



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

(86) **Date de dépôt PCT/PCT Filing Date:** 2022/10/28
(87) **Date publication PCT/PCT Publication Date:** 2023/05/11
(85) **Entrée phase nationale/National Entry:** 2024/04/30
(86) **N° demande PCT/PCT Application No.:** US 2022/048261
(87) **N° publication PCT/PCT Publication No.:** 2023/081070
(30) **Priorité/Priority:** 2021/11/02 (US63/274,667)

(51) **Cl.Int./Int.Cl. C12N 9/22** (2006.01),
C12N 9/78 (2006.01)
(71) **Demandeur/Applicant:**
UNIVERSITY OF MASSACHUSETTS MEDICAL
SCHOOL, US
(72) **Inventeurs/Inventors:**
SONTHEIMER, ERIK, US;
XUE, WEN, US;
ZHANG, HAN, US;
BAMIDELE, NATHAN, US;
DONG, XIAOLONG, US
(74) **Agent:** OSLER, HOSKIN & HARCOURT LLP

(54) **Titre : PROTEINES DE FUSION A DOMAINE NME2CAS9 INCRUSTE**
(54) **Titre: NME2CAS9 INLAID DOMAIN FUSION PROTEINS**

Fig. 1A



Fig. 1B

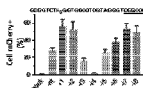


Fig. 1C



Fig. 1D

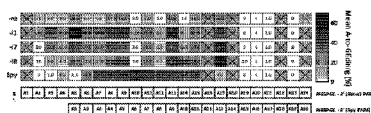


Fig. 1E



(57) **Abrégé/Abstract:**

The present invention is related to the field of gene editing. In particular, the gene editing is directed toward single nucleotide base editing. For example, such single nucleotide base editing results in a conversion of a mutated base pair to a wild type base pair. The high accuracy and precision of the presently disclosed single nucleotide base gene editor is accomplished by a fusion protein including an NmeCas9 nuclease and an inlaid nucleotide deaminase protein domain. The Nme2Cas9 protospacer interacting domain may be replaced with an SmuCas9 protospacer interacting domain.



Date Submitted: 2024/04/30

CA App. No.: 3236778

Abstract:

The present invention is related to the field of gene editing. In particular, the gene editing is directed toward single nucleotide base editing. For example, such single nucleotide base editing results in a conversion of a mutated base pair to a wild type base pair. The high accuracy and precision of the presently disclosed single nucleotide base gene editor is accomplished by a fusion protein including an NmeCas9 nuclease and an inlaid nucleotide deaminase protein domain. The Nme2Cas9 protospacer interacting domain may be replaced with an SmuCas9 protospacer interacting domain.

Nme2Cas9 Inlaid Domain Fusion Proteins

Field Of The Invention

The present invention is related to the field of gene editing. In particular, the gene editing is directed toward single nucleotide base editing. For example, such single nucleotide base editing results in a conversion of a mutated base pair to a wild type base pair. The high accuracy and precision of the presently disclosed single nucleotide base gene editor is accomplished by a fusion protein including an NmeCas9 nuclease and an inlaid nucleotide deaminase protein domain. The Nme2Cas9 protospacer interacting domain may be replaced with an SmuCas9 protospacer interacting domain.

Background

Many human diseases arise due to the mutation of a single base. The ability to correct such genetic aberrations is paramount in treating these genetic disorders. Clustered regularly interspaced short palindromic repeats (CRISPR) along with CRISPR associated (Cas) proteins comprise an RNA-guided adaptive immune system in archaea and bacteria. These systems provide immunity by targeting and inactivating nucleic acids that originate from foreign genetic elements.

SpyCas9 base editing platforms cannot be used to target all single-base mutations due to their limited editing windows. The editing window is constrained in part by the requirement for an NGG PAM and by the requirement that the edited base(s) be a very precise distance from the PAM. SpyCas9 is also intrinsically associated with high off-targeting effects in genome editing.

Gene editing using CRISPR-Cas9 technologies has advanced genetic research and promises to revolutionize gene therapy. Cytosine and adenine base editors (CBEs and ABEs) were developed as a way to precisely correct point mutations without inducing double-strand breaks or requiring a DNA donor. Base editors are comprised of a catalytically impaired Cas9 domain that is completely inactive or cleaves only one strand (a.k.a. dead/dCas9 or nickase/nCas9, respectively) fused to one or more cytosine deaminase (CBE) or adenine deaminase (ABE) domains. For efficient base editing to occur, the Cas9 base editor fusion must recognize a short sequence motif, called a PAM, adjacent to the target site, and a target adenine within an "editing window" upstream of PAM. The PAM and editing window are defined by the Cas domain, deaminase, and the type of fusion between the two effectors.

What is needed in the art is a highly accurate Cas9 single base editing platform having a programmable target specificity due to recognition of a diverse population of PAM sites and a superior gene editing efficiency to permit consistent clinical genetic mutation reversion to a wild type genotype.

5

Summary Of The Invention

The present invention is related to the field of gene editing. In particular, the gene editing is directed toward single nucleotide base editing. For example, such single nucleotide base editing results in a conversion of a C•G base pair to a T•A base pair or an A•T base pair to a G•C base pair.. The high accuracy and precision of the presently disclosed single nucleotide base gene editor is accomplished by a fusion protein including an NmeCas9 nuclease and an inlaid nucleotide deaminase protein domain. The Nme2Cas9 protospacer interacting domain may be replaced with an SmuCas9 protospacer interacting domain.

In one embodiment, the present invention contemplates a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain. In one embodiment, the inlaid NBE domain is an adenine base editor (ABE) domain. In one embodiment, the inlaid ABE domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein domain (ABE8e). In one embodiment, the inlaid NBE domain is a cytidine base editor (CBE) domain. In one embodiment, the inlaid CBE domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, said fusion protein further comprises a nuclear localization signal protein that includes, but is not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the protein further comprises a uracil glycosylase inhibitor. In one embodiment, the Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, the Nme2Cas9 (D16A) is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid domain protein. In one

embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain. In one embodiment, the inlaid NBE domain is an adenine base editor (ABE) domain. In one embodiment, the inlaid ABE domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid NBE domain is a cytidine base editor (CBE) domain. In one embodiment, the inlaid CBE protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid domain protein. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of a genetic disease; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of the genetic disease is reduced. In one embodiment, the genetic disease is caused by a gene with a mutated single base, wherein said gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In

one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to the N₄CC nucleotide sequence or N₄C nucleotide sequence. In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid NBE domain is an adenine base editor (ABE) domain.

5 In one embodiment, the inlaid ABE domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid NBE domain is a cytidine base editor (CBE) domain. In one embodiment, the inlaid CBE protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein includes, but is not limited to, 10 evoFERNY or rAPOBEC1. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further 15 comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further 20 comprises a linker that flanks at least one inlaid domain protein. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag. In one embodiment, the genetic disease includes, but is not limited to tyrosinemia, muscular dystrophy, Rett syndrome, Batten disease or amyotrophic lateral sclerosis (ALS).

25 In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of tyrosinemia; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of tyrosinemia is reduced. In one 30 embodiment, the patient further comprises a *Fah* gene with a mutated single base, wherein said *Fah* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one

embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to the N₄CC nucleotide sequence or N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an inlaid adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, said linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of muscular dystrophy; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of muscular dystrophy is reduced. In one embodiment, the patient further comprises a *Dmd* gene with a mutated single base, wherein said *Dmd* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or said N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine

deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an inlaid adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine nucleotide deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of Rett's syndrome; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of Rett's syndrome is reduced. In one embodiment, the patient further comprises a *MeCP2* gene with a mutated single base, wherein said *MeCP2* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an

adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one
5 embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

10 In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of Batten disease; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of Batten disease is
15 reduced. In one embodiment, the patient further comprises a *CLN3* gene with a mutated single base, wherein said *CLN3* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID.
20 In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine
25 deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase
30 inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A)

protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

- 5 In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient exhibiting at least one symptom of amyotrophic lateral sclerosis (ALS); and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of ALS is
- 10 reduced. In one embodiment, the patient further comprises a *SOD1* gene with a mutated single base, wherein said *SOD1* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID.
- 15 In one embodiment, the treating replaces said mutated single base with a wild type single base. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e). In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine
- 20 deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase
- 25 inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length.
- 30 In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a gene with a mutated single base, wherein said gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and a genetic disease does not develop. In one embodiment, the genetic disease includes, but is not limited to tyrosinemia, muscular dystrophy, Rett syndrome, Batten disease or ALS. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1, In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag. . In one embodiment, the gene includes, but is not limited to, a *Fah* gene, a *Dmd* gene, a *MeCP2* gene, a *CLN3* gene and an *SOD1* gene.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a *Fah* gene with a mutated single base, wherein said mutated *Fah* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an

Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and tyrosinemia does not develop. In one embodiment, the mutated *Fah* gene causes tyrosinemia. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a *Dmd* gene with a mutated single base, wherein said mutated *Dmd* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and muscular dystrophy does not develop. In one embodiment, the mutated *Dmd* gene causes muscular dystrophy. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9

PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a *MeCP2* gene with a mutated single base, wherein said mutated *MeCP2* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and Rett's syndrome does not develop. In one embodiment, the mutated *MeCP2* gene causes Rett's syndrome. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated

virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a *CLN3* gene with a mutated single base, wherein said mutated *CLN3* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and Batten disease does not develop. In one embodiment, the mutated *CLN3* gene causes Batten disease. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further

comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

In one embodiment, the present invention contemplates a method, comprising: a) providing; i) a patient comprising a *SOD1* gene with a mutated single base, wherein said mutated *SOD1* gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain; b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and amyotrophic lateral sclerosis (ALS) does not develop. In one embodiment, the mutated *MeCP2* gene causes Rett's syndrome. In one embodiment, the Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence. In one embodiment, the PID is an Nme2Cas9 PID or an SmuCas9 PID. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain. In one embodiment, the inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain. In one embodiment, the cytosine deaminase protein domain includes, but is not limited to, evoFERNY or rAPOBEC1. In one embodiment, said AAV is an adeno-associated virus 8. In one embodiment, said AAV is an adeno-associated virus 6. In one embodiment, said fusion protein further comprises a nuclear localization signal protein including, but not limited to, nucleoplasmin (NLS) and/or SV40 NLS and/or C-myc NLS. In one embodiment, the fusion protein further comprises a uracil glycosylase inhibitor. In one embodiment, said Nme2Cas9 protein further comprises a mutation. In one embodiment, said mutation is a D16A mutation. In one embodiment, said Nme2Cas9 (D16A) protein is a nickase. In one embodiment, said Nme2Cas9 protein further comprises a linker that flanks at least one inlaid protein domain. In one embodiment, the linker is approximately 73 amino acids in length. In one embodiment, the linker is approximately 20 amino acids in length. In one embodiment, said linker is a 3xHA-tag.

Definitions

To facilitate the understanding of this invention, a number of terms are defined below. Terms defined herein have meanings as commonly understood by a person of ordinary skill in the areas relevant to the present invention. Terms such as “a”, “an” and “the” are not intended to refer to only a singular entity, but include the general class of which a specific example may be used for illustration. The terminology herein is used to describe specific embodiments of the invention, but their usage does not delimit the invention, except as outlined in the claims.

As used herein, the term “edit” “editing” or “edited” refers to a method of altering a nucleic acid sequence of a polynucleotide (e.g., for example, a wild type naturally occurring nucleic acid sequence or a mutated naturally occurring sequence) by selective deletion of a specific genomic target. Such a specific genomic target includes, but is not limited to, a chromosomal region, a gene, a promoter, an open reading frame or any nucleic acid sequence.

As used herein, the term “single base” refers to one, and only one, nucleotide within a nucleic acid sequence. When used in the context of single base editing, it is meant that the base at a specific position within the nucleic acid sequence is replaced with a different base. This replacement may occur by many mechanisms, including but not limited to, substitution or modification.

As used herein, the term “target” or “target site” refers to a pre-identified nucleic acid sequence of any composition and/or length. Such target sites include, but is not limited to, a chromosomal region, a gene, a promoter, an open reading frame or any nucleic acid sequence. In some embodiments, the present invention interrogates these specific genomic target sequences with complementary sequences of gRNA.

The term “on-target binding sequence” as used herein, refers to a subsequence of a specific genomic target that may be completely complementary to a programmable DNA binding domain and/or a single guide RNA sequence.

The term “off-target binding sequence” as used herein, refers to a subsequence of a specific genomic target that may be partially complementary to a programmable DNA binding domain and/or a single guide RNA sequence.

The term “effective amount” as used herein, refers to a particular amount of a pharmaceutical composition comprising a therapeutic agent that achieves a clinically beneficial

result (i.e., for example, a reduction of symptoms). Toxicity and therapeutic efficacy of such compositions can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it can be expressed as the ratio LD₅₀/ED₅₀. Compounds that exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and additional animal studies can be used in formulating a range of dosage for human use. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The term “symptom”, as used herein, refers to any subjective or objective evidence of disease or physical disturbance observed by the patient. For example, subjective evidence is usually based upon patient self-reporting and may include, but is not limited to, pain, headache, visual disturbances, nausea and/or vomiting. Alternatively, objective evidence is usually a result of medical testing including, but not limited to, body temperature, complete blood count, lipid panels, thyroid panels, blood pressure, heart rate, electrocardiogram, tissue and/or body imaging scans.

The term “disease” or “medical condition”, as used herein, refers to any impairment of the normal state of the living animal or plant body or one of its parts that interrupts or modifies the performance of the vital functions. Typically manifested by distinguishing signs and symptoms, it is usually a response to: i) environmental factors (as malnutrition, industrial hazards, or climate); ii) specific infective agents (as worms, bacteria, or viruses); iii) inherent defects of the organism (as genetic anomalies); and/or iv) combinations of these factors.

The terms "reduce," "inhibit," "diminish," "suppress," "decrease," "prevent" and grammatical equivalents (including “lower,” “smaller,” etc.) when in reference to the expression of any symptom in an untreated subject relative to a treated subject, mean that the quantity and/or magnitude of the symptoms in the treated subject is lower than in the untreated subject by any amount that is recognized as clinically relevant by any medically trained personnel. In one embodiment, the quantity and/or magnitude of the symptoms in the treated subject is at least 10%

lower than, at least 25% lower than, at least 50% lower than, at least 75% lower than, and/or at least 90% lower than the quantity and/or magnitude of the symptoms in the untreated subject.

The term "attached" as used herein, refers to any interaction between a medium (or carrier) and a drug. Attachment may be reversible or irreversible. Such attachment includes, but is not limited to, covalent bonding, ionic bonding, Van der Waals forces or friction, and the like. A drug is attached to a medium (or carrier) if it is impregnated, incorporated, coated, in suspension with, in solution with, mixed with, *etc.*

The term "drug" or "compound" as used herein, refers to any pharmacologically active substance capable of being administered which achieves a desired effect. Drugs or compounds can be synthetic or naturally occurring, non-peptide, proteins or peptides, oligonucleotides or nucleotides, polysaccharides or sugars.

The term "administered" or "administering", as used herein, refers to any method of providing a composition to a patient such that the composition has its intended effect on the patient. An exemplary method of administering is by a direct mechanism such as, local tissue administration (*i.e.*, for example, extravascular placement), oral ingestion, transdermal patch, topical, inhalation, suppository *etc.*

The term "patient" or "subject", as used herein, is a human or animal and need not be hospitalized. For example, out-patients, persons in nursing homes are "patients." A patient may comprise any age of a human or non-human animal and therefore includes both adult and juveniles (*i.e.*, children). It is not intended that the term "patient" connote a need for medical treatment, therefore, a patient may voluntarily or involuntarily be part of experimentation whether clinical or in support of basic science studies.

The term "affinity" as used herein, refers to any attractive force between substances or particles that causes them to enter into and remain in chemical combination. For example, an inhibitor compound that has a high affinity for a receptor will provide greater efficacy in preventing the receptor from interacting with its natural ligands, than an inhibitor with a low affinity.

The term "pharmaceutically" or "pharmacologically acceptable", as used herein, refer to molecular entities and compositions that do not produce adverse, allergic, or other untoward reactions when administered to an animal or a human.

The term, "pharmaceutically acceptable carrier", as used herein, includes any and all solvents, or a dispersion medium including, but not limited to, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils, coatings, isotonic and absorption delaying agents, liposome, commercially available cleansers, and the like. Supplementary bioactive ingredients also can be incorporated into such carriers.

The term "viral vector" encompasses any nucleic acid construct derived from a virus genome capable of incorporating heterologous nucleic acid sequences for expression in a host organism. For example, such viral vectors may include, but are not limited to, adeno-associated viral vectors, lentiviral vectors, SV40 viral vectors, retroviral vectors, adenoviral vectors. Although viral vectors are occasionally created from pathogenic viruses, they may be modified in such a way as to minimize their overall health risk. This usually involves the deletion of a part of the viral genome involved with viral replication. Such a virus can efficiently infect cells but, once the infection has taken place, the virus may require a helper virus to provide the missing proteins for production of new virions. Preferably, viral vectors should have a minimal effect on the physiology of the cell it infects and exhibit genetically stable properties (e.g., do not undergo spontaneous genome rearrangement). Most viral vectors are engineered to infect as wide a range of cell types as possible. Even so, a viral receptor can be modified to target the virus to a specific kind of cell. Viruses modified in this manner are said to be pseudotyped. Viral vectors are often engineered to incorporate certain genes that help identify which cells took up the viral genes. These genes are called marker genes. For example, a common marker gene confers antibiotic resistance to a certain antibiotic.

As used herein, the term "genetic disease" refers to any medical condition having a primary causative factor of a mutated gene. The gene mutation may comprise a nucleic acid sequence wherein at least one, if not more, nucleotides are not wild type.

As used herein the term "*Dmd* gene" refers to a genetic locus that, when mutated, is believed to result in symptoms of muscular dystrophy.

As used herein the term "*Fah* gene" refers to a genetic locus that, when mutated, is believed to result in symptoms of tyrosinemia.

As used herein, the term "MeCP2 gene" refers to a genetic locus that, when mutated, is believed to result in symptoms of Rett's syndrome.

As used herein the “*ROSA26* gene” or “*Rosa26* gene” refers to a human or mouse (respectively) locus that is widely used for achieving generalized expression in the mouse. Targeting to the *ROSA26* locus may be achieved by introducing a desired gene into the first intron of the locus, at a unique XbaI site approximately 248 bp upstream of the original gene trap line. A construct may be constructed using an adenovirus splice acceptor followed by a gene of interest and a polyadenylation site inserted at the unique XbaI site. A neomycin resistance cassette may also be included in the targeting vector.

As used herein the “*PCSK9* gene” or “*Pcsk9* gene” refers to a human or mouse (respectively) locus that encodes a PCSK9 protein. The *PCSK9* gene resides on chromosome 1 at the band 1p32.3 and includes 13 exons. This gene may produce at least two isoforms through alternative splicing.

The term “proprotein convertase subtilisin/kexin type 9” and “PCSK9” refers to a protein encoded by a gene that modulates low density lipoprotein levels. Proprotein convertase subtilisin/kexin type 9, also known as PCSK9, is an enzyme that in humans is encoded by the *PCSK9* gene. Seidah et al., "The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation" *Proc. Natl. Acad. Sci. U.S.A.* 100 (3): 928–933 (2003). Similar genes (orthologs) are found across many species. Many enzymes, including PSCK9, are inactive when they are first synthesized, because they have a section of peptide chains that blocks their activity; proprotein convertases remove that section to activate the enzyme. PSCK9 is believed to play a regulatory role in cholesterol homeostasis. For example, PCSK9 can bind to the epidermal growth factor-like repeat A (EGF-A) domain of the low-density lipoprotein receptor (LDL-R) resulting in LDL-R internalization and degradation. Clearly, it would be expected that reduced LDL-R levels result in decreased metabolism of LDL-C, which could lead to hypercholesterolemia.

The term “hypercholesterolemia” as used herein, refers to any medical condition wherein blood cholesterol levels are elevated above the clinically recommended levels. For example, if cholesterol is measured using low density lipoproteins (LDLs), hypercholesterolemia may exist if the measured LDL levels are above, for example, approximately 70 mg/dl. Alternatively, if cholesterol is measured using free plasma cholesterol, hypercholesterolemia may exist if the measured free cholesterol levels are above, for example, approximately 200-220 mg/dl.

As used herein, the term “CRISPRs” or “Clustered Regularly Interspaced Short Palindromic Repeats” refers to an acronym for DNA loci that contain multiple, short, direct repetitions of base sequences. Each repetition contains a series of bases followed by 30 or so base pairs known as “spacer DNA”. The spacers are short segments of DNA from a virus and may serve as a ‘memory’ of past exposures to facilitate an adaptive defense against future invasions.

As used herein, the term “Cas” or “CRISPR-associated (cas)” refers to genes often associated with CRISPR repeat-spacer arrays.

As used herein, the term “Cas9” refers to a nuclease from Type II CRISPR systems, an enzyme specialized for generating double-strand breaks in DNA, with two active cutting sites (the HNH and RuvC domains), one for each strand of the double helix. Jinek combined tracrRNA and spacer RNA into a “single-guide RNA” (sgRNA) molecule that, mixed with Cas9, could find and cleave DNA targets through Watson-Crick pairing between the guide sequence within the sgRNA and the target DNA sequence.

As used herein, the term “N-terminal domain” refers to the fusion of a first peptide or protein at the N-terminal end of a second peptide or protein. For example, a nucleotide deaminase protein may be “N-terminally” fused to the last amino acid of a Cas9 nuclease protein.

As used herein, the term “inlaid domain” refers to the fusion of a first peptide or protein between the C-terminal and N-terminal ends of a second peptide or protein. For example, a nucleotide deaminase protein is an “inlaid domain” when inserted between the C-terminal and N-terminal ends of a Cas9 nuclease protein.

The term “protospacer adjacent motif” (or PAM) as used herein, refers to a DNA sequence that may be required for a Cas9/sgRNA to form an R-loop to interrogate a specific DNA sequence through Watson-Crick pairing of its guide RNA with the genome. The PAM specificity may be a function of the DNA-binding specificity of the Cas9 protein (e.g., a “protospacer adjacent motif recognition domain” at the C-terminus of Cas9).

As used herein, the term “sgRNA” refers to single guide RNA used in conjunction with CRISPR associated systems (Cas). sgRNAs are a fusion of crRNA and tracrRNA and contain nucleotides of sequence complementary to the desired target site. Jinek et al., “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity” *Science* 337(6096):816-821 (2012) Watson-Crick pairing of the sgRNA with the target site permits R-loop formation, which

in conjunction with a functional PAM permits DNA cleavage or in the case of nuclease-deficient Cas9 allows binds to the DNA at that locus.

As used herein, the term “fluorescent protein” refers to a protein domain that comprises at least one organic compound moiety that emits fluorescent light in response to the appropriate wavelengths. For example, fluorescent proteins may emit red, blue and/or green light. Such proteins are readily commercially available including, but not limited to: i) mCherry (Clontech Laboratories): excitation: 556/20 nm (wavelength/bandwidth); emission: 630/91 nm; ii) sfGFP (Invitrogen): excitation: 470/28 nm; emission: 512/23 nm; iii) TagBFP (Evrogen): excitation 387/11 nm; emission 464/23 nm.

As used herein, the term “sgRNA” refers to single guide RNA used in conjunction with CRISPR associated systems (Cas). sgRNAs contains nucleotides of sequence complementary to the desired target site. Watson-crick pairing of the sgRNA with the target site recruits the nuclease-deficient Cas9 to bind the DNA at that locus.

As used herein, the term “orthogonal” refers targets that are non-overlapping, uncorrelated, or independent. For example, if two orthogonal nuclease-deficient Cas9 gene fused to different effector domains were implemented, the sgRNAs coded for each would not cross-talk or overlap. Not all nuclease-deficient Cas9 genes operate the same, which enables the use of orthogonal nuclease-deficient Cas9 gene fused to a different effector domains provided the appropriate orthogonal sgRNAs.

As used herein, the term “phenotypic change” or “phenotype” refers to the composite of an organism's observable characteristics or traits, such as its morphology, development, biochemical or physiological properties, phenology, behavior, and products of behavior. Phenotypes result from the expression of an organism's genes as well as the influence of environmental factors and the interactions between the two.

"Nucleic acid sequence" and "nucleotide sequence" as used herein refer to an oligonucleotide or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand.

The term "an isolated nucleic acid", as used herein, refers to any nucleic acid molecule that has been removed from its natural state (e.g., removed from a cell and is, in a preferred embodiment, free of other genomic nucleic acid).

The terms "amino acid sequence" and "polypeptide sequence" as used herein, are interchangeable and to refer to a sequence of amino acids.

As used herein the term "portion" when in reference to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid.

The term "portion" when used in reference to a nucleotide sequence refers to fragments of that nucleotide sequence. The fragments may range in size from 5 nucleotide residues to the entire nucleotide sequence minus one nucleic acid residue.

As used herein, the terms "complementary" or "complementarity" are used in reference to "polynucleotides" and "oligonucleotides" (which are interchangeable terms that refer to a sequence of nucleotides) related by the base-pairing rules. For example, the sequence "C-A-G-T," is complementary to the sequence "G-T-C-A." Complementarity can be "partial" or "total." "Partial" complementarity is where one or more nucleic acid bases is not matched according to the base pairing rules. "Total" or "complete" complementarity between nucleic acids is where each and every nucleic acid base is matched with another base under the base pairing rules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods which depend upon binding between nucleic acids.

The terms "homology" and "homologous" as used herein in reference to nucleotide sequences refer to a degree of complementarity with other nucleotide sequences. There may be partial homology or complete homology (i.e., identity). A nucleotide sequence which is partially complementary, i.e., "substantially homologous," to a nucleic acid sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid sequence. The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence to a target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target

sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

The terms "homology" and "homologous" as used herein in reference to amino acid sequences refer to the degree of identity of the primary structure between two amino acid sequences. Such a degree of identity may be directed to a portion of each amino acid sequence, or to the entire length of the amino acid sequence. Two or more amino acid sequences that are "substantially homologous" may have at least 50% identity, preferably at least 75% identity, more preferably at least 85% identity, most preferably at least 95%, or 100% identity.

An oligonucleotide sequence which is a "homolog" is defined herein as an oligonucleotide sequence which exhibits greater than or equal to 50% identity to a sequence, when sequences having a length of 100 bp or larger are compared.

Low stringency conditions comprise conditions equivalent to binding or hybridization at 42°C in a solution consisting of 5 x SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄·H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.1% SDS, 5x Denhardt's reagent {50x Denhardt's contains per 500 ml: 5 g Ficoll (Type 400, Pharmacia), 5 g BSA (Fraction V; Sigma)} and 100 µg/ml denatured salmon sperm DNA followed by washing in a solution comprising 5x SSPE, 0.1% SDS at 42°C when a probe of about 500 nucleotides in length is employed. Numerous equivalent conditions may also be employed to comprise low stringency conditions; factors such as the length and nature (DNA, RNA, base composition) of the probe and nature of the target (DNA, RNA, base composition, present in solution or immobilized, etc.) and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate, polyethylene glycol), as well as components of the hybridization solution may be varied to generate conditions of low stringency hybridization different from, but equivalent to, the above listed conditions. In addition, conditions which promote hybridization under conditions of high stringency (e.g., increasing the temperature of the hybridization and/or wash steps, the use of formamide in the hybridization solution, etc.) may also be used.

As used herein, the term "hybridization" is used in reference to the pairing of complementary nucleic acids using any process by which a strand of nucleic acid joins with a complementary strand through base pairing to form a hybridization complex. Hybridization and the strength of hybridization (i.e., the strength of the association between the nucleic acids) is

impacted by such factors as the degree of complementarity between the nucleic acids, stringency of the conditions involved, the T_m of the formed hybrid, and the G:C ratio within the nucleic acids.

As used herein the term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0 t or R_0 t analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized to a solid support (e.g., a nylon membrane or a nitrocellulose filter as employed in Southern and Northern blotting, dot blotting or a glass slide as employed in in situ hybridization, including FISH (fluorescent in situ hybridization)).

DNA molecules are said to have "5' ends" and "3' ends" because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotide is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring. An end of an oligonucleotide is referred to as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of another mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5' of the "downstream" or 3' elements. This terminology reflects the fact that transcription proceeds in a 5' to 3' fashion along the DNA strand. The promoter and enhancer elements which direct transcription of a linked gene are generally located 5' or upstream of the coding region. However, enhancer elements can exert their effect even when located 3' of the promoter element and the coding region. Transcription termination and polyadenylation signals are located 3' or downstream of the coding region.

The term "transfection" or "transfected" refers to the introduction of foreign DNA into a cell.

As used herein, the terms "nucleic acid molecule encoding", "DNA sequence encoding," and "DNA encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these deoxyribonucleotides determines the order of amino

acids along the polypeptide (protein) chain. The DNA sequence thus codes for the amino acid sequence.

As used herein, the term "gene" means the deoxyribonucleotide sequences comprising the coding region of a structural gene and including sequences located adjacent to the coding region on both the 5' and 3' ends for a distance of about 1 kb on either end such that the gene corresponds to the length of the full-length mRNA. The sequences which are located 5' of the coding region and which are present on the mRNA are referred to as 5' non-translated sequences. The sequences which are located 3' or downstream of the coding region and which are present on the mRNA are referred to as 3' non-translated sequences. The term "gene" encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region interrupted with non-coding sequences termed "introns" or "intervening regions" or "intervening sequences." Introns are segments of a gene which are transcribed into heterogeneous nuclear RNA (hnRNA); introns may contain regulatory elements such as enhancers. Introns are removed or "spliced out" from the nuclear or primary transcript; introns therefore are absent in the messenger RNA (mRNA) transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent polypeptide.

In addition to containing introns, genomic forms of a gene may also include sequences located on both the 5' and 3' end of the sequences which are present on the RNA transcript. These sequences are referred to as "flanking" sequences or regions (these flanking sequences are located 5' or 3' to the non-translated sequences present on the mRNA transcript). The 5' flanking region may contain regulatory sequences such as promoters and enhancers which control or influence the transcription of the gene. The 3' flanking region may contain sequences which direct the termination of transcription, posttranscriptional cleavage and polyadenylation.

The term "label" or "detectable label" are used herein, to refer to any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Such labels include biotin for staining with labeled streptavidin conjugate, magnetic beads (e.g., Dynabeads[®]), fluorescent dyes (e.g., fluorescein, texas red, rhodamine, green fluorescent protein, and the like), radiolabels (e.g., ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and calorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include, but are not limited to, U.S. Pat. Nos.

3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241 (all herein incorporated by reference). The labels contemplated in the present invention may be detected by many methods. For example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light.

5 Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting, the reaction product produced by the action of the enzyme on the substrate, and calorimetric labels are detected by simply visualizing the colored label.

Brief Description Of The Drawings

10 The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

Fig. 1 presents an illustrative approach by which to design inlaid domains within a Cas9 protein.

15 Fig. 1A: Different views of Nme1Cas9 are presented as a ternary complex (PDB:6JDV). Black spheres = N/C-termini. Colored spheres = Domain insertion sites. Legend: Inlaid domain locations within the Nme1Cas9 protein.

Fig. 1B: Exemplary gene editing activities of Nme2-ABE8e constructs in HEK293T cells using an ABE mCherry reporter system activated upon A-to-G editing by plasmid transfection and measured by flow cytometry. Protospacer with target adenine (red), and PAM (bold, underlined). n=3 biological replicates, data represent mean $\pm$ SD.

20

Fig. 1C: Exemplary gene editing efficiency in HEK293T cells at eight (8) dual PAM genomic target sites. A-to-G editing at PAM-matched endogenous HEK293T genomic loci between four (4) Nme2-ABE8e constructs, each having a different inlaid domain position, was compared to an SpyCas9-ABE8e fusion protein. Maximally edited adenine for each target was plotted. Gene editing activities were measured by amplicon sequencing. n=3 biological replicates, data represent mean $\pm$ SD.

25

Fig. 1D: Exemplary data showing mean gene editing window and associated efficiency in HEK293T cells using eight (8) dual PAM genomic target sites. The

30

data is presented as a summary of mean A-to-G editing activities and editing windows for Spy- and Nme2-ABE8e constructs. Each position in the protospacer represents the mean A-to-G editing efficiency across eight (8) PAM-matched endogenous target sites measured via amplicon sequencing. Crossed out boxes represent no adenine in the targets tested. n=3 biological replicates.

Fig. 1E: Exemplary data showing a measured maximum gene editing efficiency in HEK293T cells using eight (8) dual PAM genomic target sites. Each data point represents the maximum A-to-G editing rate of an individual dual PAM target site in accordance with the Spy and Nme1 constructs presented in Fig. 1C as measured by amplicon sequencing. n=3 biological replicates. Data represent mean $\pm$ SEM.

Fig. 2 presents exemplary embodiments of specific inlaid domain Cas9 constructs

Fig. 2A: Schematic representation of eight (8) representative Nme2^{D16A}Cas9-ABE constructs and their respective inlaid domain insertion sites.

Fig. 2B: Exemplary data showing a mean gen editing window and associated efficiency in HEK293T cells using fifteen (15) genomic target sites. The data is presented as a summary of mean A-to-G editing activities and editing windows for the eight (8) Nme2-ABE8e constructs depicted in Fig. 2A. Each position in the protospacer represents the mean A-to-G editing efficiency across fifteen (15) endogenous target sites in HEK293T cells as measured via amplicon sequencing. n=3 biological replicates.

Fig. 2C: Exemplary data showing a scatter plot of the maximum A-to-G editing rate observed for an individual target site as measured by amplicon sequencing. The line and error bars represents the mean and SEM of efficiency observed across all target sites. Fold improvement relative to the n-terminal fusion shown below. n=3 biological replicates. Two-way ANOVA analysis: ns, $p > 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$

Fig. 3 illustrates embodiments of an NmeCas9 deaminase fusion protein single base editor.

Fig. 3A: An YE1-BE3-nNme2Cas9 (D16A)-UGI construct.

Fig. 3B: An ABE7.10 nNme2Cas9 (D16A) construct.

