CLAIMS

What is claimed is:

1. A composition comprising:
a) a CasY polypeptide, or a nucleic acid molecule encoding the CasY polypeptide; and
b) a CasY guide RNA, or one or more DNA molecules encoding the CasY guide RNA.

2. The composition of claim 1, wherein the CasY polypeptide comprises an amino acid sequence having 50% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

3. The composition of claim 1 or claim 2, wherein the CasY guide RNA comprises a nucleotide sequence having 80% or more identity with the crRNA sequence set forth in any one of SEQ ID NOs: 11-15.

4. The composition of claim 1 or claim 2, wherein the CasY polypeptide is fused to an NLS sequence.

5. The composition of any one of claims 1-4, wherein the composition comprises a lipid.

6. The composition of any one of claims 1-4, wherein a) and b) are within a liposome.

7. The composition of any one of claims 1-4, wherein a) and b) are within a particle.

8. The composition of any one of claims 1-7, comprising one or more of: a buffer, a nuclease inhibitor, and a protease inhibitor.

9. The composition of any one of claims 1-8, wherein the CasY polypeptide comprises an amino acid sequence having 85% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

10. The composition of any one of claims 1-9, wherein the CasY polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

11. The composition of any one of claims 1-9, wherein the CasY polypeptide is a catalytically inactive CasY Polypeptide (dCasY).

12. The composition of claim 10 or claim 11, wherein the CasY polypeptide comprises one or more mutations at a position corresponding to those selected from: D672, E769, and D935 of SEQ ID NO: 1.

13. The composition of any one of claims 1-12, further comprising a DNA donor template.

14. A CasY fusion polypeptide comprising: a CasY polypeptide fused to a heterologous polypeptide.

15. The CasY fusion polypeptide of Claim 14, wherein the CasY polypeptide comprises an amino acid sequence having 50% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

16. The CasY fusion polypeptide of Claim 14, wherein the CasY polypeptide comprises an amino acid sequence having 85% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

17. The CasY fusion polypeptide of any one of claims 14-16, wherein the CasY polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

18. The CasY fusion polypeptide of any one of claims 14-17, wherein the CasY polypeptide is a catalytically inactive CasY Polypeptide (dCasY).

19. The CasY fusion polypeptide of claim 17 or claim 18, wherein the CasY polypeptide comprises one or more mutations at a position corresponding to those selected from: D672, E769, and D935 of SEQ ID NO: 1.

20. The CasY fusion polypeptide of any one of claims 14-19, wherein the heterologous polypeptide is fused to the N-terminus and/or the C-terminus of the CasY polypeptide.

21. The CasY fusion polypeptide of any one of claims 14-20, comprising an NLS.

22. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide is a targeting polypeptide that provides for binding to a cell surface moiety on a target cell or target cell type.

23. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA.

24. The CasY fusion polypeptide of claim 23, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

25. The CasY fusion polypeptide of claim 24, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

26. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

27. The CasY fusion polypeptide of claim 26, wherein the heterologous polypeptide exhibits histone modification acitivity.

28. The CasY fusion polypeptide of claim 26 or claim 27, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

29. The CasY fusion polypeptide of claim 28, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

30. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide is an endosomal escape polypeptide.

31. The CasY fusion polypeptide of claim 30, wherein the endosomal escape polypeptide comprises an amino acid sequence selected from: GLFXALLXLLXSLWXLLLXA (SEQ ID NO:94), and GLFHALLHLLHSLWHLLLHA (SEQ ID NO:95), wherein each X is independently selected from lysine, histidine, and arginine.

32. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide is a chloroplast transit peptide.

33. The CasY fusion polypeptide of claim 32, wherein the chloroplast transit peptide comprises an amino acid sequence selected from
MASMISSSAVTTVSRASRGQSAAMAPFGGLKSMTGFPVRKVNTDITSITSNGGRVKCMQVWPPI GKKKFETLSYLPPLTRDSRA (SEQ ID NO:83);
MASMISSSAVTTVSRASRGQSAAMAPFGGLKSMTGFPVRKVNTDITSITSNGGRVKS (SEQ ID NO:84);
MASSMLSSATMVASPAQATMVAPFNGLKSSAAFPATRKANNDITSITSNGGRVNCMQVWPPIEK KKFETLSYLPDLTDSGGRVNC (SEQ ID NO:85);
MAQVSRICNGVQNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWGLKKSGMTLIGSELRPL KVMSSVSTAC (SEQ ID NO:86);
MAQVSRICNGVWNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWGLKKSGMTLIGSELRPL KVMSSVSTAC (SEQ ID NO:87);
MAQINNMAQGIQTLNPNSNFHKPQVPKSSSFLVFGSKKLKNSANSMLVLKKDSIFMQLFCSFRIS ASVATAC (SEQ ID NO:88);
MAALVTSQLATSGTVLSVTDRFRRPGFQGLRPRNPADAALGMRTVGASAAPKQSRKPHRFDRR CLSMVV (SEQ ID NO:89);
MAALTTSQLATSATGFGIADRSAPSSLLRHGFQGLKPRSPAGGDATSLSVTTSARATPKQQRSVQ RGSRRFPSVVVC (SEQ ID NO:90);
MASSVLSSAAVATRSNVAQANMVAPFTGLKSAASFPVSRKQNLDITSIASNGGRVQC (SEQ ID

