Claims

1. A protein having an amino acid sequence selected from:

(i) as shown in any one of SEQ ID NO: 1, 2, 3;

(ii) a sequence having an amino acid sequence which has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity compared to any one of SEQ ID NOs: 1, 2, and 3;

the protein is an effector protein in the CRISPR/Cas system.

2. The protein of claim 1, comprising a sequence selected from the following, or consists of a sequence selected from the following:

(A) (i) as shown in SEQ ID NO: 1;

(ii) a sequence having an amino acid sequence which has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity compared to SEQ ID NOs: 1;

for example, the protein has an amino acid sequence as shown in SEQ ID NO:1; or

(B) (i) as shown in SEQ ID NO: 2;

(ii) a sequence having an amino acid sequence which has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity compared to SEQ ID NOs: 2;

for example, the protein has an amino acid sequence as shown in SEQ ID NO:2; or

(C) (i) as shown in SEQ ID NO: 3;

(ii) a sequence having an amino acid sequence which has at least at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity compared to SEQ ID NOs: 3;

for example, the protein has an amino acid sequence as shown in SEQ ID NO:3.

3. A conjugate comprising the protein of claim 1 and a modified portion,

wherein

(1) the modified portion is optionally selected from an additional protein or polypeptide, a detectable label, and any combinations thereof, wherein the additional protein or polypeptide is optionally selected from an epitope tag, a reporter gene sequence, a nuclear localization signal (NLS) sequence, a targeting moiety, a transcription activation domain (for example, VP64), a transcription repression domain (for example, KRAB domain or SID domain), a nuclease domain (for example, Fok1), a domain having an activity selected from: nucleotide deaminase, methylase activity, demethylase, transcription activation activity, transcription inhibition activity, transcription release factor activity, histone modification activity, nuclease activity, single-stranded RNA cleavage activity, double-stranded RNA cleavage activity, single-stranded DNA cleavage activity, double-stranded DNA cleavage activity, and nucleic acid binding activity; and any combinations thereof; and/or

(2) the modified portion is optionally connected to the N-terminus or C-terminus of the protein through a linker; and/or

(3) the conjugate optionally comprises an epitope tag; and/or

(4) the conjugate comprises an NLS sequence;

for example, the NLS sequence is shown in SEQ ID NO: 19;

for example, the NLS sequence is located at, near, or close to the end of the protein (e.g., N-terminal or C-terminal).

4. A fusion protein comprising the protein of claim 1 and an additional protein or polypeptide.

5. The fusion protein of claim 4, wherein

(1) the additional protein or polypeptide is linked to the N-terminus or C-terminus of the protein through a linker; and/or

(2) the additional protein or polypeptide is selected from an epitope tag, a reporter gene sequence, a nuclear localization signal (NLS) sequence, a targeting moiety, a transcription activation domain (for example, VP64), a transcription repression domain (for example, KRAB domain or SID domain), a nuclease domain (for example, Fok1), a domain having an activity selected from: a nucleotide deaminase, methylase activity, a demethylase, transcription activation activity, transcription inhibition activity, transcription release factor activity, histone modification activity, nuclease activity, single-stranded RNA cleavage activity, double-stranded RNA cleavage activity, single-stranded DNA cleavage activity, double-stranded DNA cleavage activity, and nucleic acid binding activity ; and any combinations thereof; and/or

(3) the fusion protein comprises an epitope tag; and/or

(4) the fusion protein comprises an NLS sequence;

for example, the NLS sequence is shown in SEQ ID NO: 19;

for example, the NLS sequence is located at, near, or close to the end of the protein (e.g., N-terminal or C-terminal); and/or

(5) the fusion protein has an amino acid sequence as shown in SEQ ID NO:20.

6. An isolated nucleic acid molecule comprising a sequence selected from the following or consisting of a sequence selected from the following:

(i) a sequence as shown in SEQ ID NO: 7 or 13;

(ii) a sequence having at least 90%, at least 95% sequence identity with the sequence as shown in SEQ ID NO:7 or 13;

(iii) a sequence that hybridizes to the sequence as described in any one of (i) to (ii) under stringent conditions; or

(iv) a complementary sequence of the sequence as described in any one of (i)-(ii);

in addition, the sequence as described in any one of (ii)-(iv) substantially retains the biological function of the sequence from which it is derived; the isolated nucleic acid molecule is a direct repeat sequence in the CRISPR/Cas system, which forms a complex with the protein of claim 1;

for example, the isolated nucleic acid molecule is RNA.

