packaged into a single AAV vector for *in vivo* delivery.

The GATA1 enhancer motif in the human BCL11A enhancer region (SEQ ID NO: 748) is an ideal target for disruption using RNA guided nucleases (RGNs) to reduce BCL11A expression with concurrent re-expression of HbF in adult human erythrocytes (Wu et al. (2019) Nat Med 387:2554). Several PAM sequences compatible with RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 are readily apparent at the genetic locus surrounding this GATA1 site.

An expression cassette encoding a human codon optimized RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, or LPG10208 (set forth as SEQ ID NOs: 750-769) is produced, similar to those described in Example 4. Expression cassettes which express guide RNAs for RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 are also produced. These guide RNAs comprise: 1) a protospacer sequence that is complementary to either the non-coding or coding DNA strand within the BCL11A enhancer locus (the target sequence) and 2) an RNA sequence required for association of the guide RNA with the RGN. Because several potential PAM sequences for targeting by RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 surround the BCL11A GATA1 enhancer motif, several potential guide RNA constructs are produced to determine the best protospacer sequence that produces robust cleavage and NHEJ mediated disruption of the BCL11A GATA1 enhancer sequence. Non-limiting exemplary spacers would be complementary to the non-coding DNA strand of the target sequences set forth in Table 14 and non-limiting examples of sgRNA molecules include those comprising such spacers and the crRNA repeat and tracrRNA sequences disclosed herein for the

corresponding RGN. sGRNAs that target the target genomic sequences in Table 14 are evaluated to direct the RGN to this locus.

Table 14: Target Sequences for BCL11A GATA1 enhancer locus

RGN	Target genomic sequence (SEQ ID NO.)
LPG10165	770
LPG10165	771
LPG10165	772
LPG10166	773
LPG10166	774
LPG10166	775
LPG10167	776
LPG10167	777
LPG10167	778
LPG10168	776
LPG10168	777
LPG10168	778
LPG10207	776
LPG10207	777
LPG10207	778
LPG10169	779
LPG10169	780
LPG10169	781
LPG10171	782
LPG10171	783
LPG10171	784
LPG10186	785
LPG10186	786
LPG10186	787
LPG10190	788
LPG10190	789
LPG10190	790
LPG10191	791
LPG10191	792
LPG10191	793
LPG10194	794
LPG10194	795
LPG10194	796
LPG10195	797
LPG10195	798
LPG10195	799
LPG10196	800
LPG10196	801

RGN	Target genomic sequence (SEQ ID NO.)
LPG10196	802
LPG10197	803
LPG10197	804
LPG10197	805
LPG10198	806
LPG10198	807
LPG10198	808
LPG10200	809
LPG10200	810
LPG10200	811
LPG10203	812
LPG10203	813
LPG10203	814
LPG10204	815
LPG10204	816
LPG10204	817
LPG10205	815
LPG10205	816
LPG10205	817
LPG10208	815
LPG10208	816
LPG10208	817

To evaluate the efficiency with which RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 generate insertions or deletions that disrupt the BCL11A enhancer region, human cell lines such as human embryonic kidney cells (HEK cells) are used. A DNA vector comprising an RGN expression cassette (for example, as described in Example 4) is produced. A separate vector comprising an expression cassette comprising a coding sequence for a guide RNA sequence of Table 16 is also produced. Such an expression cassette may further comprise a human RNA polymerase III U6 promoter (SEQ ID NO: 173), as described in Example 4. Alternatively, a single vector comprising expression cassettes of both the RGN and guide RNA may be used. The vector is introduced into HEK cells using standard techniques such as those described in Example 4, and the cells are cultured for 1-3 days. Following this culture period, genomic DNA is isolated and the frequency of insertions or deletions is determined by using T7 Endonuclease I digestion and/or direct DNA sequencing, as described in Example 4.

A region of DNA encompassing the target BCL11A region is amplified by PCR with primers containing Illumina Nextera XT overhang sequences. These PCR amplicons are either examined for

NHEJ formation using T7 Endonuclease I digestion, or undergo library preparation following the Illumina 16S Metagenomic Sequencing Library protocol or a similar Next Generation Sequencing (NGS) library preparation. Following deep sequencing, the reads generated are analyzed by CRISPResso to calculate rates of editing. Output alignments are hand-curated to confirm insertion and deletion sites. This analysis identifies the preferred RGN and the corresponding preferred guide RNA (sgRNA). The analysis may result in any of RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, and LPG10208 being equally preferred. Additionally, the analysis may determine there is more than one preferred guide RNA, or that all target genomic sequences in Table 14 are equally preferred.

Example 9.2: Assay for expression of fetal hemoglobin

In this example, RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, or LPG10208-generated insertions or deletions disrupting the BCL11A enhancer region are assayed for expression of fetal hemoglobin. Healthy human donor $CD34^+$ hematopoietic stem cells (HSCs) are used. These HSCs are cultured and vector(s) comprising expression cassettes comprising the coding regions of the preferred RGN and the preferred sgRNA are introduced using methods similar to those described in Example 8.1. Following electroporation, these cells are differentiated *in vitro* into erythrocytes using established protocols (for example, Giarratana et al. (2004) Nat Biotechnology 23:69-74, herein incorporated by reference). The expression of HbF is then measured using western blotting with an anti-human HbF antibody, or quantified via High Performance Liquid Chromatography (HPLC). It is expected that successful disruption of the BCL11A enhancer locus will lead to an increase in HbF production when compared to HSCs electroporated with only the RGN but no guide.

Example 9.3: Assay for decreased sickle cell formation

In this example, RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, or LPG10208-generated insertions or deletions disrupting the BCL11A enhancer region are assayed for decreased sickle-cell formation. Donor $CD34^+$ hematopoietic stem cells (HSCs) from patients afflicted with sickle cell disease are used. These HSCs are cultured and vector(s) comprising expression cassettes comprising the coding regions of preferred RGN and the preferred sgRNA are introduced using methods similar to those described in Example 8.1. Following electroporation, these cells are differentiated *in vitro* into erythrocytes using established protocols (Giarratana et al. (2004) Nat Biotechnology 23:69-74). The expression of HbF is then measured using western blotting with an anti-human HbF antibody, or

quantified via High Performance Liquid Chromatography (HPLC). It is expected that successful disruption of the BCL11A enhancer locus will lead to an increase in HbF production when compared to HSCs electroporated with only the RGN but no guide.

Sickle cell formation is induced in these differentiated erythrocytes by the addition of metabisulfite. The numbers of sickled vs normal erythrocytes are counted using a microscope. It is expected that the numbers of sickled cells are less in cells treated with RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, or LPG10208 plus sgRNAs than with cells untreated, or treated with RGNs alone.

Example 9.4: Disease treatment validation in a murine model

To evaluate the efficacy of using RGNs LPG10165, LPG10166, LPG10167, LPG10168, LPG10207, LPG10169, LPG10171, LPG10186, LPG10190, LPG10191, LPG10194, LPG10195, LPG10196, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, or LPG10208 for disruption of the BCL11A locus, suitable humanized mouse models of sickle cell anemia are used. Expression cassettes encoding for the preferred RGN and for the preferred sgRNA are packaged into AAV vectors or adenovirus vectors. In particular, adenovirus type Ad5/35 is effective at targeting HSCs. A suitable mouse model containing a humanized HBB locus with sickle cell alleles is chosen, such as B6;FVB-Tg(LCR-HBA2,LCR-HBB*E26K)53Hhb/J or B6.Cg-*Hbatm1Paz Hbbtm1Tow* Tg(HBA-HBBs)41Paz/HhbJ. These mice are treated with granulocyte colony-stimulating factor alone or in combination with plerixafor to mobilize HSCs into circulation. AAVs or adenoviruses carrying the RGN and guide plasmid are then injected intravenously, and the mice are allowed to recover for a week. Blood obtained from these mice is tested in an *in vitro* sickling assay using metabisulfite, and the mice are followed longitudinally to monitor mortality rates and hematopoietic function. It is expected that treatment with AAVs or adenoviruses carrying an RGN and guide RNA will reduce sickling, mortality, and improve hematopoietic function when compared to mice treated with viruses lacking both expression cassettes, or with viruses carrying the RGN expression cassette alone.

Example 10: LPG10167 and LPG10169 are able to robustly edit multiple target loci in mammalian cells when delivered as plasmids

Guides targeting several genes were designed for robustness testing in HEK293T cells by plasmid delivery. These guides were tested under optimal editing conditions for HEK293T cells in a 96-well plate which involved seeding 10,000 cells on day 1, lipofection using Lipofectamine 3000 on day 2, and genomic DNA extraction on day 4. Lipofections were performed by mixing 160 ng of LPG10167- or LPG10169-encoding plasmid and 40 ng of guide RNA-encoding plasmid, and 0.4 µl of p3000 in Opti-MEM. Then, 0.3 µl of Lipofectamine 3000 diluted in Opti-MEM was added to the

plasmid mixture and incubated for 15 mins before adding to cells. Regions around the editing site for each guide were amplified from the extracted genomic DNA and sequenced. The editing efficiency of each nuclease with each guide RNA is reported as percent editing (percent of INDELS at the target site). Out of 35 guide RNAs tested, LPG10167 exhibited editing greater than 20% with 20% of the guide RNAs (7/35 guide RNAs) and had editing over 10% with 37% of guide RNAs tested (13/35 guide RNAs). Out of 35 guide RNAs tested, LPG10169 exhibited editing greater than 20% with 74% of the guide RNAs (29/35 guide RNAs) and had editing over 10% with 83% of guide RNAs tested (26/35 guide RNAs).