Fig. 4 presents exemplary data of the electroporation of HEK293T cells with DNA plasmids comprising a YE1-BE3-nNme2Cas9 (D16A)-UGI fusion protein.

Fig. 4A: The TS25 endogenous target site. GN23 sgRNA base-pairs with the target DNA strand, leaving the displaced DNA strand for cytidine deaminase to edit.

Fig. 4B: Sequencing data showing a doublet nucleotide peak (7th position from 5' end; arrow) demonstrating the successful single base editing of a cytidine to a thymidine (e.g., a C•G base pair conversion to a T•A base pair).

Fig. 4C: A quantitation of the data shown in Fig. 2B plotting the percent conversion of C → T single base editing. The percentage of C converted to T is about 40% in the base editor- and sgRNA-treated sample (p-value = 6.88×10^{-6}). The “no sgRNA” control displays the background noise due to Sanger sequencing. EditR (Kluesner et al., 2018) was used to perform the analysis.

Fig. 5 presents specific UGI target sites that were respectively integrated into YE1-BE3-nNme2Cas9/D16A mutant fusion proteins and co-expressed with enhanced green fluorescent protein (EGFP) in a stable K562-derived cell line. Converted bases are highlighted in orange color. Background signals were filtered using negative control samples (YE1-BE3-nNme2Cas9 nucleofected K562 cells without sgRNA constructs). N₄CC PAMs are boxed. The percentage of total reads exhibiting mutations in base-editor-targeted sites is shown in the right column.

Fig. 5A: EGFP-Site 1

Fig. 5B: EGFP-Site 2

Fig. 5C: EGFP-Site 3

Fig. 5D: EGFP-Site 4

Fig. 6 presents an exemplary alignment of the wildtype *Fah* gene with the tyrosinemia *Fah* mutant gene showing an A-G single base gene editing target site (position 9). The respective SpyCas9 single PAM site and NmeCas9 double PAM sites are indicated for demonstrating the suboptimal targeting window relative to the SpyCas9 PAM site.

Fig. 7 illustrates that the closely related *Neisseria meningitidis* 1, 2 and 3 Cas9 orthologs that have distinct PAMs.

Fig. 7A shows an exemplary schematic showing mutated residues (orange spheres) between Nme2Cas9 (left) and Nme3Cas9 (right) mapped onto the predicted structure of Nme1Cas9, revealing the cluster of mutations in the PID (black).

Fig. 7B shows an exemplary experimental workflow of the in vitro PAM discovery assay with a 10-bp randomized PAM region. Following in vitro digestion, adapters were ligated to cleaved products for library construction and sequencing.

Fig. 7C shows exemplary sequence logos resulting from in vitro PAM discovery reveal the enrichment of a N₄GATT PAM for Nme1Cas9, consistent with its previously established specificity.

Fig. 7D shows exemplary sequence logos indicating that Nme1Cas9 with its PID swapped with that of Nme2Cas9 (left) or Nme3Cas9 (right) requires a C at PAM position 5. The remaining nucleotides were not determined with high confidence due to the modest cleavage efficiency of the PID-swapped protein chimeras (see Fig. 6C).

Fig. 7E shows an exemplary sequence logo showing that full-length Nme2Cas9 recognizes an N₄CC PAM, based on efficient substrate cleavage of a target pool with a fixed C at PAM position 5, and with PAM nts 1-4 and 6-8 randomized.

Fig. 8 presents a characterization of *Neisseria meningitidis* Cas9 orthologs with rapidly-evolving PIDs, as related to Fig. 7.

Fig. 8A shows an exemplary unrooted phylogenetic tree of NmeCas9 orthologs that are >80% identical to Nme1Cas9. Three distinct branches emerged, with the majority of mutations clustered in the PID. Groups 1 (blue), 2 (orange), and 3 (green) have PIDs with >98%, ~52%, and ~86% identity to Nme1Cas9, respectively. Three representative Cas9 orthologs (one from each group) (Nme1Cas9, Nme2Cas9 and Nme3Cas9) are indicated.

Fig. 8B shows an exemplary schematic showing the CRISPR-*cas* loci of the strains encoding the three Cas9 orthologs (Nme1Cas9, Nme2Cas9, and Nme3Cas9) from (A). Percent identities of each CRISPR-Cas component with *N. meningitidis* 8013 (encoding Nme1Cas9) are shown. Blue and red arrows denote pre-crRNA and tracrRNA transcription initiation sites, respectively.

Fig. 8C shows an exemplary normalized read counts (% of total reads) from cleaved DNAs from the in vitro assays for intact Nme1Cas9 (grey), for chimeras with Nme1Cas9's PID swapped with those of Nme2Cas9 and Nme3Cas9 (mixed colors),

and for full-length Nme2Cas9 (orange), are plotted. The reduced normalized read counts indicate lower cleavage efficiencies in the chimeras.

Fig. 8D shows an exemplary sequence logos from the in vitro PAM discovery assay on an NNNNCNNN PAM pool by Nme1Cas9 with its PID swapped with those of Nme2Cas9 (left) or Nme3Cas9 (right).

Fig. 9 presents exemplary data showing that Nme2Cas9 uses a 22-24 nt spacer to edit sites adjacent to an N₄CC PAM. All experiments were done in triplicate, and error bars represent the standard error of the mean (s.e.m.).

Fig. 9A shows an exemplary schematic diagram depicting transient transfection and editing of HEK293T TLR2.0 cells, with mCherry⁺ cells detected by flow cytometry 72 hours after transfection.

Fig. 9B shows an exemplary Nme2Cas9 editing of the TLR2.0 reporter. Sites with N₄CC PAMs were targeted with varying efficiencies, while no Nme2Cas9 targeting was observed at an N₄GATT PAM or in the absence of sgRNA. SpyCas9 (targeting a previously validated site with an NGG PAM) and Nme1Cas9 (targeting N₄GATT) were used as positive controls.

Fig. 9C shows an exemplary effect of spacer length on the efficiency of Nme2Cas9 editing. An sgRNA targeting a single TLR2.0 site, with spacer lengths varying from 24 to 20 nts (including the 5'-terminal G required by the U6 promoter), indicate that highest editing efficiencies are obtained with 22-24 nt spacers.

Fig. 9D shows an exemplary Nme2Cas9 dual nickase can be used in tandem to generate NHEJ- and HDR-based edits in TLR2.0. Nme2Cas9- and sgRNA-expressing plasmids, along with an 800-bp dsDNA donor for homologous repair, were electroporated into HEK293T TLR2.0 cells, and both NHEJ (mCherry⁺) and HDR (GFP⁺) outcomes were scored by flow cytometry. HNH nickase, Nme2Cas9^{D16A}, RuvC nickase, Nme2Cas9^{H588A}. Cleavage sites 32 bp and 64 bp apart were targeted using either nickase. The HNH nickase (Nme2Cas9^{D16A}) yielded efficient editing, particularly with the cleavage sites that were separated by 32 bp, whereas the RuvC nickase (Nme2Cas9^{H588A}) was not effective. Wildtype Nme2Cas9 was used as a control.

Fig. 10 presents exemplary data showing PAM, spacer, and seed requirements for Nme2Cas9 targeting in mammalian cells, as related to Fig. 7. All experiments were done in triplicate and error bars represent s.e.m.

Fig. 10A shows an exemplary Nme2Cas9 targeting at N₄CD sites in TLR2.0, with editing estimated based on mCherry+ cells. Four sites for each non-C nucleotide at the tested position (N₄CA, N₄CT and N₄CG) were examined, and an N₄CC site was used as a positive control.

Fig. 10B shows an exemplary Nme2Cas9 targeting at N₄DC sites in TLR2.0 [similar to (A)].

Fig. 10C shows exemplary guide truncations on a TLR2.0 site with a N₄CCA PAM, revealing similar length requirements as those observed at the other site.

Fig. 10D shows exemplary Nme2Cas9 targeting efficiency is differentially sensitive to single-nucleotide mismatches in the seed region of the sgRNA. Data show the effects of walking single-nucleotide sgRNA mismatches along the 23-nt spacer in a TLR2.0 target site.

Fig. 11 presents exemplary data showing Nme2Cas9 genome editing at endogenous loci in mammalian cells via multiple delivery methods. All results represent 3 independent biological replicates, and error bars represent s.e.m.

Fig. 11A shows an exemplary Nme2Cas9 genome editing of endogenous human sites in HEK293T cells following transient transfection of Nme2Cas9- and sgRNA-expressing plasmids. 40 sites were screened initially (Table 1); the 14 sites shown (selected to include representatives of varying editing efficiencies, as measured by TIDE) were then re-analyzed in triplicate. An Nme1Cas9 target site (with an N₄GATT PAM) was used as a negative control.

Fig. 11B shows exemplary data charts: Left panel: Transient transfection of a single plasmid expressing both Nme2Cas9 and sgRNA (targeting the *Pcsk9* and *Rosa26* loci) enables editing in Hepa1-6 mouse cells, as detected by TIDE. Right panel: Electroporation of sgRNA plasmids into K562 cells stably expressing Nme2Cas9 from a lentivector results in efficient indel formation.

Fig. 9C shows exemplary Nme2Cas9 can be electroporated as an RNP complex to induce genome editing. 40 picomoles Cas9 along with 50 picomoles of in vitro-

transcribed sgRNAs targeting three different loci were electroporated into HEK293T cells. Indels were measured after 72h using TIDE.

Fig. 12 presents exemplary data showing dose dependence and segmental deletions by Nme2Cas9, as related to Fig. 9.

5 Fig. 12A shows exemplary increasing the dose of electroporated Nme2Cas9 plasmid (500 ng, vs. 200 ng in Fig. 3A) improves editing efficiency at two sites (TS16 and TS6). Data provided in yellow are re-used from Fig. 11A.

Fig. 12B shows exemplary Nme2Cas9 can be used to create precise segmental deletions. Two TLR2.0 targets with cleavage sites 32 bp apart were targeted

10 simultaneously with Nme2Cas9. The majority of lesions created were deletions of exactly 32 bp (blue).

Fig. 13 presents exemplary data showing that Nme2Cas9 is subject to inhibition by a subset of type II-C anti-CRISPR families in vitro and in cells. All experiments were done in triplicate and error bars represent s.e.m.

15 Fig. 13A: In vitro cleavage assay of Nme1Cas9 and Nme2Cas9 in the presence of five previously characterized anti-CRISPR proteins (10:1 ratio of Acr:Cas9). Top: Nme1Cas9 efficiently cleaves a fragment containing a protospacer with an N₄GATT PAM in the absence of an Acr or in the presence of a negative control Acr (AcrE2). All five previously characterized type II-C Acr families inhibited

20 Nme1Cas9, as expected. Bottom: Nme2Cas9 inhibition mirrors that of Nme1Cas9, except for the lack of inhibition by AcrIIC5_{Smu}.

Fig. 13B: Genome editing in the presence of the five previously described anti-CRISPR families. Plasmids expressing Nme2Cas9 (200 ng), sgRNA (100 ng) and each respective Acr (200 ng) were co-transfected into HEK293T cells, and genome

25 editing was measured using Tracking of Indels by Decomposition (TIDE) 72 hr post transfection. Consistent with our in vitro analyses, all type II-C anti-CRISPRs except AcrIIC5_{Smu} inhibited genome editing, albeit with different efficiencies.

Fig. 13C: Acr inhibition of Nme2Cas9 is dose-dependent with distinct apparent potencies. Nme2Cas9 is fully inhibited by AcrIIC1_{Nme} and AcrIIC4_{Hpa} at 2:1 and

30 1:1 mass ratios of cotransfected Acr and Nme2Cas9 plasmids, respectively.

Fig. 14 presents exemplary data showing that a Nme2Cas9 PID swap renders Nme1Cas9 insensitive to AcrIIC5_{Smu} inhibition, as related to Fig. 11. In vitro cleavage by the Nme1Cas9-Nme2Cas9PID chimera in the presence of previously characterized Acr proteins (10 uM Cas9-sgRNA + 100 uM Acr).

Fig. 15 presents exemplary data showing orthogonality and relative accuracy of Nme2Cas9 and SpyCas9 at dual target sites, as related to Fig. 12.

Fig. 15A shows exemplary Nme2Cas9 and SpyCas9 guides are orthogonal. TIDE results show the frequencies of indels created by both nucleases targeting DS2 with either their cognate sgRNAs or with the sgRNAs of the other ortholog.

Fig. 15B shows exemplary Nme2Cas9 and SpyCas9 exhibiting comparable on-target editing efficiencies as assessed by GUIDE-seq. Bars indicate on-target read counts from GUIDE-Seq at the three dual sites targeted by each ortholog. Orange bars represent Nme2Cas9 and black bars represent SpyCas9.

Fig. 15C shows an exemplary SpyCas9's on-target vs. off-target read counts for each site. Orange bars represent the on-target reads while black bars represent off-targets.

Fig. 15D shows exemplary Nme2Cas9's on-target vs. off-target reads for each site.

Fig. 15E bar graphs showing exemplary indel efficiencies (measured by TIDE) at potential off-target sites predicted by CRISPRSeek. On- and off-target site sequences are shown on the left, with the PAM region underlined and sgRNA mismatches and non-consensus PAM nucleotides given in red.

Fig. 16 presents exemplary data showing that Nme2Cas9 exhibits little or no detectable off-targeting in mammalian cells.

Fig. 16A shows an exemplary schematic depicting dual sites (DSs) targetable by both SpyCas9 and Nme2Cas9 by virtue of their non-overlapping PAMs. The Nme2Cas9 PAM (orange) and SpyCas9 PAM (blue) are highlighted. A 24nt Nme2Cas9 guide sequence is indicated in yellow; the corresponding guide sequence for SpyCas9 would be 4nt shorter at the 5' end.

Fig. 16B shows an exemplary Nme2Cas9 and SpyCas9 that both induce indels at DSs. Six DSs in *VEGFA* (with GN₃GN₁₉NGGNCC sequences) were selected for direct comparisons of editing by the two orthologs. Plasmids expressing each

Cas9 (with the same promoter, linkers, tags and NLSs) and its cognate guide were transfected into HEK293T cells. Indel efficiencies were determined by TIDE 72 hrs post transfection. Nme2Cas9 editing was detectable at all six sites and was marginally or significantly more efficient than SpyCas9 at two sites (DS2 and DS6, respectively). SpyCas9 edited four out of the six sites (DS1, DS2, DS4 and DS6), with two sites showing significantly higher editing efficiencies than Nme2Cas9 (DS1 and DS4). DS2, DS4 and DS6 were selected for GUIDE-Seq analysis as Nme2Cas9 was equally efficient, less efficient and more efficient than SpyCas9, respectively, at these sites.

Fig. 16C shows exemplary Nme2Cas9 genome editing that is highly accurate in human cells. Numbers of off-target sites detected by GUIDE-Seq for each nuclease at individual target sites are shown. In addition to dual sites, we analyzed TS6 (because of its high on-target editing efficiency) and *Pcsk9* and *Rosa26* sites in mouse Hepa1-6 cells (to measure accuracy in another cell type).

Fig. 16D shows an exemplary targeted deep sequencing to detect indels in edited cells confirms the high Nme2Cas9 accuracy indicated by GUIDE-seq.

Fig. 16E shows an exemplary sequence for the validated off-target site of the *Rosa26* guide, showing the PAM region (underlined), the consensus CC PAM dinucleotide (bold), and three mismatches in the PAM-distal portion of the spacer (red).

Fig. 17 presents exemplary data showing Nme2Cas9 genome editing in vivo via all-in-one AAV delivery.

Fig. 17A shows exemplary workflow for delivery of AAV8.sgRNA.Nme2Cas9 to lower cholesterol levels in mice by targeting *Pcsk9*. Top: schematic of the all-in-one AAV vector expressing Nme2Cas9 and the sgRNA (individual genome elements not to scale). BGH, bovine growth hormone poly(A) site; HA, epitope tag; NLS, nuclear localization sequence; h, human-codon-optimized. Bottom: Timeline for AAV8.sgRNA.Nme2Cas9 tail-vein injections (4×10^{11} GCs), followed by cholesterol measurements at day 14 and indel, histology and cholesterol analyses at day 28 post-injection.

Fig. 17B shows an exemplary TIDE analysis to measure indels in DNA extracted from livers of mice injected with AAV8.Nme2Cas9+sgRNA targeting *Pcsk9* and *Rosa26* (control) loci. Indel efficiency at the lone off-target site identified by GUIDE-seq for these two sgRNAs (*Rosa26*|OT1) were also assessed by TIDE.

Fig. 17C shows an exemplary reduced serum cholesterol levels in mice injected with the *Pcsk9*-targeting guide compared to the *Rosa26*-targeting controls. *P* values are calculated by unpaired two-tailed t-test.

Fig. 18 presents exemplary data showing PCSK9 knockdown and liver histology following Nme2Cas9 AAV delivery and editing, related to Fig. 15.

Fig. 18A shows exemplary Western blotting using anti-PCSK9 antibody reveals strongly reduced levels of PCSK9 in the livers of mice treated with sg*Pcsk9*, compared to mice treated with sg*Rosa26*. 2ng of recombinant PCSK9 was used as a mobility standard (left-most lane), and a cross-reacting band in the liver samples is indicated by an asterisk. GAPDH was used as loading control (bottom panel).

Fig. 18B shows exemplary H&E staining from livers of mice injected with AAV8.Nme2Cas9+sg*Rosa26* (left) or AAV8.Nme2Cas9+sg*Pcsk9* (right) vectors. Scale bars, 25 μ m.

Fig. 19 presents exemplary data showing *Tyr* editing ex vivo in mouse zygotes, related to Fig. 18.

Fig. 19A shows an exemplary two sites in *Tyr*, each with N₄CC PAMs, were tested for editing in Hepa1-6 cells. The sg*Tyr2* guide exhibited higher editing efficiency and was selected for further testing.

Fig. 19B shows an exemplary seven mice that survived post-natal development, and each exhibited coat color phenotypes as well as on-target editing, as assayed by TIDE.

Fig. 19C shows an exemplary Indel spectra from tail DNA of each mouse from (B), as well as an unedited C57BL/6NJ mouse, as indicated by TIDE analysis. Efficiencies of insertions (positive) and deletions (negative) of various sizes are indicated.

Fig. 20 presents exemplary data showing Nme2Cas9 genome editing ex vivo via all-in-one AAV delivery.

Fig. 20A shows an exemplary workflow for single-AAV Nme2Cas9 editing ex vivo to generate albino C57BL/6NJ mice by targeting the *Tyr* gene. Zygotes are cultured in KSOM containing AAV6.Nme2Cas9:sg*Tyr* for 5-6 hours, rinsed in M2, and cultured for a day before being transferred to the oviduct of pseudo-pregnant recipients.

Fig. 20B shows exemplary albino (left) and chinchilla or variegated (middle) mice generated by 3×10^9 GCs, and chinchilla or variegated mice (right) generated by 3×10^8 GCs of zygotes with AAV6.Nme2Cas9:sg*Tyr*.

Fig. 20C shows an exemplary summary of Nme2Cas9:sg*Tyr* single-AAV ex vivo *Tyr* editing experiments at two AAV doses.

Fig. 21 presents exemplary data showing gene editing differences between fusion proteins of NmeCas9 and SpyCas9 nuclease with an N-terminally fused adenine deaminase domain.

Fig. 21A: Schematic representation of the ABE reporter cell line.

Fig. 21B: Schematic representation of the Nme2Cas9-ABE constructs.

Fig. 21C: Comparison of Nme2Cas9-ABE editing efficiency to those of SpyCas9-ABEs in the ABE reporter cell line ($n = 3$ biological replicates).

Fig. 21D: Summary of editing windows and comparison of editing efficiencies for the ABEs at endogenous genomic loci. Each data point represents the mean A-to-G editing efficiency at the indicated position in the spacer measured by amplicon deep sequencing across 12 Nme2Cas9 target sites and 8 target SpyCas9 target sites, respectively. ($n = 3$ biological replicates).

Fig. 21E: Comparison of Nme2Cas9-ABE8e mismatch tolerance to that of SpyCas9-ABE8e in the ABE reporter cell line. The activities of the effectors with the mismatched guides are normalized to that of the perfectly complementary (WT) guide. Red, mismatched nucleotides; green, PAM sequence ($n = 3$ biological replicates).

Fig. 22 presents exemplary data showing a summary of the individual A-to-G conversion efficiency at twelve target sites for Nme2Cas9-ABE8e, which include eight dual-target sites (DS 2 – 28) and four Nme2Cas9-specific target sites (Nm2 1-4), and eight dual-target sites for SpCas9-ABE7.10 and SpyCas9 ABE8e. Each data point represents the A-to-G conversion efficiency at the indicated nucleotide position measured by amplicon deep sequencing ($n = 3$ biological replicates).

Fig. 23 presents exemplary data showing single base mutation reversion and exon skipping strategy by a fusion protein comprising an Nme2Cas9-ABE8e construct in *MeCP2* and *Dmd* genes.

Fig. 23A: Schematic representation of a nonsense mutation in the human *MeCP2* gene (c. 502 C>T; p.R168X) that causes Rett Syndrome. The black underline denotes the target sequence of an Nme2Cas9-ABE8e guide for reverting the mutant A to G (wildtype) position 10 (red, bold). The PAM region is underlined in green. A bystander edit at position 16 (orange) can generate a missense mutation (c. 496 T>C; p.S166P).

Fig. 23B: Amplicon deep sequencing quantifying the editing efficiency in Rett patient-derived fibroblasts transfected with the Nme2Cas9-ABE8e mRNA and the synthetic sgRNA_*MeCP2* noted in a (n = 3 biological replicates).

Fig. 23C: Schematic representation of the exon skipping strategy that restores the reading frame of the mouse *Dmd* transcript. Deletion of exon 51 (Δ Ex51) can alter the reading frame and generate a premature stop codon in exon 52 (red). Adenine base editing at the splice site of exon 50 (red) by Nme2Cas9-ABE8e can cause exon 50 skipping (gray) and restore the *Dmd* reading frame. The PAM region is underlined in green.

Fig. 23D: Amplicon deep sequencing quantifying the editing efficiency at the target site in the mouse N2a cells transfected with the Nme2Cas9-ABE8e and sgRNA_*Dmd* expression plasmid (n = 3 biological replicates).

Fig. 24 presents exemplary data evaluating nuclear localization signal protein delivery of fusion proteins comprising an N-terminally fused Nme2Cas9-ABE8e protein.

Fig. 24A: Schematic representation of single AAV constructs with different NLS configurations.

Fig. 24B: Comparison of different NLS configurations by plasmid transfection in the ABE reporter cell line.

Fig. 24C: Comparison of U6 and miniU6 promoters for sgRNA expression in the 2xBP_SV40 NLS construct targeting the ABE reporter site (left) or endogenous human (middle) and mouse (right) genomic sites by plasmid transfection in cultured cells followed by amplicon deep sequencing (n = 3 biological replicates).

Fig. 25 presents exemplary data confirming *Fah* gene mutation reversion with the AAV delivery plasmid/vectors.

Fig. 25A: Schematic representation of U6 or miniU6 AAV N-terminally fused Nme2Cas9-ABE8e constructs.

Fig. 25B: Anti-FAH IHC staining showing FAH⁺ hepatocytes, before NEBC withdrawal, in the *Fah*^{PM/PM} mouse injected with AAV9 expressing Nme2Cas9-ABE8e with a sgRNA targeting either the *Fah* gene, or the *Rosa26* gene that serves as a negative control. Scale bar, 500 μm. The bar graph quantifies the percentage of FAH⁺ hepatocytes detected by IHC (n=4 mice per group).

Fig. 25C: Quantification of the editing efficiency by amplicon deep-sequencing of the genomic DNA extracted from AAV9 injected mouse livers harvested before NTBC withdrawal (n=4 mice per group).

Fig. 26 presents exemplary data showing single base mutation reversion and exon 8 skipping strategy in the *Fah* gene.

Fig. 26A: Illustration of the pathogenic point mutation in the *Fah*^{PM/PM} mouse model that causes exon 8 skipping of the *Fah* gene, and the guide design for Nme2Cas9-ABE8e to correct the point mutation. Red and bold, target adenine; orange, other bystander adenines; green and underlined, PAM.

Fig. 26B: Illustration of constructs of the single AAV vector plasmids used in in vivo studies.

Fig. 26C: Editing efficiencies at the *Fah* mutant site by AAV plasmid electroporation in MEF cells derived from the *Fah*^{PM/PM} mouse. Data are from amplicon deep sequencing (n = 2 biological replicates).

Fig. 26D: Anti-FAH immunohistochemistry (IHC) staining showing FAH⁺ hepatocytes, before NTBC withdrawal, in the *Fah*^{PM/PM} mouse hydrodynamically injected with the indicated plasmid. The bar graph quantifies the percentage of FAH⁺ hepatocytes detected by IHC. Scale bars, 500 μm.

Fig. 26E: Body weight plot of mice injected with the single-AAV vector plasmid showing gradual weight gain over a month after NTBC withdrawal.

Fig. 26F: RT-PCR analysis of the plasmid- or PBS-injected mouse livers using primers that span exons 5 and 9. The wild-type amplicon is 405 bp and exon 8 skipped amplicon is 305 bp.

Fig. 26G: Representative Sanger sequencing trace of the 405 bp RT-PCR band.

5 Fig. 26H: Anti-FAH IHC staining showing expansion of FAH⁺ hepatocytes 40 days post NTBC withdrawal. Scale bars, 500 μ m.

Fig. 26I: Quantification of the editing efficiency by amplicon deep sequencing of genomic DNA of the treated mouse livers harvested 40 days post NTBC withdrawal. (d-i, n = 2 mice per group).

10 Fig. 27 presents exemplary data showing flow cytometry gating strategy for the ABE reporter cell line.

Fig. 28 presents exemplary data validating N-terminally fused ABE domain construct stability subsequent to AAV delivery at a target site.

15 Fig. 28A: Alkaline gel electrophoresis of AAV9 genomic DNAs targeted to the *Rosa26* locus.

Fig. 28B: Quantification of the gene editing at the *Rosa26* locus by amplicon deep sequencing using mouse livers injected with indicated AAV9. (n = 3 mice per group).

20 Fig. 29 presents an off-target analysis for the N-terminally fused Nme2Cas9-ABE8e domain construct at the *Fah* gene.

Fig. 29A: Sequence of the *Fah* on-target site and two top-rated Cas-OFFinder predicted off-target sites for Nme2Cas9-ABE8e. Bases that are different from the on-target site are labeled in red. PAM, green, underlined.

25 Fig. 29B: Representative amplicon deep sequencing reads at the predicted off-target sites in mouse injected with AAV9 expressing Nme2Cas9-ABE8e and sgRNA-*Fah*.

Fig. 30 presents an illustrative three-dimensional representation of an induced separation of an inlaid nucleotide deaminase protein domain and the N-terminus of a Cas9 protein.

Fig. 31 presents one embodiment of a Cas9 protein with an inlaid nucleotide deaminase domain.

Fig. 31A: One embodiment of an inlaid nucleotide deaminase domain insert. Orange: Nucleotide Deaminase protein; Purple: N-terminal linker; Blue: C-terminal linker.

Fig. 31B: An illustration of several candidate inlaid nucleotide deaminase domain insertion sites in the NmeCas9 protein (as indicated by colored lines). TadA8e Deaminase was inserted into regions of a RUV-C Nme2Cas9 (D16A) nickase. The insertion sites based on the criteria in red, and were based on NmeCas9 crystal structures (PDB: 6jDV; Sun et al. Mol Cell. 2019).

Fig. 31C: Proposed three-dimensional locations within the NmeCas9 PDB 6JDV of the inlaid nucleotide deaminase protein domains illustrated in Fig. 30B (color matched). Site 1: Q291-RECII (red); Site 2: D328-RECII (orange); Site 3: K339-RECII (taupe); Site 4: R643-HNH (green); Site 5: E659-Linker 2 (light blue); Site 6: V715-RUVCIII (dark blue); Site 7: E761-RUVCIII (purple); and Site 8: P795-RUVCIII (pink).

Fig. 32 presents exemplary data showing gene editing activity for the ABE inlaid domain Cas9 locations defined in Fig. 30.

Fig. 32A: A schematic of the ABE mCherry reporter system for identifying gene editing activity. The ABE reporter is stably integrated into the genome of HEK293T cells

Fig. 32B: Representative photomicrographs of gene editing activity at various ABE inlaid domain Cas9 locations as indicated by the red fluorescence intensity. Fluorescent images of ABE reporter cells 72hrs post transfection with plasmids that express Nme2Cas9-ABE and the guide RNA to correct the mCherry stop codon.

Fig. 33 presents exemplary data of Sanger sequencing of ABE mCherry reporter data in Fig. 32 after editing with Nme2Cas9-ABE8e variants. The positive control is N-term fused Nme2Cas9-ABE8e. The dashed black line represents the target adenine base.

Fig. 34 presents exemplary data showing estimated gene editing data based upon the mCherry system. Quantification of editing rates for the inlaid Nme2Cas9-ABE variants is compared to the N-terminal fused Nme2Cas9-ABE8e as a positive control (gray bar).

Fig. 34A: Flow cytometry of ABE mCherry reporter cells 72hrs post transfection with an ABE effector and guide RNA.

Fig. 34B: Amplicon sequencing of the targeted mCherry locus, showing % reads with an A to G conversion at the target adenine.

Fig. 35 presents exemplary data showing gene editing activity of three endogenous loci using the eight ABE inlaid domain location described in Fig. 30. Editing Rates of Nme2Cas9-ABE variants at three endogenous genomic loci 72hrs post transient transfection. Data analyzed by sanger sequencing and EditR tool that quantifies nucleotide frequency in a pool of PCR amplicons. The spacer sequence is on the X-axis and the observed nucleotide frequency is on the Y-Axis.

Fig. 35A: LINC01588-DS12

Fig. 35B: FANCF-DS28

Fig. 35C: MECP2-G2

Fig. 36 presents exemplary data comparing N-terminal domain Nme2-ABE with inlaid domain Nme2-ABE gene editing of DMD mutations. Sanger sequencing quantifies the editing efficiency at the target site in mouse N2a cells transfected with Nme2-ABE8e -nt, -i1, -i7 and -i8 expressed in an sgRNA_Dmd expression plasmid (n = 2 biological replicates, 2 technical replicates).

Fig. 37 presents illustrative guide RNA sequences with slice donors and acceptors that target CLN3 exon 5 used with Nme2-ABE constructs to treat Batten disease

Fig. 37A: Guide mRNA sequence targeting mouse CLN3 exon 5.

Fig. 37B: Guide mRNA sequence targeting human CLN3 exon 5.

Fig. 38 presents illustrative splice donor and splice acceptor target sequences in CLN3 exon 5 to treat Batten disease. for human and mouse. Target sequences are from 5' to 3' with target adenine (red), and PAM (bold, underlined). Measured via amplicon sequencing (n = 3).

Fig. 39 presents exemplary data showing CLN3 exon 5 gene editing efficiency with the Nme2-ABE-i1 construct.

Fig. 39A: A-to-G editing with Nme2-ABE8-i1 effectors at the splice donor or acceptor of mouse CLN3 in N2A cells with single AAV vector plasmid. Measured via sanger sequencing (n = 3). Exon Skipping validated by RT-PCR .

Fig. 39B: A-to-G editing with Nme2-ABE8e -nt, i1 and -i1^{V106W} constructs at the splice donor or acceptor of human CLN3 with vector plasmids expressing effector or guide. Measured via amplicon sequencing (n = 3).

Fig. 40 presents exemplary data showing mouse CLN3 gene targeting with Nme2-iABE8e_1 to generate exon 5 skipping in cultured N2a cells.

Fig. 40A: Illustration of mouse CLN3 exon 5 sequence alignment with splice acceptor/donor positions and three Nme2-Cas9 guides.

Fig. 40B: Exemplary data of CLN3 exon 5 mutation conversion with Nme2-Cas9-ABE administered using different guide constructs depicted in Fig. 22(5)A.

Fig. 40C: A representative gel electrophotograph showing RT-PCR on mCLN3 transcript from transfected N2a cells. NC: negative control; 1: Nme2-iABE_1-mCLN3_G1; 2: Nme2-iABE_1-mCLN3_G4.

Fig. 40D: Exemplary Sanger sequencing base calling data showing CLN3 exon 5 skipping subsequent to Nme2-Cas9-ABE gene editing by the adjacent location of CLN3 exon 4 and exon 6.

Fig. 41 presents exemplary data showing single AAV delivery of Nme2-iABE_1-sgRNA targeting mouse brain CLN3 genes.

Fig. 41A: Exemplary data of CLN3 exon 5 editing efficiency in mouse cortex, striatum, hippocampus and thalamus using the different Nme2-Cas9 guide constructs in accordance with Fig. 22(5).

Fig. 41B: Exemplary data showing total RNA (RT-PCR) in mouse striatal tissue subsequent to a high dose AAV delivery of Nme2-iABE_1-sgRNA and Nme2-iABE_4-sgRNA.

Fig. 41C: Exemplary Sanger sequencing base calling data showing CLN3 exon 5 skipping subsequent to Nme2-Cas9-ABE gene editing by the adjacent location of CLN3 exon 4 and exon 6.

Fig. 42 presents exemplary mouse exon 5 Nme2-ABE editing data comparing plasmid injections into neonatal intracerebral ventricles (ICV) with adult intrastriatal (IS).

Fig. 42A: Illustrative AAV plasmid constructs.

Fig. 42B: Adult B6 mouse IS, 8×10^9 GC/mouse, 8 weeks incubation, deep sequencing (target A-to-G).

Fig. 42C: P1 B6 mouse neonate ICV, 1.5×10^{10} GC/mouse, 4 weeks incubation, deep sequencing (target A-to-G).

Fig. 42D: Exemplary data showing gene editing in mouse striatum

Fig. 42E: Exemplary data showing gene editing in mouse liver.

Fig. 43 presents exemplary photomicrographs showing brain regional Nme2-1ABE8e_1 mRNA transcript expression in transverse mouse brain slices.

Fig. 43A: Adult IS injection of AAV9-Nme2-iABE8e_1-sgRNA

5 8-week mouse, bilateral IS injection, 9×10^9 vg per side.

Fig. 43B: Neonatal ICV injection of AAV9-Nme2-iABE8e_1-sgRNA

P1 mouse, bilateral ICV injection, 3×10^{10} vg per mouse.