7. The isolated nucleic acid molecule of claim 6, wherein

(1) the nucleic acid molecule comprises one or more stem loops or optimized secondary structures;

for example, the sequence as described in any one of (ii) to (iv) retains the secondary structure of the sequence from which it is derived; and/or

(2) the nucleic acid molecule comprises a sequence selected from the following, or consists of a sequence selected from the following:

(a) a nucleotide sequence as shown in SEQ ID NOs: 7 or 13;

(b) a sequence that hybridizes to the sequence as described in (a) under stringent conditions; or

(c) a complementary sequence of the nucleotide sequence as shown in SEQ ID NO: 7 or 13.

8. A complex comprising:

(i) a protein component, which is selected from: the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, and any combinations thereof; and

(ii) a nucleic acid component, which comprises a direct repeat sequence and a targeting sequence capable of hybridizing to the target sequence from 5' to 3' direction, the direct repeat sequence can form a complex with the protein of claim 1;

wherein the protein component and the nucleic acid component combine with each other to form a complex;

for example, the nucleic acid component is a guide RNA in the CRISPR/Cas

system;

for example, the nucleic acid molecule is RNA;

for example, the complex does not contain trans-activating crRNA (tracrRNA),

wherein the targeting sequence is optionally attached to the 3' end of the nucleic acid molecule and/or the targeting sequence optionally comprises a complementary sequence of the target sequence.

9. An isolated nucleic acid molecule comprising:

(i) a nucleotide sequence encoding the protein of claim 1 or the fusion protein of any one of claims 4-5;

(ii) a nucleotide sequence encoding a direct repeat sequence in the CRISPR/Cas system, which can form a complex with the protein of claim 1; and/or

(iii) a nucleotide sequence containing (i) and (ii);

for example, the nucleotide sequence described in any one of (i) to (iii) is codon-optimized for expression in a prokaryotic cell or an eukaryotic cell.

10. A vector comprising the isolated nucleic acid molecule of claim 9.

11. A host cell comprising the isolated nucleic acid molecule of claim 9 or the vector of claim 10.

12. A composition comprising:

(i) a first component, which is selected from: the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, a nucleotide sequence encoding the protein or fusion protein, and any combinations thereof; and

(ii) a second component, which is a nucleotide sequence containing a guide RNA, or a nucleotide sequence encoding the nucleotide sequence containing a guide RNA;

wherein the guide RNA includes a direct repeat sequence and a targeting

sequence from the 5' to 3', and the targeting sequence can hybridize with the target sequence;

the targeting RNA can form a complex with the protein, conjugate or fusion protein as described in (i);

for example, the direct repeat sequence is an isolated nucleic acid molecule as defined in any one of claims 6-7;

for example, the composition does not contain a trans-activating crRNA (tracrRNA).

13. A composition comprising one or more vectors comprising:

(i) a first nucleic acid, which is a nucleotide sequence encoding a protein of claim 1 or the fusion protein any one of claims 4-5; optionally, the first nucleic acid is operationally linked to a first regulatory element; and

(ii) a second nucleic acid, which encodes a nucleotide sequence comprising a guide RNA; optionally the second nucleic acid is operationally linked to a second regulatory element;

wherein:

the first nucleic acid and the second nucleic acid are present on the same or different vectors;

the guide RNA includes a direct repeat sequence and a targeting sequence from the 5' to 3', and the targeting sequence can hybridize with the target sequence;

the guide RNA can form a complex with the effector protein of claim 1 or fusion protein as described in (i);

for example, the direct repeat sequence is an isolated nucleic acid molecule as defined in any one of claims 6-7;

for example, the composition does not contain a trans-activating crRNA (tracrRNA),

wherein the first regulatory element and/or the second regulatory element is optionally a promoter, such as an inducible promoter.