Table 15: Editing of 3 genes by LPG10167 by plasmid delivery in HEK293T cells

sgRNA	sgRNA backbone seq	sgRNA Target sequence SEQ ID NO	Percent Editing	Gene
SGN005914	SEQ ID NO: 65	831	15.6902031	Gene A
SGN005915	SEQ ID NO: 65	832	2.74297473	Gene A
SGN005915	SEQ ID NO: 65	832	29.9488671	Gene A
SGN005917	SEQ ID NO: 65	833	1.35942023	Gene A
SGN005917	SEQ ID NO: 65	833	1.30778333	Gene A
SGN005810	LPG10167_altv1	834	1.84378551	Gene B
SGN005810	LPG10167_altv1	834	2.03935874	Gene B
SGN005814	LPG10167_altv1	835	21.4255711	Gene B
SGN005809	LPG10167_altv1	836	31.326861	Gene B
SGN005809	LPG10167_altv1	836	30.0722332	Gene B
SGN005811	LPG10167_altv1	837	25.7482184	Gene B
SGN005811	LPG10167_altv1	837	35.987452	Gene B
SGN005812	LPG10167_altv1	838	12.1722701	Gene B
SGN005812	LPG10167_altv1	838	18.6820282	Gene B
SGN005813	LPG10167_altv1	839	2.57028656	Gene B
SGN005815	LPG10167_altv1	840	28.4798485	Gene B
SGN005920	SEQ ID NO: 65	841	30.9580842	Gene A
SGN005920	SEQ ID NO: 65	841	27.8602401	Gene A
SGN005822	LPG10167_altv1	842	0.24640757	Gene C
SGN005822	LPG10167_altv1	842	0.52382121	Gene C
SGN003670	LPG10167_altv1	843	1.95824267	Gene C
SGN003670	LPG10167_altv1	843	1.60709531	Gene C
SGN005817	LPG10167_altv1	844	5.0984934	Gene C
SGN005817	LPG10167_altv1	844	4.08188669	Gene C
SGN005819	LPG10167_altv1	845	0	Gene C
SGN005819	LPG10167_altv1	845	0	Gene C
SGN005820	LPG10167_altv1	846	0.52905944	Gene C

SGN005820	LPG10167_altv1	846	0.1935342	Gene C
SGN005821	LPG10167_altv1	847	3.21900097	Gene C
SGN005821	LPG10167_altv1	847	2.33903006	Gene C
SGN005818	LPG10167_altv1	848	0.59027286	Gene C
SGN005818	LPG10167_altv1	848	0.59988664	Gene C

Table 16. Editing of 3 additional genes by LPG10167 by plasmid delivery in HEK293T cells

sgRNA	Percent Editing	Gene
SGN005954	4.80391545	Gene D
SGN005954	5.0663098	Gene D
SGN005950	0.66249071	Gene D
SGN005950	0.9527471	Gene D
SGN005952	6.67256243	Gene D
SGN005952	5.76765811	Gene D
SGN005951	5.57846706	Gene D
SGN005951	4.87628096	Gene D
SGN005953	5.70167496	Gene D
SGN005953	3.2888097	Gene D
SGN005949	9.91554797	Gene D
SGN005949	11.5282082	Gene D
SGN005955	8.8758457	Gene D
SGN005955	5.88584917	Gene D
SGN005986	18.6878186	Gene E
SGN005986	16.1567291	Gene E
SGN005987	17.5418784	Gene E
SGN005987	13.8706581	Gene E
SGN005984	23.0083559	Gene E
SGN005984	22.3948067	Gene E
SGN005985	29.6956919	Gene E
SGN005985	26.2228692	Gene E
SGN006018	6.83062904	Gene F
SGN006018	6.38400407	Gene F
SGN006019	0.19555778	Gene F
SGN006019	0.55690726	Gene F
SGN006021	8.66222435	Gene F
SGN006021	5.61647981	Gene F
SGN006020	0.29773582	Gene F
SGN006020	0.17068595	Gene F
SGN006017	1.63186849	Gene F
SGN006017	1.06170828	Gene F
SGN006016	0.96019362	Gene F

SGN006016	1.01881817	Gene F

All sgRNAs in Table 16 had a wild type sgRNA backbone (i.e., used the crRNA and tracrRNA as set forth in Table 1).

Table 17: Editing of 3 genes by LPG10169 by plasmid delivery in HEK293T cells

sgRNA	sgRNA target sequence SEQ ID NO	Percent editing	Gene
SGN005924	849	14.9602258	Gene A
SGN005924	849	14.8238567	Gene A
SGN005925	850	17.3862072	Gene A
SGN005925	850	20.4117764	Gene A
SGN005923	851	40.4018298	Gene A
SGN005923	851	38.7923458	Gene A
SGN005926	852	33.6043014	Gene A
SGN005843	853	47.1458927	Gene B
SGN005843	853	50.2070439	Gene B
SGN005841	854	58.8185724	Gene B
SGN005841	854	49.2075575	Gene B
SGN005783	855	47.6077834	Gene B
SGN005783	855	37.614559	Gene B
SGN005838	856	22.8976127	Gene B
SGN005838	856	25.9730383	Gene B
SGN005781	857	39.6777187	Gene B
SGN005781	857	39.3404799	Gene B
SGN005775	858	25.2095538	Gene B
SGN005775	858	30.9052822	Gene B
SGN005599	859	57.7777778	Gene B
SGN005599	859	48.7960344	Gene B
SGN006003	860	38.0171972	Gene C
SGN006003	860	31.7637247	Gene C
SGN006001	861	52.8669926	Gene C
SGN006001	861	51.8285216	Gene C
SGN006004	862	34.8762091	Gene C
SGN006004	862	34.6849171	Gene C
SGN005998	863	1.43178644	Gene C
SGN005998	863	1.23180538	Gene C
SGN005999	864	2.17040806	Gene C
SGN005999	864	2.94154505	Gene C
SGN006002	865	31.1098001	Gene C

SGN006002	865	29.7411603	Gene C
SGN006000	866	22.8669102	Gene C
SGN006000	866	21.4160599	Gene C
SGN005773	867	44.4561818	Gene D
SGN005773	867	33.2241976	Gene D
SGN005769	868	15.4975381	Gene D
SGN005769	868	15.3210356	Gene D
SGN005771	869	26.3042673	Gene D
SGN005771	869	21.7449558	Gene D
SGN005768	870	41.3729237	Gene D
SGN005768	870	41.3457484	Gene D
SGN005774	871	34.9218687	Gene D
SGN005774	871	37.6106616	Gene D
SGN005772	872	54.3926545	Gene D
SGN005772	872	29.6654487	Gene D
SGN005770	873	45.9437597	Gene D
SGN005770	873	39.5701028	Gene D

All sgRNAs in Table 17 used the LPG10169_altv1 sgRNA backbone sequence set forth as SEQ ID NO: 1042.

Table 18. Editing of 2 additional genes by LPG10169 by plasmid delivery in HEK293T cells

sgRNA	Percent Editing	Gene
SGN005957	30.358198	Gene E
SGN005957	29.0760167	Gene E
SGN005960	26.7547061	Gene E
SGN005960	29.0880437	Gene E
SGN005956	1.21155968	Gene E
SGN005956	1.3489439	Gene E
SGN005962	32.2322546	Gene E
SGN005962	26.0777019	Gene E
SGN005959	5.05967949	Gene E
SGN005959	6.03141626	Gene E
SGN005958	31.9462067	Gene E
SGN005958	28.1011266	Gene E
SGN005961	3.63621808	Gene E
SGN005961	3.39061936	Gene E
SGN006022	0	Gene F
SGN006022	0	Gene F
SGN006023	36.3461784	Gene F
SGN006023	38.1796424	Gene F

SGN006024	65.3302941	Gene F
SGN006024	66.0638266	Gene F

All sgRNAs in Table 18 used the LPG10169_altv1 sgRNA backbone sequence set forth as SEQ ID NO: 1042.

Example 11: LPG10169 is able to edit multiple target loci in mammalian cells when delivered as a ribonucleoprotein

To evaluate LPG10169 editing reproducibility, LPG10169 was delivered as ribonucleoproteins (RNP: RNA-guided nuclease complexed with single guide RNAs) in T cells through nucleofection. For delivery of LPG10169 as ribonucleoproteins in T cells, nucleofection was utilized, which is an electroporation-based method that allows for the transfer of DNA, RNA, and/or protein into the cell nucleus. Isolated peripheral blood mononuclear cells (PBMC) were thawed, activated with CD3/CD28 beads, and incubated at 37°C with 5% CO_2. After 3 days, the beads were removed from the cells, which resulted in an enriched population of activated T cells. The activated T cells were then concentrated via centrifugation and resuspended in nucleofection buffer (P3 Solution Lonza) so that $1x10^6$ cells could be used for each reaction. The RNP (60 pmol RGN and 120 pmol guide RNA) were incubated at room temperature for at least 10 minutes prior to addition of the PBMC. The T cells+LPG10169+sgRNA reactions were nucleofected according to the manufacturer's instructions (Lonza). After nucleofection, the cells rested for 10 minutes, followed by incubation in media for 1-4 days. To determine the editing efficiency, the nucleofected T cells were incubated with an appropriate fluorophore-conjugated antibody followed by flow cytometry. Editing efficiency is reported as percent knockout of the gene of interest. An unedited control was included to ensure staining efficiency was complete. Here, multiple sgRNA targeting different genes were utilized to show the reproducibility of LPG10169 editing with all guides returning a coefficient of variation (CV) value of less than 50%.