Fig. 43C: Phosphate buffered saline control injection.

10 Fig. 44 presents exemplary data showing Nme2Cas9-ABE conversion of Rett Syndrome : c.502 C>T mutation.

Fig. 44A: A schematic illustration of a portion of Mecp2 exon 4 in Rett patient-derived fibroblasts (PDFs). Nonsense mutation: c.502 C>T; p.R168X (red); Potential bystander edits (orange).

15 Fig. 44B: Exemplary data of A-to-G editing of Mecp2 c.502 C>T in the RETT-PDF cell line in accordance with Fig. 22(1+)A, with Nme2-ABE8e effectors delivered as mRNA with synthetic sgRNAs. Protospacer with target adenine (red), bystander adenine (orange), and PAM (bold, underlined). n=3 biological replicates, data represent mean $\pm$ SD.

20 Fig. 45 presents exemplary data showing Nme2Cas9-ABE conversion of Rett Syndrome c.916 C>T mutation.

Fig. 45A: A schematic illustration of a portion of Mecp2 exon 4 in RETT patient derived fibroblasts. Missense mutation: c.916 C>T; p.R306C (red). Potential bystander edits (orange).

25 Fig. 45B: Exemplary data of A-to-G editing of Mecp2 916C>T in RETT-PDF cell line in accordance with Fig. 22(2+)A. with Nme2-ABE8e effectors delivered as mRNA with synthetic sgRNAs. Protospacer with target adenine (red), bystander adenine (orange), and PAM (bold, underlined). n=3 biological replicates, data represent mean $\pm$ SD.

30 Fig. 46 presents exemplary data showing Nme2Cas9-ABE conversion of Rett Syndrome c.763C>T mutation.

Fig. 46A: A schematic illustration of a portion of Mecn2 exon 4 in RETT piggyBac cells, Missense mutation: c.763 C>T; p.R255X (red).

Fig. 46B: Exemplary data of A-to-G editing of Mecn2 763C>T in RETT-PiggyBac cell line in accordance with Fig. 22(3+)A with Nme2-ABE8e effectors and sgRNA delivered as plasmids. Protospacer with target adenine (red), and PAM (bold, underlined). n=2 biological replicates, data represent mean $\pm$ SD.

Fig. 47 presents exemplary data showing Nme2Cas9-ABE conversion of Rett Syndrome c.808C>T mutation.

Fig. 47A: A schematic illustration of a portion of Mecn2 exon 4 in RETT piggyBac cells, Missense mutation: c.808C>T; p.R270X (red) and potential bystander edits. (orange).

Fig. 47B: Exemplary data of A-to-G editing of Mecn2 : c.808C>T in RETT-PiggyBac cell line in accordance with Fig. 22(4+)A with Nme2-ABE8e effectors and sgRNA delivered as mRNA with synthetic sgRNA. Protospacer with target adenine (red), and PAM (bold, underlined). n=2 biological replicates, data represent mean $\pm$ SD.

Fig. 48 presents exemplary data showing in vivo gene editing with AAV9 comparing: i) an Nme2-ABE8e-terminal ABE domain (nt) construct; ii) an Nme2-ABE8e-inlaid ABE domain (il) construct; and iii) an Nme2-ABE8e-mutated inlaid ABE domain (il^{V106W}) construct.

Fig. 48A: Illustrative schematics of the Nme2-ABE8e -nt, -il and -il^{V106W} AAV constructs.

Fig. 48B: Exemplary data showing in vivo gene editing with the -nt, il and il^{V106W} AAV Nme2-ABE constructs in mouse liver (left panel) and striatum (right panel).

Liver data is expressed as a quantification of the editing efficiency at the *Rosa26* locus as measured by amplicon deep sequencing using liver genomic DNA from tail-vein injected mice at 4×10^{11} vg/mouse (n = 3 mice per group). Striatal data is expressed as a quantification of the editing efficiency at the *Rosa26* locus as measured by amplicon deep sequencing from intrastrially injected mice at 1×10^{10} vg/side (n = 3 mice per group). Data is represented mean $\pm$ SD and analyzed using a two-way ANOVA analysis: ns, p > 0.05; ***p < 0.05; **p < 0.01; ***p < 0.001 ****p < 0.0001.

Fig. 48C: Exemplary data showing in vivo off-target gene editing with the -nt, il and il^{V106W} AAV Nme2-ABE constructs in mouse liver relative to a *Rosa26* on-target site protospacer and a previously validated Nme2-ABE8e off-target site (OT1). Upper Panel: A *Rosa26* protospacer sequence annotated with: i) target adenines (red); ii) OT1 mismatches (orange); and iii) PAMs (bold, underlined). Lower Pane: A bar graph showing representative data of the quantification of A-to-G editing as measured by amplicon deep sequencing reads at the *Rosa26*-OT1 site by mice tail vein AAV injection. Data represent mean $\pm$ SD. two-way ANOVA analysis: ns, $p > 0.05$; **** $p < 0.0001$.

Fig. 49 presents exemplary data showing sensitivity of guide-dependent ABE domain Cas9 constructs.

Fig. 49A: Upper Panel: A 5' to 3' overlapping target sequence. Red: target adenine; Bold/Underline = PAM. Lower Panel: Exemplary bar graph data showing a comparison of on-target activity of Spy-ABE8e, Nme2-ABE8e -nt and Nme2-ABE8e-i1 constructs using the ABE mCherry reporter system by plasmid transfection and flow cytometry.

Fig. 49B: Exemplary sequences show Spy-ABE8e construct mismatch (mm) nucleotide (orange) tolerance in ABE mCherry reporter cells with the overlapping target site depicted in Fig. 27(2+)A. Bar Graph: Single guide RNAs with mismatched (mm) nucleotides (orange) are normalized to the activity of fully complementary guides. n=3 biological replicates, Bars: Mean $\pm$ SD.

Fig. 49C: Nme2-ABE8e-nt construct mismatch tolerance in ABE mCherry reporter cells as in Fig.27(2)B at the overlapping target site as in Fig. 27(2)A.

Fig. 49D: Nme2-ABE8e-i1 construct mismatch tolerance in ABE mCherry reporter cells as in Fig.27(2)B at the overlapping target site as in Fig. 27(2)A.

Fig. 49E: Ratios of on-target/off-target editing for the Spy, Nme2-nt, Nme2-i1 and Nme2-i1^{V106W} ABE constructs tested at the overlapping Linc01588 target site (S2A) and the orthogonal SaCas9 R-loops (S2B). On-target efficiency for Spy-ABE8e is derived from the mean editing within its editing window as to not skew the ratio when compared to the wider on-target editing window of the three (3) Nme2-ABE8e constructs. n=3 biological replicates, data represent mean $\pm$ SD. two-way

ANOVA analysis: ns, $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$.

Fig. 50 presents exemplary data showing sensitivity of guide-independent ABE Cas9 constructs.

- 5 Fig. 50A: Exemplary data showing orthogonal R-Loop off-Target activity of guide-independent DNA A-to-G editing at orthogonal SaCas9 R-loops with Spy, Nme2-nt, Nme-i1 and Nme-i1^{V106W} ABE constructs as measured via amplicon sequencing. HNH nicking of the SaCas9 protein increased editing sensitivity at the orthogonal R-loops. n=3 biological replicates, data represent mean $\pm$ SD.
- 10 50B: Exemplary data showing on-target activity at a dual PAM Site (DS12) at the indicated target site (Upper Panel). Lower Panel: On-target activity of the Spy, Nme2-nt, Nme-i1 and Nme-i1^{V106W} ABE constructs were tested for the R-loop activity at a target site with overlapping PAMs as measured via amplicon sequencing. Box: Spy-ABE8e editing window. Overlapping target site sequence
- 15 from 5' to 3' with adenines (red), and Spy- and Nme2- PAMs bold and underlined. n=3 biological replicates per off-target R-loop, data represent mean $\pm$ SD.

Fig. 51 presents representative Nme2Cas9 embodiments of inlaid TadA7.10 deaminase domains (e.g., ABE7.10).

- 20 Fig. 51A: Upper Panel: 5' – 3' target sequence. Targeted adenine's (red). PAM (bold, underlined). Lower Panel: Exemplary data of A-to-G editing at an endogenous HEK293T Linc01588 loci with Nme2-ABE8e-i1, Nme2-ABE7-i1, Nme2-ABE7-i7 and Nme2-ABE7-i8 constructs. Measured via amplicon sequencing (n = 3).
- 25 Fig. 51B: Upper Panel: 5' – 3' target sequence. Targeted adenine's (red). PAM (bold, underlined). Lower Panel: Exemplary data of A-to-G editing at an endogenous HEK293T MeCP2 loci with Nme2-ABE8e-i1, Nme2-ABE7-i1, Nme2-ABE7-i7 and Nme2-ABE7-i8 constructs. Measured via amplicon sequencing (n = 3).
- 30 Fig. 51C: Upper Panel: 5' – 3' target sequence. Targeted adenine's (red). PAM (bold, underlined). Lower Panel: Exemplary data of A-to-G editing at an endogenous HEK293T Pcsk9 loci with Nme2-ABE8e-i1, Nme2-ABE7-i1, Nme2-

ABE7-i7 and Nme2-ABE7-i8 constructs. Measured via amplicon sequencing (n = 3).

Fig. 52 illustrates various embodiments of an Nme2Cas9 cytidine base editor (CBE) domain constructs.

- 5 Fig. 52A: Schematic representation of exemplary Nme2Cas9-CBE constructs: i) Nme2Cas9-CBE-(nt) N-terminal domain; ii) Nme2Cas9-CBE-(i1) inlaid domain; iii) Nme2Cas9-CBE-(i7) inlaid domain; and iv) Nme2Cas9-CBE-(i8) inlaid domain. The CBE may be a cytidine deaminase including, but not limited to, evoFERNY or rAPOBEC1.
- 10 Fig. 52B: Exemplary data of C-to-T editing at endogenous HEK293T genomic loci with Nme2-evoFERNY -nt, i1, i7 and i8 constructs (top) and Nme2-rAPOBEC1-nt, i1, i7 and i8 constructs (bottom) as measured by amplicon sequencing. n=3 biological replicates, data represent mean $\pm$ SD.
- 15 Fig. 52C: Exemplary data showing rAPOBEC1 editing window and editing efficiency in HEK293T cells using three (3) genomic target sites. Data is expressed as a summary of mean C-to-T editing at the three endogenous HEK293T genomic target with Nme2-rAPOBEC1 -nt, -i1, i7 and i8 constructs. Crossed out boxes denote no cytidine at the position within the target's tested. Measured via amplicon sequencing. n=3 biological replicates.
- 20 Fig. 53 presents exemplary data showing PID Chimera's expand the targeting scope of Nme2Cas9 base editors.
- Fig. 53A: Exemplary data showing expanded PAM scope of PID chimeric Nme2Cas9 nucleases.
- Fig. 53B: Cartoon schematic of chimeric Nme2-ABE8e effectors with SmuCas9 PAM interacting domains
- 25 Fig. 53C: Examples of A-to-G editing with PID chimeric Nme2-ABE8e effectors at endogenous HEK293T genomic loci with N4CN PAMs by plasmid transfection and measured by amplicon sequencing. n=3 biological replicates, data represent mean $\pm$ SD.

Fig. 53D: Each data point represents the mean A-to-G editing rates of an individual target site, measured by amplicon sequencing. The line represents the mean efficiency observed across all target sites. n=3 biological replicates.

Fig. 53E: A-to-G editing with PID chimeric Nme2-ABE8e effectors at Linc01588 endogenous HEK293T genomic loci with N4CN PAMs by plasmid transfection and measured by amplicon sequencing. Target spacers 5' to 3' with PAM bold underlined. n=3 biological replicates, data represent mean $\pm$ SD.

Fig. 54 presents exemplary data showing conversion of the c.502 C>T (RETT-PDF) mutation with a chimeric Nme2Cas9Smu construct.

Fig. 54A: Schematic of a portion of Mecp2 exon 4, highlighting the (c.502 C>T; p.R168X) nonsense mutation (red) and potential bystander edits (orange), in RETT patient derived fibroblasts,

Fig. 54B: A-to-G editing of Mecp2 502C>T in RETT patient fibroblasts in (A), with Nme2-ABE8e effectors delivered as mRNA with synthetic sgRNAs.

Protospacer with target adenine (red), bystander adenine (orange), and PAM (bold, underlined). n=3 biological replicates, data represent mean $\pm$ SD.

Fig. 55 presents exemplary data showing conversion of the c.916 C>T (RETT-PDF) mutation with a chimeric Nme2Cas9Smu construct.

Fig. 55A: Schematic of a portion of Mecp2 exon 4, highlighting the (c.916 C>T; p.R306C) missense mutation (red) and potential bystander edits (orange), in RETT patient derived fibroblasts.

Fig. 55B: A-to-G editing of Mecp2 916C>T in RETT-PDF cell line in (A), with Nme2-ABE8e effectors delivered as mRNA with synthetic sgRNAs. Protospacer with target adenine (red), bystander adenine (orange), and PAM (bold, underlined).

n=3 biological replicates, data represent mean $\pm$ SD.

Fig. 56 presents exemplary data showing conversion of the c.763C>T (RETT-PiggyBac) mutation with a chimeric Nme2Cas9Smu construct.

Fig. 56A: Schematic of a portion of Mecp2 exon 4, highlighting the (c.763 C>T; p.R255X) missense mutation (red) in RETT piggyBac cells.

Fig. 56B: A-to-G editing of Mecp2 763C>T in RETT-PiggyBac cell line in (A), with Nme2-ABE8e effectors and sgRNA delivered as plasmids. Protospacer with

target adenine (red), and PAM (bold, underlined). n=2 biological replicates, data represent mean $\pm$ SD.

Fig. 57 presents exemplary data showing conversion of the c.808C>T (RETT-PiggyBac) mutation with a chimeric Nme2Cas9Smu construct.

5 Fig. 57A: Schematic of a portion of Mecp2 exon 4, highlighting the (c.808C>T; p.R270X) missense mutation (red) in RETT PiggyBac cells and potential bystander edits. (orange).

Fig. 57B: A-to-G editing of Mecp2 808C>T in RETT-PiggyBac cell line in (A), with Nme2-ABE8e effectors delivered as mRNA with synthetic sgRNAs.

10 Protospacer with target adenine (red), bystander adenines (orange) and PAM (bold, underlined). n=2 biological replicates, data represent mean $\pm$ SD.

Fig. 58 illustrates the expanded PAM scope and increased candidate targets provided by chimeric Cas9^{Smu} nucleases having single C PAMs.

15 Fig. 58A: A summary of reported respective target scope comparing dinucleotide C PAMs to single C PAMs.

Fig. 58B: An illustration of reported change in Cas9 nuclease PAM preference as a result of a PID swap.

Fig. 58C: A reported characterization of the SmuCas9 PID.

20 Fig. 59 presents exemplary data showing in vivo ameliation of ALS symptoms following inlaid domain Nme2Cas9-ABE administration.

Fig. 59A: Dual AAV9 vector design.

Fig 59B: Survival curve.

25 Fig. 59C: Representative L5 ventral root cross sections in P110 mice (e.g., lifespan midpoint) showing ALS-mediated cell breakdown reversal subsequent to gene editing.

Fig. 60 presents a representative illustration of SOD1 exon 2 skipping by editing splicing sites. Intron residues are in lower case.

(A) Gene editing of the intron 1 splice acceptor. The N₄CC PAM places the target A at A15. The single-C PAM places the target B at A15.

30 (B) Gene editing the intron 2 splice donor. Additional G residues (for single-C PAMs) are highlighted.

Fig. 61 presents a portion of the SOD1 exon 2 loci to identify potential Nme2Cas9-ABE fusion proteins and their respective guide RNAs to correct an SOD1^{G37R} mutation. Missense bystander nucleotides are in blue. Frameshift mutation nucleotides are in red.

Fig. 62 presents an exemplary Nme2Cas9^{Smu} construct: (A) amino acid sequence; (B) nucleic acid sequence; BPSV40-NLS (purple), Nme2Cas9 – delta PID (blue), SmuCas9-PID (orange), Linkers (black).

Fig. 63 presents an exemplary Nme2Cas9^{Smu} construct: (A) amino acid sequence; (B) nucleic acid sequence; SV40-NLS (purple), nucleoplasmin-NLS (green), Nme2Cas9 – delta PID (blue), SmuCas9 PID (orange), *3xHA* (*italicized*), unlabeled-NLS (red), Linkers (black).

Fig. 64 presents an exemplary Nme2^{smu}-ABE8e-i1 inlaid domain construct. (A) amino acid sequence; (B) nucleic acid sequence; BPSV40-NLS (purple), Nme2Cas9 – delta PID (blue), TadA8e (red), SmuCas9 PID (orange), Linkers (black).

Fig. 65 presents an exemplary Nme2^{smu}-ABE8e-i7 inlaid domain construct. (A) amino acid sequence; (B) nucleic acid sequence; BPSV40-NLS (purple), Nme2Cas9 – delta PID (blue), TadA8e (red), SmuCas9 PID (orange), Linkers (black).

Fig. 66 presents an exemplary Nme2^{smu}-ABE8e-i8 inlaid domain construct. (A) amino acid sequence; (B) nucleic acid sequence; BPSV40-NLS (purple), Nme2Cas9 – delta PID (blue), TadA8e (red), SmuCas9 PID (orange), Linkers (black).

Fig. 67 presents an exemplary Nme2^{smu}-ABE8e-nt N-terminal domain construct. (A) amino acid sequence; (B) nucleic acid sequence; BPSV40-NLS (purple), Nme2Cas9 – delta PID (blue), TadA8e (red), SmuCas9 PID (orange), Linkers (black).

Detailed Description Of The Invention

The present invention is related to the field of gene editing. In particular, the gene editing is directed toward single nucleotide base editing. For example, such single nucleotide base editing results in a conversion of a mutated base pair to a wild type base pair. The high accuracy and precision of the presently disclosed single nucleotide base gene editor is accomplished by a fusion protein including an NmeCas9 nuclease and an inlaid nucleotide deaminase protein domain. The Nme2Cas9 protospacer interacting domain may be replaced with an SmuCas9 protospacer interacting domain.

In one embodiment, the present invention contemplates a Cas9 protein contemplating an exogenous inlaid domain. In one embodiment, the exogenous inlaid domain is inserted in the Cas9 REC domain. In one embodiment, the exogenous inlaid domain is inserted in the Cas9 HNH domain. In one embodiment, the exogenous inlaid domain is inserted in the Cas9 RuvC domain. In one embodiment, the exogenous inlaid domain is a nucleotide base editor. In one embodiment, the nucleotide base editor is an adenine base editor (ABE). In one embodiment, the nucleotide base editor is a cytidine base editor (CBE). For example, an inlaid domain Nme2Cas9-ABE fusion protein comprises greater gene editing efficiency as compared to a N-terminal domain NmWcas9-ABE fusion protein. See, Figs 1A-D.

The insertion of an inlaid domain may be placed in a variety of positions within the Cas9 protein, each of which has superior gene editing activity as compared to the N-terminal domain construct. See, Figs 2A-C.

I. CRISPR Cas9 Gene Editors

A. N-Terminal Cas9 Deaminase Fusion Proteins

Fusion proteins have been reported comprising an Nme2Cas9 and an N-terminal deaminase protein. See, Fig. 3A and Fig. 3B. For example, the deaminase protein is Apobec1 (YE1-BE3). Kim et al., "Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions". *Nature Biotechnology* 35 (2017). The YE1-BE3-nNme2Cas9 (D16A)-UGI construct has the sequence of:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLLYEINWGGRHSIWRHTSQNTNKH
VEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGECSRAITEFLSRYPHVTLFIYIARLYHHAD
PENRQGLRDLISSGVTIQIMTEQESGYCWRNRFVNYSNPSNEAHWPYPHLLWVRLYVLELYC

IILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESATPESMA
 AFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRFVFERAEVPKTGDSLAMARRLARSV
 RRLTRRRRAHRLLRARLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLL
 HLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEEKESGHIRNQ
 5 RGDYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGGLKEGIEITLLMTQRPALSGDAVQKMLGHC
 TFEPAPKAAKNNTYTAERFIWLTKLNNLRILEQGSERPLDTERATLMDEPYRKSCLTYAQARKL
 LGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFS
 LFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKALRRIVPLMEQGKRYDEACAEIYGDH
 YGKKNTTEEKIYLPPIPADEIRNPVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEI
 10 EKROEENRKDREKAAAKFREYFPNFVGEPSKSDILKLRLYEQQHKGKCLYSGKEINLVRNEKGY
 VEIDAALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSFRFPR
 SKKQORILLQKFDEEDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGF
 WGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTH
 FPQPWEFFAQEVMRVFGKPDGKPEFEEADTPEKLRITLLAEKLSSRPEAVHEYVTPLFVSRAPN
 15 RKMSGAHKDTLRSKRFRVKNHEKISVSRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAY
 GGNKQAFDPKDNPFYKKGGLVKAARVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCVKV
 DKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYI
 NCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVRSRGGSTNL
 SDIEKETGKQLVQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAP
 20 EYKPWALVIQDSNGENKIKMLSGGSPKKRKV*

YE1-BE3 (underlined); linker (bold), nNme2Cas9 (italics), UGI (bold/underlined), SV40 NLS (unannotated*).

Another example comprises an Nme2Cas9 and a terminal TadA adenine deaminase protein
 25 (e.g., ABE7.10). The ABE7.10-nNme2Cas9 (D16A) construct has the following sequence:

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHA
 EIMALROGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSL
 MDVLHHPGMNHRVEITEGILADECAALLSDFRMRREQEIKAQKKAQSSTD.SGGSSGGSSG
 30 SETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLN
NRVIGEGWNRRAIGLHDPTAHAEIMALROGGLVMQNYRLIDATLYVTLEPCVMCAGA
MIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPROVFNAQKKAQSSTD.SGGSSGGSSGSETPGTSESATPESSGGSSGGSMAAFKPNPINYILG
 LAIGIASVGWAMVEIDEEENPIRLIDLGVRFVFERAEVPKTGDSLAMARRLARSVRRLTRRRRAHRL
 35 RARLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR
 KNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKD
 LQAEILLLFEKQKEFGNPHVSGGLKEGIEITLLMTQRPALSGDAVQKMLGHCTFEPAPKAAKN
 TYTAERFIWLTKLNNLRILEQGSERPLDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGL
 40 RYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLK
 DRVQPEILEALLKHISFDKQVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTTEEKIYLP
 PIPADEIRNPVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKROEENRKDRE
 KAAAKFREYFPNFVGEPSKSDILKLRLYEQQHKGKCLYSGKEINLVRNEKGYVEIDHALPFSRT
 WDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSFRFPRSKKQORILLQKF

DEDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAEND
 RHHALDAVVVACSTVAMQOKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQE
 VMIRVFGKPDGKPEFEEADTPEKLRITLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMMSGAHKDTL
 5 RSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPK
 DNPFFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPI
 YAWQVAENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAW
 HDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVREDKRPAAATKKAGQAACKKK
 *

TadA (underlined), TadA 7.10 (underlined/bold), *linker* (bold), nNme2Cas9 (italics),

10 Nucleoplasmin NLS (unannotated*).

Electroporation of HEK293T cells with DNA plasmids comprising a YE1-BE3-Nme2Cas9
 with a terminally fused nucleotide deaminase protein demonstrated single-base editing of a C•G
 base pair to a T•A base pair at an endogenous target site (TS25). See, Fig.s 4A-C. Four other
 YE1-BE3-nNme2Cas9/D16A mutant N-terminal fusion proteins were co-expressed with enhanced
 15 green fluorescent protein (EGFP) in a stable K562-derived cell line. Each YE1-BE3-
 nNme2Cas9/D16A mutant N-terminal fusion protein had a specific UGI target site. See, Fig.s 5A-
 D. Deep-sequencing analysis indicates YE1-BE3-nNme2Cas9 converts C residues to T residues at
 each of the four EGFP target sites. The percentage of editing ranged from 0.24% to 2%. The
 potential base editing window is from nucleotides 2-8 in the displaced DNA strand, counting the
 20 nucleotide at the 5' (PAM-distal) end as nucleotide #1.

The expression of an ABE7.10-nNme2Cas9 (D16A) N-terminal fusion protein for base
 editing may be an effective treatment for tyrosinemia by reversing a G-to-A point mutation in the
Fah gene. G-to-A mutation (red) at the last nucleotide of exon 8 in *Fah* gene, causing exon
 skipping. Exon skipping provides benefit by eliminating a pathogenic mutation from a mature
 25 mRNA, restoring reading frame to compensate for a disease-causing frameshift mutation, or
 inactivating the expression of a gene that contributes to disease (by inducing an out-of-frame
 splicing event, or deleting an essential gene region from the mature mRNA, or both). The latter
 allows ABEs to induce gene knockouts, given that ABEs cannot be used to introduce nonsense
 mutations - the ABE precursor to either G-containing stop codon (UAG or UGA) is UAA, which is
 30 already a stop codon. ABEs can convert AG splice acceptor site dinucleotides to GG and can also
 convert GU splice donor site dinucleotides (AC on the opposite strand) to GC.

FAH deficiency leads to toxin accumulation and severe liver damage. The position of a
 SpyCas9 PAM (black rectangular box) downstream of the mutation is not optimal for designing
 the sgRNA since the A mutation is out of the efficient base editing window of ABE7.10, which is

4-7th nt at the 5' (PAM-distal) end (underlined) (Gaudelli et al., 2017). However, there are two Nme2Cas9 PAMs (red rectangular box) in the downstream sequences that can potentially correct the mutation and revert DNA sequence to wildtype via ABE7.10-nNme2Cas9 (D16A). This figure serves as a potential example of a site where Nme2Cas9 with N-terminal adenine deaminase proteins could overcome limitations of existing base editors. See, Fig. 6. It is further believed that Nme2Cas9 base editors can perform precise base editing that cannot be achieved with conventional SpyCas9-derived base editors due to a suboptimal base editing window relative to available PAMs nearby.

B. Nme2Cas9 PAM Interacting Domains

Protospacer adjacent motif (PAM) recognition by Cas9 orthologs occurs predominantly through protein-DNA interactions between the PAM Interacting Domain (PID) and the nucleotides adjacent to the protospacer (Jiang and Doudna, 2017). PAM mutations often enable phage escape from type II CRISPR immunity (Paez-Espino et al., 2015), placing these systems under selective pressure not only to acquire new CRISPR spacers, but also to evolve new PAM specificities via PID mutations. In addition, some phages and MGEs express anti-CRISPR (Acr) proteins that inhibit Cas9 (Pawluk et al., 2016; Hynes et al., 2017; Rauch et al., 2017). PID binding is an effective inhibitory mechanism adopted by some Acrs (Dong et al., 2017; Shin et al., 2017; Yang and Patel, 2017), suggesting that PID variation may also be driven by selective pressure to escape Acr inhibition. Cas9 PIDs can evolve such that closely-related orthologs recognize distinct PAMs, as illustrated recently in two species of *Geobacillus*. The Cas9 encoded by *G. stearothermophilus* recognizes a N₄CRAA PAM, but when its PID was swapped with that of strain LC300's Cas9, its PAM requirement changed to N₄GMAA (Harrington et al., 2017b).

In one embodiment, the present invention contemplates a plurality of *N. meningitidis* Cas9 orthologs with divergent PIDs that recognize different PAMs. In one embodiment, the present invention contemplates a Cas9 protein with a high sequence identity (>80% along their entire lengths) to that of NmeCas9 strain 8013 (Nme1Cas9) (Zhang et al., 2013). Nme1Cas9 also has a small size and naturally high accuracy as discussed above. (Lee et al., 2016; Amrani et al., 2018). Alignments revealed three clades of meningococcal Cas9 orthologs, each with >98% identity in the N-terminal ~820 amino acid (aa) residues, which includes all regions of the protein other than the PID. See, Fig. 7A and Fig. 8A.

All of these Cas9 orthologs are 1,078-1,082 aa in length. The first clade (group 1) includes orthologs in which the >98% aa sequence identity with Nme1Cas9 extends through the PID. In contrast, the other two groups had PIDs that were significantly diverged from that of Nme1Cas9, with group 2 and group 3 orthologs averaging ~52% and ~86% PID sequence identity with Nme1Cas9, respectively. One meningococcal strain was selected from each group: i) De11444 from group 2; and ii) 98002 from group 3 for detailed analysis, which are referred to herein as Nme2Cas9 (1,082 aa) and Nme3Cas9 (1,081 aa), respectively. The CRISPR-*cas* loci from these two strains have repeat sequences and spacer lengths that are identical to those of strain 8013. See, Fig. 8B. This strongly suggested that their mature crRNAs also have 24nt guide sequences and 24 nt repeat sequences (Zhang et al., 2013). Similarly, the tracrRNA sequences of De11444 and 98002 were 100% identical to the 8013 tracrRNA. See, Fig. 8B. These observations imply that the same sgRNA sequence scaffold can guide DNA cleavage by all three Cas9s.

To determine whether these Cas9 orthologs have distinct PAMs, the PID of Nme1Cas9 was replaced with that of either Nme2Cas9 or Nme3Cas9. To identify the corresponding PAM requirements, these protein chimeras were expressed in *Escherichia coli*, purified, and used for in vitro PAM identification (Karvelis et al., 2015; Ran et al., 2015; Kim et al., 2017). Briefly, a pool of DNA fragments containing a protospacer followed by a 10-nt randomized sequence was cleaved in vitro using recombinant Cas9 and a cognate, in vitro-transcribed sgRNA. See, Fig. 7B. Only those DNAs containing a Cas9 PAM sequence were expected to be cleaved. Cleavage products were then sequenced to identify the PAMs. See, Figs 7C-D.

The expected N₄GATT PAM consensus was validated in the recovered full-length Nme1Cas9. See, Fig. 7C. Chimeric PID-swapped derivatives exhibited a strong preference for a C residue in the 5th position in place of the G recognized by Nme1Cas9. See, Fig. 7D. Any remaining PAM nucleotides could not be confidently assigned due to the low cleavage efficiencies of the chimeric proteins under the conditions used. See, Fig. 8C. To further resolve the PAMs, in vitro assays were performed on a library with a 7-nt randomized sequence possessing an invariant C at the 5th PAM position (e.g., 5'-NNNNC(NNN)-3' on the sgRNA-noncomplementary strand). This strategy yielded a much higher cleavage efficiency and the results indicated that the Nme2Cas9 and Nme3Cas9 PIDs recognize NNNNCC(A) and NNNNCAAA PAMs, respectively. See, Figs 8C-D. The Nme3Cas9 consensus is similar to that of GeoCas9 (Harrington et al., 2017b).

These tests were repeated using a full-length Nme2Cas9 (rather than a PID-swapped chimera) with the NNNNCNNN DNA pool, and again a NNNNCC(A) consensus was recovered. See, Fig. 7E. It was noted that this test had more efficient cleavage. See, Fig. 8C. These data suggest that one or more of the 15 amino acid changes in Nme2Cas9 (relative to Nme1Cas9) outside of the PID support efficient DNA cleavage activity. See, Fig. 8C. Because the unique, 2-3 nt PAM of Nme2Cas9 affords a higher density of potential target sites than the previously described compact Cas9 orthologs, it was selected for further analyses.

To test the efficacy of Nme2Cas9 in human genome editing, a full-length (e.g., not PID-swapped) human-codon-optimized Nme2Cas9 construct was cloned into a mammalian expression plasmid with appended nuclear localization signals (NLSs) and linkers validated previously for Nme1Cas9 (Amrani et al., 2018). For initial tests, a modified, fluorescence-based Traffic Light Reporter (TLR2.0) was used (Certo et al., 2011). Briefly, a disrupted GFP is followed by an out-of-frame T2A peptide and mCherry cassette. When DNA double-strand breaks (DSBs) are introduced in the broken-GFP cassette, a subset of non-homologous end joining (NHEJ) repair events leave +1-frameshifted indels, placing mCherry in frame and yielding red fluorescence that can be easily quantified by flow cytometry See, Fig. 9A. Homology-directed repair (HDR) outcomes can also be scored simultaneously by including a DNA donor that restores the functional GFP sequence, yielding a green fluorescence (Certo et al., 2011). Because some indels do not introduce a +1 frameshift, the fluorescence readout generally provides an underestimate of the true editing efficiency. Nonetheless, the speed, simplicity, and low cost of the assay makes it useful as an initial, semi-quantitative measure of genome editing in HEK293T cells carrying a single TLR2.0 locus incorporated via lentivector.

For initial tests, Nme2Cas9 plasmid was transiently co-transfected with one of fifteen sgRNA plasmids carrying spacers that target TLR2.0 sites with N₄CC PAMs. No HDR donor was included, so only NHEJ-based editing (mCherry) was scored. Most sgRNAs were in a G23 format (i.e. a 5'-terminal G to facilitate transcription, followed by a 23nt guide sequence), as used routinely for Nme1Cas9 (Lee et al., 2016; Pawluk et al., 2016; Amrani et al., 2018; Ibraheim et al., 2018). No sgRNA and an sgRNA targeting an N₄GATT PAM were used as negative controls, and SpyCas9+sgRNA and Nme1Cas9+sgRNA co-transfections (targeting NGG and N₄GATT protospacers, respectively) were included as positive controls. Editing by SpyCas9 and Nme1Cas9 was readily detectable (~28% and 10% mCherry, respectively). See, Fig. 7B.

For Nme2Cas9, all 15 targets with N₄CC PAMs were functional, though to various extents ranging from 4% to 20% mCherry. These fifteen sites include examples with each of the four possible nucleotides in the 7th PAM position (e.g., after the CC dinucleotide), indicating that a slight preference for an A residue was observed in vitro (Fig. 7E) does not reflect a PAM requirement for editing applications in human cells. The N₄GATT PAM control yielded mCherry signal similar to no-sgRNA control. See, Fig. 9B.

To determine whether both C residues in the N₄CC PAM are involved in editing, a series of N₄DC (D = A, T, G) and N₄CD PAM sites were tested in TLR2.0 reporter cells. See, Figs 10A and 10B. No detectable editing was found at any of these sites, providing an initial indication that both C residues of the N₄CC PAM consensus are required for efficient Nme2Cas9 activity.