NO:91);
MESLAATSVFAPSRVAVPAARALVRAGTVVPTRRTSSTSGTSGVKCSAAVTPQASPVISRSAAAA (SEQ ID NO:92); and
MGAAATSMQSLKFSNRLVPPSRRLSPVPNNVTCNNLPKSAAPVRTVKCCASSWNSTINGAAATT NGASAASS (SEQ ID NO:93).

34. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide is protein that increases or decreases transcription.

35. The CasY fusion polypeptide of claim 34, wherein the heterologous polypeptide is a transcriptional repressor domain.

36. The CasY fusion polypeptide of claim 34, wherein the heterologous polypeptide is a transcriptional activation domain.

37. The CasY fusion polypeptide of any one of claims 14-21, wherein the heterologous polypeptide is a protein biding domain.

38. A nucleic acid molecule encoding the CasY fusion polypeptide of any one of clamis 14-37.

39. The nucleic acid molecule of Claim 38, wherein the nucleotide sequence encoding the CasY fusion polypeptide is operably linked to a promoter.

40. The nucleic acid molecule of Claim 39, wherein the promoter is functional in a eukaryotic cell.

41. The nucleic acid molecule of Claim 40, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

42. The nucleic acid molecule of any one of Claims 39-41, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

43. The nucleic acid molecule of any one of Claims 38-42, wherein the DNA molecule is a recombinant expression vector.

44. The nucleic acid molecule of Claim 43, wherein the recombinant expression vector is a recombinant adenoassociated viral vector, a recombinant retroviral vector, or a recombinant lentiviral vector.

45. The nucleic acid molecule of Claim 39, wherein the promoter is functional in a prokaryotic cell.

46. The nucleic acid molecule of Claim 38, wherein the nucleic acid molecule is an mRNA.

47. One or more nucleic molecules encoding:
(a) a CasY guide RNA; and
(b) a CasY polypeptide.

48. The one or more nucleic acid molecules of claim 47, wherein the CasY polypeptide comprises an amino acid sequence having 50% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

49. The one or more nucleic acid molecules of claim 47, wherein the CasY polypeptide comprises an amino acid sequence having 85% or more identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

50. The one or more nucleic acid molecules of any one of claims 47-49, wherein the CasY guide RNA comprises a nucleotide sequence having 80% or more identity with the crRNA sequence set forth in any one of SEQ ID NOs: 11-15.

51. The one or more nucleic acid molecules of any one of claims 47-50, wherein the CasY polypeptide is fused to an NLS sequence.

52. The one or more nucleic acid molecules of any one of claims 47-51, wherein said one or more nucleic acid molecules comprises a nucleotide sequence encoding the CasY guide RNA that is operably linked to a promoter.

53. The one or more nucleic acid molecules of any one of claims 47-52, wherein said one or more nucleic acid molecules comprises a nucleotide sequence encoding the CasY polypeptide that is operably linked to a promoter.

54. The one or more nucleic acid molecules of Claim 52 or Claim 53, wherein the promoter operably linked to the nucleotide sequence encoding the CasY guide RNA, and/or the promoter operably linked to the nucleotide sequence encoding the CasY polypeptide, is functional in a eukaryotic cell.

55. The one or more nucleic acid molecules of Claim 54, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

56. The one or more nucleic acid molecules of any one of Claims 53-55, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

57. The one or more nucleic acid molecules of any one of Claims 47-56, wherein the one or more nucleic acid molecules is one or more recombinant expression vectors.

58. The one or more nucleic acid molecules of Claim 57, wherein the one or more recombinant expression vectors are selected from: one or more adenoassociated viral vectors, one or more recombinant retroviral vectors, or one or more recombinant lentiviral vectors.

