CLAIMS

1. A variant of a reference CasX protein (CasX variant), wherein:

a. the CasX variant comprises at least one modification in the reference CasX protein, wherein the reference CasX protein comprises the sequence of SEQ ID NO: 1, SEQ ID NO: 2, or SEQ ID NO: 3; and

b. the CasX variant exhibits at least one improved characteristic as compared to the reference CasX protein.

2. The CasX variant of claim 1, wherein the improved characteristic of the CasX variant is selected from the group consisting of: improved folding of the CasX variant; improved binding affinity to a guide nucleic acid (gNA); improved binding affinity to a target DNA; improved ability to utilize a greater spectrum of one or more PAM sequences, including ATC, CTC, GTC, or TTC, in the editing of target DNA; improved unwinding of the target DNA; increased editing activity; improved editing efficiency; improved editing specificity; increased nuclease activity; increased target strand loading for double strand cleavage; decreased target strand loading for single strand nicking; decreased off-target cleavage; improved binding of non-target DNA strand; improved protein stability; improved protein solubility; improved protein:gNA complex (RNP) stability; improved protein:gNA complex solubility; improved protein yield; improved protein expression; improved fusion characteristics or a combination thereof.

3. The Cas X variant of claim 1 or 2, wherein the at least one modification comprises:

a. at least one amino acid substitution in a domain of the CasX variant;

b. at least one amino acid deletion in a domain of the CasX variant;

c. at least one amino acid insertion in a domain of the CasX variant;

d. a substitution of all or a portion of a domain from a different CasX;

e. a deletion of all or a portion of a domain of the CasX variant; or

f. any combination of (a)-(e).

4. The CasX variant of any one of claims 1-3, wherein the at least one modification is in a domain selected from:

a. a non-target strand binding (NTSB) domain;

b. a target strand loading (TSL) domain;

c. a helical I domain;

d. a helical II domain;

e. an oligonucleotide binding domain (OBD); or

f. a RuvC DNA cleavage domain.

5. The CasX variant of claim 4, wherein:

the at least one modification in the TSL domain comprises an amino acid substitution of one or more of amino acids Y857, S890, or S932 of SEQ ID NO: 2;

the at least one modification in the helical I domain comprises an amino acid substitution of one or more of amino acids S219, L249, E259, Q252, E292, L307, or D318 of SEQ ID NO: 2;

the at least one modification in the helical II domain comprises an amino acid substitution of one or more of amino acids D361, L379, E385, E386, D387, F399, L404, R458, C477, or D489 of SEQ ID NO: 2;

the at least one modification in the OBD comprises an amino acid substitution of one or more of amino acids F536, E552, T620, or I658 of SEQ ID NO: 2; and/or

the at least one modification in the RuvC DNA cleavage domain comprises an amino acid substitution of one or more of amino acids K682, G695, A708, V711, D732, A739, D733, L742, V747, F755, M771, M779, W782, A788, G791, L792, P793, Y797, M799, Q804, S819, or Y857 or a deletion of amino acid P793 of SEQ ID NO: 2.

6. The CasX variant of claim 5, wherein the modification results in an increased ability to edit the target DNA.

7. The CasX variant of any one of the preceding claims, wherein the CasX variant is capable of forming a ribonuclear protein complex (RNP) with a guide nucleic acid (gNA).

8. The CasX variant of any one of the preceding claims, wherein the CasX variant comprises two or more modifications in one domain, or wherein the CasX variant comprises modifications in two or more domains.

9. The CasX variant of any one of claims 1-4, wherein the at least one modification compared to the reference CasX sequence of SEQ ID NO: 2 is selected from one or more of:

a. an amino acid substitution of L379R;
b. an amino acid substitution of A708K;
c. an amino acid substitution of T620P;
d. an amino acid substitution of E385P;
e. an amino acid substitution of Y857R;
f. an amino acid substitution of I658V;
g. an amino acid substitution of F399L;
h. an amino acid substitution of Q252K;
i. an amino acid substitution of L404K; and
j. an amino acid deletion of P793.