The length of the spacer in the crRNA differs among Cas9 orthologs and can affect on- vs. off-target activity (Cho et al., 2014; Fu et al., 2014). SpyCas9's optimal spacer length is 20 nts, with truncations down to 17 nts tolerated (Fu et al., 2014). In contrast, Nme1Cas9 usually has 24-nt spacers (Hou et al., 2013; Zhang et al., 2013), and tolerates truncations down to 18-20 nts (Lee et al., 2016; Amrani et al., 2018). To test spacer length requirements for Nme2Cas9, guide RNA plasmids were created for each targeted single TLR2.0 site, but with varying spacer lengths. See, Fig. 9C and Fig. 10C. Comparable activities were observed with G23, G22 and G21 guides, but significantly decreased activity upon further truncation to G20 and G19 lengths. See, Fig. 9C. These results validate Nme2Cas9 as a genome editing platform, with 22-24 nt guide sequences, at N₄CC PAM sites in cultured human cells.

B. HDR And HNH Cas9 Nickases

Cas9 enzymes use their HNH and RuvC domains to cleave the guide-complementary and non-complementary strand of the target DNA, respectively. Cas9 nickases (nCas9s), in which either the HNH or RuvC domain is mutationally inactivated, have been used to induce homology-directed repair (HDR) and to improve genome editing specificity via DSB induction by dual nickases (Mali et al., 2013a; Ran et al., 2013).

To test the efficacy of Nme2Cas9 as a nickase, a Nme2Cas9^{D16A} (HNH nickase) and Nme2Cas9^{H588A} (RuvC nickase) were created, which possess alanine mutations in catalytic residues of the RuvC and HNH domains, respectively (Esvelt et al., 2013; Hou et al., 2013; Zhang et al., 2013). TLR2.0 cells, along with a GFP donor dsDNA, were used to determine whether Nme2Cas9-induced nicks can induce precise edits via HDR. Target sites within TLR2.0 were

used to test the functionality of each nickase using guides targeting cleavage sites spaced 32 bp and 64 bp apart. See, Fig. 9D. Wildtype Nme2Cas9 targeting a single site showed efficient editing, with both NHEJ and HDR as outcomes of repair. For nickases, cleavage sites 32 bp and 64 bp apart showed editing using the Nme2Cas9^{D16A} (HNH nickase), but neither target pair worked with Nme2Cas9^{H588A}. These results suggest that Nme2Cas9 HNH nickase can be used for efficient genome editing, as long as the sites are in close proximity.

Studies in previously characterized Cas9s have identified a specific region proximal to the PAM where Cas9 activity is highly sensitive to sequence mismatches. This 8 to 12-nt region is known as the seed sequence and has been observed among all Cas9s characterized to date (Gorski et al., 2017). To determine whether Nme2Cas9 also possesses a seed sequence, a series of transient transfections was performed, each targeting the same locus in TLR2.0, but with a single-nucleotide mismatch at different positions of the guide. See, Fig. 10D. A significant decrease in the number of mCherry-positive cells was observed for mismatches in the first 10-12 nts proximal to the PAM, suggesting that Nme2Cas9 possesses a seed sequence in this region.

C. Cas9 Plasmid/Vector Cell Transfection

Nme2Cas9's ability to function in different mammalian cell lines was tested using various delivery methods. As an initial test, forty (40) different sites (29 with a N₄CC PAM, and 11 sites were tested with a N₄CD PAM). Several loci were selected (*AAVS1*, *VEGFA*, etc.), and target sites with N₄CC PAMs were randomly chosen for editing with Nme2Cas9. Editing (%) was determined by transiently transfecting 150 ng of Nme2Cas9 along with 150 ng of sgRNA plasmids followed by TIDE analysis 72 hours post-transfection. A subset of target sites and their respective TIDE primer sets exhibiting a range of editing efficiencies in this initial screen was selected for repeat analyses in triplicate. See, Fig. 11A; Table 1 and Table 2.

Table 1. Exemplary Endogenous human genome editing sites targeted by Nme2Cas9 (bolded nts).

No.	Site Name	Spacer Seq	PAM	Locus	Editing (%)
1	TS1	GGTTCTGGGTACTTTTATCTGTCC	CCTCCACC	<i>AAVS1</i>	ND
2	TS4	GTCTGCCTAACAGGAGGTGGGGGT	TAGACGAA	<i>AAVS1</i>	11
3	TS5	GAATATCAGGAGACTAGGAAGGAG	GAGGCCTA	<i>AAVS1</i>	15
4	TS6	GCCTCCCTGCAGGGCTGCTCCC	CAGCCCAA	<i>LINC01588</i>	20
5	TS10	GAGCTAGTCTTCTTCTCCAACCC	GGGCCCTA	<i>AAVS1</i>	3.5
6	TS11	GATCTGTCCCCTCCACCCACAGT	GGGGCCAC	<i>AAVS1</i>	9

7	TS12	GGCCCAAATGAAAGGAGTGAGAGG	TGACCCGA	<i>AAVSI</i>	10
8	TS13	GCATCCTCTTGCTTTCTTTGCCTG	GACACCCCA	<i>AVSI</i>	2
9	TS16	GGAGTCGCCAGAGGCCGGTGGTGG	ATTTCCCTC	<i>LINC01588</i>	28
10	TS17	GCCCAGCGGCCGGATATCAGCTGC	CACGCCCG	<i>LINC01588</i>	ND
11	TS18	GGAAGGGAACATATTACTATTGC	TTTCCCTC	<i>CYBB</i>	1
12	TS19	GTGGAGTGCCCTGCTATCAGCTAC	CTATCCAA	<i>CYBB</i>	6
13	TS20	GAGGAAGGGAACATATTACTATTG	CTTTCCCT	<i>CYBB</i>	11.2
14	TS21	GTGAATTCTCATCAGCTAAAATGC	CAAGCCTT	<i>CYBB</i>	1
15	TS25	GCTCACTCACCCACACAGACACAC	ACGTCCCTC	<i>VEGFA</i>	15.6
16	TS26	GGAAGAATTTCAATTCTGTTCTCAG	TTTTCCCTG	<i>CFTR</i>	2
17	TS27	GCTCAGTTTTCTGGATTATGCCT	GGCACCAT	<i>CFTR</i>	4
18	TS31	GCGTTGGAGCGGGGAGAAGGCCAG	GGGTCACT	<i>VEGFA</i>	9
19	TS34	GGGCCGCGGAGATAGCTGCAGGGC	GGGGCCCC	<i>LINC01588</i>	ND
20	TS35	GCCCACCCGCGCGGCCCTCCCTGC	AGGGCTGC	<i>LINC01588</i>	ND
21	TS36	GCGTGCGAGCTGATATCCGGCCGC	TGGGCGTC	<i>LINC01588</i>	ND
22	TS37	GCCGCGGCGCGACGTGGAGCCAGC	CCCGCAA	<i>LINC01588</i>	ND
23	TS38	GTGCTCCCCAGCCCAAACCGCCGC	GGCGCGAC	<i>LINC01588</i>	2
24	TS41	GTCAGATTGGCTTGCTCGGAATTG	CCAGCCAA	<i>AGA</i>	3
25	TS44	GCTGGGTGAATGGAGCGAGCAGCG	TCTTCGAG	<i>VEGFA</i>	3
26	TS45	GTCTGGAGTGACCCCTGGCCTTC	TCCCCGCT	<i>VEGFA</i>	7.4
27	TS46	GATCCTGGAGTGACCCCTGGCCTT	CTCCCCGC	<i>VEGFA</i>	6
28	TS47	GTGTGTCCCTCTCCCCACCCGTCC	CTGTCCGG	<i>VEGFA</i>	23.1
29	TS48	GTTGGAGCGGGGAGAAGGCCAGGG	GTCCTCC	<i>VEGFA</i>	2
30	TS49	GCGTTGGAGCGGGGAGAAGGCCAG	GGGTCACT	<i>VEGFA</i>	4
31	TS50	GTACCCTCCAATAATTGGCTGGC	AATTCCGA	<i>AGA</i>	6
32	TS51	GATAATTTGGCTGGCAATTCCGAG	CAAGCCAA	<i>AGA</i>	4.5
33	TS58 (DS1)	GCAGGGGCCAGGTGTCCTTCTCTG	GGGGCCTC	<i>VEGFA</i>	5
34	TS59 (DS2)	GAATGGCAGGCGGAGGTTGTA CTG	GGGGCCAG	<i>VEGFA</i>	11.5
35	TS60 (DS3)	GAGTGAGAGAGTGAGAGAGAGAC A	CGGGCCAG	<i>VEGFA</i>	3
36	TS61 (DS4)	GTGAGCAGGCACCTGTGCCAACAT	GGGCCCCG	<i>VEGFA</i>	3.5
37	TS62 (DS5)	GCGTGGGGGCTCCGTGCCCCACGC	GGGTCCAT	<i>VEGFA</i>	3.4
38	TS63 (DS6)	GCATGGGCAGGGGCTGGGGTGAC	AGGCCCAG	<i>VEGFA</i>	16
39	TS64	GAAAATTGTGATTTCCAGATCCAC	AAGCCCAA	<i>FANCI</i>	7
40	TS65	GAGCAGAAAAAATTGTGATTTCC	AGATCCAC	<i>FANCI</i>	ND

Table 2: Exemplary TIDE Primers

No.	Target Site	Primer name	Forward primer	Reverse primer
1	TS1	AAVS1_TIDE1	TGGCTTAGCACCTCT CCAT	AGAACTCAGGACCAACT TATTCTG
2	TS4	AAVS1_TIDE1	TGGCTTAGCACCTCT CCAT	AGAACTCAGGACCAACT TATTCTG
3	TS5	AAVS1_TIDE1	TGGCTTAGCACCTCT CCAT	AGAACTCAGGACCAACT TATTCTG
4	TS6	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
5	TS10	AAVS1_TIDE1	TGGCTTAGCACCTCT CCAT	AGAACTCAGGACCAACT TATTCTG
6	TS11	AAVS1_TIDE1	TGGCTTAGCACCTCT CCAT	AGAACTCAGGACCAACT TATTCTG
7	TS12	AAVS1_TIDE2	TCCGTCTTCCTCCAC TCC	TAGGAAGGAGGAGGCCT AAG
8	TS13	AAVS1_TIDE2	TCCGTCTTCCTCCAC TCC	TAGGAAGGAGGAGGCCT AAG
9	TS16	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
10	TS17	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
11	TS18	NTS55_TIDE	TAGAGAACTGGGTA GTGTG	CCAATATTGCATGGGAT GG
12	TS19	NTS55_TIDE	TAGAGAACTGGGTA GTGTG	CCAATATTGCATGGGAT GG
13	TS20	NTS55_TIDE	TAGAGAACTGGGTA GTGTG	CCAATATTGCATGGGAT GG
14	TS21	NTS55_TIDE	TAGAGAACTGGGTA GTGTG	CCAATATTGCATGGGAT GG
15	TS25	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
16	TS26	hCFTR_TIDE1	TGGTGATTATGGGAG AACTGGAGC	ACCATTGAGGACGTTTGT CTCAC
17	TS27	hCFTR_TIDE1	TGGTGATTATGGGAG AACTGGAGC	ACCATTGAGGACGTTTGT CTCAC
18	TS31	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
19	TS34	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
20	TS35	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
21	TS36	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA

22	TS37	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
23	TS38	LINC01588_TIDE	AGAGGAGCCTTCTGA CTGCTGCAGA	ATGACAGACACAACCAG AGGGCA
24	TS41	AGA_TIDE1	GGCATAAGGAAATC GAAGGTC	CATGTCCTCAAGTCAAG AACAAG
25	TS44	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
26	TS45	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
27	TS46	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
28	TS47	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
29	TS48	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
30	TS49	VEGF_TIDE3	GTACATGAAGCAACT CCAGTCCCA	ATCAAATTCCAGCACCG AGCGC
31	TS50	AGA_TIDE1	GGCATAAGGAAATC GAAGGTC	CATGTCCTCAAGTCAAG AACAAG
32	TS51	AGA_TIDE1	GGCATAAGGAAATC GAAGGTC	CATGTCCTCAAGTCAAG AACAAG
33	TS58 (DS1)	VEGF_TIDE4	ACACGGGCAGCATG GGAATAGTC	GCTAGGGGAGAGTCCCA CTGTCCA
34	TS59 (DS2)	VEGF_TIDE5	CCTGTGTGGCTTTGC TTTGGTC	GGTAGGGTGTGATGGGA GGCTAAGC
35	TS60 (DS3)	VEGF_TIDE5	CCTGTGTGGCTTTGC TTTGGTC	GGTAGGGTGTGATGGGA GGCTAAGC
36	TS61 (DS4)	VEGF_TIDE5	CCTGTGTGGCTTTGC TTTGGTC	GGTAGGGTGTGATGGGA GGCTAAGC
37	TS62 (DS5)	VEGF_TIDE6	GGAGGAAGAGTAGC TCGCCGAGG	AGACCGAGTGGCAGTGA CAGCAAG
38	TS63 (DS6)	VEGF_TIDE7	AGGGAGAGGGAAGT GTGGGGAAGG	GTCTTCCTGCTCTGTGCG CACGAC
39	TS64	FancJ_TIDE5	GTTGGGGGCTCTAAG TTATGTAT	CTTCATCTGTATCTTCAG GATCA
40	TS65	FancJ_TIDE5	GTTGGGGGCTCTAAG TTATGTAT	CTTCATCTGTATCTTCAG GATCA

HEK293T cells were used to support transient transfections and at 72-hours post transfection the, cells were harvested, followed by genomic DNA extraction and selective amplification of the targeted locus. TIDE analysis was used to measure indel efficiency at each locus (Brinkman et al., 2014). Nme2Cas9 editing was detectable at most of these sites, even though efficiencies varied

depending on the target sequence. Interestingly, Nme2Cas9 induced indels at several genomic sites with N₄CD PAMs, albeit less consistently and at lower levels. Table 1. Fourteen (14) sites with N₄CC PAMs were analyzed in triplicate, and consistent editing was observed. See, Fig. 11A. In addition, editing efficiency could be improved significantly by increasing the quantity of the Nme2Cas9 plasmid delivered, and this high efficiency could be extended to precise segmental deletion with two guides. See, Fig.s 12A and 12B.

The ability of Nme2Cas9 to function was tested in mouse Hepa1-6 cells (hepatoma-derived). For Hepa1-6 cells, a single plasmid encoding both Nme2Cas9 and an sgRNA (targeting either *Rosa26* or *Pcsk9*) was transiently transfected and indels were measured after 72 hrs. Editing was readily observed at both sites. See, Fig. 11B, left. Nme2Cas9's functionality was also tested when stably expressed in human leukemia K562 cells. To this end, a lentiviral construct was created expressing Nme2Cas9 and transduced cells to stably express Nme2Cas9 under the control of the SFFV promoter. This stable cell line did not show any visible differences with respect to growth and morphology in comparison to untransduced cells, suggesting that Nme2Cas9 is not toxic when stably expressed. These cells were transiently electroporated with plasmids expressing sgRNAs and analyzed by TIDE after 72 hours to measure indel efficiencies. Efficient (>50%) editing was observed at all three sites tested, validating Nme2Cas9's ability to function upon lentiviral delivery in K562 cells. See, Fig. 11B.

Ribonucleoprotein (RNP) delivery of Cas9 and its sgRNA is also useful for some genome editing applications, and the greater transience of Cas9's presence can minimize off-target editing (Kim et al., 2014; Zuris et al., 2015). Moreover, some cell types (e.g. certain immune cells) are recalcitrant to DNA transfection-based editing (Schumann et al., 2015). To test whether Nme2Cas9 is functional by RNP delivery, a 6xHis-tagged Nme2Cas9 (fused to three NLSs) was cloned into a bacterial expression construct and the recombinant protein was purified. The recombinant protein was then loaded with T7 RNA polymerase-transcribed sgRNAs targeting three previously validated sites. Electroporation of the Nme2Cas9:sgRNA complex induced successful editing at each of the three target sites in HEK293T cells, as detected by TIDE. See, Fig. 11C. Collectively these results indicate that Nme2Cas9 can be delivered effectively via plasmid or lentivirus, or as an RNP complex, in multiple cell types.

D. Anti-CRISPR (Acr) Regulation

To date, five families of Acr proteins from diverse bacterial species have been shown to inhibit Nme1Cas9 in vitro and in human cells (Pawluk et al., 2016; Lee et al., 2018, submitted). Considering the high sequence identity between Nme1Cas9 and Nme2Cas9, at least some of these Acr families should inhibit Nme2Cas9. To test this, all five families of recombinant Acrs were expressed, purified and tested for Nme2Cas9's ability to cleave a target in vitro in the presence of a member of each family (10:1 Acr:Cas9 molar ratio). An inhibitor was used for the type I-E CRISPR system in *E. coli* (AcrE2) as a negative control, while Nme1Cas9 was used as a positive control. (Pawluk et al., 2014); (Pawluk et al., 2016). As expected, all 5 families inhibited Nme1Cas9, while AcrE2 failed to do so. See, Fig. 13A, top. AcrIIC1_{Nme}, AcrIIC2_{Nme}, AcrIIC3_{Nme}, and AcrIIC4_{Hpa} completely inhibited Nme2Cas9. Strikingly, however, AcrIIC5_{Smu} which has been previously reported as the most potent of the Nme1Cas9 inhibitors (Lee et al., 2018), did not inhibit Nme2Cas9 in vitro even at a 10-fold molar excess. This suggests that it likely inhibits Nme1Cas9 by interacting with its PID.

To further test this, a Nme1Cas9/Nme2Cas9 chimera with the PID of Nme2Cas9 was tested. See, Fig. 7D and Fig. 8D. Due to the reduced activity of this hybrid, a ~30x higher concentration of Cas9 was used to achieve a similar cleavage efficiency while maintaining the 10:1 Cas9:Acr molar ratio. No inhibition was observed by AcrIIC5_{Smu} on this protein chimera. See, Fig. 14. This data provides further evidence that AcrIIC5_{Smu} likely interacts with the PID of Nme1Cas9. Regardless of the mechanistic basis for the differential inhibition by AcrIIC5_{Smu}, these results indicate that Nme2Cas9 is subject to inhibition by the other four type II-C Acr families.

Based on the above in vitro data, it was hypothesized that AcrIIC1_{Nme}, AcrIIC2_{Nme}, AcrIIC3_{Nme}, and AcrIIC4_{Hpa} could be used as off-switches for Nme2Cas9 genome editing. To test this, Nme2Cas9/sgRNA plasmid transfections (150 ng of each plasmid) targeting TS16 were performed in HEK293T cells in the presence or absence of Acr expression plasmids, as it has been reported that most Acrs inhibited Nme1Cas9 at those plasmid ratios (Pawluk et al., 2016). As expected, AcrIIC1_{Nme}, AcrIIC2_{Nme}, AcrIIC3_{Nme} and AcrIIC4_{Hpa} inhibited Nme2Cas9 genome editing, while AcrIIC5_{Smu} had no effect. See, Fig. 13B. Complete inhibition was observed by AcrIIC3_{Nme} and AcrIIC4_{Hpa}, suggesting that they have high potency against Nme2Cas9 as compared to AcrIIC1_{Nme} and AcrIIC2_{Nme}. To further compare the potency of AcrIIC1_{Nme} and AcrIIC4_{Hpa}, we repeated the experiments at various ratios of Acr plasmid to Cas9 plasmid. See,

Fig. 13C. The data show that the AcrIIC4_{Hpa} plasmid is especially potent against Nme2Cas9. Together, these data suggest that several Acr proteins can be used as off-switches for Nme2Cas9-based applications.

E. NmeCas9 Gene Editing Efficiency

Nme1Cas9 demonstrates remarkable editing fidelity in cells and mouse models (Lee et al., 2016; Amrani et al., 2018; Ibraheim et al., 2018). Furthermore, the similarity of Nme2Cas9 to Nme1Cas9 over most of its length suggests that it may likewise be hyper-accurate. However, the higher number of sites sampled in the genome as a result of the CC dinucleotide PAM could create more opportunities for Nme2Cas9 off-targeting in comparison with Nme1Cas9 and its less frequently encountered 4-nucleotide PAM.

To assess the off-target profile of Nme2Cas9, genome-wide, unbiased identification of double-stranded breaks enabled by sequencing (GUIDE-seq) was used to identify potential off-target sites empirically and in an unbiased fashion (Tsai et al., 2014). Even the best off-target prediction algorithms are prone to false negatives necessitating empirical target site profiling methods (Bolukbasi et al., 2015b; Tsai and Joung, 2016; Tycko et al., 2016). GUIDE-seq relies on the incorporation of double-stranded oligodeoxynucleotides (dsODNs) into DNA double-stranded break sites throughout the genome. These insertion sites are then detected by amplification and high-throughput sequencing.

Because SpyCas9 is a well-characterized Cas9 ortholog it is useful for multiplexed applications with other Cas9s, and as a benchmark for their editing properties (Jiang and Doudna, 2017; Komor et al., 2017). SpyCas9 and Nme2Cas9 were cloned into identical plasmid backbones, with the same UTRs, linkers, NLSSs, and promoters, for parallel transient transfections (along with similarly matched sgRNA-expressing plasmids) into HEK293T cells. First, it was confirmed that the RNA guides for SpyCas9 and Nme2Cas9 are orthogonal, i.e. that Nme2Cas9 sgRNAs do not direct editing by SpyCas9, and vice versa. See, Fig. 15A. This was in contrast to earlier reported results with Nme1Cas9 (Esvelt et al., 2013; Fonfara et al., 2014).

Next, to identify a use of SpyCas9 as a benchmark for GUIDE-seq, because SpyCas9 and Nme2Cas9 have non-overlapping PAMs its can therefore potentially edit any dual site (DS) flanked by a 5'-NGGNCC-3' sequence, which simultaneously fulfills the PAM requirements of both Cas9's. This permits side-by-side comparisons of off-targeting with RNA guides that facilitate an edit of the exact same on-target site. See, Fig. 16A. Six (6) DSs in *VEGFA* were

targeted, each of which also has a G at the appropriate positions 5' of the PAM such that both SpyCas9 and Nme2Cas9 guides (driven by the U6 promoter) were 100% complementary to the target site. Seventy-two (72) hours after transfection, a TIDE analysis was performed on these sites targeted by each nuclease. Nme2Cas9 induced indels at all six sites, albeit at low efficiencies at two of them, while SpyCas9 induced indels at four of the six sites. See, Fig. 16B. At two of the four sites (DS1 and DS4) at which SpyCas9 was effective, it induced ~7-fold more indels than Nme2Cas9, while Nme2Cas9 induced a ~3-fold higher frequency of indels than SpyCas9 at DS6. Both Cas9 orthologs edited DS2 with approximately equal efficiency.

For GUIDE-seq, the DS2, DS4 and DS6 target sites were selected to sample off-target cleavage with Nme2Cas9 guides that direct on-target editing as efficiently, less efficiently, or more efficiently than the corresponding SpyCas9 guides, respectively. In addition to the three dual sites, TS6 was added as it has been observed to be an efficiently edited Nme2Cas9 target sites, having an approximate 30-50% indel efficiency depending on the cell type. See, Figs 11A and 12A. Similar data is seen with the mouse *Pcsk9* and *Rosa26* Nme2Cas9 sites. See, Fig. 11B.

Plasmid transfections were performed for SpyCas9 and Nme2Cas9 along with their cognate sgRNAs and the dsODNs. Subsequently, GUIDE-seq libraries were prepared as described previously (Amrani et al., 2018). A GUIDE-seq analysis revealed efficient on-target editing for both Cas9 orthologs, with relative efficiencies (as reflected by GUIDE-seq read counts) that are similar to those observed by TIDE. Fig. 15A-E; and Tsai et al., 2014; Zhu et al., 2017. For off-target identification, the analysis revealed that the DS2, DS4, and DS6 SpyCas9 sgRNAs appeared to direct editing at 93, 10, and 118 candidate off-target sites, respectively, in the normal range of off-targets when plasmid-based SpyCas9 editing is analyzed by GUIDE-seq (Fu et al., 2014; Tsai et al., 2014). In striking contrast, the DS2, DS4, and DS6 Nme2Cas9 sgRNAs appeared to direct editing at 1, 0, and 1 off-target sites, respectively. Fig. 16C. When compared to the GUIDE-seq read counts for the SpyCas9 off-targets, those of Nme2Cas9 were very low, further suggesting that Nme2Cas9 is highly specific. Fig. 15C *cf.* Fig. 15D. Nme2Cas9 GUIDE-seq analyses with the TS6, *Pcsk9*, and *Rosa26* yielded similar results (0, 0, and 1 off-target sites, respectively, with a modest read count for the *Rosa26*-OT1 off-target site). Fig. 15C, and Fig. 16D.

To validate the off-target sites detected by GUIDE-seq, a targeted deep sequencing was performed to measure indel formation at the top off-target loci following GUIDE-seq-independent editing (i.e. without co-transfection of the dsODN). While SpyCas9 showed considerable editing at most off-target sites tested and, in some instances, was more efficient than that at the corresponding on-target site, Nme2Cas9 exhibited no detectable indels at the lone DS2 and DS6 candidate off-target sites. See, Fig. 16D. With the *Rosa26* sgRNA, Nme2Cas9 induced ~1% editing at the *Rosa26*-OT1 site in Hepa1-6 cells, compared to ~30% on-target editing. See, Fig. 16D. It is noteworthy that this off-target site has a consensus Nme2Cas9 PAM (ACTCCCT) with only 3 mismatches at the PAM-distal end of the guide-complementary region (i.e. outside of the seed). See, Fig. 16E. These data support and reinforce our GUIDE-seq results indicating a high degree of accuracy for Nme2Cas9 genome editing in mammalian cells.

To further corroborate the above GUIDE-Seq results, CRISPRseek was used to computationally predict potential off-target sites for two active Nme2Cas9 sgRNAs that targeted TS25 and TS47, both of which are also in *VEGFA*. See, Fig. 11A; (Zhu et al., 2014). Three (TS25) or four (TS47) of the most closely matched predicted sites, five with N₄CC PAMs and two with N₄CA PAMs; each had 2-5 mismatches, mostly in their PAM-distal, non-seed regions. See, Fig. 15E. On- vs. off-target editing was compared after Nme2Cas9+sgRNA plasmid transfections into HEK293T cells by targeted amplification of each locus, followed by TIDE analysis. Consistently, no indels could be detected at those off-target sites for either sgRNA by TIDE, while efficient on-target editing was readily detected in DNA from the same populations of cells. Taken together, our data indicate that Nme2Cas9 is a naturally hyper-accurate genome editing platform in mammalian cells.

F. Adeno-Associated Virus Nme2Cas9 Delivery

Clustered, regularly interspaced, short, palindromic repeats (CRISPR) along with CRISPR-associated (Cas) proteins constitute bacterial and archaeal adaptive immune pathways against phages and other mobile genetic elements (MGEs) (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008). In Type II CRISPR systems, CRISPR RNA (crRNA) is bound to a trans-activating crRNA (tracrRNA) and loaded onto a Cas9 effector protein that cleaves MGE nucleic acids complementary to the crRNA (Garneau et al., 2010; Deltcheva et al., 2011; Sapranaukas et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012). The crRNA:tracrRNA hybrid

can be fused into a single-guide RNA (sgRNA) (Jinek et al., 2012). The RNA programmability of Cas9 endonucleases has made it a powerful genome editing platform in biotechnology and medicine (Cho et al., 2013; Cong et al., 2013; Hwang et al., 2013; Jiang et al., 2013; Jinek et al., 2013; Mali et al., 2013b).

In addition to sgRNA, Cas9 target recognition is usually associated with a 1-5 nucleotide signature downstream of the complementary DNA sequence, called a protospacer adjacent motif (PAM) (Deveau et al., 2008; Mojica et al., 2009). Cas9 orthologs exhibit considerable diversity in PAM length and sequence. Among Cas9 orthologs that have been characterized, *Streptococcus pyogenes* Cas9 (SpyCas9) is the most widely used, in part because it recognizes a short NGG PAM (Jinek et al., 2012) (N represents any nucleotide) that affords a high density of targetable sites. Nevertheless, Spy's relatively large size (i.e., 1,368 amino acids) makes this Cas9 difficult to package (along with sgRNA and promoters) into a single recombinant adeno-associated virus (rAAV). This has been shown to be a drawback for therapeutic applications given the promise shown by AAV vectors for in vivo gene delivery (Keeler et al., 2017). Moreover, SpyCas9 and its RNA guides have required extensive characterization and engineering to minimize the tendency to edit near-cognate, off-target sites. (Bolukbasi et al., 2015b; Tsai and Joung, 2016; Tycko et al., 2016; Chen et al., 2017; Casini et al., 2018; Yin et al., 2018). To date, subsequent engineering efforts have not overcome these size limitations.

Several Cas9 orthologs of less than 1,100 amino acids in length obtained from diverse species have been validated for mammalian genome editing, including strains of *N. meningitidis* (NmeCas9, 1,082 aa) (Esvelt et al., 2013; Hou et al., 2013), *Staphylococcus aureus* (SauCas9, 1,053 aa) (Ran et al., 2015), *Campylobacter jejuni* (CjeCas9, 984 aa) (Kim et al., 2017), and *Geobacillus stearothermophilus* (GeoCas9, 1,089 aa) (Harrington et al., 2017b). NmeCas9, CjeCas9, and GeoCas9 are representatives of type II-C Cas9s (Mir et al., 2018), most of which are <1,100 aa. With the exception of GeoCas9, each of these shorter sequence orthologs has been successfully deployed for in vivo editing via all-in-one AAV delivery (in which a single vector expresses both guide and effector) (Ran et al., 2015; Kim et al., 2017; Ibraheim et al., 2018, submitted). Furthermore, NmeCas9 and CjeCas9 have been shown to be naturally resistant to off-target editing (Lee et al., 2016; Kim et al., 2017; Amrani et al., 2018, submitted). However, the PAMs that are recognized by compact Cas9s are usually longer than that of SpyCas9, substantially reducing the number of targetable sites at or near a given locus; for example, i)

N₄GAYW/N₄GYTT/N₄GTCT for NmeCas9 (Esvelt et al., 2013; Hou et al., 2013; Lee et al., 2016; Amrani et al., 2018); ii) N₂GRRT for SauCas9 (Ran et al., 2015); iii) N₄RYAC for CjeCas9 (Kim et al., 2017); and iv) N₄CRAA/N₄GMAA for GeoCas9s (Harrington et al., 2017b) (Y = C, T; R = A, G; M = A, C; W = A, T). A smaller subset of target sites is advantageous for highly accurate and precise gene editing tasks including, but not limited to: i) editing of small targets (e.g. miRNAs); ii) correction of mutations by base editing which alters a very narrow window of bases relative to the PAM (Komor et al., 2016; Gaudelli et al., 2017); or iii) precise editing via homology-directed repair (HDR) which is most efficient when the rewritten bases are close to the cleavage site (Gallagher and Haber, 2018). Because of PAM restrictions, many editing sites cannot be targeted using all-in-one AAV vectors for in vivo delivery even with these shorter Cas9 proteins. For example, A SauCas9 mutant (SauCas9^{KKH}) has been developed that has reduced PAM constraints (N₃RRT), though this increase in targeting range often comes at the cost of reduced on-target editing efficacy, and off-target edits are still observed. (Kleinstiver et al., 2015).

Safe and effective CRISPR-based therapeutic gene editing will be greatly enhanced by Cas9 orthologs and variants that are highly active in human cells, resistant to off-targeting, sufficiently compact for all-in-one AAV delivery, and capable of accessing a high density of genomic sites. In one embodiment, the present invention contemplates a compact, hyper-accurate Cas9 (Nme2Cas9) from a distinct strain of *N. meningitidis*. In one embodiment, the present invention contemplates a method for single-AAV delivery of Nme2Cas9 and its sgRNA to perform efficient genome editing in vivo and/or ex vivo. Although it is not necessary to understand the mechanism of an invention, it is believed that this ortholog functions efficiently in mammalian cells and recognizes an N₄CC PAM that affords a target site density identical to that of wild-type SpyCas9 (e.g., every 8 bp on average, when both DNA strands are considered).

The compact size, small PAM, and high fidelity of Nme2Cas9 offer major advantages for in vivo genome editing using Adeno-Associated Virus (AAV) delivery. To test whether effective Nme2Cas9 genome editing can be achieved via single-AAV delivery, Nme2Cas9 was cloned with its sgRNA and their promoters (U1a and U6, respectively) into an AAV vector backbone. See, Fig. 17A. An all-in-one AAV was prepared with an sgRNA-Nme2Cas9 packaged into a hepatotropic AAV8 capsid to target two genes in the mouse liver: i) *Rosa26* (a commonly used safe harbor locus for transgene insertion) (Friedrich and Soriano, 1991) as a

negative control; and ii) *Pcsk9*, a major regulator of circulating cholesterol homeostasis (Rashid et al., 2005), as a phenotypic target.

SauCas9- or Nme1Cas9-induced indels in *Pcsk9* in the mouse liver results and reduced cholesterol levels providing a useful and easy-to-score in vivo benchmark for new editing platforms (Ran et al., 2015; Ibraheim et al., 2018). The Nme2Cas9 RNA guides were the same as those used above. See, Fig. 11B, Fig. 15D, and Fig. 16. As *Rosa26*-OT1 was the only Nme2Cas9 off-target site that has been validated in cultured mammalian cells, the *Rosa26* guide also provided us with an opportunity to assess on- vs. off-target editing in vivo. See, Figs 16D-E. The tail veins of two groups of mice ($n = 5$) were injected with 4×10^{11} AAV8.sgRNA.Nme2Cas9 genome copies (GCs) targeting either *Pcsk9* or *Rosa26*. Serum was collected at 0, 14 and 28 days post-injection for cholesterol level measurement. Mice were sacrificed at 28 days post-injection and liver tissues were harvested. See, Fig. 17A. Targeted deep sequencing of each locus revealed ~38% and ~46% indel induction at the *Pcsk9* and *Rosa26* editing sites, respectively, in the liver. See, Fig. 17B. Because hepatocytes constitute only 65-70% of total cellular content in the adult liver, Nme2Cas9 AAV-induced hepatocyte editing efficiencies with sg*Pcsk9* and sg*Rosa* were approximately 54-58% and 66-71%, respectively (Racanelli and Rehmann, 2006).

