MODIFIED CPF1 GUIDE RNA

INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ELECTRONICALLY

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This patent application claims priority to U.S. Provisional Patent Application No. 62/567,123, filed on October 2, 2017; U.S. Provisional Patent Application No. 62/617,138, filed on January 12, 2018; and U.S. Provisional Application No. 62/697,327, file on July 12, 2018, the entire disclosures of which are hereby incorporated by referenced.

BACKGROUND OF THE INVENTION

[0002] RNA-guided endonucleases have proven to be an effective tool for genome engineering in multiple cell types and microorganisms. RNA-guided endonucleases generate site-specific double-stranded DNA breaks or single-stranded DNA breaks within target nucleic acids. When the cleavage of a target nucleic acid occurs within a cell, the break in the nucleic acid can be repaired by non-homologous end joining (NHEJ) or homology directed repair (HDR).

[0003] Direct delivery of RNA-guided endonucleases and their gene-editing components (e.g., guide RNA) into cells both *in vitro* and *in vivo* has tremendous potential as a therapeutic strategy for treating genetic diseases. Currently, however, direct delivery of these components into cells with reasonable efficiency is challenging.

[0004] Therefore, there is a need to identify new compositions and related methods to improve cellular delivery and other properties of RNA-guided endonucleases that will enhance genome engineering. This invention provides such compositions and related methods.

BRIEF SUMMARY

[0005] The disclosure is a nucleic acid comprising of a Cpf1 crRNA with an extension sequence. In one aspect, the present disclosure refers to a nucleic acid comprising a Cpf1 crRNA, an extension sequence 5' of the crRNA, and, optionally, a processing sequence between the crRNA and the extension sequence, wherein the processing sequence is a sequence that is self-cleaved by Cpf1. In another aspect, the present disclosure refers to a composition comprising the

nucleic acid as described herein and a carrier, and optionally further comprising a Cpf1 protein. In one example, the nucleic acid comprises a Cpf1 crRNA and an extension sequence on the 5' end of the Cpf1 crRNA, wherein the extension sequence comprises fewer than about 60 nucleotides. In another example, the nucleic acid comprises a Cpf1 crRNA, a processing sequence 5' of the Cpf1 crRNA, and an extension sequence 5' of the processing sequence. Also provided is a composition comprising the nucleic acid, a carrier, and optionally a Cpf1 protein or vector encoding the protein.

[0006] The disclosure also provides a method for genetically modifying a eukaryotic target cell. The method involves contacting the eukaryotic target cell with the nucleic acid containing the Cpf1 crRNA as described herein.

[0007] These and other examples of the disclosure are described in greater detail in the following sections.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S)

[0008] Figure 1A is a graph comparing delivery of unmodified crRNA in complex with Cpf1 using either lipofectamine or electroporation (nucleofection).

[0009] Figures 1B and 1C are graphs comparing the lipofectamine mediated delivery and NHEJ generation of Cpf1-crRNA complexes comprising unmodified crRNA (41 nucleotides long) and extended crRNA.

[0010] Figure 1D is a schematic illustration of the structures of 41-nucleotide (nt) unmodified crRNA and extended crRNAs. The arrows represent Cpf1 cleavage sites.

[0011] Figure 1E is a graph showing NHEJ efficiency for the crRNA constructs illustrated in Fig. 1D.

[0012] Figure 1F is a graph depicting the cellular delivery of Cpf1 RNP using lipofectamine and crRNAs with different length labeled with fluorescent dye.

[0013] Figure 2 provides a graph (right panel) illustrating gene editing efficiency as a function of GFP knockdown for 5' extended crRNA delivered to GFP-HEK cells via electroporation, and a schematic illustration of the crRNA (left panel).

[0014] Figure 3A is a schematic illustration of in vivo studies in AI9 mice.

[0015] Figure 3B is a schematic illustration of gastrocnemius muscle injection site and imaging sections.

CLAIMS:

1. A nucleic acid comprising a Cpf1 crRNA, an extension sequence 5’ of the crRNA, and, optionally, a processing sequence between the crRNA and the extension sequence, wherein the processing sequence is a sequence that is self-cleaved by Cpf1.

2. The nucleic acid of claim 1, wherein the nucleic acid comprises a processing sequence, and the processing sequence comprises a fragment of a direct repeat sequence of a Cpf1 array, wherein the direct repeat sequence comprises a crRNA sequence portion and a processing portion positioned 5’ of the crRNA sequence portion, and the fragment comprises least 5 contiguous nucleotides of the processing portion of the direct repeat sequence.

3. The nucleic acid of claim 2, wherein the processing sequence comprises a fragment of at least 10 nucleotides of the processing portion of the direct repeat sequence.

4. The nucleic acid of claim 2, wherein the processing sequence comprises the entire processing portion of the direct repeat sequence.

5. The nucleic acid of any of claims 1-4, wherein the extension sequence does not comprise the processing sequence or the sequence of the crRNA.

6. The nucleic acid of any of claims 1-5, wherein the extension sequence comprises at least 2 nucleotides.

7. The nucleic acid of any of claims 1-6, wherein the extension sequence comprises 10 to 100 nucleotides.

8. The nucleic acid of any of claims 1-7, wherein the nucleic acid contains only a single Cpf1 crRNA sequence.

9. The nucleic acid of any of claims 1-8, further comprising a second processing sequence 5’ of the extension sequence, and a second extension sequence 5’ of the second processing sequence.