Table 19. Editing of 2 genes by LPG10169 by RNP delivery in T cells

sgRNA	sgRNA backbone seq	sgRNA target sequence SEQ ID NO	Percent Editing	Gene
SGN005782	APG01604_altv1	992	0	Gene A
SGN005782	APG01604_altv1	992	0	Gene A
SGN005782	APG01604_altv1	992	0	Gene A
SGN005782	APG01604_altv1	992	0	Gene A
SGN005782	APG01604_altv1	992	0	Gene A
SGN005782	APG01604_altv1	992	0	Gene A

SGN005777	APG01604_altv1	993	58.7235151	Gene A
SGN005777	APG01604_altv1	993	59.8450226	Gene A
SGN005777	APG01604_altv1	993	58.3732381	Gene A
SGN005777	APG01604_altv1	993	57.4838508	Gene A
SGN005777	APG01604_altv1	993	57.547783	Gene A
SGN005777	APG01604_altv1	993	55.9661206	Gene A
SGN005783	LPG10169_altv1	994	3.94383289	Gene A
SGN005783	LPG10169_altv1	994	5.81863884	Gene A
SGN005783	LPG10169_altv1	994	5.55818727	Gene A
SGN005783	LPG10169_altv1	994	3.841975	Gene A
SGN005783	LPG10169_altv1	994	2.55979693	Gene A
SGN005783	LPG10169_altv1	994	2.34298474	Gene A
SGN005781	LPG10169_altv1	995	0.48979667	Gene A
SGN005781	LPG10169_altv1	995	0.34600111	Gene A
SGN005781	LPG10169_altv1	995	0.12481467	Gene A
SGN005781	LPG10169_altv1	995	0.04235935	Gene A
SGN005781	LPG10169_altv1	995	0.13790916	Gene A
SGN005781	LPG10169_altv1	995	0.07981781	Gene A
SGN005776	APG01604_altv1	996	53.2908207	Gene A
SGN005776	APG01604_altv1	996	53.7604188	Gene A
SGN005776	APG01604_altv1	996	55.0354163	Gene A
SGN005776	APG01604_altv1	996	52.3445185	Gene A
SGN005776	APG01604_altv1	996	50.9454218	Gene A
SGN005776	APG01604_altv1	996	53.4400559	Gene A
SGN004748	APG01604_altv1	997	17.2240528	Gene A
SGN004748	APG01604_altv1	997	14.1702725	Gene A
SGN004748	APG01604_altv1	997	12.7553689	Gene A
SGN004748	APG01604_altv1	997	13.3994387	Gene A
SGN004748	APG01604_altv1	997	11.5806321	Gene A
SGN005780	APG01604_altv1	998	3.67326235	Gene A
SGN005780	APG01604_altv1	998	3.73146582	Gene A
SGN005780	APG01604_altv1	998	3.63697053	Gene A
SGN005780	APG01604_altv1	998	4.24748929	Gene A
SGN005780	APG01604_altv1	998	3.57239421	Gene A
SGN005779	APG01604_altv1	999	68.4804172	Gene A
SGN005779	APG01604_altv1	999	73.8171045	Gene A
SGN005779	APG01604_altv1	999	54.615976	Gene A
SGN005779	APG01604_altv1	999	48.6238283	Gene A
SGN005779	APG01604_altv1	999	60.9002569	Gene A
SGN005779	APG01604_altv1	999	57.1299903	Gene A
SGN005773	LPG10169_altv1	1000	84.7583661	Gene B
SGN005773	LPG10169_altv1	1000	92.3016714	Gene B

SGN005773	LPG10169_altv1	1000	75.8830753	Gene B
SGN005773	LPG10169_altv1	1000	81.6155928	Gene B
SGN005773	LPG10169_altv1	1000	79.6614921	Gene B
SGN005773	LPG10169_altv1	1000	85.0834861	Gene B
SGN005769	LPG10169_altv1	1001	0.19490896	Gene B
SGN005769	LPG10169_altv1	1001	0	Gene B
SGN005769	LPG10169_altv1	1001	0	Gene B
SGN005769	LPG10169_altv1	1001	0.18070379	Gene B
SGN005769	LPG10169_altv1	1001	0.11359151	Gene B
SGN005769	LPG10169_altv1	1001	0.13972122	Gene B
SGN005771	LPG10169_altv1	1002	0.44167316	Gene B
SGN005771	LPG10169_altv1	1002	0.47722078	Gene B
SGN005771	LPG10169_altv1	1002	0.07050686	Gene B
SGN005771	LPG10169_altv1	1002	0.73890895	Gene B
SGN005771	LPG10169_altv1	1002	0.65421247	Gene B
SGN005771	LPG10169_altv1	1002	0.52317665	Gene B
SGN005765	APG01604_altv1	1003	4.65913098	Gene B
SGN005765	APG01604_altv1	1003	5.27222029	Gene B
SGN005765	APG01604_altv1	1003	2.54049051	Gene B
SGN005765	APG01604_altv1	1003	2.84028893	Gene B
SGN005765	APG01604_altv1	1003	5.7169896	Gene B
SGN005766	APG01604_altv1	1004	0.14281903	Gene B
SGN005766	APG01604_altv1	1004	0.06928507	Gene B
SGN005766	APG01604_altv1	1004	0	Gene B
SGN005766	APG01604_altv1	1004	0	Gene B
SGN005766	APG01604_altv1	1004	0	Gene B
SGN005766	APG01604_altv1	1004	0	Gene B
SGN005768	LPG10169_altv1	1005	1.73997495	Gene B
SGN005768	LPG10169_altv1	1005	3.71015723	Gene B
SGN005768	LPG10169_altv1	1005	2.02763059	Gene B
SGN005768	LPG10169_altv1	1005	2.28036298	Gene B
SGN005768	LPG10169_altv1	1005	3.28816821	Gene B
SGN005768	LPG10169_altv1	1005	1.94649221	Gene B
SGN005774	LPG10169_altv1	1006	39.4160209	Gene B
SGN005774	LPG10169_altv1	1006	37.4843685	Gene B
SGN005774	LPG10169_altv1	1006	24.3723105	Gene B
SGN005774	LPG10169_altv1	1006	27.6958085	Gene B
SGN005774	LPG10169_altv1	1006	21.5484699	Gene B
SGN005774	LPG10169_altv1	1006	26.7920212	Gene B
SGN005772	LPG10169_altv1	1007	46.2335416	Gene B
SGN005772	LPG10169_altv1	1007	37.9421513	Gene B
SGN005772	LPG10169_altv1	1007	18.7995158	Gene B

SGN005772	LPG10169_altv1	1007	31.5397509	Gene B
SGN005772	LPG10169_altv1	1007	35.4858142	Gene B
SGN005772	LPG10169_altv1	1007	32.8500034	Gene B
SGN005767	APG01604_altv1	1008	0	Gene B
SGN005767	APG01604_altv1	1008	0	Gene B
SGN005767	APG01604_altv1	1008	0	Gene B
SGN005767	APG01604_altv1	1008	0.16998575	Gene B
SGN005767	APG01604_altv1	1008	0	Gene B
SGN005767	APG01604_altv1	1008	0.04165963	Gene B
SGN005770	LPG10169_altv1	1009	0	Gene B
SGN005770	LPG10169_altv1	1009	0.64652568	Gene B
SGN005770	LPG10169_altv1	1009	0.86978327	Gene B
SGN005770	LPG10169_altv1	1009	0.73410084	Gene B
SGN005770	LPG10169_altv1	1009	0.97032105	Gene B
SGN005770	LPG10169_altv1	1009	0.70931921	Gene B

The sgRNA backbone sequences APG01604_altv1 and LPG10169_altv1 are set forth as SEQ ID NOs: 1041 and 1042, respectively.

Example 12: LPG10196 is able to edit multiple target loci in mammalian cells when delivered as plasmids or an mRNA

Guide RNAs targeting several genes were designed for robustness testing in HEK293T cells by plasmid delivery. These guides were tested under optimal editing conditions for HEK293T cells in a 96-well plate which involved seeding 10,000 cells on day 1, lipofection using Lipofectamine 3000 on day 2, and genomic DNA extraction on day 4. Lipofections were performed by mixing 160 ng of LPG10196-encoding plasmid and 40 ng of guide RNA-encoding plasmid, and 0.4 µl of p3000 in Opti-MEM. Then, 0.3 µl of Lipofectamine 3000 diluted in Opti-MEM was added to the plasmid mixture and incubated for 15 mins before adding to cells. Regions around the editing site for each guide are amplified from the extracted genomic DNA and sequenced by NGS as described in Example 4. The editing efficiency of each nuclease with each guide RNA is reported as percent editing (percent of INDELS at the target site). Out of 40 guide RNAs tested, no guide RNAs exhibited editing greater than 10%.