Only 2.25% liver indels overall (~3-3.5% in hepatocytes) were detected at the *Rosa26*-OT1 off-target site, comparable to the 1% editing that we observed at this site in transfected Hepal-6 cells. Fig. 17B *cf* Fig. 16D. At both 14 and 28 days post-injection, *Pcsk9* editing was accompanied by a ~44% reduction in serum cholesterol levels, whereas mice treated with the sg*Rosa26*-expressing AAV maintained normal level of cholesterol throughout the study. See, Fig. 17C. The ~44% reduction in serum cholesterol in the Nme2Cas9/sg*Pcsk9* AAV-treated mice compares well with the ~40% reduction reported with SauCas9 all-in-one AAV when targeting the same gene (Ran et al., 2015).

Western blotting was performed using an anti-PCSK9 antibody to estimate PCSK9 protein levels in the livers of mice treated with sg*Pcsk9* and sg*Rosa26*. Liver PCSK9 was below the detection limit in mice treated with sg*Pcsk9*, whereas sg*Rosa26*-treated mice exhibited normal levels of PCSK9. See, Fig. 18A. Hematoxylin and eosin (H&E) staining and histology revealed no signs of toxicity or tissue damage in either group after Nme2Cas9 expression. See,

Fig. 18B. These data validate Nme2Cas9 as a highly effective genome editing system in vivo, including when delivered by single-AAV vectors.

AAV vectors have recently been used for the generation of genome-edited mice, without the need for microinjection or electroporation, simply by soaking the zygotes in culture medium containing AAV vector(s), followed by reimplantation into pseudopregnant females (Yoon et al., 2018). Editing was obtained previously with a dual-AAV system in which SpyCas9 and its sgRNA were delivered in separate vectors (Yoon et al., 2018). To test whether Nme2Cas9 could perform accurate and efficient editing in mouse zygotes with an all-in-one AAV delivery system, *Tyrosinase* (*Tyr*) was targeted. A bi-allelic inactivation of *Tyr* disrupts melanin production resulting in an albino phenotype (Yokoyama et al., 1990).

An efficient *Tyr* sgRNA was validated that cleaves the *Tyr* locus only seventeen (17) bp from the site of the classic *albino* mutation in Hepa1-6 cells by transient transfections. See, Fig. 19A. Next, C57BL/6NJ zygotes were incubated for 5-6 hours in culture medium containing 3×10^9 or 3×10^8 GCs of an all-in-one AAV6 vector expressing Nme2Cas9 along with the *Tyr* sgRNA. After overnight culture in fresh media, those zygotes that advanced to the two-cell stage were transferred to the oviduct of pseudopregnant recipients and allowed to develop to term. See, Fig. 20A. Coat color analysis of pups revealed mice that were albino, chinchilla (indicating a hypomorphic allele of *Tyrosinase*), or that had variegated coat color composed of albino and chinchilla spots but lacking black pigmentation. See, Figs 19B-C. These results suggest a high frequency of biallelic mutations since the presence of a wild-type *Tyrosinase* allele should render black pigmentation. A total of five pups (10%) were born from the 3×10^9 GCs experiment. All of them carried indels; phenotypically, two were albino, one was chinchilla, and two had variegated pigmentation, indicating mosaicism.

From the 3×10^8 GCs experiment, four (4) pups (14%) were obtained, two of which died at birth, preventing a coat color or genome analysis. Coat color analysis of the remaining two pups revealed one chinchilla and one mosaic pup. These results indicate that single-AAV delivery of Nme2Cas9 and its guide can be used to generate mutations in mouse zygotes without microinjection or electroporation.

To measure on-target indel formation in the *Tyr* gene, DNA was isolated from the tails of each mouse, the locus was amplified and upon which a TIDE analysis was performed. All mice had high levels of on-target editing by Nme2Cas9, varying from 84% to 100%. See, Figs

19B-C. Most lesions in albino mouse 9-1 were either a 1- or a 4-bp deletion, suggesting either mosaicism or trans-heterozygosity, but albino mouse 9-2 exhibited a uniform 2-bp deletion. See, Fig. 19C.

The data is inconclusive as to whether there was no mosaicism in mouse 9-2, or that additional alleles were absent from mouse 9-1, because only tail samples were sequenced and other tissues could have distinct lesions. Analysis of tail DNA from chinchilla mice revealed the presence of in-frame mutations that are potentially the cause of the chinchilla coat color. The limited mutational complexity suggests that editing occurred early during embryonic development in these mice. These results provide a streamlined route toward mammalian mutagenesis through the application of a single AAV vector, in this case delivering both Nme2Cas9 and its sgRNA.

G. Cas9 Nucleases As Base Editors

Point mutations represent the largest class of known human pathogenic genetic variants. Base editors (BEs), which comprise a single-guide RNA (sgRNA) loaded onto a Cas9 nickase fused to a deaminase enzyme, enables precise installation of A•T to G•C substitutions, in the case of adenine base editor (ABE), or C•G to T•A substitutions, in the case of cytidine base editor (CBE). In contrast to traditional nuclease-dependent genome editing approaches, base editors do not generate double-stranded DNA breaks (DSBs), do not require a DNA donor template, and are more efficient in editing non-dividing cells, making them attractive agents for in vivo therapeutic genome editing.

While robust editing has been achieved in many cultured mammalian cell systems, safe and effective in vivo delivery of the base editors remains a major challenge. To date, both non-viral and viral delivery methods have been reported to deliver base editors for in vivo therapeutic purposes in rodents and primates which hold great promise. For example, in vivo delivery using adeno-associated virus (AAV) has achieved efficient editing in a wide range of tissue and cell types including liver, heart, muscle, retina, and CNS. However, the large coding sizes (5.2 kb) of most well-characterized *Streptococcus pyogenes* Cas9 (SpCas9) containing BEs well exceed the packaging limit of AAV (5 kb). Currently, in vivo delivery of base editors by AAV has been approached by splitting the base editors into two AAVs and relying on the use of intein trans-splicing for the reconstitution of the full-length effector. Although effective, this approach requires simultaneous entry of both AAVs in the target cell and successful in trans reconstitution of the two

intein halves, which may compromise the on-target efficiency. Furthermore, the requirement of delivering two AAV vectors for each disease target site increases the total viral dosage needed for a treatment, which raises safety concerns and adds burdens to AAV manufacturing.

Compact Cas9 orthologs are ideal candidates for engineering base editors suitable for single-AAV delivery. For example, the Hewitt group has achieved single-AAV delivery of a domain-inlaid *Staphylococcus aureus* Cas9 (SaCas9) ABE in cultured HEK293 cells. Previously, we characterized *Neisseria meningitidis* Cas9, Nme2Cas9, for in vivo genome editing.

Nme2Cas9 is a compact, intrinsically accurate Cas9 with a distinct N4CC PAM specificity. In data presented below, N-terminal ABEs fused to a Nme2Cas9 were developed, and their editing efficiencies, editing windows, and off-target activities were defined in comparison with the widely applied SpyCas9-BEs in cultured cells. Next, N-terminal Nme2Cas9-ABE was shown to edit multiple therapeutically relevant loci, including one of the common mutations occurring in Rett syndrome patients that cannot be targeted by the SpCas9-ABEs, because of the PAM restrictions. Lastly, by optimizing the promoter and the nuclear localization signals, we show that Nme2Cas9-ABE can be packaged in a single AAV for in vivo delivery. One systematic administration of the single AAV encoding Nme2Cas9-ABE readily corrects the disease mutation and phenotype in an adult mouse model of hereditary tyrosinemia type 1 (HT1).

First, to quickly evaluate the base editing efficiency, an ABE reporter cell line was developed, where a G-to-A mutation in an mCherry coding sequence generates a nonsense mutation. Adenine base editing can reverse the mutation and recover the red fluorescence, and the editing efficiency can be readily measured by fluorescent-activated cell sorting (FACS). Initially, an Nme2Cas9-ABE7.10 was constructed by linking TadA-TadA7.10 dimer from the SpCas9-ABE7.10 to the N-terminus of the Nme2Cas9 HNH nickase. However, by plasmid transient transfection, Nme2Cas9-ABE7.10 showed very low to no activity in the ABE reporter cell line. Preliminary data comparing NmeCas9 and SypCas9 nucleases with N-terminally fused nucleotide deaminase domains demonstrated several differences between the two constructs. For example, an ABE reporter cell line was constructed to test the gene editing characteristics between the two orthologs. See, Fig. 21A. The NmeCas9 constructs are shown with either an N-terminal fusion of an ABE7.10 control domain, or an N-terminal fusion of an ABE8e domain. Because the evolved TadA8e is highly active and compatible with a wide range of Cas9s, Nme2Cas9-ABE8e was engineered by linking TadA8e to the N-terminus of the Nme2Cas9 HNH nickase. Similar

constructs were created that replaced the NmeCas9 with the SpyCas9. See, Fig. 21B. The data show that both Cas9 orthologs had greater gene editing efficiency when N-terminally fused to the ABE8e domain as compared to the ABE7.10 control domain. Further, the SpyCas9 ortholog was observed to have greater gene editing efficiency than the NmeCas9 ortholog. See, Fig. 21C. Next, to define the editing window and editing efficiency of Nme2Cas9-ABE8e, and to compare to those of SpCas9-ABE7.10 and SpCas9-ABE8e, plasmids were transfected expressing the ABE along with sgRNAs targeting 12 human genomic loci for Nme2Cas9-ABE8e (including 8 dual-target sites (target sites followed by NGGNCC PAMs) and 4 Nme2Cas9 specific target sites), and 8 dual-target sites for SpCas9-ABEs. The data shows that Nme2Cas9-ABE8e has an editing window of 2-18 (where position 1 is the first nucleotide of the protospacer and the PAM is at positions 25–30, wider than those of the SpCas9-ABEs, and overall lower efficiency. See, Fig. 21D.

Off-target effects of the Nme2Cas9-ABE8e were then evaluated. It has been shown that the major source of DNA off-target base editing is Cas9-dependent, which is caused by Cas9 binding and unwinding at near-cognate sequences. Because Nme2Cas9 is intrinsically highly accurate, it was hypothesized that Nme2Cas9-ABE8e will show a lower Cas9-dependent off-target effect than SpCas9-ABE8e. However, the overall low on-target efficiency and the limited number of potential genome-wide off-target sites for Nme2Cas9-ABE8e makes it difficult to detect and compare the off-target effect to that of SpCas9-ABE8e. Alternatively, a systematic investigation of the tolerance of nucleotide mismatches between the sgRNAs and the target sequence was undertaken. To do this, panel of guides was designed targeting the ABE reporter with single- or di-nucleotide mismatches to the target sequence for both Nme2Cas9-ABE8e and SpCas9-ABE8e and measured their activities by plasmid transfection and FACS. See, Fig. 21E. In contrast to the gene editing efficiencies, however, the N-terminally fused Nme2Cas9-ABE8e construct had less sgRNA nucleotide mismatches as compared to the SpyCas9 construct. This observation is consistent with the observation that the Nme2Cas9 nucleases, in general, are hyper-accurate as compared to other Cas9 nucleases (*infra*). See, Fig. 21E. Considering the differences in on-target efficiencies between the two effectors, the activities of the mismatched guides were normalized to the perfectly complementary guides for each effector. The Nme2Cas9-ABE8e has a significantly lower off-target effect than SpCas9-ABE8e: while single-nucleotide mismatches in the seed region (guide nucleotide 17-24 for Nme2Cas9, and 10-20 for SpCas9) and the majority of di-nucleotide mismatches significantly compromised the efficiency of Nme2Cas9-ABE8e, they were mostly

well-tolerated by SpCas9-ABE8e. A target-specific summary presentation of these A-G conversion efficiencies confirms these differences. See, Fig. 22.

II. Therapeutic Applications Of N-Terminal Cas9 Fusion Constructs

A. Rett Syndrome

Because of this confirmed hyper-accuracy of the N-terminal Nme2Cas9-ABE8e construct, its ability to revert single base gene mutations that result in genetic diseases was evaluated. For example, an sgRNA was created to guide the N-terminal Nme2Cas9-ABE8e to single base mutations in positions 10 and 16 of the *MeCP2* gene, which are known to result in Rett syndrome. An N-terminal Nme2Cas9-ABE8e mRNA was electroporated with a synthetic sgRNA into a Rett syndrome patient-derived fibroblast cell line that possesses this mutation. By amplicon deep sequencing, the data showed that Nme2Cas9-ABE8e generates $17.82 \pm 5.07\%$ editing at the target adenine (A10). An inefficient bystander editing ($4.6 \pm 1.36\%$) at an upstream adenine (A16) will cause a missense mutation (c.296 T>C; p.S166P). Because S166 has been shown subject to phosphorylation in mice and is conserved from *X. laevis* to humans, the bystander editing at A16 may impede functional rescue of edited cells. See, Fig. 23A. The data show 15 - 20% A-G conversion efficiency at position 10 and a 5% A-G conversion efficiency at position 16. See, Fig. 23B.

B. Duchenne Muscular Dystrophy (DMD)

SgRNAs were also created to guide an N-terminal Nme2Cas9-ABE8e to single base mutations at positions 3, 7, 9, 16 and 19 of the *Dmd* gene, which are known to result in muscular dystrophy. Also determined was that this gene editing strategy resulted in the skipping of exon 50 to restore the wild type reading frame. A disease-suppressing mutation was generated that has been shown to reverse phenotypes of a validated Duchenne muscular dystrophy (DMD) mouse model (Δ Ex51). The Δ Ex51 mouse model was generated by deletion of the exon 51 in the *Dmd* gene, resulting in a downstream premature stop codon in exon 52, causing the production of a nonfunctional truncated dystrophin protein. Previously, it has been shown that the *Dmd* reading frame can be restored by skipping exon 50 by adenine base editing. However, in vivo base editing in those studies using ABEmax-SpCas9-NG delivered by dual-AAV vectors was limited to local muscle injection due to the high viral dosage required to achieve therapeutic effects. An sgRNA design for an N-terminally fused Nme2Cas9-ABE8e was created to target the adenine (A7) within

the splicing donor site of exon 50. See, Fig. 23C. By plasmid transfection in the mouse N2a cell line and amplicon deep sequencing, it was found that N-terminally fused Nme2Cas9-ABE8e can generate $17.67 \pm 4.57\%$ editing at A7. While multiple bystander adenines were edited efficiently, these adenines are either within exon 50 or the intron, which will not be expressed. Overall, the data show a range of 5 - 15% A-G conversion efficiency at these positions. See, Fig. 23D.

C. AAV Delivery of N-Terminal Cas9 ABE Constructs

The clinical administration of N-terminal Cas9-ABE8e constructs were evaluated as to their compatibility with adeno-associated virus (AAV) delivery. As detailed further below, AAV delivery of Nme2Cas9 constructs, in particular, have numerous advantages over other Cas9 nuclease orthologs due to their smaller size. Previously, it has been shown that Nme2Cas9 with a sgRNA can be packaged into a single AAV and support efficient editing in vivo. Because of the compact sizes of Nme2Cas9 and TadA8e, Nme2Cas9-ABE8e with a sgRNA can, in theory, be packaged into a single AAV for in vivo base editing. To test this idea, Nme2Cas9 was replaced with an N-terminally fused Nme2Cas9-ABE8e in an all-in-one AAV vector. One cMyc NLS sequence was attached on each terminus of Nme2Cas9-ABE8e while retaining the original promoters for Nme2Cas9-ABE8e and sgRNA expression. By plasmid transient transfection, the single-AAV vector showed $9.12 \pm 0.69\%$ editing efficiency in the ABE reporter cell line. To further improve editing efficiency, three different NLS configurations were tested: 1) one cMyc NLS on the N-terminus and two cMyc NLS on the C-terminus; 2) one Ty1 NLS, which derived from the yeast Ty1 retrotransposon that supports robust nuclear localization in dPSPCas13b fusion proteins, on the N-terminus; and 3) one bipartite SV40 NLS (BP_SV40) on each terminus of Nme2Cas9-ABE8e. When transfecting the vector plasmid into the ABE reporter cell line, the construct having BP_SV40 NLS on each terminus showed an editing efficiency of $20.5 \pm 3.2\%$. As the in vivo delivery of AAV plasmid/vector payloads are assisted by appending nuclear localization signal proteins (e.g., NLS), several Nme2Cas9-ABE8e N-terminally fused constructs were created and tested for efficacy using the standard m-Cherry reporter system. See, Figs 24A and 24B.

As the 2XBP-SV40 NLS construct was observed to have optimal performance, this construct was chosen to evaluate gene editing efficiency between the U6 and miniU6 promoters at the D12 and Rosa26 test target sites. The total length of the vector construct with the BP-SV40 NLS, hereafter Nme2Cas9-ABE8e-U6, is 4998 bp, just below the packaging limit of AAV. To test

if the vector size could be further reduced without significantly compromising the editing efficiency, a “miniU6” promoter, which has been shown to support sgRNA expression to achieve a similar level of CRISPR editing efficiency was compared to the U6 promoter. When replacing the U6 promoter with miniU6 promoter, hereafter Nme2Cas9-ABE8e-miniU6, a vector was generated with a total length of 4860 bp, well below the packaging limit of $\Delta\Delta V$. (infra). When performing transient plasmid transfection, the construct showed reduced yet significant editing in the ABE reporter cell line. The data show that the U6 promoter consistently showed greater gene editing efficiency. To avoid potential ABE reporter-specific effect, both single-AAV vectors were tested at two endogenous target sites: 1) one of the human dual-target sites, DS12, and 2) a previously reported Nme2Cas9 target site in the mouse *Rosa26* gene. By plasmid transfection in human HEK293T or mouse N2a cells, significant editing was observed at these loci by both vectors, although the Nme2Cas9-ABE8e-miniU6 vector was less efficient. See, Fig. 24C. These AAV N-terminally fused Nme2-ABE8e constructs were used to confirm the above *Fah* gene editing data. See, Fig. 25A-C.

The liver disease HT1 was used to test in vivo editing efficiency and the therapeutic potential of the single-AAV constructs. HT1 is caused by mutations in fumarylacetoacetate hydrolase (*Fah*) gene, which catalyzes the tyrosine catabolic pathway. FAH deficiency leads to accumulations of toxic fumarylacetoacetate and succinyl acetoacetate, causing liver, kidney, and CNS damage. The *Fah*^{PM/PM} mouse model possesses a G•C to A•T point mutation in the last nucleotide of exon 8, which causes skipping of exon 8 and FAH deficiency. See, Fig. 26A. Without treatment, an individual will rapidly lose weight and eventually die. The *Fah*^{PM/PM} mouse can be treated with 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (NTBC), an inhibitor of an enzyme upstream in the tyrosine degradation pathway, to prevent toxin accumulation. A hepatocyte in which the *Fah* gene has been repaired has growth advantage and expands after NTBC withdrawal. Repair of 1 in 100,000 hepatocytes was reported to rescue the phenotype of *Fah*^{PM/PM} mice. Previously, in vivo gene-editing tools have been tested to treat the *Fah*^{PM/PM} mouse model, including Cas9-directed HDR, base editing, microhomology-directed end joining, and prime editing. Multiple approaches including AAV, lipid nanoparticle (LNP), and hydrodynamic tail vein injection of plasmids have been used to deliver the gene-editing agents into this mouse model. The *Fah*^{PM/PM} mouse model will facilitate comparisons between different genome-editing platforms.

Exon-skipping strategies was further evaluated using single base mutations in the *Fah* gene, which is known to cause tyrosinemia. An sgRNA was created that targets the point mutation by electroporation of the single-AAV vector plasmids into the mouse embryonic fibroblasts (MEFs) isolated from the *Fah*^{PM/PM} mouse. This sgRNA was designed to guide an N-terminal Nme2-Cas9-ABE8e construct to positions 5, 10 13, 16 and 17 within exon 8 of the *Fah* gene. One construct incorporated a U6 promoter and a second construct incorporated a miniU6 promoter. See, Fig. 26A and Fig. 26B. A low but significant gene editing ($2.94 \pm 0.11\%$ for N-terminal Nme2Cas9-ABE8e-U6, and $1.23 \pm 0.30\%$ for Nme2Cas9-ABE8e-miniU6) was detected at the target adenine at position 13 (A13). Also observed was a significantly higher bystander editing (6.52 $\pm$ 0.53% for Nme2Cas9-ABE8e-U6, and 2.51 $\pm$ 0.52% for Nme2Cas9-ABE8e-miniU6) at A16. Nevertheless, A16 is within the intron region downstream of the splicing donor and is likely to be harmless. In general, the data show that the constructs with the U6 promoter had higher A-G conversion efficiencies than the miniU6 promoter. See, Fig. 26C.

To test single-AAV vectors of N-terminal Nme2Cas9-ABE8e in vivo, hydrodynamic tail vein injection of the AAV-vector plasmids was performed into 10-week-old HT1 mice, or PBS for the negative control group. Also injected were plasmids expressing SpCas9-RA6.3, which is a codon-optimized SpCas9-ABE with increased efficiency, as a positive control. Seven days post-injection, gene editing efficiency was measured before hepatocyte expansion, and withdrawal NTBC for the rest of the mice for long-term phenotypic study. Before NTBC withdrawal, anti-FAH immunohistochemistry (IHC) staining showed $4.58 \pm 1.1\%$ FAH+ hepatocytes from the group that injected with the Nme2Cas9-ABE8e-U6 plasmid, and $3.04 \pm 0.07\%$ from the group injected with the Nme2Cas9-ABE8e-miniU6 plasmid. The mouse injected with SpCas9-RA6.3 plasmid showed 4.5% FAH+ hepatocytes, consistent with the reported data. See, Fig. 26D. After NTBC withdrawal, body weight changes were monitored. While the PBS injected group rapidly lost body weight within the first 3 days after NTBC withdrawal and was thus euthanized, the mice that were injected with either the Nme2Cas9-ABE8e-U6 plasmid or the Nme2Cas9-ABE8e-miniU6 plasmid gradually gained body weight over 40 days. See, Fig. 25E. These immunohistochemical analysis and body weight data confirmed mutation reversion by demonstrating the reappearance of *Fah*⁺ hepatocytes subsequent to gene editing and improved body weight gain in the U6 promoter constructs.

To determine whether the N-terminal Nme2Cas9-ABE8e successfully corrects the *Fah* gene splicing defect, total RNA was extracted from livers and reverse transcription PCR (RT-PCR) was performed using primers that spanned exons 5 and 9. In contrast to the PBS-injected mice, which only showed a 305 bp PCR band that corresponds to the truncated mRNA lacking exon 8, the majority of the treated mouse livers were observed to have the 405 bp PCR band that contained exon 8. See, Fig. 26F. Sanger sequencing of the 405 bp bands further confirmed the presence of the corrected 'G' at position 13. See, Fig. 26G. When performing anti-FAH immunohistochemical (IHC) staining, a significant expansion of FAH⁺ hepatocytes were observed in the groups that were injected with either of the single-AAV vector plasmids. See, Fig. 25H. Amplicon deep sequencing on the liver genomic DNA from treated mice showed $12.21 \pm 1.25\%$ editing for Nme2Cas9-ABE8e-U6 plasmid treated mice, and $10.1 \pm 0.647\%$ for Nme2Cas9-ABE8e-miniU6 plasmid treated mice at A13. See, Fig. 26I. These data indicate that N-terminal Nme2Cas9-ABE8e single-AAV vector plasmids successfully corrected the disease genotype and phenotype of the *Fah* PM/PM mice. These gene editing results for fusion proteins comprising an Nme2Cas9 and an N-terminal adenine deaminase domain were validated by flow cytometry gating. See, Fig. 27.

In vivo therapeutic base editing was then assessed using an N-terminal Nme2Cas9-ABE8e delivered by AAV9 in an adult HT1 mouse model. In particular, AAV9 were packaged with the Nme2Cas9-ABE8e-U6 construct or a the Nme2Cas9-ABE8e-miniU6 construct. Next, AAV genome integrity was confirmed by DNA extraction and alkaline gel electrophoresis, where the data did not show any sign of genome truncation. See, Fig. 28A. 8-week-old *Fah* PM/PM mice were the injected in the tail with 4×10^{13} vg/kg and maintained on NTCB for one month prior to analyzing gene editing efficiency. Because gene expression from the single-strand AAV, unlike from plasmid, requires second-strand synthesis, the mice were kept on NTBC for one month before analyzing the editing efficiency. One month after AAV injection and before NTBC withdrawal, IHC staining was performed using an anti-FAH antibody. The negative control groups injected with AAV9 expressing N-terminal Nme2Cas9-ABE8e and a sgRNA targeting the *Rosa26* gene did not show any FAH⁺ hepatocytes. In contrast, $6.49 \pm 2.08\%$ FAH⁺ hepatocytes were observed in the AAV9-Nme2Cas9-ABE8e-U6-*Fah* treated group, and $0.98 \pm 0.49\%$ FAH⁺ hepatocytes in AAV9-Nme2Cas9-ABE8e-miniU6-*Fah* treated group. See, Fig. 28B. The percentage of edited hepatocytes by AAV9-Nme2Cas9-ABE8e-U6-*Fah* was significantly higher

than what has been achieved previously by other genome editing strategies. By targeted deep-sequencing, the editing efficiency at the target adenine (A13) in the AAV9-Nme2Cas9-ABE8e-U6 treated group is $0.34 \pm 0.14\%$, and $0.08 \pm 0.09\%$ in AAV9-Nme2Cas9-ABE8e-miniU6 treated group. The reason for the higher frequency of FAH⁺ hepatocytes than the frequency of editing at the DNA level is because of the polyploidy of the hepatocytes, and the presence of genomic DNA from nonparenchymal cells. Similar results were also observed in previous studies using this mouse model.

To evaluate potential Cas9-dependent off-target effect in AAV9-injected mice, genome-wide off-target sites for Nme2Cas9 were identified using Cas9-OFFiner, allowing for up to 6 mismatches. Amplicon deep sequencing was performed in AAV9-Nme2Cas9-ABE8e-Fah treated livers (n = 8) at the two top-ranking potential off-target sites, each including 5 mismatches. Above-the-background A•T to G•C editing was not detected at these sites. See, Fig. 29.

In conclusion, the data suggest that while a fusion protein comprising an Nme2Cas9 comprising an N-terminal ABE8e domain is capable of hyperaccurate gene editing to revert single base gene mutations and provide therapeutic efficacy, the gene editing efficiency was believed too low to provide a clinically optimal treatment platform.

III. Nme2Cas9 Single Base Editing

Cas9 is a programmable nuclease that uses a guide RNA to create a double-stranded break at any desired genomic locus. This programmability has been harnessed for biomedical and therapeutic approaches. However, Cas9-induced breaks often lead to imprecise repair by the cellular machinery, hindering its therapeutic application for single-base corrections as well as uniform and precise gene knock-outs. Moreover, it is extremely challenging to combine Cas9-induced DNA double strand breaks and a repair template for homologous directed repair (HDR) for correcting genetic mutations in post-mitotic cells (e.g. neuronal cells).

Single nucleotide base editing is a genome editing approach where a nuclease-dead or -impaired Cas9 (e.g., dead Cas9 (dCas9) or nickase Cas9 (nCas9)) is fused to another enzyme capable of base-editing nucleotides without causing DNA double strand breaks. To date, two broad classes of Cas9 base editors have been developed: i) cytidine deaminase (edits a C•G base pair to a T•A base pair) SpyCas9 fusion protein; and ii) adenosine deaminase (edits a A•T base pair to a G•C base pair) SpyCas9. Liu et al., "Nucleobase editors and uses thereof" US 2017/0121693; and

Lui et al., “Fusions of cas9 domains and nucleic acid-editing domains” US 2015/0166980; (both herein incorporated by reference).

In one embodiment, the present invention contemplates a deaminase fusion protein with a compact and hyper-accurate Nme2Cas9 (*Neisseria meningitidis* spp.). This Nme2Cas9 has 1,082 amino acids as compared to SpyCas9 that has 1,368 amino acids. This Nme2Cas9 ortholog functions efficiently in mammalian cells, recognizes an N₄CC PAM, and is intrinsically hyper-accurate. Edraki et al., *Mol Cell*. (in preparation). Although it is not necessary to understand the mechanism of an invention, it is believed that the compactness and hyperaccuracy of an NmeCas9 base editor targets single-base mutations that could not be reached previously by other Cas9 platforms currently known in the art. It is further believed that the NmeCas9 base editors contemplated herein target pathogenic mutations that are not feasible via current base editor platforms, and with an increased base editing accuracy.

First generation base editors did include deaminases fused to the N- or C- termini of a Cas domain. Komor et al. *Nature*. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage; and Gaudelli et al. *Nature*. (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. Newer iterations of base editors have led to improved editing efficiencies and widened or shifted editing windows. These developments relied on efforts to engineer improved deaminases, and/or Cas domains. Richter et al. *Nature Biotech*. (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity; Wang et al. *Signal Transduction and Targeted Therapy*. (2019) BE-PIGS: a base-editing tool with deaminases inlaid into Cas9 PI domain significantly expanded the editing scope; Tran et al. *Nature Commun*. (2020); Engineering domain-inlaid SaCas9 adenine base editors with reduced RNA off-targets and increased on-target DNA editing; and Chu et al. *The Crispr Journal*. (2021) Rationally Designed Base Editors for Precise Editing of the Sickle Cell Disease Mutation.

IV. Inlaid Domain Cas9 Fusion Protein Constructs

Domain-inlaid Nme2Cas9 nucleotide base editors are shown herein to improve editing efficiencies and modulate editing windows. Multiple crystal structures Nme2Cas9 reveal that the N-terminus is on the protein face, opposite that of the edited DNA strand, suggesting that N-terminal deaminase fusions are likely not optimally positioned. Surface loops that are closer to potential paths of the displaced DNA strand were identified and linker-flanked TadA8e were inserted into these internal sites (versions i1 - i8). The data shown herein demonstrates that all eight positions of these inlaid domains were active at editing endogenous loci, having editing efficiencies consistently exceeding that of the N-terminal fusion. In particular, the i1, i2, i7 and i8 effectors are consistently 2- to 2.5-fold more active than the N-terminal Nme2-ABE. These results demonstrate that, in general, inlaid domains of Nme2Cas9-ABEs have enhanced editing efficiencies as compared to the N-terminal domain construct.

In one embodiment, the present invention contemplates a fusion protein comprising an NmeCas9 protein and an inlaid nucleotide deaminase protein domain. In one embodiment, the NmeCas9 protein is a Nme2Cas9 protein. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid adenine deaminase protein domain. In one embodiment, the inlaid nucleotide deaminase protein domain is an inlaid cytidine deaminase protein domain. In one embodiment, the inlaid adenine deaminase protein domain is an inlaid adenine deaminase8e (ABE8e) protein domain.

Several approaches to improve the above observed gene editing efficiencies of nucleotide deaminase domains that are N-terminally fused to an NmeCas9 protein may include, but are not limited to: i) move the deaminase domain closer to the R loop; ii) tune the editing window; or iii) increase the deaminase activity. Possible techniques to accomplish these goals may include, but are not limited to: i) use of alternative linkers and location within the Cas9 protein (e.g. shorter/rigid, N- vs. C-termini fusion); iii) alternative deaminase domain insertion into the Cas9 protein; and iii) circular permutation of the Cas9 protein which would create new N- and C-termini. In regards to moving the deaminase domain, both SaCas9 and SpyCas9 have been reported to contain an inlaid adenine deaminase protein domain. Li et al., “Docking Sites Inside Cas9 For Adenine Base Editing Diversification And Off-Target Elimination”; Tran et al., “Engineering Domain-Inlaid SaCas9 Adenine Base Editors With Reduced RNA Off-Targets And Increased On-Target DNA Editing”; Chu et al., “Rationally Designed (Spy) Base Editors

(ABE7.10) for Precise Editing of the Sickle Cell Disease Mutation' The CRISPR Journal (2021). It is envisioned that an inlaid insertion of a nucleotide deaminase protein may three-dimensionally separate the deaminase DNA target site from the Cas9 N-terminus. See, Fig. 30.

These inlaid domain Nme2Cas9 nucleotide base editors enhanced activity and editing window preferences with no significant optimization of the flanking linkers. In one embodiment, the linker flanks each end of the nucleotide base editor. In one embodiment, the linker is flexible and approximately twenty (20) amino acids in length. In one embodiment, linker length may maximize activity gains and improve control over editing windows. For example, a reduced linker length may improve AAV packaging efficiency/titre. In one embodiment, a linker has a length including twenty amino acid linker

The inlaid domain constructs Nme2-ABE i1, i7 and i8 were found to be the most active and exhibited the greatest advantages for favoring PAM-distal (i1) vs PAM-proximal (i7, i8) editing window control.

A. Inlaid Adenine Base Editors (ABEs)

In one embodiment, the present invention contemplates an inlaid adenine deaminase protein domain comprising an adenine deaminase, an N-terminal linker and a C-terminal linker. See, Fig. 31A. For example, the inserted domain may be approximately 206 amino acids in length, wherein the deaminase is TadA8e (166 amino acids in length) and the N-terminal and C-terminal linkers are both about 20 amino acids in length. Oakes (2018) and Chu (2021). Several criteria for inlaid nucleotide deaminase insertion sites were selected including, but not limited to: i) surface exposed amino acids; ii) regions of high conformational flexibility; and iii) apparent proximity to the Non-Target Strand. Based on these selection criteria several potential insertion sites were located along the NmeCas9 protein. See, Fig. 31B. Three dimensional modeling of the Cas9 protein predicted the respective locations of these candidate inlaid insertion locations. See, Fig. 31C.

The gene editing capability was determined for each of the eight (8) inlaid locations for the fusion protein comprising an NmeCas9 protein and an inlaid adenine deaminase8e (ABE8e) protein domain shown in in Fig.s 31B and 31C. The assay was performed using a construct comprising the mCherry reporter system. See, Fig. 32A. Briefly, the mCherry reporter system utilizes the following components; i) a premature stop codon inhibits mCherry translation; ii) ABE editing converts an internal STOP codon (e.g., TAG) to a GLN codon (e.g., CAG); iii) the sgRNA

targets a distal PAM adenosine position and a proximal PAM bystander adenosine position. The data shows inlaid domain locations having gene editing activity by the appearance of a red fluorescence. See, Fig. 32B. Sanger sequencing was then performed to validate the gene editing activity indicated by the mCherry reporter system analysis. See, Fig. 33. Next, based on the mCherry reporter system data, gene editing efficiencies were then estimated. See, Figs 34A and 34B. A general pattern was observed that gene editing was least efficient when the ABE domain was inlaid in the central region of the Cas9 protein. In most cases, the inlaid ABE protein domain had superior gene editing activity as compared to the N-terminally fused ABE protein domain. Gene editing rates were also determined at each of the eight inlaid domain locations at three endogenous gene loci: i) LINC01588-DS12; ii) FANCF-DS28; and iii) MECP2-G2. See, Figs 35A, 35B and 35C, respectively.