10. The nucleic acid of any of claims 1-9, further comprising a donor nucleic acid hybridized or covalently linked thereto.

11. The nucleic acid of claim 10, wherein the donor nucleic acid is covalently linked 5’ of the processing sequence, or 5’ of the extension sequence.

12. The nucleic acid of claim 11, wherein the donor nucleic acid is linked to the processing sequence or extension sequence by a linker group.

13. The nucleic acid of claim 10, wherein the donor nucleic acid is hybridized to the extension sequence and/or processing sequence.

14. The nucleic acid of any of claims 1-13 further comprising a targeting nucleic acid 3’ of the Cpf1 crRNA.

15. The nucleic acid of any of claims 1-14, wherein the extension sequence comprises fewer than about 60 nucleotides.

16. The nucleic acid of claim 15, wherein the extension sequence comprises fewer than about 20 nucleotides.

17. The nucleic acid of claim 15, wherein the extension sequence comprises about 2-20 nucleotides.

18. The nucleic acid of claim 1, wherein the nucleic acid does not comprise a processing sequence.

19. The nucleic acid of claim 18, wherein the nucleic acid further comprises a donor nucleic acid covalently linked 5’ of the extension sequence.

20. The nucleic acid of claim 19, wherein the donor nucleic acid is linked by a linker group.

21. The nucleic acid of claim 18, wherein the nucleic acid further comprises a donor nucleic acid hybridized to the extension sequence.

22. The nucleic acid of any of claims 18-21 further comprising a targeting nucleic acid 3’ of the Cpf1 crRNA.

23. The nucleic acid of any of claims 1-22, wherein the extension sequence comprises a self-hybridizing sequence.

24. The nucleic acid of any of claims 1-23, wherein the extension sequence comprises a semi-stable hairpin structure, a stable hairpin structure, a pseudoknot structure, a G-quadraplex structure, a bulge loop structure, an internal loop structure, a branch loop structure, or a combination thereof.

25. The nucleic acid of claim 24, wherein the extension sequence comprises a repeating trinucleotide motif.

26. The nucleic acid of any of claims 25, wherein the repeating trinucleotide motif is CAA, UUG, AAG, CUU, CCU, CCA, UAA, or a combination thereof.

27. The nucleic acid of any of claims 25, wherein the repeating trinucleotide motif is CAU, CUA, UUA, AUG, UAG, or a combination thereof.

28. The nucleic acid of any of claims 25, wherein the repeating trinucleotide motif is CGA, CGU, CGG, CAG, CUG, CCG, or a combination thereof.

29. The nucleic acid of any of claims 25, wherein the repeating trinucleotide motif is a CNG motif, or combination of CNG motifs, optionally with CGA or CGU.

30. The nucleic acid of any of claims 25, wherein the repeating trinucleotide motif is AGG, UGG, or combination thereof.

31. The nucleic acid of any of claims 1-25, wherein the extension sequence comprises a combination of the repeating trinucleotide motifs of any of claims 26-30.

32. The nucleic acid of any of claims 1-31, wherein the extension sequence or portion thereof is resistant to nuclease degradation.

33. The nucleic acid of any of claims 1-32, wherein the extension sequence comprises one or more modified internucleotide bonds.

34. The nucleic acid of claim 33, wherein a region of 4 or more contiguous nucleotides of the extension sequence, or the entire extension sequence, has modified internucleotide bonds.

35. The nucleic acid of claim 33 or 34, wherein the modified internucleotide bonds comprise phosphorothioate, phosphorodithioate, methylphosphonate, phosphoramidate, 2'-O-methyl, 2'-O-methoxyethyl, 2'-fluoro, bridged nucleic acid (BNA), or phosphotriester modified bonds, or a combination thereof.

36. The nucleic acid of any of claims 1-35, wherein the extension sequence comprises one or more xeno nucleic acids (XNA).

37. The nucleic acid of any of claims 1-36, wherein the nucleic acid further comprises a biotin and/or avidin or streptavidin molecule attached to the 5' terminus of the extension sequence.

38. The nucleic acid of claim 37, wherein the nucleic acid further comprises a biotin molecule attached to the 5' terminus of the extension sequence, an avidin or streptavidin molecule conjugated to the biotin molecule, and, optionally, a targeting construct comprising a biotin molecule conjugated to the avidin or streptavidin.

39. The nucleic acid of claim 38, wherein the targeting construct is a peptide comprising a biotin molecule attached thereto.

40. A composition comprising the nucleic acid of any of claims 1-39 and a carrier, and optionally further comprising a Cpf1 protein.

41. The composition of claim 40, wherein the composition is substantially free of a divalent metal ion that promotes Cpf1 cleavage.

42. The composition of claim 40 or 41, wherein the composition comprises cationic lipids.

43 The composition of any of claims 40-42, wherein the nucleic acid is in a liposome.

44. The composition of any of claims 40-42, wherein the nucleic acid is partially or totally encapsulated by a polymer nanoparticle, or attached to a metal or polymer nanoparticle.

45. The nucleic acid of any one of claims 1-39 or composition of any of claims 40-44 for use in genetically modifying a eukaryotic target cell.