Table 20: Editing of 3 genes by LPG10196 by plasmid delivery in HEK293T cells

sgRNA	sgRNA target sequence SEQ ID NO	Percent Editing	Gene
SGN008751	1010	2.31119542	Gene A
SGN008751	1010	2.07930219	Gene A
SGN008752	1011	1.53116594	Gene A
SGN008752	1011	1.02926413	Gene A

SGN008752	1011	0.26787925	Gene A
SGN008752	1011	0.24384033	Gene A
SGN008753	1012	0	Gene A
SGN008753	1012	0	Gene A
SGN008754	1013	0.34027866	Gene B
SGN008754	1013	0.16095991	Gene B
SGN008754	1013	0	Gene B
SGN008755	1014	2.10475191	Gene B
SGN008755	1014	1.18748882	Gene B
SGN008756	1015	1.00449529	Gene B
SGN008756	1015	0.61317269	Gene B
SGN008757	1016	1.72424183	Gene B
SGN008757	1016	1.56266594	Gene B
SGN008758	1017	0.47270321	Gene B
SGN008758	1017	0.21425377	Gene B
SGN008759	1018	0.79871802	Gene B
SGN008759	1018	0.71299172	Gene B
SGN008759	1018	0.37832124	Gene B
SGN008759	1018	0.36686486	Gene B
SGN008760	1019	0.15655804	Gene B
SGN008760	1019	0	Gene B
SGN008761	1020	1.27660997	Gene B
SGN008761	1020	0.60405209	Gene B
SGN008762	1021	4.06040657	Gene B
SGN008763	1022	0	Gene B
SGN008763	1022	0	Gene B
SGN008771	1023	7.67238324	Gene B
SGN008771	1023	6.90172821	Gene B
SGN008772	1024	0.05511203	Gene B
SGN008772	1024	0.03041783	Gene B
SGN008773	1025	1.59124182	Gene B
SGN008773	1025	1.47677043	Gene B
SGN008774	1026	0.04540704	Gene B
SGN008774	1026	0.04108932	Gene B
SGN008775	1027	0.0880911	Gene B
SGN008775	1027	0.04038421	Gene B
SGN008776	1028	0	Gene B
SGN008776	1028	0	Gene B
SGN008776	1028	0	Gene B
SGN008777	1029	0.03742341	Gene B
SGN008777	1029	0	Gene B

SGN008777	1029	0	Gene B
SGN008778	1030	0.29135427	Gene A
SGN008778	1030	0.14321118	Gene A
SGN008779	1031	4.32838036	Gene A
SGN008779	1031	3.05104187	Gene A
SGN008780	1032	2.20102518	Gene A
SGN008780	1032	1.91115982	Gene A
SGN008781	1033	0.13457689	Gene C
SGN008781	1033	0	Gene C
SGN008782	1034	1.17617006	Gene C
SGN008782	1034	0.32513646	Gene C
SGN008782	1034	0	Gene C
SGN008783	1035	0.09980008	Gene C
SGN008783	1035	0	Gene C
SGN008783	1035	0	Gene C
SGN008784	1036	6.04145714	Gene C
SGN008784	1036	4.60291001	Gene C
SGN008785	1037	2.33054617	Gene C
SGN008786	1038	2.17424483	Gene C
SGN008786	1038	1.7169595	Gene C
SGN008787	1039	0.73367513	Gene C

Table 21: Editing of 2 additional genes by LPG10196 by plasmid delivery in HEK293T cells

sgRNA	Percent editing	Gene
SGN008764	6.40775496	Gene D
SGN008764	6.36429188	Gene D
SGN008765	5.85900556	Gene D
SGN008765	4.93704309	Gene D
SGN008766	1.62050983	Gene D
SGN008766	1.20221383	Gene D
SGN008768	1.02012565	Gene D
SGN008768	0.49157556	Gene D
SGN008769	2.52760451	Gene D
SGN008770	1.19088503	Gene D
SGN008770	0.75523987	Gene D
SGN008770	0	Gene D
SGN008788	5.02264123	Gene E
SGN008788	4.66534168	Gene E
SGN008788	0	Gene E

SGN008789	0.07087023	Gene E
SGN008789	0	Gene E
SGN008789	0	Gene E

To demonstrate editing of mammalian cells with an alternative delivery format of LPG10196, mRNA nucleofection delivery was tested with primary T-cells. Purified CD3+ T-cells or peripheral blood mononuclear cells (PBMCs) were thawed, activated using CD3/CD28 beads (ThermoFisher) for 3 days, then nucleofected using the Lonza 4D-Nucleofector X unit and Nucleocuvette strips. The P3 Primary Cell kit was used for mRNA delivery. Cells were transfected using the EO-115 program for mRNA delivery. Cells were cultured in CTS OpTimizer T cell expansion medium (ThermoFisher) containing IL-2, IL-7, and IL-15 (Miltenyi Biotec) for 4 days post nucleofection before being harvested using a Nucleospin Tissue genomic DNA isolation kit (Machery Nagel). Amplicons surrounding the editing sites were generated by PCR and subjected to NGS sequencing using the Illumina Nextera platform following the method in Example 4.

Table 22: Editing of 2 genes by LPG10196 by mRNA delivery in T cells

sgRNA	sgRNA target sequence SEQ ID NO	Percent Editing	Gene
SGN005324	1046	0.26644188	Gene A
SGN005326	1047	0.21256232	Gene A
SGN005328	1048	0.11126232	Gene A
SGN005330	1049	0	Gene B
SGN005329	1050	0	Gene B
SGN005333	1051	0	Gene B
SGN005332	1052	0	Gene B
SGN005328	1048	0	Gene A
SGN005327	1053	0	Gene A
SGN005333	1051	0	Gene B
SGN005332	1052	0	Gene B
SGN005331	1054	0	Gene B
SGN005330	1049	0	Gene B
SGN005329	1050	0	Gene B
SGN005327	1053	0	Gene A
SGN005326	1047	0	Gene A
SGN005325	749	0	Gene A
SGN005324	1046	0	Gene A
SGN005333	1051	0	Gene B

SGN005332	1052	0	Gene B
SGN005331	1054	0	Gene B
SGN005330	1049	0	Gene B
SGN005329	1050	0	Gene B
SGN005328	1048	0	Gene A
SGN005327	1053	0	Gene A
SGN005325	749	0	Gene A

Table 23: Editing of an additional gene by LPG10196 by mRNA delivery in T cells

sgRNA	Percent Editing
SGN008450	0.04128176
SGN008449	0.04118244
SGN008449	0.03958039
SGN008450	0.03853235
SGN008450	0.03516029
SGN008449	0

Example 13: LPG10190, LPG10194, LPG10195, LPG10197, LPG10198, LPG10200, LPG10203, LPG10204, LPG10205, LPG10207, and LPG10208 are able to edit multiple target loci in mammalian cells when delivered as plasmids

Guide RNAs targeting several genes were designed for robustness testing of multiple RGNs in HEK293T cells by plasmid delivery. These guides were tested under optimal editing conditions for HEK293T cells in a 96-well plate which involved seeding 10,000-15,000 cells on day 1, lipofection using Lipofectamine 3000 on day 2, and genomic DNA extraction on day 4. Lipofections were performed by mixing 160 ng of LPG10196-encoding plasmid and 40 ng of guide RNA-encoding plasmid, and 0.4 µl of p3000 in Opti-MEM. Then, 0.3 µl of Lipofectamine 3000 diluted in Opti-MEM was added to the plasmid mixture and incubated for 15 mins before adding to cells. Regions around the editing site for each guide are amplified from the extracted genomic DNA and sequenced by NGS as described in Example 4. Results are shown in Table 24.

Table 24: Editing of multiple genes by various RGNs by plasmid delivery in HEK293T cells

RGN	sgRNA	sgRNA target sequence SEQ ID NO	Indel Rate	Gene
LPG10190	SGN005942	874	0.36701908	Gene A