B. Therapeutic Applications of Inlaid Domain Nme2-Cas9 Fusion Proteins

1. Duchenne's Muscular Dystrophy

In one embodiment, the present invention contemplates a method to treat DMD with an Nme2Cas9 inlaid ABE construct. For example, several A → G conversion have been successfully converted a mutated DMD gene to wild type. See Figure 36. These data were collected comparing Nme2Cas9-ABE -nt, i1, - i1, i7 and i8 constructs. The inlaid domain constructs were observed to have a higher editing efficiency than the N-terminal construct for the A16G and A19G conversions.

2. Rett Syndrome

In one embodiment, the present invention contemplates a method to treat Rett Syndrome with an Nme2Cas9-ABE construct. In one embodiment, the ABE is a terminal domain. In one embodiment, the ABE is an inlaid domain. For example, Rett syndrome mutations in exon 4 of the MeCP2 gene are targeted by Nme2-ABEs. The empirical nature of this treatment is shown by attempts to correct known Rett syndrome mutations in the HEK293T Rett-PiggyBac cell line by plasmid transfections or editing in Rett- patient derived fibroblasts (PDF) with mRNAs and synthetic gRNA. See, Table 3.

Table 3: Representative data showing ABE-conversion of MeCP2 mutations

Mutation	Protein Change	Patient Population with mutation (%)	Bystander Adenine's	RETT-PiggyBac	RETT-PDF
c.502C>T	R168X	7.63	yes	✓	✓
c.763C>T	R255X	6.68	no	✓	X
c.808C>T	R270X	5.80	yes	✓	X
c.916C>T	R306X	5.17	yes	✓	✓

✓ = Successful

X = In Progress

5

The following MeCP2 target site sequences and their respective PAM were tested in attempt to correct Rett syndrome mutations. See, Table 4.

Table 4: MeCP2 Target Site Sequences Associated With Rett Syndrome Mutations

Guide	Spacer 5'-3'	PAM	Target A.	Bystander A's.
502-G8	GTGGTTTCTGCTCTCACCGGGAGG	GGCTCC	16	22
502-G7	TGGTTTCTGCTCTCACCGGGAGGG	GCTCCC	15	21
502-G8	TCTGCTCTCACCGGGAGGGGCTCC	CTCTCC	10	16
502-G9	AGGTGGTTTCTGCTCTCACCGGGA	GGGGCT	18	24
502-G10	GGTTTCTGCTCTCACCGGGAGGGG	CTCCCT	14	20
763-G1	GGCCTCAGCTTTTCACTTCCTGCC	GGGGCG	15	7
763-G2	CTTTCACTTCCTGCCGGGGCGTT	TGATCA	7	n/a
763-G3	TTTCACTTCCTGCCGGGGCGTTTG	ATCACC	5	n/a
763-G4	TTCACTTCCTGCCGGGGCGTTTGA	TCACCA	4	n/a
808-G1	CCACACTCCCCGGCTTTAGCCCC	GTTTCT	19	3, 5
808-G2	GGCTTTAGCCCCGTTTCTTGGGA	ATGGCC	8	24
808-G3	GCTTTAGCCCCGTTTCTTGGGAA	TGGCCT	7	23, 24
916-G1	CGGGTCTTGCACTTCTTGATGGGG	AGTACG	11	19
916-G2	TCTTGCACTTCTTGATGGGGAGTA	CGGTCT	7	15, 21, 24
916-G3	TTGCACTTCTTGATGGGGAGTACG	GTCTCC	5	13, 19, 22
916-G4	TGCACTTCTTGATGGGGAGTACGG	TCTCCT	4	12, 19, 23

10

The target adenine position (e.g., Target A) in the spacer sequence is determined by counting from the 5' terminal base. Bystander adenines are also identified in the same manner.

Representative target sites were screened in RETT-PiggyBac cells and Rett- patient derived fibroblasts (PDFs).

Conversion of four (4) different McCp2 exon 4 mutations have been successfully performed with the presently disclosed Nme2Cas9-ABE constructs. In particular, the N-terminal (-nt) construct was compared to three (3) inlaid domain (i1, i7 and i8) constructs. The inlaid domain constructs demonstrated higher c.502 C>T conversion in most A→G edited sites. See, Figures 44A- 44B. The inlaid domain constructs demonstrated higher c.916 C>T conversion in some A→G edited sites. See, Figures 45A- 45B. The inlaid domain constructs demonstrated higher c.763C>T conversion in the A→G edited site. See, Figures 46A- 46B. The inlaid domain constructs demonstrated higher c.808C>T conversion in some of the A→G edited sites. See, Figures 47A- 47B.

3. Batten Disease

In one embodiment, the present invention contemplates a method of treating Batten disease with Nme2Cas9 base editing of the CLN3 Δ ex7/8 mutation and concomitant exon 5 skipping. In one embodiment, the method further comprises a guide mRNA to target the CLN3 mutation. See, Figures 37A, 37B and 38. Significant gene editing activity was observed with inlaid domain constructs using these mRNA sequences, and were superior to that seen with N-terminal constructs. See, Figures 39A and 39B.

Exon 5 skipping was observed. See, Figures 40A – 40D. ASO-induced exon 5 skipping of mouse CLN3 transiently restores reading frame in Cln3 Δ ex7/8 mice and ameliorates Batten disease phenotypes. Nme2-ABE editing of exon 5 splice sites could potentially induce long-term exon skipping. The above data demonstrates the targeting of the WT allele in mouse Neuro2a cells, which express Cln3, using Nme2-ABE-i1 and candidate guide RNAs. These guide mRNAs yielded efficient editing at each splice site (acceptor site and receptor site). Splice site mutations can sometimes lead to activation of nearby cryptic splice sites rather than exon skipping. To assess this, RT-PCR analysis of total RNA was performed using primers for exons 4 and 6. A smaller gel band appeared in the edited samples. See, Figure 40C. Sanger sequencing confirmed its identity as an exon-5-skipped species. See, Figure 40D. These results confirm that inlaid domain Nme2-ABEs can induce exon skipping via splice site editing.

Gene editing and exon skipping was also observed in brain structures such as cortex, striatum, hippocampus and thalamus using AAV delivery of the inlaid domain and N-terminal

domain Nme2-Cas9 constructs. See, Figures 41A-41C. CLN3 gene editing was observed in specific brain regions of both adults and neonates, as well as in liver. See, Figures 42A – 42E. The regional distribution of CLN3 mutation conversions was documented by brain slice mRNA transcript imaging. See, Figure 43A – 43C.

5 Batten disease is an autosomal recessive fatal neurological disorder caused by mutations in the CLN3 gene. Patients are often blind by later childhood and gradually develop seizure and movement abnormalities, and most patients only live into early adulthood. Approximately 85% of Batten disease patients carry a 1.02 kb deletion spanning exons 7 and 8 of CLN3 61, which results in a premature termination codon in exon 9 and loss of a C-terminal lysosomal targeting sequence.

10 The homozygous Cln3 Δ ex7/8 mouse model (JAX #017895) exhibits pathological changes. Administering AAVrh.10 expressing human wild-type CLN3 protein to newborn Cln3 Δ ex7/8 mice partially corrects neurological lysosomal storage defects. Homozygous Cln3 Δ ex7/8 mice can also be treated by skipping exon 5 of the Cln3 gene via splice-switching antisense oligos (ASOs) to bring the C-terminus of the CLN3 ORF back in frame, suggesting that BE-induced CLN3 exon 5

15 skipping could have durable therapeutic value. Kytälä et al., “Two motifs target Batten disease protein CLN3 to lysosomes in transfected nonneuronal and neuronal cells” Mol Biol Cell 2004 Mar;15(3):1313–1323; Pontikis et al., “Late onset neurodegeneration in the Cln3 Δ mouse model of juvenile neuronal ceroid lipofuscinosis is preceded by low level glial activation” Brain Res 2004 Oct 15;1023(2):231–242; Osório et al., “Neurodevelopmental delay in the Cln3Deltaex7/8 mouse

20 model for Batten disease” Genes Brain Behav 2009 Apr;8(3):337–345; Burkovetskaya et al., “Evidence for aberrant astrocyte hemichannel activity in Juvenile Neuronal Ceroid Lipofuscinosis (JNCL)” PLoS One 2014 Apr 15;9(4):e95023; and Sondhi et al., “Partial correction of the CNS lysosomal storage defect in a mouse model of juvenile neuronal ceroid lipofuscinosis by neonatal CNS administration of an adeno-associated virus serotype rh.10 vector expressing the human

25 CLN3 gene” Hum Gene Ther 2014 Mar;25(3):223–239.

4. Amyotrophic Lateral Sclerosis (ALS)

In one embodiment, the present invention contemplates a method of treating ALS with Nme2Cas9 base editing of a SOD1 mutation.

30 ALS is a neurodegenerative disease in which loss of motor neurons results in progressive muscle weakness, paralysis, and death, typically within 2-5 years of onset. Only two FDA-approved drugs are available, with modest delays (~three months) in ALS progression. There is an

unmet clinical need to develop treatments for ALS. Approximately 90% of ALS cases are sporadic and 10% of cases are familial. Mutations in the free-radical scavenger gene SOD1 (Cu-Zn superoxide dismutase 1) are the second most common genetic cause of ALS. ALS-associated, dominant SOD1 mutations destabilize the protein, causing aberrant misfolding and aggregation that likely contribute to cell death. A SOD1G93A (a gain-of-function mutation) transgenic mouse model exhibits motor neuron loss and a shortened lifespan (5 to 6 months). SOD1G37R mice carrying another toxic SOD1 mutation develop similar ALS symptoms. Sod1^{-/-} knockout mice are normal and healthy, while in rare instances humans devoid of SOD1 exhibit neurodevelopmental defects. These findings point to the therapeutic potential of depleting mutant SOD1 to treat ALS in adults. Gurney et al., “Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation” *Science* 1994 Jun 17;264(5166):1772–1775; Wong et al., “An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria” *Neuron* 1995 Jun;14(6):1105–1116; Taylor et al., “Decoding ALS: from genes to mechanism” *Nature* 2016 Nov 10;539(7628):197–206; Liu et al., “Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria” *Neuron* 2004 Jul 8;43(1):5–17; Park et al., “SOD1 deficiency: a novel syndrome distinct from amyotrophic lateral sclerosis” *Brain* 2019 Aug 1;142(8):2230–2237; and Ezer et al., “Infantile SOD1 deficiency syndrome caused by a homozygous SOD1 variant with absence of enzyme activity” *Brain* 2022 Apr 29;145(3):872–878. Previous studies demonstrated that repressing SOD1 by RNAi ameliorates ALS in animal models and increases survival, and potential clinical benefit of SOD1 suppression has been reported in humans. Borel et al., “Therapeutic rAAVrh10 Mediated SOD1 Silencing in Adult SOD1(G93A) Mice and Nonhuman Primates” *Hum Gene Ther* 2016 Jan;27(1):19–31; Miller et al., “Phase 1-2 Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS” *N Engl J Med* 2020 Jul 9;383(2):109–119; and Mueller et al., “SOD1 Suppression with Adeno-Associated Virus and MicroRNA in Familial ALS. *N Engl J Med* 2020 Jul 9;383(2):151–158.

A Cas9 strategy for targeting SOD1 is appealing because of the ability to cause permanent genetic alteration, eliminating the need for repeated dosing. It has been shown that AAV9 delivery of Cas9 and guide RNA in vivo can deplete mutant SOD1 and prolong survival in SOD1G93A mice, though DSB-induced editing efficiencies were very low in these studies. Lee et al., “Imaging Net Retrograde Axonal Transport In Vivo: A Physiological Biomarker” *Ann Neurol.* 2022

May;91(5):716–729; and Gaj et al., “In vivo genome editing improves motor function and extends survival in a mouse model of ALS” *Sci Adv* 2017 Dec;3(12):eaar3952.

In one embodiment, the present invention contemplates a method comprising an inlaid domain Nm32Cas9-ABE to perform SOD1 gene editing. Previous data show that AAV-Cas9 treatment prolongs survival in SOD1G93A mice with a SOD1 exonic sgRNA targeting both WT and G93A human alleles. Consequently, two AAV9 vectors, clinically validated for motor neuron transduction were generated: i) AAV9.sgSOD1 expressing a guide RNA, and ii) AAV9.Cas9 expressing SpyCas9. See, Fig. 59A; and Mendell et al., “Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy” *N Engl J Med* 2017 Nov 2;377(18). In HEK293T cells, dual plasmid backbone transfections reduced SOD1 protein levels. In addition, sgSOD1+Cas9 mice showed significant improvement in rotarod performance. Lee et al., “Imaging Net Retrograde Axonal Transport In Vivo: A Physiological Biomarker” *Ann Neurol* 2022 May;91(5):716–729. To test editing in vivo, the dual AAV9 to transgenic SOD1G93A were delivered to mouse P1 neonates via ICV injections. The mice were monitored throughout their lifespan for ALS onset and symptoms. Compared to all control groups, AAV9 sgSOD1+Cas9 mice exhibited increased survival. See, Figure 59B. To determine if sgSOD1 preserves axonal architecture of the ventral spinal root in ALS mice, L5 dorsal root axon cross sections were analyzed. sgSOD1 mice had significantly more intact axons than sgControl as measured by distributions of axon area. See, Figure 59C. Together, these data indicate that dual AAV-Cas9 treatment has therapeutic benefits for ALS conditions.

In one embodiment, the present invention contemplates a method to induce SOD1 exon 2 skipping to treat ALS. Both N₄CC and N₄CD PAMs can be considered for disrupting splice sites and inducing SOD1 exon 2 skipping. Exon 2 skipping leads to a frameshift and loss of SOD1 function leading to at least two SOD1 mutations, e.g., G37R and G93A. An Nme2-ABE sgRNA targeting the intron 1 splice acceptor has been identified. See, Figure 60A. For Nme2Cas9-ABEs, the A15 position of the protospacer (e.g., target A) is the “AG” splice acceptor and targets a dual-C PAM. For Nme2Cas9^{Smm}-ABEs, the A14 position (e.g., target B) of the protospacer is the “AG” splice acceptor and targets a single-C PAM. Similarly, the intron 2 splicing donor is targeted, including single-G (C on the opposite strand) PAMs. See, Figure 60B.

In one embodiment, the present invention contemplates an inlaid domain Nme2Cas9-ABE to correct a mutated SOD1 G37R allele and treat ALS symptoms, Although splice site disruption

and exon 2 skipping has potential therapeutic benefit, base editors are more efficient for correction of disease-causing SNPs, such as the common SOD1^{G37R} G-to-A mutation. 17; and Figure 61. Several missense mutations may occur in the SOD1. See, Figure 61 (blue nts). While it is not yet known if mutations in the residues cause ALS, their tolerance for mutation (though possible) cannot be assumed. There are three PAMs (one N₄CC and two N₄CD) within this loci that could be targeted with inlaid Nme2-ABEs and Nme2^{Smi}-ABEs, respectively. The N₄CC PAM places the target at A15 (red).

C. AAV Nme2Cas9 Inlaid Domain Constructs

Inlaid domain Nme2Cas9-ABE constructs are shown herein to be compatible with an in vivo single-AAV delivery platform. The all-in-one AAV vector for N-terminal Nme2-ABE was validated with the U1a and U6 promoters driving effector and sgRNA expression (respectively), is 4998 bp including the ITRs. Two analogous AAV9 versions (both 4996 nts) were generated with an inlaid deaminase domain at the il site: a first construct comprised the TadA8e domain, while the second construct comprised a TadA8e^{V106W} mutant that greatly reduces unintended A-to-G conversion in RNA molecules while minimally affecting on-target DNA deamination activity^{69,70}. All three vectors were targeted a common site in Rosa26 and administered via tail vein injection of AAV9 at 4 x 10¹¹ vg (vector genomes) in adult mice.

The data presented herein demonstrates significant adenine editing activity at Rosa26 mutations with the editing with AAV9 integration the Nme2Cas9-ABE -nt, il and il^{V106W} constructs as compared to a physiological buffer saline (PBS) control. See, Figures 48A – 48C. These results demonstrate the in vivo efficacy of single-AAV delivery systems for enhanced-efficiency, domain-inlaid Nme2-ABEs.

D. Inlaid Cytosine Base Editors (CBEs)

In one embodiment, the present invention contemplates an Nme2Cas9-CBE construct. In one embodiment, the Nme2Cas9-CBE construct comprises an N-terminal CBE. In one embodiment, the Nme2Cas9-CBE construct comprises an inlaid domain CBE. In one embodiment, the inlaid domain includes, but is not limited to Nme2Cas9-CBE-(i1), Nme2Cas9-CBE-(i7) and Nme2Cas9-CBE-(i8). In one embodiment, the CBE is a cytidine deaminase. In one embodiment, the cytidine deaminase includes, but is not limited to, evoFERNY or rAPOBEC1. See, Figure 52A.

The data presented herein shows that inlaid domain Nme2Cas9-CBE constructs have cytidine editing activity that is either comparable to, or superior to, the N-terminal domain construct. See, Figure 52B – 52C.

5 **V. Inter-Cas9 Protospacer Interacting Domain (PID) Swapping**

The therapeutic promise of base editing systems is believed to hinge upon improvements in editing efficiency, limiting bystander edits (or their consequences), maximizing PAM-dependent targeting scope, and minimizing immunogenicity, toxicity, and prolonged deaminase expression (which can compromise editing efficiency and lead to safety risks such as hepatotoxicity and the
10 accumulation of unwanted edits). Advances to overcome these issues are encompassed by the presently disclosed chimeric Cas9 nucleases which encompass a cross-species PID. It is believed that these improved base editing constructs have increased effectiveness, targeting scope, utility, and safety.

Gene editing using CRISPR-Cas9 technologies has advanced genetic research and
15 promises to revolutionize gene therapy. For efficient editing to occur, a Cas9 recognizes a sequence motif, called a protospacer adjacent motif (PAM), adjacent to the target site. Different Cas9 homologs have distinct PAM sequences; most are 2-5 nucleotides long. The shorter the PAM, the less restrictive the PAM sequence requirement for editing, and the higher the density of candidate target sites. One example discussed above, is a compact Type II-C Cas9 ortholog from
20 *Neisseria meningitidis* (Nme2Cas9). Nme2Cas9 exhibits a unique DNA targeting motif (PAM of N₄CC), high accuracy, and the ability to mediate efficient ex vivo and in vivo gene editing in mammalian cells. Edraki et al., Mol Cell (2019) “A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing”.

Also as disclosed herein, base editors (BEs) can be fused to Nme2Cas9 either as an N-
25 terminal or inlaid domain. Zhang et al., GEN Biotechnology (2022) “Adenine Base Editing In Vivo with a Single Adeno-Associated Virus Vector”. Base editors include, but are not limited to adenine base editors (ABEs) or cytidine base editors (CBEs) facilitated by Nme2Cas9 fusion with an adenosine deaminase (ABE8e) or a cytosine deaminase. Richter et al., Nature Biotechnology (2020) “Phage-assisted evolution of an adenine base editor with improved Cas domain
30 compatibility and activity”; and Chatterjee et al. Nature Commun. (2020) “A Cas9 with PAM recognition for adenine dinucleotides”. These Nme2Cas9-ABE/CBE constructs comprise a wild-

type PAM-interacting domain (PID) that recognizes a CC dinucleotide. For example if a potential editing site is not positioned an appropriate distance from a CC dinucleotide, these Nme2Cas9 base editors are unable to bind and cleave at that site.

In one embodiment, the present invention contemplates an Nme2_Cas9-ABE/CBE that has undergone a protospacer interacting domain (PID) -swapping that alters the PAM specificity of the Nme2Cas9 protein. In one embodiment, a PID is removed from a first Cas9 nuclease and inserted into a second Cas9 nuclease. Closely related type II-C Cas9 orthologs have been reported to recognize diverse PAMs. doi.org/10.7554/eLife.77825.

The general concept of PID-swapping was demonstrated between closely related *Neisseria meningitidis* orthologues. (Edraki et al., 2019). Distantly related Cas9 orthologues were also shown to be tolerant of PID swapping. Hu et al. *Nucleic Acids Res.* (2021) "Discovery and engineering of small SlugCas9 with broad targeting range and high specificity and activity"; and Schmidt et al. The SmuCas9 PAM (N₄C) of SmuCas9 was characterized using in vitro cleavage assays. Lee et al., *mBio* (2018) Potent Cas9 Inhibition in Bacterial and Human Cells by AcrIIC4 and AcrIIC5 Anti-CRISPR Proteins. More recently, SmuCas9's N₄C PAM was confirmed using a cell-based assay and an Nsp2Cas9-SmuCas9 PID hybrid nuclease was developed. An Nme2Cas9-SmuCas9 PID hybrid nuclease was also reported. Wei et al., *eLife* (2022) 11, e77825.

In one embodiment, the present invention contemplates a chimeric Nme2Cas9 fusion protein comprising an SmuCas9 protospacer interacting domain (PID) and a nucleotide deaminase protein. In one embodiment, the nucleotide deaminase protein is an adenosine deaminase. In one embodiment, the adenosine deaminase is ABE8e. In one embodiment, the nucleotide deaminase protein is a cytosine deaminase. (Huang et al., *Nature Biotechnology* 2022; dx.doi.org/10.1038/s41587-022-01410-2. In one embodiment, the nucleotide deaminase protein is an N-terminal domain of the chimeric Nme2^{Smu}Cas9 fusion protein. In one embodiment, the nucleotide deaminase protein is an inlaid domain of the chimeric Nme2^{Smu}Cas9 fusions protein. In one embodiment, the SmuCas9 PID replaces a wild type Nme2Cas9 PID. Although it is not necessary to understand the mechanism of the invention, it is believed that Nme2^{Smu}Cas9 chimeric proteins as disclosed herein have a predicted DNA targeting motif (e.g., an N₄C PAM) and can mediate gene editing with N₄C PAM-targeting guide RNAs.

In one embodiment, the present invention contemplates a composition comprising a chimeric Nme2^{Smu}Cas9-ABE8e fusion proteins and an sgRNA. In one embodiment, the sgRNA

targets a human MeCP2 gene mutation. In one embodiment, the MeCP2 gene causes Rett syndrome.

In one embodiment, the present invention contemplates a method of treating Rett syndrome comprising administering to a patient a composition comprising a chimeric Nme2^{Snu}Cas9-ABE8e fusion protein and an sgRNA, wherein the sgRNA targets an MeCP2 gene mutation.

In one embodiment, the composition converts the MeCP2 gene mutation into a wild type sequence. For example, MeCP2 gene mutation-directed sgRNAs were administered with a chimeric Nme2^{Snu}Cas9-ABE83 fusion protein to a HEK293T cell line (Rett-PiggyBac) and/or Rett patient derived fibroblast cells (Rett-PDFs) harboring pathogenic MeCP2 mutant alleles including, but not limited to: i) c.502 C>T; p.R168X with four (4) sgRNAs; ii) c.763 C>T; p.R255X with two (2) sgRNAs; iii) c. 808 C>T; p.R270X with two (2) sgRNAs; and/or iv) c.916 C>T; p.R306C with two (2) sgRNAs.

Although it is not necessary to understand the mechanism of an invention, it is believed that PID chimeric Nme2Cas9 base editors comprising a cross-species (e.g., exogenous) PID have an expanded targeting scope as compared to an Nme2Cas9 base editor with a wild type (e.g., endogenous) PID. See, Figure 53A. For example, several chimeric Nme2Cas9^{Snu} constructs were created that also contain inlaid domains of a nucleotide base editor. See, Fig. 53B. These Nme2Cas9^{Snu} constructs were shown to have significant base editing activity and efficiency. See, Fig. 53C and Fig. 53D. Gene editing activity was observed the at Linc01588 endogenous HEK293T genomic loci targeted to N₄CN PAMs. See, Fig. 53D.

Conversion of the Rett syndrome mutations were observed with the chimeric Nme2Cas9^{Snu} constructs: i) c.502 C>T (RETT-PDF); See, Figures 54A and 54B; ii) c.916 C>T (RETT-PDF); See, Figures 55A and 55B; iii) c.763C>T (RETT-PiggyBac); See, Figures 56A and 56B; and iv) c.808C>T (RETT-PiggyBac); See, Figures 57A and 57B.

Although it is not necessary to understand the mechanism of an invention, it is believed that an expanded PAM scope for chimeric Cas9^{Snu} nucleases concomitantly increase the number of candidate targets as compared to a wild type Cas9 nuclease. See, Figure 58A. It has been reported that PID replacement (e.g., swap) concomitantly changes the Cas9 nuclease PAM preference. See, Figure 58B. It has been reported that the SmuCas9 PID does interact with a PAM comprising a single C nucleotide. See, Figure 58C.

It is believed that no compact base editing platforms with single-nt PAMs have been previously reported. SpyCas9 and its BEs have been engineered to SpyCas9-NG, SpRY, and other versions with reduced PAM requirements that (on average) enable targeting every 2-4 nts. Nonetheless, size and off-targeting remain issues with those platforms. SmuCas9 was identified with an apparent single-nt (N4C) PAM requirement, but its native tracrRNA sequence has not been available, and its activity with tracrRNA sequences from related Cas9s was poor. Lee et al., “Potent Cas9 Inhibition in Bacterial and Human Cells by AcrIIC4 and AcrIIC5 Anti-CRISPR Proteins” 2018 Dec 4;9(6):e02321–18. However, further exploration of this single-nt PAM capability to fully exploit its usefulness is necessary, as described below. For example, PAM minimization not only allows more sites to be accessed but can also enable target nts to be edited by multiple guides. This can allow the definition of a guide register within the local sequence that avoids bystander editing.

A. A Nme2Cas9 Nuclease With A Single-Cytidine PAM Compatibility

Nme2Cas9 has been found to be effective at N₄CC PAMs across a broad editing window (nts ~2 to ~19 of the 24nt protospacer), with maximal activity between nts 6-17. Zhang et al., “Adenine Base Editing In Vivo with a Single Adeno-Associated Virus Vector” GEN Biotechnol 2022 Jun 1;1(3):285–299; and Davis et al., “Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors” Nat Biomed Eng. 2022 Jul 28;1–12.

A relatively wide editing window imparts a trade-off: more editing sites can be accessed, but bystander editing can also be exacerbated. Anzalone et al., “Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors” Nat Biotechnol 2020 Jul;38(7):824–844; and Rees et al., “Base editing: precision chemistry on the genome and transcriptome of living cells” Nat Rev Genet 2018 Dec;19(12):770–788. The ability to control Nme2-ABE bystander editing would be very useful for therapeutic and other applications. Finally, although the N₄CC PAM enables Nme2-ABE to access target sites that other AAV-validated compact Cas9s cannot, targeting range improvement via PAM minimization would be tremendously beneficial, including for minimizing bystander editing. The data presented herein address these three areas – efficiency improvement, bystander modulation, and PAM minimization – and indicate that all are achievable.

In one embodiment, the present invention contemplates Nme2-ABEs with an inserted SmuCas9 PAM-interacting domain (PID) that replaces the wild type Nme2 PID and enables

editing of sites via N₄C PAMs. An in vitro analyses showed that SmuCas9 has strong PAM preference of a single cytidine residue. Lee et al., “Tissue-restricted genome editing in vivo specified by microRNA-repressible anti-CRISPR proteins” RNA 2019 Nov;25(11):1421–1431. Further, PIDs from other Cas9 homologs, besides Smu can be transplanted into Nme1Cas9 (98% identical to Nme2Cas9 outside of the PID) to functionally reprogram PAM requirements. Edraki et al., “A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing” Mol Cell 2019 Feb 21;73(4):714–726.e4.

Results confirm that inserting a SmuCas9 PID into Nme2Cas9 reduces its PAM requirement from two cytidines to one cytidine. While Nme2-ABEs are largely inactive at N₄CA, N₄CG, and N₄CT PAMs, the i1, i7, and i8 inlaid Nme2Smu-ABEs efficiently edit many such sites. Fig. 53C. These results confirm that Nme2-ABEs remain the platforms of choice at N4CC PAM sites, but that Nme2Smu-ABEs can expand the range of targetable sites to N4CD (D = not C) PAMs.

B. A Single-AAV Nme2Cas9^{Smu}-ABE Construct

As shown above, the Nme2^{Smu}-ABE construct is functionally validated, but the SmuCas9 PID is 8 aa larger than the native Nme2Cas9 PID, increasing vector size by 24 nts (5,020 nts total). It is possible that even a modest 24 nt increase in cargo size might compromise packaging efficiency and integrity as well as delivery efficiency.

In one embodiment, the single AAV Nme2Cas9^{Smu}-ABE construct comprises an EFS promoter (212 bp). Data has been reported the EFS promoter is effective in driving expression from both Nme2-ABE and Nme2Smu-ABE vectors (4,957 and 4,981 nts, respectively). Davis et al., “Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors” Nat Biomed Eng 2022 Jul 28;1–12.

In one embodiment,, an all-in-one AAV9 construct includes, but is not limited to, Nme2^{Smu}-ABE [-i1(V106W), -i7(V106W), and -i8(V106W)] that target mouse N₄C PAM sites and measure editing activity in multiple tissues in vivo.

VI. Cas9 Therapeutic Plasmid/Vector Constructs

Although compact Cas9 orthologs have been previously validated for genome editing, including via single-AAV delivery, their longer PAMs have restricted therapeutic development due to target site frequencies that are lower than that of the more widely adopted SpyCas9. In

addition, SauCas9 and its KKH variant with relaxed PAM requirements (Kleinstiver et al., 2015) are prone to off-target editing with some sgRNAs (Friedland et al., 2015; Kleinstiver et al., 2015). These limitations are exacerbated with target loci that require editing within a narrow sequence window, or that require precise segmental deletion.

As described above, Nme2Cas9 has been identified as a compact and highly accurate Cas9 with a less restrictive dinucleotide PAM for genome editing by AAV delivery in vivo. The development of Nme2Cas9 greatly expands the genomic scope of in vivo editing, especially via viral vector delivery. The Nme2Cas9 all-in-one AAV delivery platform, can in principle, be used to target as wide a range of sites as SpyCas9 (due to the identical densities of optimal N₄CC and NGG PAMs), but without the need to deliver two separate vectors to the same target cells. The availability of a catalytically dead version of Nme2Cas9 (dNme2Cas9) also promises to expand the scope of applications such as CRISPRi, CRISPRa, base editing, and related approaches (Dominguez et al., 2016; Komor et al., 2017). Moreover, Nme2Cas9's hyper-accuracy enables precise editing of target genes, potentially ameliorating safety issues resulting from off-target activities. Perhaps counterintuitively, the higher target site density of Nme2Cas9 (compared to that of Nme1Cas9) does not lead to a relative increase in off-target editing for the former. Similar results have been reported recently with SpyCas9 variants evolved to have shorter PAMs (Hu et al., 2018). Type II-C Cas9 orthologs are generally slower nucleases in vitro than SpyCas9 (Ma et al., 2015; Mir et al., 2018); interestingly, enzymological principles indicate that a reduced apparent k_{cat} (within limits) can improve on- vs. off-target specificity for RNA-guided nucleases (Bisaria et al., 2017).

The discovery of Nme2Cas9 and Nme3Cas9 hinged on unexplored Cas9s that are highly related (outside of the PID) to an ortholog that was previously validated for human genome editing (Esvelt et al., 2013; Hou et al., 2013; Lee et al., 2016; Amrani et al., 2018). The relatedness of Nme2Cas9 and Nme3Cas9 to Nme1Cas9 brought an added benefit, namely that they use the exact same sgRNA scaffold, circumventing the need to identify and validate functional tracrRNA sequences for each. In the context of natural CRISPR immunity, the accelerated evolution of novel PAM specificities could reflect selective pressure to restore targeting of phages and MGEs that have escaped interference through PAM mutations (Deveau et al., 2008; Paez-Espino et al., 2015). Representative examples of plasmids/vectors and associated sequences that are compatible with NmeCas9 constructs are shown below. See, Table 5 and Table 6.