14. The composition of any one of claims 12-13, wherein

(1) at least one component of the composition is non-naturally occurring or modified; and/or

(2) the targeting sequence is connected to the 3' end of the direct repeat sequence; and/or

(3) the targeting sequence comprises a complementary sequence of the target sequence; and/or

(4) when the target sequence is DNA, the target sequence is located at the 3' end of the protospacer adjacent motif (PAM), and the PAM has a sequence shown by 5'-TTN, wherein N is selected from A, G, T, C; when the target sequence is RNA, the target sequence does not have PAM domain restrictions; and/or

(5) the target sequence is a DNA or RNA sequence derived from a prokaryotic cell or an eukaryotic cell; or, the target sequence is a non-naturally occurring DNA or RNA sequence; and/or

(6) the target sequence is present in a cell;

for example, the target sequence is present in a nucleus or in a cytoplasm (e.g., organelle);

for example, the cell is an eukaryotic cell;

for example, the cell is a prokaryotic cell; and/or

(7) the protein is linked to one or more NLS sequences, or the conjugate or fusion protein comprises one or more NLS sequences;

for example, the NLS sequence is connected to the N-terminus or C-terminus of the protein;

for example, the NLS sequence is fused to the N-terminus or C-terminus of the protein.

15. A kit comprising one or more components selected from: the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, the

isolated nucleic acid molecule of any one of claims 6-7, the complex of claim 8, the isolated nucleic acid molecule of claim 9, the vector of claim 10, the composition of any one of claims 12-14;

for example, the kit comprises the composition of claim 17, and instructions for using the composition;

for example, the kit includes the composition of claim 18, and instructions for using the composition.

16. A delivery composition comprising a delivery vehicle and one or more selected from the group consisting of: the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, the isolated nucleic acid molecule of any one of claims 6-7, the complex of claim 8, the isolated nucleic acid molecule of claim 9, the vector of claim 10, the composition of any one of claims 12-14;

for example, the delivery vehicle is a particle;

for example, the delivery vehicle is selected from a lipid particle, sugar particle, metal particle, protein particle, liposome, exosome, microvesicle, gene gun, or viral vector (e.g., replication defective retrovirus, lentivirus, adenovirus or adeno-associated virus).

17. A non-diagnostic and non-therapeutic method for modifying a target gene, comprising: contacting the complex of claim 8 or the composition of any one of claims 12-14 with the target gene, or delivering that to a cell containing the target gene; the target sequence is present in the target gene.

18. The method of claim 17, wherein

the modification refers to a break in the target sequence, such as a double-strand break in DNA or a single-strand break in RNA;

for example, the modification also includes the insertion of an exogenous nucleic acid into the break.

19. A non-diagnostic and non-therapeutic method for altering the expression of a gene product, comprising: combining the complex of claim 8 or the composition of any one of claims 12-14 with a nucleic acid molecule encoding the gene product, or delivering that to a cell containing the nucleic acid molecule in which the target sequence is present, wherein the expression of the gene product is optionally altered, preferably the gene product is a protein.

20. The method of any one of claims 17-19,

(1) wherein the nucleic acid molecule is present in a cell;

for example, the cell is a prokaryotic cell;

for example, the cell is an eukaryotic cell, such as a mammalian cell (such as a human cell) or a plant cell; or.

the nucleic acid molecule is present in a nucleic acid molecule (e.g., a plasmid) *in vitro*; and/or

(2) wherein the protein, conjugate, fusion protein, isolated nucleic acid molecule, complex, vector or composition is contained in a delivery vehicle;

for example, the delivery vehicle is selected from a lipid particle, sugar particle, metal particle, protein particle, liposome, exosome, viral vector (such as replication-defective retrovirus, lentivirus, adenovirus or adeno-associated virus); and/or

(3) which is used to change one or more target sequences in a target gene or a nucleic acid molecule encoding a target gene product to modify a cell, cell line, or organism.

21. An in vitro, isolated or *in vivo* cell or cell line or progeny thereof, the cell or cell line or the progeny thereof comprises: the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, the isolated nucleic acid molecule of any one of claims 6-7, the complex of claim 8, the isolated nucleic acid

molecule of claim 9, the vector of claim 10, the composition of any one of claims 12-14;

for example, the cell is an eukaryotic cell;

for example, the cell is an animal cell (for example, a mammalian cell, such as a human cell) or a plant cell;

for example, the cell is a stem cell or stem cell line.

22. Use of the protein of claim 1, the conjugate of claim 3, the fusion protein of any one of claims 4-5, the isolated nucleic acid molecule of any one of claims 6-7, the complex of claim 8, the isolated nucleic acid molecule of claim 9, the vector of claim 10, the composition of any one of claims 12-14 or the kit of claim 15 in the preparation of a preparation for:

(i) the isolated gene or genome editing;

(ii) the detection of an isolated single-stranded DNA;

(iii) editing the target sequence in the target locus to modify a biological or non-human organism;

(iv) treating the disease caused by defects in the target sequence in the target locus.