LPG10190	SGN005948	875	0.2889043	Gene A
LPG10190	SGN005945	876	8.66876033	Gene A
LPG10190	SGN005944	877	0.04169017	Gene A
LPG10190	SGN005900	878	0	Gene B
LPG10190	SGN005904	879	0	Gene B
LPG10190	SGN005902	880	0	Gene B
LPG10190	SGN005906	881	1.45647858	Gene B
LPG10190	SGN005903	882	0	Gene B
LPG10190	SGN005905	883	6.13382844	Gene B
LPG10190	SGN005901	884	0	Gene B
LPG10190	SGN006014	885	0.01352876	Gene C
LPG10190	SGN006012	886	0	Gene C
LPG10190	SGN006013	887	0.86899015	Gene C
LPG10190	SGN006015	888	0.04033643	Gene C
LPG10190	SGN005907	889	0	Gene D
LPG10190	SGN005908	890	0	Gene D
LPG10190	SGN005913	891	0.06573541	Gene D
LPG10190	SGN005911	892	0	Gene D
LPG10190	SGN005912	893	0	Gene D
LPG10190	SGN005909	894	0	Gene D
LPG10190	SGN005910	895	0	Gene D
LPG10194	SGN005312	896	0.030428	Gene B
LPG10194	SGN005309	897	21.0121199	Gene B
LPG10194	SGN005313	898	2.20160132	Gene B
LPG10194	SGN005311	899	7.09320917	Gene B
LPG10194	SGN005310	900	0.19347422	Gene B
LPG10194	SGN005308	901	0.53503148	Gene D
LPG10194	SGN005306	902	0	Gene D
LPG10194	SGN005304	903	0.35365696	Gene D
LPG10194	SGN005307	904	0.83313993	Gene D
LPG10194	SGN005305	905	0.15847329	Gene D
LPG10195	SGN005321	906	1.60788133	Gene B
LPG10195	SGN005322	907	7.9384461	Gene B
LPG10195	SGN005319	908	1.16046009	Gene B
LPG10195	SGN005323	909	1.35810358	Gene B
LPG10195	SGN005320	910	1.98190455	Gene B
LPG10195	SGN005314	911	0.48531898	Gene D
LPG10195	SGN005316	912	0.06789522	Gene D
LPG10195	SGN005317	913	0.033033	Gene D
LPG10195	SGN005315	914	3.20645014	Gene D
LPG10195	SGN005318	915	0	Gene D
LPG10197	SGN005341	916	1.42612491	Gene B

LPG10197	SGN005342	917	2.89469264	Gene B
LPG10197	SGN005339	918	0.65474111	Gene B
LPG10197	SGN005343	919	2.3066702	Gene B
LPG10197	SGN005340	920	0.11537443	Gene B
LPG10197	SGN005336	921	0	Gene D
LPG10197	SGN005334	922	1.06043066	Gene D
LPG10197	SGN005337	923	0.96785633	Gene D
LPG10197	SGN005338	924	0	Gene D
LPG10197	SGN005335	925	0.02076962	Gene D
LPG10198	SGN005346	926	0	Gene D
LPG10198	SGN005344	927	0.01407238	Gene D
LPG10198	SGN005347	928	0	Gene D
LPG10198	SGN005348	929	0.02833437	Gene D
LPG10198	SGN005345	930	0	Gene D
LPG10200	SGN005354	931	0.49286011	Gene D
LPG10200	SGN005355	932	0.03016978	Gene D
LPG10200	SGN005356	933	0	Gene D
LPG10200	SGN005357	934	0.0899422	Gene D
LPG10200	SGN005358	935	0.01369655	Gene D
LPG10200	SGN005359	936	0	Gene B
LPG10200	SGN005360	937	2.3326179	Gene B
LPG10200	SGN005361	938	0.88360097	Gene B
LPG10200	SGN005362	939	4.12895472	Gene B
LPG10200	SGN005363	940	27.8668507	Gene B
LPG10203	SGN005364	941	3.22685247	Gene D
LPG10203	SGN005365	942	13.9042672	Gene D
LPG10203	SGN005366	943	14.2677892	Gene D
LPG10203	SGN005367	944	15.0606662	Gene D
LPG10204	SGN005373	945	23.5520127	Gene B
LPG10204	SGN005374	946	37.4986793	Gene B
LPG10204	SGN005376	947	7.37352091	Gene D
LPG10204	SGN005377	948	9.22081113	Gene D
LPG10204	SGN005378	949	2.69392448	Gene D
LPG10204	SGN005379	950	1.51490663	Gene D
LPG10204	SGN005380	951	2.17714153	Gene B
LPG10204	SGN005381	952	0.04262038	Gene B
LPG10204	SGN005382	953	25.5481725	Gene B
LPG10205	SGN005383	954	23.8069402	Gene B
LPG10205	SGN005384	955	43.8835176	Gene B
LPG10205	SGN005386	956	4.19188596	Gene D
LPG10205	SGN005387	957	5.0923384	Gene D
LPG10205	SGN005388	958	0.9325359	Gene D

LPG10205	SGN005389	959	0.51241228	Gene D
LPG10205	SGN005390	960	1.09905382	Gene B
LPG10205	SGN005391	961	0.10620702	Gene B
LPG10205	SGN005392	962	17.2982248	Gene B
LPG10205	SGN005393*	963	0.09406381	Gene B
LPG10205	SGN005394*	964	0.27666136	Gene B
LPG10205	SGN005396*	965	0	Gene D
LPG10205	SGN005397*	966	0	Gene D
LPG10205	SGN005398*	967	0	Gene D
LPG10205	SGN005399*	968	0	Gene D
LPG10205	SGN005400*	969	0	Gene B
LPG10205	SGN005401*	970	0	Gene B
LPG10205	SGN005402*	971	0.30758559	Gene B
LPG10207	SGN005403	972	6.8683675	Gene D
LPG10207	SGN005404	973	6.14261875	Gene D
LPG10207	SGN005405	974	10.9052765	Gene D
LPG10207	SGN005406	975	15.9374664	Gene D
LPG10207	SGN005407	976	21.4552493	Gene D
LPG10207	SGN005408	977	28.4481502	Gene B
LPG10207	SGN005409	978	35.3689448	Gene B
LPG10207	SGN005410	979	38.9217997	Gene B
LPG10207	SGN005411	980	21.7166823	Gene B
LPG10207	SGN005412	981	17.189728	Gene B
LPG10208	SGN005373	982	37.4768411	Gene B
LPG10208	SGN005374	983	47.1793232	Gene B
LPG10208	SGN005375	984	0.92766506	Gene D
LPG10208	SGN005376	985	13.6968015	Gene D
LPG10208	SGN005377	986	25.6968565	Gene D
LPG10208	SGN005378	987	3.3814505	Gene D
LPG10208	SGN005379	988	3.00935013	Gene D
LPG10208	SGN005380	989	3.20853702	Gene B
LPG10208	SGN005381	990	0.58667551	Gene B
LPG10208	SGN005382	991	27.5155924	Gene B

* LPG10205 sgRNAs used the LPG10205 sgRNA backbone sequence set forth as SEQ ID NO: 80, except for those that have a *, which used the LPG10205 sgRNA backbone sequence set forth as SEQ ID NO: 81

THAT WHICH IS CLAIMED:

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:1-20.

2. The nucleic acid molecule of claim 1, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of the target sequence.

3. The nucleic acid molecule of claim 1 or 2, wherein said polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to said polynucleotide.

4. The nucleic acid molecule of any one of claims 1-3, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

5. The nucleic acid molecule of any one of claims 1-4, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule upon binding.

6. The nucleic acid molecule of claim 5, wherein said RGN polypeptide is capable of generating a double-stranded break.

7. The nucleic acid molecule of claim 5, wherein said RGN polypeptide is capable of generating a single-stranded break.

8. The nucleic acid molecule of any one of claims 1-4, wherein said RGN polypeptide is nuclease inactive or is a nickase.

9. The nucleic acid molecule of any one of claims 1-8, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

10. The nucleic acid molecule of claim 9, wherein the base-editing polypeptide is a deaminase.

11. The nucleic acid molecule of any one of claims 1-10, wherein the RGN polypeptide comprises one or more nuclear localization signals.

12. The nucleic acid molecule of any one of claims 1-11, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

13. The nucleic acid molecule of any one of claims 1-12, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

14. A vector comprising the nucleic acid molecule of any one of claims 1-13.

15. The vector of claim 14, further comprising at least one nucleotide sequence encoding said gRNA capable of hybridizing to the non-target strand of said target sequence.

16. The vector of claim 15, 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: 21; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 22; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) 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: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 24; 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: 4;

e) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 26; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 28; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 29; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) 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: 51;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10;

k) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 31; 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: 11;

l) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 32; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 33; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 34; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 35; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 36; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) 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: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 17;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 38 or 39; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 40; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; 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: 41; and

ii) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62;

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

17. The vector of claim 13, 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 100% sequence identity to SEQ ID NO: 21; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 22; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 23; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 24; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 26; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 28; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 29; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 30; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 10;

k) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 31; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 32; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 33; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 34; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 35; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 36; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 37; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 38 or 39; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 40; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a guide RNA comprising:

i) a CRISPR RNA comprising a CRISPR repeat sequence having at least 100% sequence identity to SEQ ID NO: 41; and

ii) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62;

wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

18. The vector of any one of claims 14-17, where said gRNA is a single guide RNA.

19. The vector of any one of claims 15-17, wherein said gRNA is a dual-guide RNA.

20. A cell comprising the nucleic acid molecule of any one of claims 1-14 or the vector of any one of claims 14-19.

21. A plant comprising the cell of claim 20.

22. A seed comprising the cell of claim 20.

23. A method for making an RGN polypeptide comprising culturing the cell of claim 20 under conditions in which the RGN polypeptide is expressed.

24. A method for making an RGN polypeptide comprising introducing into a cell a heterologous nucleic acid molecule comprising a nucleotide sequence encoding an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20;

and culturing said cell under conditions in which the RGN polypeptide is expressed.

25. The method of claim 24, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of said target sequence.

26. The method of claim 24 or 25, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

27. The method of any one of claims 24-26, further comprising purifying said RGN polypeptide.

28. The method of any one of claims 24-27, wherein said cell further expresses one or more guide RNAs capable of binding to said RGN polypeptide to form an RGN ribonucleoprotein complex.