59. The one or more nucleic acid molecules of Claim 53, wherein the promoter is functional in a prokaryotic cell.

60. A eukaryotic cell comprising one or more of:
a) a CasY polypeptide, or a nucleic acid molecule encoding the CasY polypeptide,
b) a CasY fusion polypeptide, or a nucleic acid molecule encoding the CasY fusion polypeptide, and
c) a CasY guide RNA, or a nucleic acid molecule encoding the CasY guide RNA.

61. The eukaryotic cell of claim 60, comprising the nucleic acid molecule encoding the CasY polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

62. The eukaryotic cell of claim 60 or claim 61, wherein the eukaryotic cell is a plant cell, a mammalian cell, an insect cell, an arachnid cell, a fungal cell, a bird cell, a reptile cell, an amphibian cell, an invertebrate cell, a mouse cell, a rat cell, a primate cell, a non-human primate cell, or a human cell.

63. A cell comprising a CasY fusion polypeptide, or a nucleic acid molecule encoding the CasY fusion polypeptide.

64. The cell of claim 63, wherein the cell is a prokaryotic cell.

65. The cell of claim 63 or claim 64, comprising the nucleic acid molecule encoding the CasY fusion polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

66. A method of modifying a target nucleic acid, the method comprising contacting the target nucleic acid with:
a) a CasY polypeptide; and
b) a CasY guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid,
wherein said contacting results in modification of the target nucleic acid by the CasY polypeptide.

67. The method of claim 66, wherein said modification is cleavage of the target nucleic acid.

68. The method of claim 66 or claim 67, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

69. The method of any of claims 66-68, wherein said contacting takes place *in vitro* outside of a cell.

70. The method of any of claims 66-68, wherein said contacting takes place inside of a cell in culture.

71. The method of any of claims 66-68, wherein said contacting takes place inside of a cell *in vivo.*

72. The method of claim 70 or claim 71, wherein the cell is a eukaryotic cell.

73. The method of claim 72, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod cell, a

cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non-human primate cell, and a human cell.

74. The method of claim 70 or claim 71, wherein the cell is a prokaryotic cell.

75. The method of any one of claims 66-74, wherein said contacting results in genome editing.

76. The method of any one of claims 66-75, wherein said contacting comprises: introducing into a cell: (a) the CasY polypeptide, or a nucleic acid molecule encoding the the CasY polypeptide, and (b) the CasY guide RNA, or a nucleic acid molecule encoding the the CasY guide RNA.

77. The method of claim 76, wherein said contacting further comprises: introducing a DNA donor template into the cell.

78. The method of any one of claims 66-77, wherein the CasY guide RNA comprises a nucleotide sequence having 80% or more identity with the crRNA sequence set forth in any one of SEQ ID NOs: 11-15.

79. The method of any one of claims 66-78, wherein the CasY polypeptide is fused to an NLS sequence.

80. A method of modulating transcription from a target DNA, modifying a target nucleic acid, or modifying a protein associated with a target nucleic acid, the method comprising contacting the target nucleic acid with:
a) a CasY fusion polypeptide comprising a CasY polypeptide fused to a heterologous polypeptide; and
b) a CasY guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid.

81. The method of claim 80, wherein the CasY guide RNA comprises a nucleotide sequence having 80% or more identity with the crRNA sequence set forth in any one of SEQ ID NOs: 11-15.

82. The method of claim 80 or claim 81, wherein the CasY fusion polypeptide comprises an NLS sequence.

83. The method of any of claims 80-82, wherein said modification is not cleavage of the target nucleic acid.

84. The method of any of claims 80-83, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

85. The method of any of claims 80-84, wherein said contacting takes place *in vitro* outside of a cell.

86. The method of any of claims 80-84, wherein said contacting takes place inside of a cell in culture.

87. The method of any of claims 80-84, wherein said contacting takes place inside of a cell *in vivo.*

88. The method of claim 86 or claim 87, wherein the cell is a eukaryotic cell.

89. The method of claim 88, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod cell, a cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non-human primate cell, and a human cell.

90. The method of claim 86 or claim 87, wherein the cell is a prokaryotic cell.

91. The method of any one of claims 80-90, wherein said contacting comprises: introducing into a cell: (a) the CasY fusion polypeptide, or a nucleic acid molecule encoding the the CasY fusion polypeptide, and (b) the CasY guide RNA, or a nucleic acid molecule encoding the the CasY guide RNA.

92. The method of any one of claims 80-91, wherein the CasY polypeptide is a catalytically inactive CasY Polypeptide (dCasY).

93. The method of any one of claims 80-92, wherein the CasY polypeptide comprises one or more mutations at a position corresponding to those selected from: D672, E769, and D935 of SEQ ID NO: 1.