Table 5. Exemplary Plasmids For NmeCas9 Constructs

Plasmid #	Name	Insert description	Backbone	Purpose	Insert Sequence
1	pAE70	Nme3Cas9 PID on Nme1Cas9	pMCSG 7	Bacterial expression of Nme1Cas9 with Nme3Cas9 PID	See examples herein.
2	pAE71	Nme2Cas9 PID on Nme1Cas9	pMCSG 7	Bacterial expression of Nme1Cas9 with Nme2Cas9 PID	See examples herein.
3	pAE113	Nme2TLR1	pLKO	Targeting TLR2.0 with Nme2Cas9	GTCACCTGCCTCGT GGAATACGG
4	pAE114	Nme2TLR2	pLKO	Targeting TLR2.0 with Nme2Cas9	GCACCTGCCTCGTG GAATACGGT
5	pAE115	Nme2TLR5	pLKO	Targeting TLR2.0 with Nme2Cas9	GTTTCAGCGTGTCCG GCTTTGGC
6	pAE116	Nme2TLR11	pLKO	Targeting TLR2.0 with Nme2Cas9	GTGGTGAGCAAGG GCGAGGAGCTG
7	pAE117	Nme2TLR12	pLKO	Targeting TLR2.0 with Nme2Cas9	GGGCGAGGAGCTG TTCACCGGGGT
8	pAE118	Nme2TLR13	pLKO	Targeting TLR2.0 with Nme2Cas9	GTGAACTTGTGGCC GTTTACGTCG
9	pAE119	Nme2TLR14	pLKO	Targeting TLR2.0 with Nme2Cas9	GCGTCCAGCTCGAC CAGGATGGGC
10	pAE120	Nme2TLR15	pLKO	Targeting TLR2.0 with Nme2Cas9	GCGGTGAACAGCT CCTCGCCCTTG
11	pAE121	Nme2TLR16	pLKO	Targeting TLR2.0 with Nme2Cas9	GGGCACCACCCCG GTGAACAGCTC
12	pAE122	Nme2TLR17	pLKO	Targeting TLR2.0 with Nme2Cas9	GGCACCACCCCGGT GAACAGCTCC
13	pAE123	Nme2TLR18	pLKO	Targeting TLR2.0 with Nme2Cas9	GGGATGGGCACCA CCCCGGTGAAC
14	pAE124	Nme2TLR19	pLKO	Targeting TLR2.0 with Nme2Cas9	GCGTGTCCGGCTTT GGCGAGACAA
15	pAE125	Nme2TLR20	pLKO	Targeting TLR2.0 with Nme2Cas9	GTCCGGCTTTGGCG AGACAAATCA
16	pAE126	Nme2TLR21	pLKO	Targeting TLR2.0 with Nme2Cas9	GATCACCTGCCTCG TGGAATACGG
17	pAE149	Nme2TLR22	pLKO	Targeting TLR2.0 with Nme2Cas9	GACGCTGAACTTGT GGCCGTTTAC
18	pAE150	Nme2TLR23	pLKO	Targeting TLR2.0 with Nme2Cas9	GCCAAAGCCGGAC ACGCTGAACTT
19	pAE193	Nme2TLR13 with 23nt	pLKO	Targeting TLR2.0 with Nme2Cas9	GGAAGCTTGTGGCCG TTTACGTCG

		spacer			
20	pAE194	Nme2TLR13 with 22nt spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GAAC TTGTGGCCGT TTACGTCG
21	pAE195	Nme2TLR13 with 21nt spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GACTTGTGGCCGTT TACGTCG
22	pAE196	Nme2TLR13 with 20nt spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GCTTGTGGCCGTTT ACGTCG
23	pAE197	Nme2TLR13 with 19nt spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GTTGTGGCCGTTTA CGTCG
24	pAE213	Nme2TLR21 with G22 spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GTCACCTGCCTCGT GGAATACGG
25	pAE214	Nme2TLR21 with G21 spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GCACCTGCCTCGTG GAATACGG
26	pAE215	Nme2TLR21 with G20 spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GACCTGCCTCGTGG AATACGG
27	pAE216	Nme2TLR21 with G19 spacer	pLKO	Targeting TLR2.0 with Nme2Cas9	GCCTGCCTCGTGGA ATACGG
28	pAE90	Nme2TS1	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GGTTCTGGGTACTT TTATCTGTCC
29	pAE93	Nme2TS4	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GTCTGCCTAACAGG AGGTGGGGGT
30	pAE94	Nme2TS5	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GAATATCAGGAGA CTAGGAAGGAG
31	pAE129	Nme2TS6	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GCCTCCCTGCAGGG CTGCTCCC
32	pAE130	Nme2TS10	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GAGCTAGTCTTCTT CCTCCAACCC
33	pAE131	Nme2TS11	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GATCTGTCCCCTCC ACCCACAGT
34	pAE132	Nme2TS12	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GGCCCAAATGAAA GGAGTGAGAGG
35	pAE133	Nme2TS13	pLKO	Targeting <i>AAVS1</i> with Nme2Cas9	GCATCCTCTTGCTT TCTTTGCCTG
36	pAE136	Nme2TS16	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GGAGTCGCCAGAG GCCGGTGGTGG
37	pAE137	Nme2TS17	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GCCCAGCGGCCGG ATATCAGCTGC

38	pAE138	Nme2TS18	pLKO	Targeting <i>CYBB</i> with Nme2Cas9	GGAAGGGAACATA TTACTATTGC
39	pAE139	Nme2TS19	pLKO	Targeting <i>CYBB</i> with Nme2Cas9	GTGGAGTGGCCTGC TATCAGCTAC
40	pAE140	Nme2TS20	pLKO	Targeting <i>CYBB</i> with Nme2Cas9	GAGGAAGGGAACA TATTACTATTG
41	pAE141	Nme2TS21	pLKO	Targeting <i>CYBB</i> with Nme2Cas9	GTGAATTCTCATCA GCTAAAATGC
42	pAE144	Nme2TS25	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GCTCACTCACCCAC ACAGACACAC
43	pAE145	Nme2TS26	pLKO	Targeting <i>CFTR</i> with Nme2Cas9	GGAAGAATTTTCATT CTGTTCTCAG
44	pAE146	Nme2TS27	pLKO	Targeting <i>CFTR</i> with Nme2Cas9	GCTCAGTTTTCTG GATTATGCCT
45	pAE152	Nme2TS31	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GCGTTGGAGCGGG GAGAAGGCCAG
46	pAE153	Nme2TS34	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GGGCCGCGGAGAT AGCTGCAGGGC
47	pAE154	Nme2TS35	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GCCCACCCGGCGG CGCCTCCCTGC
48	pAE155	Nme2TS36	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GCGTGGCAGCTGAT ATCCGGCCGC
49	pAE156	Nme2TS37	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GCCGCGGCGCGAC GTGGAGCCAGC
50	pAE157	Nme2TS38	pLKO	Targeting <i>LINC01588</i> with Nme2Cas9	GTGCTCCCCAGCCC AAACCGCCGC
51	pAE159	Nme2TS41	pLKO	Targeting <i>AGA</i> with Nme2Cas9	GTCAGATTGGCTTG CTCGGAATTG
52	pAE185	Nme2TS44	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GCTGGGTGAATGG AGCGAGCAGCG
53	pAE186	Nme2TS45	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GTCCTGGAGTGACC CCTGGCCTTC
54	pAE187	Nme2TS46	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GATCCTGGAGTGAC CCCTGGCCTT
55	pAE188	Nme2TS47	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GTGTGTCCCTCTCC CCACCCGTCC
56	pAE189	Nme2TS48	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GTTGGAGCGGGGA GAAGGCCAGGG
57	pAE190	Nme2TS49	pLKO	Targeting <i>VEGFA</i> with Nme2Cas9	GCGTTGGAGCGGG GAGAAGGCCAG
58	pAE191	Nme2TS50	pLKO	Targeting <i>AGA</i> with Nme2Cas9	GTACCCTCCAATAA TTTGGCTGGC
59	pAE192	Nme2TS51	pLKO	Targeting <i>AGA</i> with Nme2Cas9	GATAATTTGGCTGG CAATTCCGAG
60	pAE232	TS64_FancJ1	pLKO	Targeting <i>FANCI</i>	GAAAATTGTGATTT

				with Nme2Cas9	CCAGATCCAC
61	pAE233	TS65_FancJ2	pLKO	Targeting <i>FANCI</i> with Nme2Cas9	GAGCAGAAAAAAT TGTGATTTC
62	pAE200	Nme2TS58 (Nme2DS1)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GCAGGGGCCAGGT GTCCTTCTCTG
63	pAE201	Nme2TS59 (Nme2DS2)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GAATGGCAGGCGG AGGTTGTACTG
64	pAE202	Nme2TS60 (Nme2DS3)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GAGTGAGAGAGTG AGAGAGAGACA
65	pAE203	Nme2TS61 (Nme2DS4)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GTGAGCAGGCACC TGTGCCAACAT
66	pAE204	Nme2TS62 (Nme2DS5)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GCGTGGGGGCTCC GTGCCCCACGC
67	pAE205	Nme2TS63 (Nme2DS6)	pLKO	Targeting DS in <i>VEGFA</i> with Nme2Cas9	GCATGGGCAGGGG CTGGGGTGCAC
68	pAE207	SpyDS1	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GGGCCAGGTGTCCT TCTCTG
69	pAE208	SpyDS2	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GGCAGGCGGAGGT TGTACTG
70	pAE209	SpyDS3	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GAGAGAGTGAGAG AGAGACA
71	pAE210	SpyDS4	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GCAGGCACCTGTGC CAACAT
72	pAE211	SpyDS5	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GGGGGCTCCGTGCC CCACGC
73	pAE212	SpyDS6	pLKO	Targeting DS in <i>VEGFA</i> with SpyCas9	GGGCAGGGGCTGG GGTGCAC
74	pAE169	hDeCas9 Wt in AAV backbone	AAV	Nme2Cas9 all-in-one AAV expression with sgRNA cassette	See examples herein.
75	pAE217	hDeCas9 wt in pMSCG7 backbone	pMCSG7	wildtype Nme2Cas9 for bacterial expression	See examples herein.

76	pAE107	2xNLS Nme2Cas9 with HA	pCdest	Nme2Cas9 CMV- driven expression plasmid	See examples herein.
77	pAE127	hDemonCas9 3X NLS in pMSCG7	pMSCG 7	Targeting endogenous loci with Nme2Cas9	See examples herein.
78	pAM172	hNme2Cas9 4X NLS with 3XHA	pCVL	Lentivector containing UCOE, SFFV driven Nme2Cas9 and Puro	See examples herein.
79	pAM174	nickase hNme2Cas9 D16A 4X NLS with 3XHA	pCVL	Lentivector containing UCOE, SFFV driven Nme2Cas9 and Puro	See examples herein.
80	pAM175	nickase hNme2Cas9 H588A 4X NLS with 3XHA	pCVL	Lentivector containing UCOE, SFFV driven Nme2Cas9 and Puro	See examples herein.
81	pAM177	dead hNme2Cas9 4X NLS with 3XHA	pCVL	Lentivector containing UCOE, SFFV driven Nme2Cas9 and Puro	See examples herein.

Table 6: Exemplary Oligonucleotide Sequences For NmeCas9 Constructs

Number	Name	Sequence	Purpose
1	AAVS1_TIDE1_FW	TGGCTTAGCACCTCTCCAT	TIDE analysis
2	LINC01588_TIDE_FW	AGAGGAGCCTTCTGACTGCTGCAGA	TIDE analysis
3	AAVS1_TIDE2_FW	TCCGTCTTCCTCCACTCC	TIDE analysis
4	NTS55_TIDE_FW	TAGAGAACTGGGTAGTGTG	TIDE analysis
5	VEGF_TIDE3_FW	GTACATGAAGCAACTCCAGTCCCA	TIDE analysis
6	hCFTR_TIDE1_FW	TGGTGATTATGGGAGAACTGGAGC	TIDE analysis
7	AGA_TIDE1_FW	GGCATAAGGAAATCGAAGGTC	TIDE analysis
8	VEGF_TIDE4_FW	ACACGGGCAGCATGGGAATAGTC	TIDE analysis
9	VEGF_TIDE5_FW	CCTGTGTGGCTTTGCTTTGGTCG	TIDE analysis
10	VEGF_TIDE6_FW	GGAGGAAGAGTAGCTCGCCGAGG	TIDE analysis
11	VEGF_TIDE7_FW	AGGGAGAGGGAAGTGTGGGGAAGG	TIDE analysis
12	AAVS1_TIDE1_RV	AGAACTCAGGACCAACTTATTCTG	TIDE analysis
13	LINC01588_TIDE_RV	ATGACAGACACAACCAGAGGGCA	TIDE analysis
14	AAVS1_TIDE2_RV	TAGGAAGGAGGAGGCCTAAG	TIDE analysis

15	NTS55_TIDE_RV	CCAATATTGCATGGGATGG	TIDE analysis
16	VEGF_TIDE3_RV	ATCAAATTCACGACCGAGCGC	TIDE analysis
17	hCFTR_TIDE1_RV	ACCATTGAGGACGTTTGTCTCAC	TIDE analysis
18	AGA_TIDE1_RV	CATGTCCTCAAGTCAAGAACAAG	TIDE analysis
19	VEGF_TIDE4_RV	GCTAGGGGAGAGTCCCAGTGTCCA	TIDE analysis
20	VEGF_TIDE5_RV	GTAGGGTGTGATGGGAGGCTAAGC	TIDE analysis
21	VEGF_TIDE6_RV	AGACCGAGTGGCAGTGACAGCAAG	TIDE analysis
22	VEGF_TIDE7_RV	GTCTTCCTGCTCTGTGCGCACGAC	TIDE analysis
23	Random PAM_FW	TAGCGGCCGCTCATGCGCGGCGCATT ACCTTTACNNNNNNNNNGGATCCT CTAGAGTCG	Protospacer with randomized PAM
24	Random PAM_RV	ACAGGAAACAGCTATGACCATGAAA GCTTGCATGCCTGCAGGTCGACTCTA GAGGATC	Protospacer with randomized PAM
25	DS2_ON_FW1	ctacacgacgctcttccgatctCCTGGAGCGTGT ACGTTGG	Targeted Deep Seq
26	SpyDS2_OT1_FW1	ctacacgacgctcttccgatctCCTGTGGTCCCAG CTACTTG	Targeted Deep Seq
27	SpyDS2_OT2_FW1	ctacacgacgctcttccgatctATCTGCGATGTCC TCGAGG	Targeted Deep Seq
28	SpyDS2_OT3_FW1	ctacacgacgctcttccgatctTGGTGTGCGCCTC TAACG	Targeted Deep Seq
29	SpyDS2_OT4_FW1	ctacacgacgctcttccgatctGGAGTCTTGCTTT GTCCTCAGA	Targeted Deep Seq
30	SpyDS2_OT5_FW1	ctacacgacgctcttccgatctAGCCTAGACCCA GTCCCAT	Targeted Deep Seq
31	SpyDS2_OT6_FW1	ctacacgacgctcttccgatctGCTGGGCATAGT AGTGGACT	Targeted Deep Seq
32	SpyDS2_OT7_FW1	ctacacgacgctcttccgatctTGGGGAGGCTGA GACACGA	Targeted Deep Seq
33	SpyDS2_OT8_FW1	ctacacgacgctcttccgatctCTTGGGAGGCTG AGGCAAG	Targeted Deep Seq
34	DS2_ON_RV1	agacgtgtgctcttccgatctCAGGAGGATGAGA GCCAGG	Targeted Deep Seq
35	SpyDS2_OT1_RV1	agacgtgtgctcttccgatctCAGGGTCTCACTC TATCACCCA	Targeted Deep Seq
36	SpyDS2_OT2_RV1	agacgtgtgctcttccgatctACTGAATGGGTTG AACTTGGC	Targeted Deep Seq
37	SpyDS2_OT3_RV1	agacgtgtgctcttccgatctGAGACAGAATCTT GCTCTGTCTCC	Targeted Deep Seq
38	SpyDS2_OT4_RV1	agacgtgtgctcttccgatctTCCCAGCTACTTG GGAGGC	Targeted Deep Seq

39	SpyDS2_OT5_RV1	agacgtgtgctcttccgatctCCTGCCCAAATAG GGAAGCAG	Targeted Deep Seq
40	SpyDS2_OT6_RV1	agacgtgtgctcttccgatctTGGCGCCTTAGTCT CTGCTAC	Targeted Deep Seq
41	SpyDS2_OT7_RV1	agacgtgtgctcttccgatctGCATGAGACACAG TTTCACTCTG	Targeted Deep Seq
42	SpyDS2_OT8_RV1	agacgtgtgctcttccgatctGAGAGAGTCTCAC TGC GTTGC	Targeted Deep Seq
43	DS4_ON_FW3	ctacacgacgctcttccgatctTCTCTCACCCACT GGGCAC	Targeted Deep Seq
44	DS4_ON_RV3	agacgtgtgctcttccgatctGCTTCCAGACGAG TGCAGA	Targeted Deep Seq
45	SpyDS4_OT1_FW1	ctacacgacgctcttccgatctAAGTTTTCAAACC AGAAGAACTACGAC	Targeted Deep Seq
46	SpyDS4_OT2_FW1	ctacacgacgctcttccgatctCCGGTATAAGTCC TGGAGCG	Targeted Deep Seq
47	SpyDS4_OT3_FW1	ctacacgacgctcttccgatctGCCAGGGAGCAA TGGCAG	Targeted Deep Seq
48	SpyDS4_OT6_FW1	ctacacgacgctcttccgatctCCTCGAATTCCAC GGGGTT	Targeted Deep Seq
49	DS16_ON_FW1	ctacacgacgctcttccgatctGTTGGTGGGAGG GAAGTGAG	Targeted Deep Seq
50	SpyDS6_OT1_FW1	ctacacgacgctcttccgatctGATGGCGGTTGT AGCGGC	Targeted Deep Seq
51	SpyDS6_OT2_FW1	ctacacgacgctcttccgatctCACATAAACCTAT GTTTCAGCAGA	Targeted Deep Seq
52	SpyDS6_OT3_FW1	ctacacgacgctcttccgatctGCTAGTTGGATTG AAGCAGGGT	Targeted Deep Seq
53	SpyDS6_OT4_FW1	ctacacgacgctcttccgatctTTGAGTGCGGCA GCTTCC	Targeted Deep Seq
54	SpyDS6_OT6_FW1	ctacacgacgctcttccgatctATAACCCTCCCAG GCAAAGTC	Targeted Deep Seq
55	SpyDS6_OT7_FW1	ctacacgacgctcttccgatctAGCCTGCACATCT GAGCTC	Targeted Deep Seq
56	SpyDS6_OT8_FW1	ctacacgacgctcttccgatctGGAGCATTGAAG TGCCTGG	Targeted Deep Seq
57	DeDS6_ON_RV1	agacgtgtgctcttccgatctCAGCCTGGGACCA CTGA	Targeted Deep Seq
58	SpyDS6_OT1_RV1	agacgtgtgctcttccgatctCATCCTCGACAGT CGCGG	Targeted Deep Seq
59	SpyDS6_OT2_RV1	agacgtgtgctcttccgatctGACTGATCAAGTA GAATACTCATGGG	Targeted Deep Seq
60	SpyDS6_OT3_RV1	agacgtgtgctcttccgatctCCCTGCCAGCACT GAAGC	Targeted Deep Seq
61	SpyDS6_OT4_Rv1	agacgtgtgctcttccgatctGGTTCCTATCTTTC	Targeted Deep Seq

		TAGACCAGGAGT	
62	SpyDS6_OT6_RV1	agacgtgtgctcttccgatctAGTGTGGAGGGCT CAGGG	Targeted Deep Seq
63	SpyDS6_OT7_RV1	agacgtgtgctcttccgatctGATGGGCAGAGGA AGGCAA	Targeted Deep Seq
64	SpyDS6_OT8_RV1	agacgtgtgctcttccgatctTCACTCTCATGAG CGTCCCA	Targeted Deep Seq
65	Nme2DS2_OT1_FW1	ctacacgacgctcttccgatctAAGGTTCTTGCG GTTCGC	Targeted Deep Seq
66	Nme2DS2_OT1_RV1	agacgtgtgctcttccgatctCGCTGCCATTGCTC CCT	Targeted Deep Seq
67	Nme2DS6_OT1_FW1	ctacacgacgctcttccgatctTCTCGCACATTCT TCACGTCC	Targeted Deep Seq
68	Nme2DS6_OT1_RV1	agacgtgtgctcttccgatctAGGAACCTTCCCG ACTTAGGG	Targeted Deep Seq
69	Rosa26_ON_FW1	ctacacgacgctcttccgatctCCCGCCCATCTTC TAGAAAGAC	Targeted Deep Seq
70	Rosa26_OT1_FW1	ctacacgacgctcttccgatctTGCCAGGTGAGG GACTGG	Targeted Deep Seq
71	Rosa26_ON_RV1	agacgtgtgctcttccgatctTCTGGGAGTTCTCT GCTGCC	Targeted Deep Seq
72	Rosa26_OT1_RV1	agacgtgtgctcttccgatctTGCCCAACCTTAG CAAGGAG	Targeted Deep Seq
73	pCSK9_ON_FW2	ctacacgacgctcttccgatcttaccttgagcaacggcg	Targeted Deep Seq
74	PCSK9_ON_RV2	agacgtgtgctcttccgatctcccaggacgaggatggag	Targeted Deep Seq
75	Tyr_500_FW3	GATAGTCACTCCAGGGGTTG	TIDE analysis
76	Tyr_500_RV3	GTGGTGAACCAATCAGTCCT	TIDE analysis

Experimental

Example I

Discovery of Cas9 orthologs with differentially diverged PIDs

5 Nme1Cas9 peptide sequence was used as a query in BLAST searches to find all Cas9 orthologs in *Neisseria meningitidis* species. Orthologs with >80% identity to Nme1Cas9 were selected for the remainder of this study. The PIDs were then aligned with that of Nme1Cas9 (residues 820-1082) using ClustalW2 and those with clusters of mutations in the PID were selected for further analysis. An unrooted phylogenetic tree of NmeCas9 orthologs was constructed using
10 FigTree (tree.bio.ed.ac.uk/software/figtree/).

Example II

In vitro PAM discovery assay

A dsDNA target library with randomized PAM sequences was generated by overlapping
15 PCR, with the forward primer containing the 10-nt randomized PAM region. The library was gel-purified and subjected to in vitro cleavage reaction by purified Cas9 along with T7-transcribed sgRNAs. 300 nM Cas9:sgRNA complex was used to cleave 300 nM of the target fragment in 1X NEBuffer 3.1 (NEB) at 37°C for 1 hr. The reaction was then treated with proteinase K at 50°C for 10 minutes and run on a 4% agarose/1xTAE gel. The cleavage product was excised, eluted, and
20 cloned using a previously described protocol (Zhang et al., 2012), with modifications. Briefly, DNA ends were repaired, non-templated 2'-deoxyadenosine tails were added, and Y-shaped adapters were ligated. After PCR, the product was quantitated with KAPA Library Quantification Kit and sequenced using a NextSeq 500 (Illumina) to obtain 75 nt paired-end reads. Sequences were analyzed with custom scripts and R.

25

Example IV

Transfections and mammalian genome editing

Human codon-optimized Nme2Cas9 was cloned by Gibson Assembly into the pCDEST2 plasmid backbone previously used for Nme1Cas9 and SpyCas9 expression (Pawluk et al., 2016;
30 Amrani et al., 2018). Transfection of HEK293T and HEK293T-TLR2.0 cells was performed as previously described (Amrani et al., 2018). For Hepa1-6 transfections, Lipofectamine LTX was

used to transfect 500ng of all-in-one AAV.sgRNA.Nme2Cas9 plasmid in 24-well plates ($\sim 10^5$ cells/well), using cells that had been cultured 24 hours before transfection. For K562 cells stably expressing Nme2Cas9 delivered via lentivector (see below), 50,000 – 150,000 cells were electroporated with 500 ng sgRNA plasmid using 10 μ L Neon tips. To measure indels in all cells 72 hr after transfections, cells were harvested and genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen). The targeted locus was amplified by PCR, Sanger-sequenced (Genewiz), and analyzed by TIDE (Brinkman et al., 2014) using the Desktop Genetics web-based interface (tide.deskgen.com).

Example V

Lentiviral transduction of K562 cells to stably express Nme2Cas9

K562 cells stably expressing Nme2Cas9 were generated as previously described for Nme1Cas9 (Amrani et al., 2018). For lentivirus production, the lentiviral vector was co-transfected into HEK293T cells along with the packaging plasmids (Addgene 12260 & 12259) in 6-well plates using TransIT-LT1 transfection reagent (Mirus Bio). After 24 hours, culture media was aspirated from the transfected cells and replaced with 1 mL of fresh DMEM. The next day, the supernatant containing the virus was collected and filtered through a 0.45 μ m filter. 10 μ L of the undiluted supernatant along with 2.5 μ g of Polybrene was used to transduce $\sim 10^6$ K562 cells in 6-well plates. The transduced cells were selected using media supplemented with 2.5 μ g/mL puromycin.

Example VI

RNP Delivery for mammalian genome editing

For RNP experiments, the Neon electroporation system was used exactly as described (Amrani et al., 2018). Briefly, 40 picomoles of 3xNLS-Nme2Cas9 along with 50 picomoles of T7-transcribed sgRNA was assembled in buffer R and electroporated using 10 μ L Neon tips. After electroporation, cells were plated in pre-warmed 24-well plates containing the appropriate culture media without antibiotics. Electroporation parameters (voltage, width, number of pulses) were 1150 V, 20 ms, 2 pulses for HEK293T cells; 1000 V, 50 ms, 1 pulse for K562 cells.

Example VII

GUIDE-seq

GUIDE-seq experiments were performed as described previously (Tsai et al., 2014), with minor modifications (Bolukbasi et al., 2015a). Briefly, HEK293T cells were transfected with 200 ng of Cas9 plasmid, 200 ng of sgRNA plasmid, and 7.5 pmol of annealed GUIDE-seq oligonucleotides using Polyfect (Qiagen). Alternatively, Hepa1-6 cells were transfected as described above. Genomic DNA was extracted with a DNeasy Blood and Tissue kit (Qiagen) 72 h after transfection according to the manufacturer's protocol. Library preparation and sequencing were performed exactly as described previously (Bolukbasi et al., 2015a). For analysis, all sequences with up to ten mismatches with the target site, as well as a C in the fifth PAM position (N₄CN), were considered potential off-target sites. Data were analyzed using the Bioconductor package GUIDEseq version 1.1.17 (Zhu et al., 2017).

Example VIII

Targeted deep sequencing and analysis

We used targeted deep sequencing to confirm the results of GUIDE-seq and to measure indel rates with maximal accuracy. We used two-step PCR amplification to produce DNA fragments for each on- and off-target site. For SpyCas9 editing at DS2 and DS6, we selected the top off-target sites based on GUIDE-seq read counts. For SpyCas9 editing at DS4, fewer candidate off-target sites were identified by GUIDE-seq, and only those with NGG (DS4|OT1, DS4|OT3, DS4|OT6) or NGC (DS4|OT2) PAMs were examined by sequencing. In the first step, we used locus-specific primers bearing universal overhangs with ends complementary to the adapters. In the first step, 2x PCR master mix (NEB) was used to generate fragments bearing the overhangs. In the second step, the purified PCR products were amplified with a universal forward primer and indexed reverse primers. Full-size products (~250 bp) were gel-purified and sequenced on an Illumina MiSeq in paired-end mode. MiSeq data analysis was performed as previously described (Pinello et al., 2016; Ibraheim et al., 2018).

Example IX

Off-target analysis using CRISPRseek

Global off-target predictions for TS25 and TS47 were performed using the Bioconductor package CRISPRseek. Minor changes were made to accommodate characteristics of Nme2Cas9 not shared with SpyCas9. Specifically, we used the following changes to: gRNA.size = 24, PAM = "NNNNCC", PAM.size = 6, RNA.PAM.pattern = "NNNNCN", and candidate off-target sites with fewer than 6 mismatches were collected. The top potential off-target sites based on the numbers and positions of mismatches were selected. Genomic DNA from cells targeted by each respective sgRNA was used to amplify each candidate off-target locus and then analyzed by TIDE.

10

Example X

Mouse strains and embryo collection

All animal experiments were conducted under the guidance of the Institutional Animal Care and Use Committee (IACUC) of the University of Massachusetts Medical School. C57BL/6NJ (Stock No. 005304). Mice were obtained from The Jackson Laboratory. All animals were maintained in a 12 h light cycle. The middle of the light cycle of the day when a mating plug was observed was considered embryonic day 0.5 (E0.5) of gestation. Zygotes were collected at E0.5 by tearing the ampulla with forceps and incubation in M2 medium containing hyaluronidase to remove cumulus cells.

20

Example XI

In vivo AAV8.Nme2Cas9+sgRNA delivery and liver tissue processing

For the AAV8 vector injections, 8-week-old female C57BL/6NJ mice were injected with 4×10^{11} genome copies per mouse via tail vein, with the sgRNA targeting a validated site in either *Pcsk9* or *Rosa26*. Mice were sacrificed 28 days after vector administration and liver tissues were collected for analysis. Liver tissues were fixed in 4% formalin overnight, embedded in paraffin, sectioned and stained with hematoxylin and eosin (H&E). Blood was drawn from the facial vein at 0, 14 and 28 days post injection, and serum was isolated using a serum separator (BD, Cat. No. 365967) and stored at -80°C until assay. Serum cholesterol level was measured using the Infinity™ colorimetric endpoint assay (Thermo-Scientific) following the manufacturer's protocol and as previously described (Ibraheim et al., 2018). For the anti-PCSK9 Western blot, 40 µg of

30

protein from tissue or 2 ng of Recombinant Mouse PCSK9 Protein (R&D Systems, 9258-SE-020) were loaded onto a MiniPROTEAN® TGX™ Precast Gel (Bio-Rad). The separated bands were transferred onto a PVDF membrane and blocked with 5% Blocking-Grade Blocker solution (Bio-Rad) for 2 hours at room temperature. Next, the membrane was incubated with rabbit anti-GAPDH (Abcam ab9485, 1:2,000) or goat anti-PCSK9 (R&D Systems AF3985, 1:400) antibodies overnight. Membranes were washed in TBST and incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit (Bio-Rad 1706515, 1:4,000), and donkey anti-goat (R&D Systems HAF109, 1:2,000) secondary antibodies for 2 hours at room temperature. The membranes were washed again in TBST and visualized using Clarity™ western ECL substrate (Bio-Rad) using an M35A XOMAT Processor (Kodak).

Example XII

Ex vivo AAV6.Nme2Cas9 delivery in mouse zygotes

Zygotes were incubated in 15 µl drops of KSOM (Potassium-Supplemented Simplex Optimized Medium, Millipore, Cat. No. MR-106-D) containing 3×10^9 or 3×10^8 GCs of AAV6.Nme2Cas9.sgTyr vector for 5-6 h (4 zygotes in each drop). After incubation, zygotes were rinsed in M2 and transferred to fresh KSOM for overnight culture. The next day, the embryos that advanced to 2-cell stage were transferred into the oviduct of pseudopregnant recipients and allowed to develop to term.

Example XIII

Ribonucleoprotein (RNP) Cas9 Delivery

For RNP experiments, the Neon electroporation system was used exactly as described (Amrani et al., 2018). Briefly, 40 picomoles of 3xNLS-Nme2Cas9 along with 50 picomoles of T7-transcribed sgRNA was assembled in buffer R and electroporated using 10 µL Neon tips. After electroporation, cells were plated in pre-warmed 24-well plates containing the appropriate culture media without antibiotics. Electroporation parameters (voltage, width, number of pulses) were 1150 V, 20 ms, 2 pulses for HEK293T cells; 1000 V, 50 ms, 1 pulse for K562 cells.

Example IX

Nme2Cas9 Plasmid Insert Construction

SV40 NLS (yellow- **BOLD**); 3X-HA-Tag (green-(**underlined/bold**); cMyc-like NLS (teal- plain); Linker (magenta - ***bold italics***) and *Nme2Cas9* (*italics*).

5
 MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPKTGDSLAMARRLA
 RSVRRLTRRRRAHRLLRARLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSA
 VLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEEKESGHIR
 10 NQRGDYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLG
 HCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQGSEPLTDTERATLMDEPYRKSCLTYAQAR
 KLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTA
 FSLFKTDEDITGRLKDRVQPEILEALLKHISFDKDFVQISLKALRRIVPLMEQKGKRYDEACAEIYGD
 HYGKKNTEEKIYLPPIPADEIRNPVVLRLALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRK
 EIEKRQEENRKDREKAAAKFREYFPNFVGEPSKSDILKLRLYEQQHKGKCLYSGKEINLVRLNEK
 15 GYVEIDHALPFSRTWDDSFNNKVLVLGSENNKGNQTPYEYFNGKDNSREWQEFKARVETSRF
 PRSKKQORILLQKFDEDDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFAVNGQITNLLRG
 FWGLRKVRAENDRHHALDAVVVACSTVAMQOKITRFVRYKEMNAFDGKTIDKETGKVLHQKT
 HFPQPWEFFAQEV MIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAP
 NRKMSGAHKDTLRSARKI'VKHNEKISVKRVWLTEIKLADLENNMVNYKNGRIEELYEALKARLEA
 20 YGGNAKQAFDPKDNPFYKKGQQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCVKV
 DKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYI
 NCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVRGTGGPK
KKRKVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSAAPAAKKKKLDFESG*

25 Example X

Nme2Cas9 AAV Insert Construction

SV40 NLS (yellow- **BOLD**); 3X-HA-Tag (green-(**underlined/bold**); Nucleoplasmin-like NLS (red-underline); c-myc NLS (teal- plain); Linker (magenta - ***bold italics***) and *Nme2Cas9* (*italics*).