29. The method of claim 28, further comprising purifying said RGN ribonucleoprotein complex.

30. An 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: 1-20.

31. The RGN polypeptide of claim 31, wherein said RGN polypeptide is capable of binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and an on-target strand, in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to the non-target strand of said target sequence.

32. The RGN polypeptide of claim 30 or 31, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

33. The RGN polypeptide of any one of claims 30-32, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule upon binding.

34. The RGN polypeptide of claim 33, wherein cleavage by said RGN polypeptide generates a double-stranded break.

35. The RGN polypeptide of claim 33, wherein cleavage by said RGN polypeptide generates a single-stranded break.

36. The RGN polypeptide of any one of claims 30-32, wherein said RGN polypeptide is nuclease inactive or a nickase.

37. The RGN polypeptide of any one of claims 30-36, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

38. The RGN polypeptide of claim 37, wherein the base-editing polypeptide is a deaminase.

39. The RGN polypeptide of any one of claims 30-38, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

40. The RGN polypeptide of any one of claims 28-36, wherein the RGN polypeptide comprises one or more nuclear localization signals.

41. A ribonucleoprotein (RNP) complex comprising the RGN polypeptide of any one of claims 30-40 and a guide RNA bound to the RGN polypeptide.

42. A nucleic acid molecule comprising a CRISPR RNA (crRNA) or a polynucleotide encoding a crRNA, wherein said crRNA comprises a spacer and a CRISPR repeat, wherein said CRISPR repeat comprises a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

43. The nucleic acid molecule of claim 42, wherein a guide RNA comprising:

a) said crRNA; and

b) a trans-activating CRISPR RNA (tracrRNA) hybridized to said CRISPR repeat of said crRNA;

is capable of hybridizing to the non-target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

44. The nucleic acid molecule of claim 42 or 43, wherein said polynucleotide encoding a crRNA is operably linked to a promoter heterologous to said polynucleotide.

45. The nucleic acid molecule of any one of claims 42-44, wherein said CRISPR repeat comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 21-41, or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, or nucleotides 1-22 of SEQ ID NO: 1044 or 1045.

46. A vector comprising the nucleic acid molecule comprising said polynucleotide encoding said crRNA of any one of claims 42-45.

47. The vector of claim 46, wherein said vector further comprises a polynucleotide encoding said tracrRNA.

48. The vector of claim 47, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 21;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 22;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 23;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 24;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 26;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 28;

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 29;

j) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 30;

k) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 31;

l) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 32;

m) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 33;

n) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 34;

o) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 35;

p) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 36;

q) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 37;

r) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 38 or 39;

s) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 40; and

t) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 41.

49. The vector of claim 47, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 21;

b) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 22;

c) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 23;

d) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 24;

e) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042;

f) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 26;

g) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045;

h) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 28;

i) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 29;

j) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 30;

k) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 31;

l) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 32;

m) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 33;

n) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 34;

o) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 35;

p) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 36;

q) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 37;

r) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 38 or 39;

s) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 40; and

t) a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 41.

50. The vector of any one of claims 42-49, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide.

51. The vector of claim 50, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 1, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 2, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 3, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 4, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 5, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 6, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO:26 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 7, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 8, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 9, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 10, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 11, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 12, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 13, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 14, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 15, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 16, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 17, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 18, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 38 or 39 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 19, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 20, wherein said CRISPR repeat has at least 90% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

52. The vector of claim 50, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 1, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 2, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 3, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 4, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 5, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 6, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO:26 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 7, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 8, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 9, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 10, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 11, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 12, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 13, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 14, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 15, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 16, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 17, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 18, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 38 or 39 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 19, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 20, wherein said CRISPR repeat has at least 100% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

53. A nucleic acid molecule comprising a trans-activating CRISPR RNA (tracrRNA) or a polynucleotide encoding a tracrRNA comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

54. The nucleic acid molecule of claim 53, wherein a guide RNA comprising:

a) said tracrRNA; and

b) a crRNA comprising a spacer and a CRISPR repeat, wherein said tracrRNA hybridizes with said CRISPR repeat of said crRNA;

is capable of hybridizing to the non-target strand of a target sequence in a target nucleic acid molecule in a sequence specific manner through the spacer of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide.

55. The nucleic acid molecule of claim 53 or 54, wherein said polynucleotide encoding a tracrRNA is operably linked to a promoter heterologous to said polynucleotide.

56. The nucleic acid molecule of any one of claims 53-55, wherein said tracrRNA comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 42-62, nucleotides 19-111 of SEQ ID NO: 1040, nucleotides 22-85 of SEQ ID NO: 1041 or 1042, nucleotides 24-138 of SEQ ID NO: 143, nucleotides 27-96 of SEQ ID NO: 1044, or nucleotides 27-95 of SEQ ID NO: 1045.

57. A vector comprising the nucleic acid molecule comprising said polynucleotide encoding said tracrRNA of any one of claims 53-56.

58. The vector of claim 57, wherein said vector further comprises a polynucleotide encoding said crRNA.

59. The vector of claim 58, wherein said crRNA comprises a CRISPR repeat selected from the group consisting of:

a) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41, wherein said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

60. The vector of claim 58, wherein said crRNA comprises a CRISPR repeat selected from the group consisting of:

a) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41, wherein said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

61. The vector of any one of claims 53-60, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide.

62. The vector of claim 61, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 1, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 2, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 3, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 4, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 5, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 6, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 7, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 8, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 9, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 10, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 11, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 12, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 13, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 14, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 15, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 16, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 17, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 18, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38, or 39 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 19, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 90% sequence identity to SEQ ID NO: 20, wherein said crRNA comprises a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 90% sequence identity to SEQ ID NO: 62.

63. The vector of claim 61, wherein said RGN polypeptide is selected from the group consisting of:

a) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 1, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 42;

b) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 2, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 43;

c) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 3, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 4, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 45;

e) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 5, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 6, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 7, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 8, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 49;

i) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 9, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 50;

j) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 10, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 51;

k) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 11, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 52;

l) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 12, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 53;

m) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 13, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 54;

n) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 14, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 55;

o) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 15, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 56;

p) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 16, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 57;

q) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 17, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 58;

r) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 18, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38, or 39 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 19, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 61; and

t) a RGN polypeptide having at least 100% sequence identity to SEQ ID NO: 20, wherein said crRNA comprises a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and said tracrRNA has at least 100% sequence identity to SEQ ID NO: 62.

64. A cell comprising the nucleic acid molecule of any one of claims 42-45 and 53-56, the vector of any one of claims 46-52 and 57-63, the single guide RNA of claim 142, or the double guide RNA of claim 143.

65. A plant comprising the cell of claim 64.

66. A seed comprising the cell of claim 64.

67. A system for binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, said system comprising:

a) one or more guide RNAs capable of hybridizing to the non-target strand of said target 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: 1-20 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide;

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 sequence.

68. The system of claim 67, 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.

69. A system for binding a target sequence in a target nucleic acid molecule, wherein said target sequence comprises a target strand and a non-target strand, said system comprising:

a) one or more guide RNAs capable of hybridizing to the non-target strand of said target 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: 1-20;

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 sequence.

70. The system of any one of claims 67-69, 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.

71. The system of any one of claims 67-70, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

72. The system of any one of claims 67-71, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

73. The system of any one of claims 67-72, wherein said target sequence is a eukaryotic target sequence.

74. The system of any one of claims 67-73, wherein said gRNA is a single guide RNA (sgRNA).

75. The system of any one of claims 67-73, wherein said gRNA is a dual-guide RNA.

76. The system of any one of claims 67-75, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 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: 4;

e) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said

RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NOs:10;

k) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 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: 11;

l) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 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: 17;

r) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60,

wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

77. The system of any one of claims 67-75, wherein said gRNA is selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides

24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NOs:10;

k) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide comprises an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

78. The system of any one of claims 67-77, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

79. The system of any one of claims 67-78, wherein the target sequence is within a cell.

80. The system of any one of claims 67-81, wherein the one or more guide RNAs is capable of hybridizing to the non-target strand of said target sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target nucleic acid molecule.

81. The system of claim 80, wherein the cleavage generates a double-stranded break.

82. The system of claim 80, wherein the cleavage generates a single-stranded break.

83. The system of any one of claims 67-79, wherein said RGN polypeptide is nuclease inactive or is a nickase.

84. The system of any one of claims 67-83, wherein the RGN polypeptide is operably linked to a base-editing polypeptide.

85. The system of claim 84, wherein the base-editing polypeptide is a deaminase.

86. The system of any one of claims 67-85, wherein the RGN polypeptide comprises one or more nuclear localization signals.

87. The system of any one of claims 67-86, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

88. The system of any one of claims 67-87, wherein said system further comprises one or more donor polynucleotides.

89. A cell comprising the system of any one of claims 67-88.

90. A plant comprising the cell of claim 89.

91. A seed comprising the cell of claim 89.

92. A pharmaceutical composition comprising the nucleic acid molecule of any one of claims 1-13, 42-45, and 53-56, the vector of any one of claims 14-19, 46-52, and 57-63, the cell of any one of claims 20, 64, and 89, the RGN polypeptide of any one of claims 30-40, the RNP complex of claim 41, or the system of any one of claims 67-88, and a pharmaceutically acceptable carrier.

