

CLAIMS

WHAT IS CLAIMED IS:

1. A variant Cas12i4 polypeptide comprising a sequence having at least 95% identity to a sequence set forth in any one of SEQ ID NOs: 3-59.
2. The variant Cas12i4 polypeptide of claim 1, wherein the variant Cas12i4 polypeptide:
 - i) is a variant of a parent polypeptide of SEQ ID NO: 2;
 - ii) comprises a substitution of Table 2;
 - iii) comprises the sequence set forth in any one of SEQ ID NOs: 3-59; and/or
 - iv) the variant Cas12i4 polypeptide comprises a RuvC domain or a split RuvC domain.
3. The variant Cas12i4 polypeptide of claim 1 or 2, comprising the sequence set forth in SEQ ID NO: 4.
4. The variant Cas12i4 polypeptide of any one of claims 1 to 3, wherein the variant Cas12i4 polypeptide exhibits increased binary complex formation with an RNA guide or increased nuclease activity, relative to a parent polypeptide, optionally wherein a binary complex comprising the variant Cas12i4 polypeptide exhibits increased stability, relative to a parent binary complex.
5. The variant Cas12i4 polypeptide of any one of claims 1 to 4, wherein the variant Cas12i4 polypeptide further comprises a substitution of Table 4, optionally wherein the substitution of Table 4 increases binary complex formation with an RNA guide or increases stability of a binary complex comprising the variant Cas12i4 polypeptide, relative to a parent polypeptide.
6. The variant Cas12i4 polypeptide of any one of claims 1 to 5, wherein the substitution is a substitution of Table 5, Table 6, Table 7, Table 8, Table 9, Table 10, and/or Table 11.
7. A composition comprising the variant Cas12i4 polypeptide of any one of claims 1 to 6, wherein the composition further comprises an RNA guide or a nucleic acid encoding the RNA guide, wherein the RNA guide comprises a direct repeat sequence and a spacer sequence.
8. The variant Cas12i4 polypeptide of any one of claims 1-7 or the composition of claim 7, wherein the variant Cas12i4 polypeptide further comprises a nuclear localization signal (NLS), a peptide tag, a fluorescent protein, a base-editing domain, a DNA methylation domain, a histone residue modification domain, a localization factor, a transcription modification factor, a light-gated control factor, a chemically inducible factor, or a chromatin visualization factor.
9. A composition comprising a nucleic acid that encodes the Cas12i4 polypeptide of any one of claims 1-7 or 8.
10. The composition of claim 9, wherein the nucleic acid is:

- i) codon-optimized for expression in a cell;
 - ii) operably linked to a promoter; and/or
 - iii) in a vector, optionally wherein the vector comprises a retroviral vector, a lentiviral vector, a phage vector, an adenoviral vector, an adeno-associated vector, or a herpes simplex vector.
11. The variant Cas12i4 polypeptide of any one of claims 1-8 or the composition of claim 9 or 10, wherein the variant Cas12i4 polypeptide or the nucleic acid encoding the variant Cas12i4 polypeptide is present in a delivery system comprising a nanoparticle (e.g., a lipid nanoparticle), a liposome, an exosome, a microvesicle, or a gene-gun.
12. The composition of claim 7, wherein the direct repeat comprises:
- a. nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - b. nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - c. nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - d. nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - e. nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - f. nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - g. nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - h. nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;

- i. nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - j. nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - k. nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - l. nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - m. nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - n. nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124; or
 - o. a sequence that is at least 90% identical to a sequence of SEQ ID NO: 61 or a portion thereof.
13. An RNA guide or a nucleic acid encoding the RNA guide, wherein the RNA guide comprises a direct repeat sequence comprising:
- a. nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - b. nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - c. nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - d. nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;

- e. nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - f. nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - g. nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - h. nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - i. nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - j. nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - k. nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - l. nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - m. nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124;
 - n. nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, or 124; or
 - o. a sequence that is at least 90% identical to a sequence of SEQ ID NO: 61 or a portion thereof.
14. The composition of any one of claims 7, 8, or 12 or the RNA guide or the nucleic acid encoding the RNA guide of claim 13, wherein the direct repeat sequence comprises

AGN₁N₂N₃N₄GUGUN₅N₆N₇CAGN₈GACN₉C (SEQ ID NO: 125), wherein N₁ is A or G, N₂ is C or U, N₃ is A or G, N₄ is U or C, N₅ is C or U, N₆ is C or U, N₇ is U, A, C, or G, N₈ is U or C, and N₉ is A or C.

15. The RNA guide or the nucleic acid encoding the RNA guide of claim 13 or 14, wherein the RNA guide further comprises a spacer sequence.
16. The composition of any one of claims 7, 8, 12, 16, or 15 or the RNA guide or the nucleic acid encoding the RNA guide of claim 15, wherein the spacer sequence:
 - i) comprises about 15 to about 35 nucleotides in length;
 - ii) recognizes a target nucleic acid; and/or
 - iii) comprises sequence complementarity to a target strand sequence of a target nucleic acid.
17. The composition of claim 16 or the RNA guide or the nucleic acid encoding the RNA guide of claim 16, wherein the target nucleic acid comprises a target sequence adjacent to a protospacer adjacent motif (PAM) sequence, optionally wherein the PAM sequence comprises a nucleotide sequence set forth as 5'-TTN-3', 5'-NTTN-3', 5'-NTN'-3', 5'-NNTN-3', 5'-VTN-3', or 5'-NVTN-3', wherein N is any nucleotide (e.g., A, G, T, or C) and V is A, G, or C.
18. A composition comprising the RNA guide or the nucleic acid encoding the RNA guide of any one of claims 13 to 17.
19. The composition of claim 18, wherein the composition is a delivery composition comprising a nanoparticle (e.g., a lipid nanoparticle), a liposome, an exosome, a microvesicle, or a gene-gun.
20. A cell comprising variant Cas12i4 polypeptide of any one of claims 1 to 6, 8 or 11, the composition of any one of claims 7 to 12, 14, or 16 to 19, or the RNA guide or the nucleic acid encoding the RNA guide of any one of claims 13 to 17.
21. The cell of claim 20, wherein the cell is a eukaryotic cell, optionally wherein the eukaryotic cell is a mammalian cell (e.g., a human cell) or a plant cell.
22. A method for editing a gene in a cell, the method comprising contacting the cell with a variant of any one of claims 1 to 6, 8 or 11; a composition of any one of claims 7 to 10, 12, 18, or 19; or an RNA guide of any one of claims 13 to 17.
23. A nucleic acid molecule encoding a Cas12i4 variant of SEQ ID NO: 4, wherein the sequence of the nucleic acid molecule is 95% identical to any one of SEQ ID NOs: 222-228.