10. The CasX variant of any one of claims 1-4, wherein the CasX variant has a sequence selected from the group consisting of the sequences of Tables 3, 8, 9, 10 and 12, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about

80%, at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, sequence identity thereto.

11. The CasX variant of any one of the preceding claims, further comprising a heterologous protein or domain thereof fused to the CasX.

12. The CasX variant of claim 13, wherein the heterologous protein or domain thereof comprises one or more nuclear localization signals (NLS) that are selected from the group of sequences consisting of PKKKRKV (SEQ ID NO: 352), KRPAATKKAGQAKKKK (SEQ ID NO: 353), PAAKRVKLD (SEQ ID NO: 354), RQRRNELKRSP (SEQ ID NO: 355), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 356), RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 357), VSRKRPRP (SEQ ID NO: 358), PPKKARED (SEQ ID NO: 35(), PQPKKKPL (SEQ ID NO: 360), SALIKKKKKMAP (SEQ ID NO: 361), DRLRR (SEQ ID NO: 362), PKQKKRK (SEQ ID NO: 363), RKLKKKIKKL (SEQ ID NO: 364), REKKKFLKRR (SEQ ID NO: 365), KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 366), RKCLQAGMNLEARKTKK (SEQ ID NO: 367), PRPRKIPR (SEQ ID NO: 368), PPRKKRTVV (SEQ ID NO: 369), NLSKKKKRKREK (SEQ ID NO: 370), RRPSRPFRKP (SEQ ID NO: 371), KRPRSPSS (SEQ ID NO: 372), KRGINDRNFWRGENERKTR (SEQ ID NO: 373), PRPPKMARYDN (SEQ ID NO: 374), KRSFSKAF (SEQ ID NO: 375), KLKIKRPVK (SEQ ID NO: 376), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 377), PKTRRRPRRSQRKRPPT (SEQ ID NO: 378), SRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 379), KTRRRPRRSQRKRPPT (SEQ ID NO: 380), RRKKRRPRRKKRR (SEQ ID NO: 381), PKKKSRKPKKKSRK (SEQ ID NO: 382), HKKKHPDASVNFSEFSK (SEQ ID NO: 383), QRPGPYDRPQRPGPYDRP (SEQ ID NO: 384), LSPSLSPLLSPSLSPL (SEQ ID NO: 385), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 386), PKRGRGRPKRGRGR (SEQ ID NO: 387), and PKKKRKVPPPPKKKRKV (SEQ ID NO: 389), wherein the one or more NLS are positioned at or near the C-terminus of the CasX protein; at or near at the N-terminus of the CasX protein; or at or near the N-terminus and at or near the C-terminus of the CasX protein.

13. The CasX variant of any one of claims 2-12, wherein one or more of the improved characteristics of the CasX variant is at least about 1.1 to about 100-fold or more improved relative to the reference CasX protein of SEQ ID NO: 1, SEQ ID NO: 2, or SEQ ID NO: 3.

14. The CasX variant of any one of the preceding claims, wherein the RNP comprising the CasX variant exhibits greater editing efficiency and/or binding of a target sequence in the target DNA when any one of the PAM sequences TTC, ATC, GTC, or CTC is located 1 nucleotide 5’ to the non-target strand of the protospacer having identity with the targeting sequence of the gNA in a cellular assay system compared to the editing efficiency and/or binding of an RNP comprising a reference CasX protein in a comparable assay system.

15. The CasX variant of any one of the preceding claims, wherein the CasX variant protein comprises a nuclease domain having nickase activity, wherein the CasX variant protein comprises a nuclease domain having double-stranded cleavage activity, or wherein the CasX protein is a catalytically inactive CasX (dCasX) protein, and wherein the dCasX and the gNA retain the ability to bind to the target DNA.

16. The CasX variant of claim 15, wherein the CasX protein is catalytically inactive and wherein the dCasX comprises a mutation at residues:

a. D672, and/or E769, and/or D935 corresponding to the CasX protein of SEQ ID NO:1; or

b. D659, and/or E756, and/or D922 corresponding to the CasX protein of SEQ ID NO: 2.