93. A method for binding a target sequence in a target nucleic acid molecule comprising delivering a system according to any one of claims 67-88, to said target sequence or a cell comprising the target sequence.

94. The method of claim 93, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target sequence.

95. The method of claim 93, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby modulating expression of a target gene comprising said target sequence.

96. A method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence, said method comprising delivering a system according to any one of claims 67-88 to said target sequence or a cell comprising the target sequence, wherein cleavage or modification of said target nucleic acid molecule occurs.

97. The method of claim 96, wherein said modified target nucleic acid molecule comprises insertion of heterologous DNA into the target nucleic acid molecule.

98. The method of claim 96, wherein said modified target nucleic acid molecule comprises deletion of at least one nucleotide from the target nucleic acid molecule.

99. The method of claim 96, wherein said modified target nucleic acid molecule comprises mutation of at least one nucleotide in the target nucleic acid molecule.

100. A method for binding a target sequence in a target nucleic acid molecule, wherein the target sequence comprises a target strand and a non-target strand, said method comprising:

a) assembling an RNA-guided nuclease (RGN) ribonucleotide complex by combining:

i) one or more guide RNAs capable of hybridizing to the non-target strand of the target sequence; and

ii) an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs:1-20;

under conditions suitable for formation of the RGN ribonucleotide complex; and

b) contacting said target nucleic acid molecule or a cell comprising said target nucleic acid molecule with the assembled RGN ribonucleotide complex;

wherein the one or more guide RNAs hybridize to the non-target strand of the target sequence, thereby directing said RGN polypeptide to bind to said target sequence.

101. The method of claim 100, wherein said method is performed *in vitro*, *in vivo*, or *ex vivo*.

102. The method of claim 100 or 101, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target sequence.

103. The method of claim 100 or 101, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby allowing for the modulation of expression of a target gene comprising said target sequence.

104. The method of claim 100 or 101, wherein said RGN polypeptide further comprises a base-editing polypeptide, thereby allowing for the modification of said target nucleic acid molecule.

105. The method of claim 104, wherein said base-editing polypeptide comprises a deaminase.

106. The method of claim 100 or 101, wherein said RGN polypeptide is capable of cleaving said target nucleic acid molecule, thereby allowing for the cleaving and/or modifying of said target nucleic acid molecule.

107. A method for cleaving and/or modifying a target nucleic acid molecule comprising a target sequence, wherein said target sequence comprises a target strand and a non-target strand, said method comprising contacting the target nucleic acid 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: 1-20; and

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

wherein the one or more guide RNAs hybridize to the non- target strand of the target sequence, thereby directing said RGN polypeptide to bind to said target nucleic acid molecule and cleavage and/or modification of said target nucleic acid molecule occurs.

108. The method of claim 107, wherein cleavage by said RGN polypeptide generates a double-stranded break.

109. The method of claim 107, wherein cleavage by said RGN polypeptide generates a single-stranded break.

110. The method of claim 107, wherein said RGN polypeptide is nuclease inactive or a nickase and is operably fused to a base-editing polypeptide.

111. The method of claim 110, wherein the base-editing polypeptide is a deaminase.

112. The method of claim 107, wherein said modified target nucleic acid molecule comprises insertion of heterologous DNA into the target nucleic acid molecule.

113. The method of claim 107, wherein said modified target nucleic acid moleculecomprises deletion of at least one nucleotide from the target nucleic acid molecule.

114. The method of claim 107, wherein said modified target nucleic acid moleculecomprises mutation of at least one nucleotide in the target nucleic acid molecule.

115. The method of any one of claims 106-113, wherein said target sequence is located adjacent to a protospacer adjacent motif (PAM).

116. The method of any one of claims 107-115, wherein said target sequence is a eukaryotic target sequence.

117. The method of any one of claims 107-116, wherein said gRNA is a single guide RNA (sgRNA).

118. The method of any one of claims 107-116, wherein said gRNA is a dual-guide RNA.

119. The method of any one of claims 107-118, wherein said RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-20.

120. The method of any one of claims 107-118, wherein:

a) said RGN has at least 90% sequence identity to SEQ ID NO: 1, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42;

b) said RGN has at least 90% sequence identity to SEQ ID NO: 2, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43;

c) said RGN has at least 90% sequence identity to SEQ ID NO: 3, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) said RGN has at least 90% sequence identity to SEQ ID NO: 4, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45;

e) said RGN has at least 90% sequence identity to SEQ ID NO: 5, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-

17 of SEQ ID NO: 1041 or 1042, a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) said RGN has at least 90% sequence identity to SEQ ID NO: 6, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) said RGN has at least 90% sequence identity to SEQ ID NO: 7, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) said RGN has at least 90% sequence identity to SEQ ID NO: 8, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49;

i) said RGN has at least 90% sequence identity to SEQ ID NO: 9, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50;

j) said RGN has at least 90% sequence identity to SEQ ID NO: 10, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51;

k) said RGN has at least 90% sequence identity to SEQ ID NO: 11, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52;

l) said RGN has at least 90% sequence identity to SEQ ID NO: 12, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53;

m) said RGN has at least 90% sequence identity to SEQ ID NO: 13, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54;

n) said RGN has at least 90% sequence identity to SEQ ID NO: 14, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55;

o) said RGN has at least 90% sequence identity to SEQ ID NO: 15, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56;

p) said RGN has at least 90% sequence identity to SEQ ID NO: 16, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57;

q) said RGN has at least 90% sequence identity to SEQ ID NO: 17, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58;

r) said RGN has at least 90% sequence identity to SEQ ID NO: 18, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 59 or 60;

s) said RGN has at least 90% sequence identity to SEQ ID NO: 19, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61; and

t) said RGN has at least 90% sequence identity to SEQ ID NO: 20, said guide RNA comprises a crRNA repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62.

121. The method of any one of claims 107-118, wherein:

a) said RGN has at least 100% sequence identity to SEQ ID NO: 1, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42;

b) said RGN has at least 100% sequence identity to SEQ ID NO: 2, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43;

c) said RGN has at least 100% sequence identity to SEQ ID NO: 3, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040;

d) said RGN has at least 100% sequence identity to SEQ ID NO: 4, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45;

e) said RGN has at least 100% sequence identity to SEQ ID NO: 5, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042;

f) said RGN has at least 100% sequence identity to SEQ ID NO: 6, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043;

g) said RGN has at least 100% sequence identity to SEQ ID NO: 7, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides

1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045;

h) said RGN has at least 100% sequence identity to SEQ ID NO: 8, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49;

i) said RGN has at least 100% sequence identity to SEQ ID NO: 9, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50;

j) said RGN has at least 100% sequence identity to SEQ ID NO: 10, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51;

k) said RGN has at least 100% sequence identity to SEQ ID NO: 11, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52;

l) said RGN has at least 100% sequence identity to SEQ ID NO: 12, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53;

m) said RGN has at least 100% sequence identity to SEQ ID NO: 13, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54;

n) said RGN has at least 100% sequence identity to SEQ ID NO: 14, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55;

o) said RGN has at least 100% sequence identity to SEQ ID NO: 15, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56;

p) said RGN has at least 100% sequence identity to SEQ ID NO: 16, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57;

q) said RGN has at least 100% sequence identity to SEQ ID NO: 17, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58;

r) said RGN has at least 100% sequence identity to SEQ ID NO: 18, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 38, or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 59 or 60;

s) said RGN has at least 100% sequence identity to SEQ ID NO: 19, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61; and

t) said RGN has at least 100% sequence identity to SEQ ID NO: 20, said guide RNA comprises a crRNA repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62.

122. The method of any one of claims 107-121, wherein the target sequence is within a cell.

123. The method of claim 122, further comprising culturing the cell under conditions in which the RGN polypeptide is expressed and cleaves and modifies the target nucleic acid molecule to produce a DNA molecule comprising a modified target nucleic acid molecule; and selecting a cell comprising said modified target nucleic acid molecule.

124. A cell comprising a modified target nucleic acid moleculeaccording to the method of claim 123.

125. A plant comprising the cell of claim 124.

126. A seed comprising the cell of claim 124.

127. A pharmaceutical composition comprising the cell of claim 124 and a pharmaceutically acceptable carrier.

128. A method for producing a genetically modified cell with a correction in a causal mutation for a genetically inherited disease, the method comprising introducing into the cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell

whereby the RGN and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

129. The method of claim 128, wherein the RGN is nuclease inactive or a nickase and is fused to a polypeptide which has base-editing activity.

130. The method of claim 129, wherein the base-editing polypeptide is a deaminase.

131. The method of any one of claims 128-129, wherein the genetically inherited disease is caused by a single nucleotide polymorphism.

132. The method of claim 131, wherein the gRNA further comprises a spacer that targets a region proximal to the causal single nucleotide polymorphism.

133. A method for producing a genetically modified cell with a deletion in a disease-causing expanded trinucleotide repeat, the method comprising introducing into the cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-20, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a first guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein the gRNA comprises a spacer that targets the 5' flank of the expanded trinucleotide repeat; and

c) a second guide RNA (gRNA) or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell, and further wherein said second gRNA comprises a spacer that targets the 3'flank of the expanded trinucleotide repeat;

whereby the RGN and the two gRNAs target to the expanded trinucleotide repeat and at least a portion of the expanded trinucleotide repeat is removed.