94. The method of any one of claims 80-93, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA.

95. The method of claim 94, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

96. The method of claim 95, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

97. The method of any one of claims 80-93, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

98. The method of claim 97, wherein the heterologous polypeptide exhibits histone modification acitivity.

99. The method of claim 97 or claim 98, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

100. The method of claim 99, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

101. The method of any one of claims 80-93, wherein the heterologous polypeptide is protein that increases or decreases transcription.

102. The method of claim 101, wherein the heterologous polypeptide is a transcriptional repressor domain.

103. The method of claim 101, wherein the heterologous polypeptide is a transcriptional activation domain.

104. The method of any one of claims 80-93, wherein the heterologous polypeptide is a protein biding domain.

105. A transgenic, multicellular, non-human organism whose genome comprises a transgene comprising a nucleotide sequence encoding one or more of:
a) a CasY polypeptide,
b) a CasY fusion polypeptide, and
c) a CasY guide RNA.

106. The transgenic, multicellular, non-human organism of claim 105, wherein the CasY polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

107. The transgenic, multicellular, non-human organism of claim 105, wherein the CasY polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

108. The transgenic, multicellular, non-human organism of any one of claims 105-107, wherein the organism is a plant, a monocotyledon plant, a dicotyledon plant, an invertebrate animal, an insect, an arthropod, an arachnid, a parasite, a worm, a cnidarian, a vertebrate animal, a fish, a reptile, an amphibian, an ungulate, a bird, a pig, a horse, a sheep, a rodent, a mouse, a rat, or a non-human primate.

109. A system comprising:
a) a CasY polypeptide and a CasY guide RNA;
b) a CasY polypeptide, a CasY guide RNA, and a DNA donor template;
c) a CasY fusion polypeptide and a CasY guide RNA;
d) a CasY fusion polypeptide, a CasY guide RNA, and a DNA donor template;
e) an mRNA encoding a CasY polypeptide, and a CasY guide RNA;
f) an mRNA encoding a CasY polypeptide; a CasY guide RNA, and a DNA donor template;
g) an mRNA encoding a CasY fusion polypeptide, and a CasY guide RNA;

h) an mRNA encoding a CasY fusion polypeptide, a CasY guide RNA, and a DNA donor template;
i) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasY polypeptide; and ii) a nucleotide sequence encoding a CasY guide RNA;
j) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasY polypeptide; ii) a nucleotide sequence encoding a CasY guide RNA; and iii) a DNA donor template;
k) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasY fusion polypeptide; and ii) a nucleotide sequence encoding a CasY guide RNA; and
l) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasY fusion polypeptide; ii) a nucleotide sequence encoding a CasY guide RNA; and a DNA donor template.

110. The CasY system of claim 109, wherein the CasY polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

111. The CasY system of claim 109, wherein the CasY polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

112. The CasY system of any of claims 109-111, wherein the donor template nucleic acid has a length of from 8 nucleotides to 1000 nucleotides.

113. The CasY system of any of claims 109-111, wherein the donor template nucleic acid has a length of from 25 nucleotides to 500 nucleotides.

114. A kit comprising the CasY system of any one of claims 109-113.

115. The kit of claim 114, wherein the components of the kit are in the same container.

116. The kit of claim 114, wherein the components of the kit are in separate containers.

117. A sterile container comprising the CasY system of any one of claims 109-116.

118. The sterile container of claim 117, wherein the container is a syringe.

119. An implantable device comprising the CasY system of any one of claims 109-116.

120. The implantable device of claim 119, wherein the CasY system is within a matrix.

121. The implantable device of claim 119, wherein the CasY system is in a reservoir.

122. A method of identifying a CRISPR RNA-guided endonuclease, the method comprising:
detecting, in a plurality of metagenomic nucleotide sequences, a nucleotide sequence encoding a Cas1 polypeptide;
detecting a CRISPR array in the vicinity of the Cas1-encoding nucleotide sequence;
cloning, from a nucleic acid sample from which the plurality of metagenomic nucleotide sequences was derived, a CRISPR locus comprising the detected CRISPR array, into an expression vector to generate a recombinant CRISPR locus expression vector;
assaying the recombinant CRISPR locus expression vector for the ability to cleave a target nucleic acid, wherein a CRISPR locus that has the ability to cleave a target nucleic acid comprises a nucleotide sequence encoding a CRISPR RNA-guided endonuclease.
identifying, in the CRISPR locus, an open reading frame encoding a polypeptide that has less than 20% amino acid sequence identity to the amino acid sequence of a known CRISPR RNA-guided endonuclease polypeptide.

123. The method of Claim 122, wherein said assaying comprises introducing the recombinant CRISPR locus expression vector and a target nucleic into a cell.