17. The CasX variant of any one of the preceding claims, wherein the CasX variant comprises a first domain from a first CasX protein and second domain from a second CasX protein different from the first CasX protein.

18. The CasX variant of any one of claim 17, wherein the first domain comprises a portion of the sequence selected from the group consisting of amino acids 1-56, 57-100, 101-191, 192-332, 333-509, 510-660, 661-824, 825-934, and 935-986 of SEQ ID NO: 1 and the second domain comprises a portion of the sequence selected from the group consisting of amino acids 1-58, 59-102, 103-192, 193-333, 334-501, 502-647, 648-812, 813-921, and 922-978 of SEQ ID NO: 2.

19. The CasX variant of any one of claims 1-18, wherein the CasX variant comprises a sequence selected from the group consisting of SEQ ID NOS: 328-336 and 4413-4415.

20. The CasX variant of any one of claims 1-16, wherein the CasX variant comprises at least one chimeric domain comprising a first part from a first CasX protein and a second part from a second CasX protein different from the first CasX protein.

21. A variant of a reference guide nucleic acid scaffold (gNA variant) capable of binding a reference CasX protein or a CasX variant, wherein:

a. the gNA variant comprises at least one modification compared to the reference guide nucleic acid scaffold sequence; and

b. the gNA variant exhibits one or more improved characteristics compared to the reference guide nucleic acid scaffold comprising a sequence selected from the group consisting of SEQ ID NOS: 4-16.

22. The gNA variant of claim 21, wherein the one or more improved characteristics is selected from the group consisting of: improved stability; improved solubility; improved transcription of the gNA; improved resistance to nuclease activity; increased folding rate of the gNA; decreased side product formation during folding; increased productive folding;

improved binding affinity to a CasX protein; improved binding affinity to a target DNA when complexed with the CasX protein; improved gene editing when complexed with the CasX protein; improved specificity of editing when complexed with the CasX protein; and improved ability to utilize a greater spectrum of one or more PAM sequences, including ATC, CTC, GTC, or TTC, in the editing of target DNA when complexed with the CasX protein.

23. The gNA variant of claim 21 or 22, wherein the at least one modification comprises:

a. at least one nucleotide substitution in a region of the gNA variant;

b. at least one nucleotide deletion in a region of the gNA variant;

c. at least one nucleotide insertion in a region of the gNA variant;

d. a substitution of all or a portion of a region of the gNA variant;

e. a deletion of all or a portion of a region of the gNA variant; or

f. any combination of (a)-(e);

wherein the region of the gNA variant is selected from the group consisting of extended stem loop, scaffold stem loop, triplex, and pseudoknot.

24. The gNA variant of claim 23, wherein:

the scaffold stem comprises a bubble;

the scaffold further comprises a triplex loop region; and/or

the scaffold further comprises a 5' unstructured region.

25. The gNA variant of claim 23 or 24, wherein the at least one modification comprises:

a. a substitution of 1 to 15 consecutive or non-consecutive nucleotides in the gNA variant in one or more regions;

b. a deletion of 1 to 10 consecutive or non-consecutive nucleotides in the gNA variant in one or more regions;

c. an insertion of 1 to 10 consecutive or non-consecutive nucleotides in the gNA variant in one or more regions;

d. a substitution of the scaffold stem loop or the extended stem loop with an RNA stem loop sequence from a heterologous RNA source with proximal 5' and 3' ends; or

e. any combination of (a)-(d).

26. The gNA variant of any one of claims 21-25, comprising an extended stem loop region comprising at least 10, at least 100, at least 500nucleotides.

27. The gNA variant of claim 26, wherein the heterologous RNA stem loop sequence is selected from MS2, Qβ, U1 hairpin II, Uvsx, or PP7 stem loops.

28. The gNA variant of any one of claims 25-27, wherein the at least one modification compared to the reference guide scaffold of SEQ ID NO: 5 is selected from one or more of:

a. a C18G substitution in the triplex loop;

b. a G55 insertion in the stem bubble;

c. a U1 deletion;

d. a modification of the extended stem loop wherein

i. a 6 nt loop and 13 loop-proximal base pairs are replaced by a Uvsx hairpin; and

ii. a deletion of A99 and a substitution of G64U that results in a loop-distal base that is fully base-paired.