134. The method of claim 133, wherein the first gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat.

135. The method of claim 134, wherein the second gRNA further comprises a spacer that targets a region within or proximal to the expanded trinucleotide repeat.

136. The method of any one of claims 133-135, wherein said RGN polypeptide has 100% sequence identity to any one of SEQ ID NOs: 1-20.

137. The method of any one of claims 133-136, wherein said gRNA, said first gRNA, said second gRNA, or said first gRNA and said second gRNA is selected from a gRNA selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 44 or nucleotides 19-

111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10;

k) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 58 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 90% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 90% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 20.

138. The method of any one of claims 133-136, wherein said gRNA, said first gRNA, said second gRNA, or said first gRNA and said second gRNA is selected from a gRNA selected from the group consisting of:

a) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 21 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 42, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 1;

b) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 22 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 43, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 2;

c) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 23 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 44 or nucleotides 19-111 of SEQ ID NO: 1040, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 3;

d) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 24 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 45, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 4;

e) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 25 or nucleotides 1-17 of SEQ ID NO: 1041 or 1042, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 46 or nucleotides 22-85 of SEQ ID NO: 1041 or 1042, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 5;

f) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 26 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 47 or nucleotides 24-138 of SEQ ID NO: 1043, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 6;

g) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 27 or nucleotides 1-22 of SEQ ID NO: 1044 or 1045, and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 48 or nucleotides 27-96 of SEQ ID NO 1044 or nucleotides 27-95 of SEQ ID NO: 1045, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 7;

h) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 28 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 49, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 8;

i) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 29 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 50, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 9;

j) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 30 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 51, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 10;

k) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 31 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 52, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 11;

l) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 32 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 53, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 12;

m) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 33 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 54, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 13;

n) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 34 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 55, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 14;

o) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 35 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 56, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 15;

p) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 36 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 57, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 16;

q) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 37 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 17;

r) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 38 or 39 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 58 or 60, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 18;

s) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 40 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 61, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 19; and

t) a gRNA comprising a CRISPR repeat having at least 100% sequence identity to SEQ ID NO: 41 and a tracrRNA having at least 100% sequence identity to SEQ ID NO: 62, wherein said RGN polypeptide has an amino acid sequence having at least 100% sequence identity to SEQ ID NO: 20.

139. A method of treating a disease, disorder, or condition, said method comprising administering to a subject in need thereof a pharmaceutical composition of claim 92 or 127.

140. The method of claim 139, wherein said disease, disorder, or condition is associated with a causal mutation and said pharmaceutical composition corrects said causal mutation.

141. The method of claim 139 or 140, wherein said subject is at risk of developing said disease, disorder, or condition.

142. A single guide RNA comprising the nucleic acid molecule comprising said crRNA of any one of claims 42-45 and the nucleic acid molecule comprising said tracrRNA of any one of claims 53-57.

143. A dual guide RNA comprising the nucleic acid molecule comprising said crRNA of any one of claims 42-45 and the nucleic acid molecule comprising said tracrRNA of any one of claims 53-57.

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2023/058160

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/22 C12N15/10 C12N15/113 C07K14/32 C12N15/11
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DANIEL H. HAFT ET AL: "A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes", PLOS COMPUTATIONAL BIOLOGY, vol. 1, no. 6, 1 January 2005 (2005-01-01) , page e60, XP055445295, US ISSN: 1553-734X, DOI: 10.1371/journal.pcbi.0010060 Identifier: WP_242196486 from Bacillus cereus group sp. BfR-BA-01492 -/--	1-143

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 November 2023	23/01/2024

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer
Bladier, Cecile

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2023/058160

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2023/058160

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:

1-143(partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	& "type II CRISPR RNA-guided endonuclease Cas9 [Bacillus cereus group sp. - Protein - NCBI", , 28 March 2022 (2022-03-28), XP093099273, Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/protein/WP_242196486 [retrieved on 2023-11-08] the whole document	
X	US 2019/264232 A1 (HOU ZHENGLIN [US] ET AL) 29 August 2019 (2019-08-29) claim 3; sequence 1042	1-143
A	DING WENTAO ET AL: "Development and Application of CRISPR/Cas in Microbial Biotechnology", FRONTIERS IN BIOENGINEERING AND BIOTECHNOLOGY, vol. 8, 1 January 2020 (2020-01-01), page 711, XP055808086, the whole document	1-143
A	WO 2019/195379 A1 (LIFEEDIT INC [US]) 10 October 2019 (2019-10-10) claims 1-47	1-143
A	WO 2016/186946 A1 (PIONEER HI BRED INT [US]) 24 November 2016 (2016-11-24) claims 1-24	1-143

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10165 (SEQ ID NO: 1) and related crRNA and tracrRNA sequences (SEQ ID NOs: 21 and 42); and related embodiments.

---

2. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10166 (SEQ ID NO: 2) and related crRNA and tracrRNA sequences (SEQ ID NOs: 22 and 43); and related embodiments.

---

3. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10167 (SEQ ID NO: 3) and related crRNA and tracrRNA sequences (SEQ ID NOs: 23 and 44); and related embodiments.

---

4. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10168 (SEQ ID NO: 4) and related crRNA and tracrRNA sequences (SEQ ID NOs: 24 and 45); and related embodiments.

---

5. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10169 (SEQ ID NO: 5) and related crRNA and tracrRNA sequences (SEQ ID NOs: 25 and 46); and related embodiments.

---

6. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10171 (SEQ ID NO: 6) and related crRNA and tracrRNA sequences (SEQ ID NOs: 26 and 47); and related embodiments.

---

7. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10186 (SEQ ID NO: 7) and related crRNA and tracrRNA sequences (SEQ ID NOs: 27 and 48); and related embodiments.

---

8. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10190 (SEQ ID NO: 8) and

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

related crRNA and tracrRNA sequences (SEQ ID NOs: 28 and 49); and related embodiments.

---

9. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10191 (SEQ ID NO: 9) and related crRNA and tracrRNA sequences (SEQ ID NOs: 29 and 50); and related embodiments.

---

10. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10194 (SEQ ID NO: 10) and related crRNA and tracrRNA sequences (SEQ ID NOs: 30 and 51); and related embodiments.

---

11. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10195 (SEQ ID NO: 11) and related crRNA and tracrRNA sequences (SEQ ID NOs: 31 and 52); and related embodiments.

---

12. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10196 (SEQ ID NO: 12) and related crRNA and tracrRNA sequences (SEQ ID NOs: 32 and 53); and related embodiments.

---

13. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10197 (SEQ ID NO: 13) and related crRNA and tracrRNA sequences (SEQ ID NOs: 33 and 54); and related embodiments.

---

14. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10198 (SEQ ID NO: 14) and related crRNA and tracrRNA sequences (SEQ ID NOs: 34 and 55); and related embodiments.

---

15. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10200 (SEQ ID NO: 15) and related crRNA and tracrRNA sequences (SEQ ID NOs: 35 and 56); and related embodiments.

---

16. claims: 1-143(partially)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

RNA-guided nuclease labeled LPG10203 (SEQ ID NO: 16) and related crRNA and tracrRNA sequences (SEQ ID NOs: 36 and 57); and related embodiments.

---

17. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10204 (SEQ ID NO: 17) and related crRNA and tracrRNA sequences (SEQ ID NOs: 37 and 58); and related embodiments.

---

18. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10205 (SEQ ID NO: 18) and related crRNA and tracrRNA sequences (SEQ ID NOs: 38, 39 and 59, 60); and related embodiments.

---

19. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10207 (SEQ ID NO: 19) and related crRNA and tracrRNA sequences (SEQ ID NOs: 40 and 61); and related embodiments.

---

20. claims: 1-143(partially)

RNA-guided nuclease labeled LPG10208 (SEQ ID NO: 20) and related crRNA and tracrRNA sequences (SEQ ID NOs: 41 and 62); and related embodiments.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IB2023/058160

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2019264232 A1	29-08-2019	CA 3091267 A1	29-08-2019
		CN 112020554 A	01-12-2020
		EP 3755792 A1	30-12-2020
		KR 20200124702 A	03-11-2020
		US 2019264232 A1	29-08-2019
		US 2022010293 A1	13-01-2022
		WO 2019165168 A1	29-08-2019
WO 2019195379 A1	10-10-2019	US 2021172008 A1	10-06-2021
		WO 2019195379 A1	10-10-2019
WO 2016186946 A1	24-11-2016	AU 2016263026 A1	09-11-2017
		AU 2016265560 A1	09-11-2017
		AU 2022259752 A1	01-12-2022
		BR 112017024534 A2	24-07-2018
		BR 112017024535 A2	24-07-2018
		CA 2985079 A1	24-11-2016
		CA 2985506 A1	24-11-2016
		CN 107709562 A	16-02-2018
		CN 107849546 A	27-03-2018
		EP 3294877 A1	21-03-2018
		EP 3294878 A1	21-03-2018
		KR 20180002852 A	08-01-2018
		KR 20180008572 A	24-01-2018
		US 2018258417 A1	13-09-2018
		US 2018346895 A1	06-12-2018
		US 2019136248 A1	09-05-2019
		US 2022356485 A1	10-11-2022
		WO 2016186946 A1	24-11-2016
		WO 2016186953 A1	24-11-2016