29. The gNA variant of claim 28, comprising the sequence of any one of SEQ ID NOS: 2236, 2237, 2238, 2241, 2244, 2248, 2249, or 2259-2280, wherein the gNA variant further comprises a targeting sequence, wherein the targeting sequence is complementary to the target DNA sequence.

30. The gNA variant of claim 29, wherein the targeting sequence has 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35 nucleotides.

31. The gNA variant of claim 29 or 30, wherein the gNA is a single-guide gNA comprising the scaffold sequence linked to the targeting sequence.

32. The gNA variant of any one of claims 21-31, wherein the one or more of the improved characteristics of the CasX variant is at least about 1.1 to about 100-fold or more improved relative to the reference gNA of SEQ ID NO: 4 or SEQ ID NO: 5.

33. The gNA variant of any one of claims 21-32, wherein the scaffold of the gNA variant sequence comprises a sequence selected from the group of sequences of SEQ ID NOS: 2101-2280, or having at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto.

34. The gNA variant of claim 33, wherein the scaffold of the gNA variant sequence consists of a sequence selected from the group of sequences of SEQ ID NOS: 2101-2280.

35. The gNA variant of any one of claims 21-34, further comprising a protein binding motif; and/or a thermostable stem loop.

36. A gene editing pair comprising a CasX protein variant of any one of claims 1-20 and a gNA comprising:

a. the gNA variant of any one of claims 21-35; and

b. a targeting sequence, wherein the targeting sequence is complementary to the target DNA.

37. A method of editing a target DNA, comprising contacting the target DNA with the gene editing pair of claim 36, wherein the contacting results in editing or modification of the target DNA.

38. The method of claim 37, comprising contacting the target DNA with a plurality of gNAs comprising targeting sequences complementary to different or overlapping regions of the target DNA.

39. The method of claim 37 or 38, wherein:

the contacting by the gene editing pair comprises binding the target DNA and results in introducing a mutation, an insertion, or a deletion in the target DNA;

the contacting introduces one or more single-stranded breaks in the target DNA and wherein the editing comprises introducing a mutation, an insertion, or a deletion in the target DNA; or

the contacting comprises introducing one or more double-stranded breaks in the target DNA and wherein the editing comprises introducing a mutation, an insertion, or a deletion in the target DNA.

40. The method of claim 39, further comprising contacting the target DNA with a nucleotide sequence of a donor template nucleic acid wherein the donor template comprises a nucleotide sequence having homology to the target DNA, wherein the donor template comprises homologous arms on the 5' and 3' ends of the donor template.

41. The method of any one of claims 37-40, wherein editing occurs in vitro outside of a cell, editing occurs in vitro inside of a cell, or in vivo inside of a cell.

42. The method of claim 41, wherein the cell is a eukaryotic cell.

43. The method of claim 41 or 42, wherein the cell is in a subject.

44. The method of claim 42 or 43, wherein the method comprises contacting the eukaryotic cell with a vector encoding or comprising the CasX protein and the gNA, and optionally further comprising the donor template.

45. The method of claim 44, wherein the vector is an Adeno-Associated Viral (AAV) vector, a lentiviral vector, a non-viral particle or a virus-like particle (VLP).

46. A polynucleotide encoding the CasX variant of any one of claims 1-20, and/or the gNA variant of any one of claims 21-35.

47. A vector comprising the polynucleotide of claim 46.

48. A vector comprising encoding the CasX variant of any one of claims 1-20 and/or the gNA variant of any one of claims 21-35.

49. The composition of claim 48, wherein the CasX protein and the gNA are associated together in a ribonuclear protein complex (RNP).

50. The composition claim 48 or 49, further comprising a donor template nucleic acid wherein the donor template comprises a nucleotide sequence having homology to a target DNA.

51. A kit, comprising the CasX variant of any one of claims 1-20, the gNA of any one of claims 21-35 or the gene editing pair of claim 36.