30 **MV****PKKKRKVED****K**RPAATKKAGQAKKKKMAAFKPNPINYILGLDIGIASVGWAMVEIDEEE
 NPIRLIDLGVRFERAEVPKTGDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFD
 ENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGV
 ANNAHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLLFEKQKEFGNP
 35 HVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRIL
 EQGSEPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKA
 YHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFD
 KFVQISLKALRRIVPLMEQKGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRLALSQ
 RKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPS
 40 KSKDILKLRLYEQQHKGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENN
 KGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQORILLQKFDEDDGFKECNLNDTRYVNR
 FLCQFVADHILLTGKGKRRVFAVNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQ

QKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEA
 DTPEKLRITLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFRVKNHNEKISVKRV
 WLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKAVRV
 EKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKG
 5 YRIDDSTYFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNL
 VLIQKYQVNELGKEIRPCRLKKRPPVREDKRPAATKKAGQAKKKKKYPYDVPDYAGYPYDV
PDYAGSYPYDVPDYA11PAAKKKKLD*

Example XI

10 Recombinant Nme2Cas9 Construction
 SV40 NLS (yellow- **BOLD**); Nucleoplasmin-like NLS (red-underline); Linker (magenta - **bold italics**) and *Nme2Cas9* (*italics*).
 15 **PKKKRKVN**AMAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPKTG
 DSLAMARRLARSVRRLTRRRHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALD
 RKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALN
 KFEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALS
 GDAVQKMLGHCTFEPAEPKAAKNITYTAERFIWLTKLNNLRILEQGSERPLTDITERATIMDIEPYR
 KSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPNLNLS
 20 SELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFFVQISLKALRRIVPLMEQ GKRY
 DEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRLALSQARKVINGVVRRYGSPARIHIETAR
 EVGKSFKDRKEIEKRQEEENRKDREKAAAFREYFPNFVGEPKSKDILKRLYEQQHKGKCLYSG
 KEINLVRNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQ
 EFKARVETSRFPRSKKQRIILLQKFDEDEGDFKECNLNDTRYVNRFLCQFVADHILLTGKGRRVFA
 25 SNGQITNLLRGFWGLRKVRAENDRHHALDAVVACSTVAMQKITRFVRYKEMNAFDGKTIDK
 ETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFFTEADTPEKLRITLLAEKLSSRPEAVHE
 YVTPLFVSRAPNRKMSGAHKDTLRSKRFRVKNHNEKISVKRVWLTEIKLADLENMVNYKNGREIE
 LYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKAVRVEKTQESGVLLNKKNAYTIADNGD
 MVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSTYFCFSLHKYDLIAFQKD
 30 EKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRP
 PVRGGGGSGGGGSGGGGSPA**AKKKKKLDGGGSKRPAATKKAGQAKKKK***

Example XII

35 RNP-Recombinant Nme2Cas9 Construction
 SV40 NLS (yellow- **BOLD**); Nucleoplasmin-like NLS (red-underline); Linker (magenta - **bold italics**) and *Nme2Cas9* (*italics*).
 40 **PKKKRKVN**AMAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPKTG
 DSLAMARRLARSVRRLTRRRHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALD

RKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALN
 KFEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALS
 GDAVQKMLGHCTFEPAPKAAKNTYTAERFIWLTCLNNLRILEQGSERPLTDTERATLMDEPYR
 KSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLS
 5 SELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFFVQISLKALRRIVPLMEQGKRY
 DEACAEIYGDHYGKKNTTEEKIYLPPIPADEIRNPVLRALSQARKVINGVVRRYGSPARIHIETAR
 EVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHKGKCLYSG
 KEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQ
 EFKARVETSRFPRSKKQORILLQKFDEDEGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFA
 10 SNGQITNLLRGFWGLRKVRAENDRRHALDAVVACSTVAMQOKITRFVRYKEMNAFDGKTIDK
 ETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRLLAEKLSSRPEAVHE
 YVTPLFVSRAPNRKMSGAHKDTLRSARFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIE
 LYEALKARLEAYGGNAKQAFDPKDNPFYKKGQVLKAVRVEKTQESGVLLNKKNAYTIADNGD
 MVRVDVFCVKVDKKGKNQYFVPIYAWQVAENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKD
 15 EKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRP
 PVRGGGGSGGGGSGGGGSPAAKKKKLDGGGSKRPAATKKAGOAKKKK*

REFERENCES

20 Each is herein incorporated by reference in their entirety:

- Amrani, N., Gao, X.D., Liu, P., Edraki, A., Mir, A., Ibraheim, R., Gupta, A., Sasaki, K.E., Wu, T., Donohoue, P.D., et al. (2018). NmeCas9 is an intrinsically high-fidelity genome editing platform. *BioRxiv*, doi.org/10.1101/172650.
- 25 Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709-1712.
- Bisaria, N., Jarmoskaite, I., and Herschlag, D. (2017). Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. *Cell Syst.* 4, 21-29.
- 30 Bolukbasi, M.F., Gupta, A., Oikemus, S., Derr, A.G., Garber, M., Brodsky, M.H., Zhu, L.J., and Wolfe, S.A. (2015a). DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat. Methods* 12, 1150-1156.
- Bolukbasi, M.F., Gupta, A., and Wolfe, S.A. (2015b). Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. *Nat. Methods* 13, 41-50.
- 35

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 42, e168.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964.

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., et al. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36, 265-271.

Certo, M.T., Ryu, B.Y., Annis, J.E., Garibov, M., Jarjour, J., Rawlings, D.J., and Scharenberg, A.M. (2011). Tracking genome engineering outcome at individual DNA breakpoints. *Nat. Methods* 8, 671-676.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407-410.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 230-232.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132-141.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607.

Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390-1400.

Dominguez, A.A., Lim, W.A., and Qi, L.S. (2016). Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* 17, 5-15.

Dong, Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., and Huang, Z. (2017). Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* 546, 436-439.

5 Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116-1121.

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lecrivain, A.L., Bzdrenga, J., Koonin, E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 2577-2590.

10 Friedland, A.E., Baral, R., Singhal, P., Loveluck, K., Shen, S., Sanchez, M., Marco, E., Gotta, G.M., Maeder, M.L., Kennedy, E.M., et al. (2015). Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol.* 16, 257.

15 Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* 5, 1513-1523.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279-284.

20 Gallagher, D.N., and Haber, J.E. (2018). Repair of a Site-Specific DNA Cleavage: Old-School Lessons for Cas9-Mediated Gene Editing. *ACS Chem. Biol.* 13, 397-405.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67-71.

25 Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* 109, E2579-2586.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471.

Ghanta, K., Dokshin, G., Mir, A., Krishnamurthy, P., Gneid, H., Edraki, A., Watts, J., Sontheimer, E., and Mello, C. (2018). 5' Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. *Biorxiv* 354480.

5 Gorski, S.A., Vogel, J., and Doudna, J.A. (2017). RNA-based recognition and targeting: sowing the seeds of specificity. *Nat. Rev. Mol. Cell Biol.* 18, 215-228.

Harrington, L.B., Doxzen, K.W., Ma, E., Liu, J.J., Knott, G.J., Edraki, A., Garcia, B., Amrani, N., Chen, J.S., Cofsky, J.C., et al. (2017a). A Broad-Spectrum Inhibitor of CRISPR-Cas9. *Cell* 170, 1224-1233.

10 Harrington, L.B., Paez-Espino, D., Staahl, B.T., Chen, J.S., Ma, E., Kyrpides, N.C., and Doudna, J.A. (2017b). A thermostable Cas9 with increased lifetime in human plasma. *Nat. Commun.* 8, 1424.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA* 110, 15644-15649.

15 Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57-63.

20 Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227-229.

Hynes, A.P., Rousseau, G.M., Lemay, M.-L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. *Nat. Microbiol.* 2, 1374-1380.

25 Ibraheim, R., Song, C.-Q., Mir, A., Amrani, N., Xue, W., and Sontheimer, E.J. (2018). All-in-One Adeno-associated Virus Delivery and Genome Editing by *Neisseria meningitidis* Cas9 in vivo. *BioRxiv*, doi.org/10.1101/295055.

Jiang, F., and Doudna, J.A. (2017). CRISPR–Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* 46, 505-529.

30 Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233-239.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed
5 genome editing in human cells. *eLife* 2, e00471.

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol.* 16, 253.

Keeler, A.M., ElMallah, M.K., and Flotte, T.R. (2017). Gene Therapy 2017: Progress and
10 Future Directions. *Clin. Transl. Sci.* 10, 242-248.

Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K.-E., Kim, K., Cho, H.-Y., Song, D.W., Lee, K.J., Jung, M.H., et al. (2017). In vivo genome editing with a small Cas9 ortholog derived from *Campylobacter jejuni*. *Nat. Commun.* 8, 14500.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided
15 genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012-1019.

Kim, B., Komor, A., Levy, J., Packer, M., Zhao, K., and Liu, D. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nature Biotechnology* 35.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Topkar, V.V., Zheng, Z., and
20 Joung, J.K. (2015). Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* 33, 1293-1298.

Kluesner, M., Nedveck, D., Lahr, W., Garbe, J., Abrahante, J., Webber, B., and Moriarity, B. (2018). EditR: A Method to Quantify Base Editing from Sanger Sequencing. *The CRISPR*
25 *Journal* 1, 239–250.

Koblan, L., Doman, J., Wilson, C., Levy, J., Tay, T., Newby, G., Maianti, J., Raguram, A., and Liu, D. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* 36, 843.

Komor, A.C., Badran, A.H., and Liu, D.R. (2017). CRISPR-Based Technologies for the
30 Manipulation of Eukaryotic Genomes. *Cell* 168, 20-36.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424.

Lee, C.M., Cradick, T.J., and Bao, G. (2016). The *Neisseria meningitidis* CRISPR-Cas9 system enables specific genome editing in mammalian cells. *Mol. Ther.* 24, 645-654.

Lee, J., Mir, A., Edraki, A., Garcia, B., Amrani, N., Lou, H.E., Gainetdinov, I., Pawluk, A., Ibraheim, R., Gao, X.D., et al. (2018). Potent Cas9 inhibition in bacterial and human cells by new anti-CRISPR protein families. *BioRxiv*, biorxiv.org/content/early/2018/2006/2020/350504.

Ma, E., Harrington, L.B., O'Connell, M.R., Zhou, K., and Doudna, J.A. (2015). Single-Stranded DNA Cleavage by Divergent CRISPR-Cas9 Enzymes. *Mol. Cell* 60, 398-407.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833-838.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. *Science* 339, 823-826.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843-1845.

Mir, A., Edraki, A., Lee, J., and Sontheimer, E.J. (2018). Type II-C CRISPR-Cas9 biology, mechanism and application. *ACS Chem. Biol.* 13, 357-365.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733-740.

Paez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B.C., Barrangou, R., and Banfield, J.F. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *mBio* 6.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., et al. (2016). Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167, 1829-1838 e1829.

Pawluk, A., Bondy-Denomy, J., Cheung, V.H., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* 5, e00896.

Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., and Yuan, G.C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat. Biotechnol.* 34, 695-697.

Racanelli, V., and Rehmann, B. (2006). The liver as an immunological organ.

5 *Hepatology* 43, S54-62.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186-191.

10 Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389.

Rashid, S., Curtis, D.E., Garuti, R., Anderson, N.N., Bashmakov, Y., Ho, Y.K., Hammer, R.E., Moon, Y.A., and Horton, J.D. (2005). Decreased plasma cholesterol and hypersensitivity to statins in mice lacking *Pcsk9*. *Proc. Natl. Acad. Sci. USA* 102, 5374-5379.

15 Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* 168, 150-158 e110.

20 Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275-9282.

Schumann, K., Lin, S., Boyer, E., Simeonov, D.R., Subramaniam, M., Gate, R.E., Haliburton, G.E., Ye, C.J., Bluestone, J.A., Doudna, J.A., et al. (2015). Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci. USA* 112, 10437-10442.

25 Shin, J., Jiang, F., Liu, J.J., Bray, N.L., Rauch, B.J., Baik, S.H., Nogales, E., Bondy-Denomy, J., Corn, J.E., and Doudna, J.A. (2017). Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* 3, e1701620.

Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet.* 17, 300-312.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2014). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187-197.

5 Tycko, J., Myer, V.E., and Hsu, P.D. (2016). Methods for optimizing CRISPR-Cas9 genome editing specificity. *Mol. Cell* 63, 355-370.

Yang, H., and Patel, D.J. (2017). Inhibition Mechanism of an Anti-CRISPR Suppressor AcrIIA4 Targeting SpyCas9. *Mol Cell* 67, 117-127 e115.

10 Yin, H., Song, C.Q., Suresh, S., Kwan, S.Y., Wu, Q., Walsh, S., Ding, J., Bogorad, R.L., Zhu, L.J., Wolfe, S.A., et al. (2018). Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat. Chem. Biol.* 14, 311-316.

Yokoyama, T., Silversides, D.W., Waymire, K.G., Kwon, B.S., Takeuchi, T., and Overbeek, P.A. (1990). Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucleic Acids Res.* 18, 7293-7298.

15 Yoon, Y., Wang, D., Tai, P.W.L., Riley, J., Gao, G., and Rivera-Perez, J.A. (2018). Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses. *Nat. Commun.* 9, 412.

Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., Vogel, J., and Sontheimer, E.J. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* 50, 488-503.

20 Zhang, Y., Rajan, R., Seifert, H.S., Mondragón, A., and Sontheimer, E.J. (2015). DNase H activity of *Neisseria meningitidis* Cas9. *Mol. Cell* 60, 242-255.

Zhang, Z., Theurkauf, W.E., Weng, Z., and Zamore, P.D. (2012). Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. *Silence* 3, 9.

25 Zhu, L.J., Holmes, B.R., Aronin, N., and Brodsky, M.H. (2014). CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. *PLoS One* 9, e108424.

Zhu, L.J., Lawrence, M., Gupta, A., Pagés, H., Kucukural, A., Garber, M., and Wolfe, S.A. (2017). GUIDEseq: a bioconductor package to analyze GUIDE-Seq datasets for CRISPR-Cas nucleases. *BMC Genomics* 18, 379.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 33, 73-80.

5 All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred
10 embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in biological control, biochemistry, molecular biology, entomology, plankton, fishery systems, and fresh water ecology, or related fields are intended to be within the scope of the following claims.

Claims

We claim:

- 5 1. A fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain.
2. The fusion protein of Claim 1, wherein said inlaid NBE domain is an adenine base editor (ABE) domain.
- 10 3. The fusion protein of Claim 2, wherein said inlaid ABE domain is an inlaid adenosine deaminase protein domain.
- 15 4. The fusion protein of Claim 3, wherein said inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein domain (ABE8e).
5. The fusion protein of Claim 1, wherein said inlaid NBE domain is a cytidine base editor (CBE) domain.
- 20 6. The fusion protein of Claim 5, wherein said inlaid CBE domain is an inlaid cytosine deaminase protein domain.
7. The fusion protein of Claim 6, wherein said cytosine deaminase protein domain is selected from the group consisting of evoFERNY and rAPOBEC1
- 25 8. The fusion protein of Claim 1, wherein said Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence.
- 30 9. The fusion protein of Claim 8, wherein said PID is an Nme2Cas9 PID or an SmuCas9 PID.

10. The fusion protein of Claim 1, wherein said fusion protein further comprises a nuclear localization signal (NLS) protein selected from the group consisting of nucleoplasmin NLS, SV40 NLS and C-myc NLS.
- 5 11. The fusion protein of Claim 1, wherein said fusion protein further comprises a uracil glycosylase inhibitor.
12. The fusion protein of Claim 1, wherein said Nme2Cas9 protein comprises a mutation.
- 10 13. The fusion protein of Claim 12, wherein said mutation is a D16A mutation.
14. The fusion protein of Claim 13, wherein said Nme2Cas9(D16A) is a nickase.
- 15 15. The fusion protein of Claim 1, wherein said Nme2Cas9 protein comprises a linker that flanks at least one inlaid domain protein.
16. The fusion protein of Claim 15, wherein said linker is approximately 73 amino acids in length.
- 20 17. The fusion protein of Claim 1, wherein said linker is approximately 20 amino acids in length.
18. The fusion protein of Claim 1, wherein said linker is a 3xHA-tag.
- 25 19. An adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain.
20. The AAV of Claim 19, wherein said inlaid NBE domain is an adenine base editor (ABE) domain.
- 30

21. The AAV of Claim 20, wherein said inlaid ABE domain is an inlaid adenosine deaminase protein domain.
22. The AAV of Claim 21, wherein said inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e).
23. The AAV of Claim 19, wherein said inlaid NBE domain is a cytidine base editor (CBE) domain.
24. The AAV of Claim 23, wherein said inlaid CBE protein domain is an inlaid cytosine deaminase protein domain.
25. The AAV of Claim 24, wherein said cytosine deaminase protein is selected from the group consisting of evoFERNY and rAPOBEC1.
26. The AAV of Claim 19, wherein said Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence.
27. The AAV of Claim 26, wherein said PID is an Nme2Cas9 PID or an SmuCas9 PID.
28. The AAV of Claim 19, wherein said AAV is an adeno-associated virus 8.
29. The AAV of Claim 19, wherein said AAV is an adeno-associated virus 6.
30. The AAV of Claim 19, wherein said fusion protein further comprises a nuclear localization signal (NLS) protein selected from the group consisting of nucleoplasmin NLS and/or SV40 NLS and/or C-myc NLS.
31. The AAV of Claim 19, wherein said fusion protein further comprises a uracil glycosylase inhibitor.

32. The AAV of Claim 19, wherein said Nme2Cas9 protein comprises a mutation.

33. The AAV of Claim 32, wherein said mutation is a D16A mutation.

34. The AAV of Claim 33, wherein said Nme2Cas9(D16A) protein is a nickase.

35. The AAV of Claim 19, wherein said Nme2Cas9 protein comprises a linker that flanks at least one inlaid domain protein.

36. The AAV of Claim 35, wherein said linker is approximately 73 amino acids in length.

37. The AAV of Claim 35, wherein said linker is approximately 20 amino acids in length.

38. The AAV of Claim 35, wherein said linker is a 3xHA-tag.

39. A method, comprising:

a) providing;

i) a patient exhibiting at least one symptom of a genetic disease; and

ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide base editor (NBE) domain;

b) treating said patient with said adeno-associated virus under conditions such that said at least one symptom of the genetic disease is reduced.

40. The method of Claim 39, wherein said genetic disease is caused by a gene with a mutated single base, wherein said gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence.

41. The method of Claim 40, wherein said Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or said N₄C nucleotide sequence.
- 5 42. The method of Claim 39, wherein said treating replaces said mutated single base with a wild type single base.
43. The method of Claim 39, wherein said inlaid NBE domain is an adenine base editor (ABE) domain.
- 10 44. The method of Claim 43, wherein said inlaid ABE domain is an inlaid adenosine deaminase protein domain.
45. The method of Claim 44, wherein said inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e).
- 15 46. The method of Claim 39, wherein said inlaid NBE domain is a cytidine base editor (CBE) domain.
- 20 47. The method of Claim 46, wherein said inlaid CBE protein domain is an inlaid cytosine deaminase protein domain.
48. The method of Claim 47, wherein said cytosine deaminase protein is evoFERNY or rAPOBEC1.
- 25 49. The method of Claim 39, wherein said Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes with an N₄CC nucleotide sequence or an N₄C nucleotide sequence.
- 30 50. The method of Claim 49, wherein said PID is an Nme2Cas9 PID or an SmuCas9 PID.

51. The method of Claim 39, wherein said AAV is an adeno-associated virus 8.
52. The method of Claim 39, wherein said AAV is an adeno-associated virus 6.
- 5 53. The method of Claim 39, wherein said fusion protein further comprises a nuclear localization signal (NLS) protein selected from the group consisting of nucleoplasmin NLS, SV40 NLS and C-myc NLS.
- 10 54. The method of Claim 39, wherein said fusion protein further comprises a uracil glycosylase inhibitor.
55. The method of Claim 39, wherein said Nme2Cas9 protein comprises a mutation.
- 15 56. The method of Claim 55, wherein said mutation is a D16A mutation.
57. The method of Claim 56, wherein said Nme2Cas9(D16A) protein is a nickase.
- 20 58. The method of Claim 39, wherein said Nme2Cas9 protein comprises a linker that flanks at least one inlaid domain protein.
59. The method of Claim 58, wherein said linker is approximately 73 amino acids in length.
60. The method of Claim 58, wherein said linker is approximately 20 amino acids in length.
- 25 61. The method of Claim 58, wherein said linker is a 3xHA-tag.
62. The method of Claim 39, wherein said genetic disease is selected from the group consisting of tyrosinemia, muscular dystrophy, Rett syndrome, Batten disease or amyotrophic lateral sclerosis (ALS).
- 30

63. A method, comprising:

a) providing;

i) a patient comprising a gene with a mutated single base, wherein said gene is flanked by an N₄CC nucleotide sequence or an N₄C nucleotide sequence; and

ii) an adeno-associated virus (AAV) comprising a vector encoding a fusion protein comprising an Nme2Cas9 protein and an inlaid nucleotide deaminase protein domain;

b) treating said patient with said adeno-associated virus under conditions such that said mutated single base is replaced with a wild type single base and a genetic disease does not develop.

64. The method of Claim 63, wherein said genetic disease is selected from the group consisting of tyrosinemia, muscular dystrophy, Rett syndrome, Batten disease and amyotrophic lateral sclerosis (ALS).

65. The method of Claim 63, wherein said Nme2Cas9 protein comprises a protospacer adjacent motif interacting domain (PID) having a peptide that hybridizes to said N₄CC nucleotide sequence or the N₄C nucleotide sequence.

66. The method of Claim 65, wherein said PID is an Nme2Cas9 PID or an SmuCas9 PID.

67. The method of Claim 63, wherein said inlaid nucleotide deaminase protein domain is an inlaid adenosine deaminase protein domain.

68. The method of Claim 67, wherein said inlaid adenosine deaminase protein domain is an adenosine deaminase8e protein (ABE8e) domain.

69. The method of Claim 63, wherein said inlaid nucleotide deaminase protein domain is an inlaid cytosine deaminase protein domain.

70. The method of Claim 69, wherein said cytosine deaminase protein domain is evoFERNY or rAPOBEC1,
- 5 71. The method of Claim 63, wherein said AAV is an adeno-associated virus 8.
72. The method of Claim 63, wherein said AAV is an adeno-associated virus 6.
73. The method of Claim 63, wherein said fusion protein further comprises a nuclear localization signal (NLS) protein selected from the group consisting of nucleoplasmin
10 NLS, SV40 NLS and C-myc NLS.
74. The method of Claim 63, wherein said fusion protein further comprises a uracil glycosylase inhibitor.
- 15 75. The method of Claim 63, wherein said Nme2Cas9 protein comprises a mutation.
76. The method of Claim 75, wherein said mutation is a D16A mutation.
77. The method of Claim 63, wherein said Nme2Cas9(D16A) protein is a nickase.
20
78. The method of Claim 63, wherein said Nme2Cas9 protein comprises a linker that flanks at least one inlaid protein domain.
79. The method of Claim 78, wherein said linker is approximately 73 amino acids in length.
25
80. The method of Claim 78, wherein said linker is approximately 20 amino acids in length.
81. The method of Claim 78, wherein said linker is a 3xHA-tag.
- 30 82. The method of Claim 63, wherein said gene is selected from the group consisting of a Fah gene, a Dmd gene, a MeCP2 gene, a CLN3 gene and an SOD1 gene.

Fig. 1

Fig. 1A

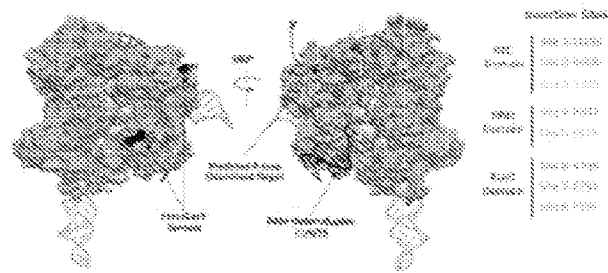


Fig. 1B

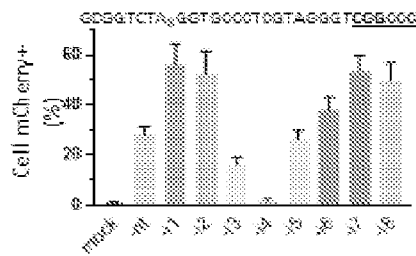


Fig. 1C

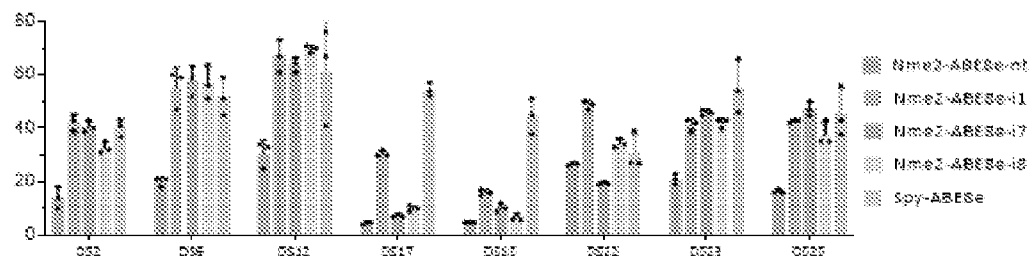


Fig. 1 (cont'd)

Fig. 1D

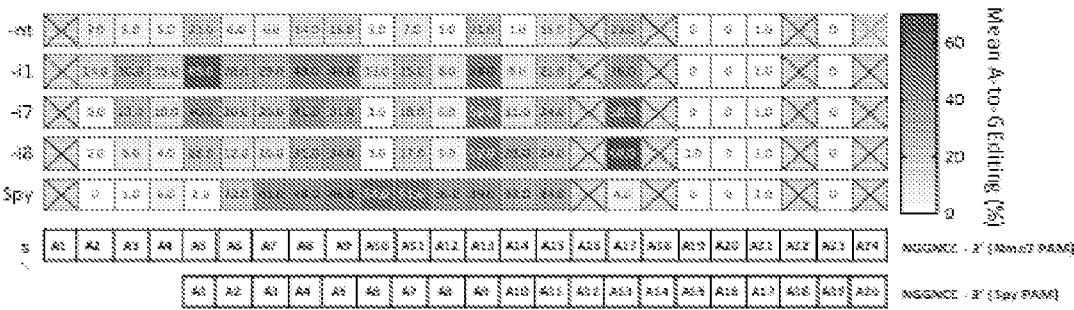


Fig. 1E

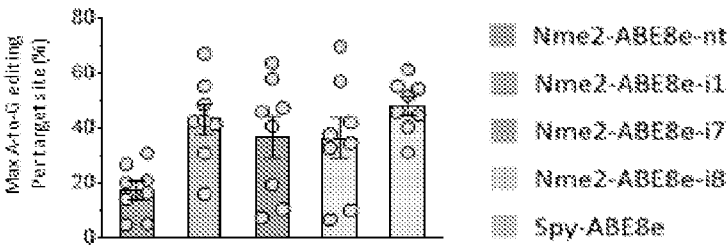


Fig. 2

Fig. 2A

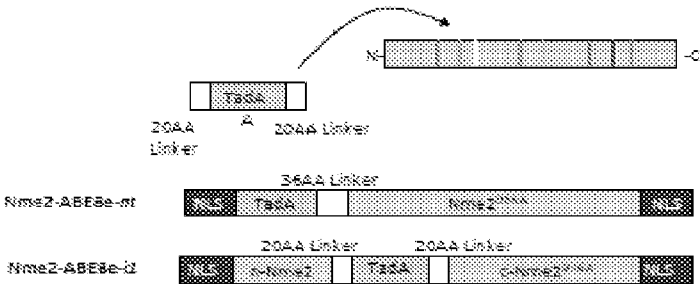


Fig. 2B

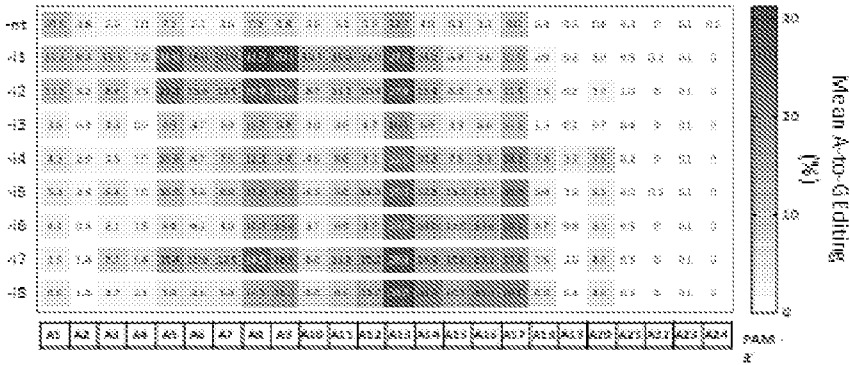


Fig. 2C

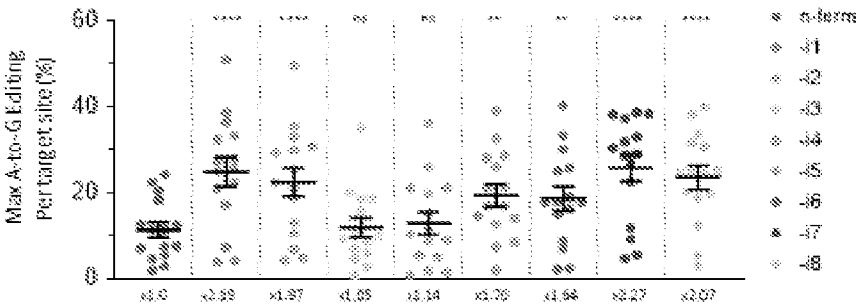


Fig. 3

Figs. 3A-B

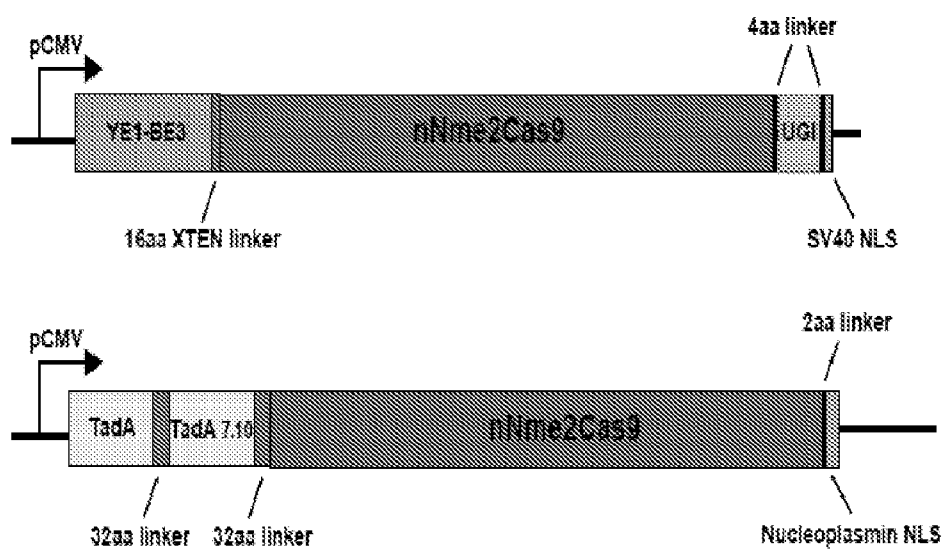


Fig. 4

Fig. 4A

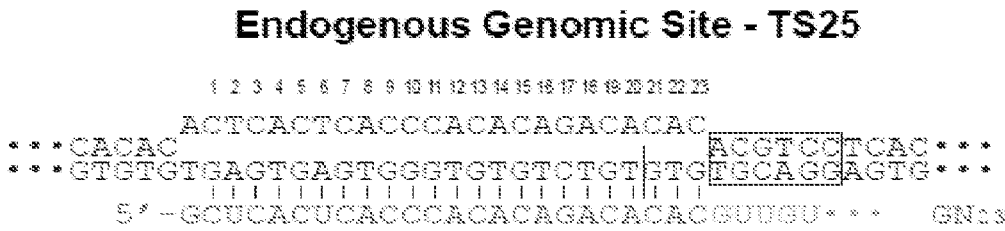


Fig. 4B

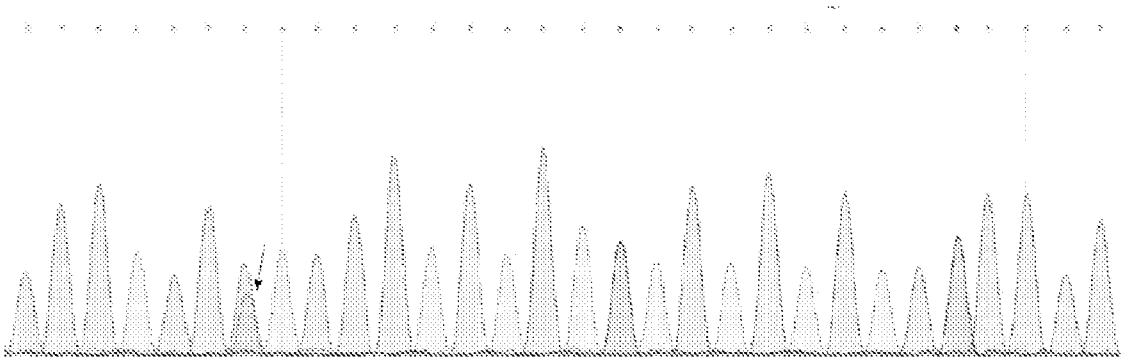


Fig. 4 (cont'd)

Fig. 4C

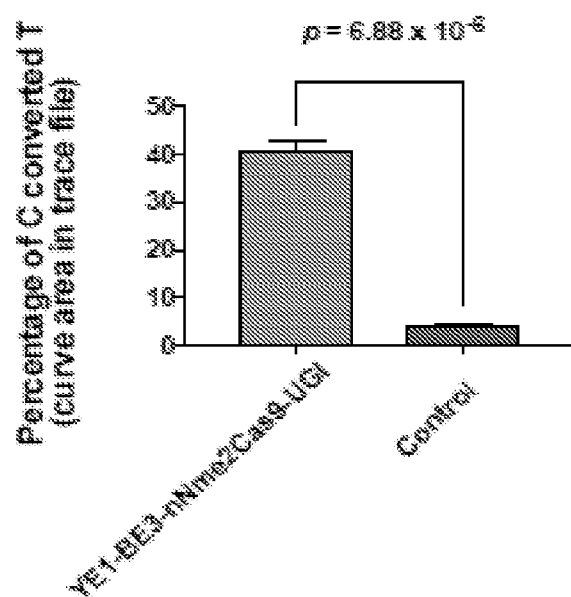


Fig. 5A

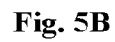


Fig. 5 (cont'd)

Fig. 5C

EGFP- site 3																							PAM							
5'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		3'					
	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G	reference
	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G	97.78%
	A	G	T	T	T	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G	0.40%
	A	G	T	T	C	A	T	T	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G	0.33%

Fig. 5D

EGFP- site 4																							PAM							
5'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		3'					
	C	G	T	G	C	C	C	T	G	G	C	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	T	reference
	C	G	T	G	C	C	C	T	G	G	C	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	T	97.93%
	C	G	T	G	C	C	T	T	G	G	C	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	T	0.35%

Fig. 6

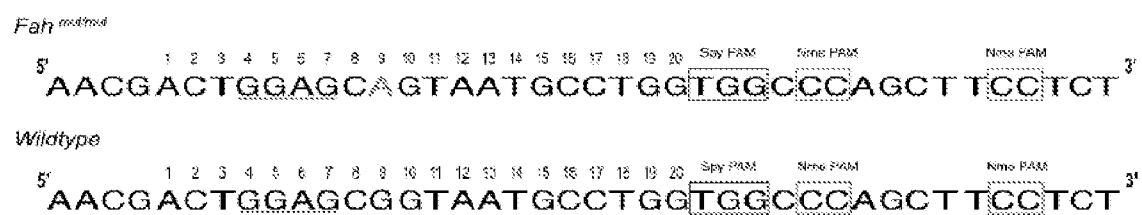


Fig. 7

Fig. 7A

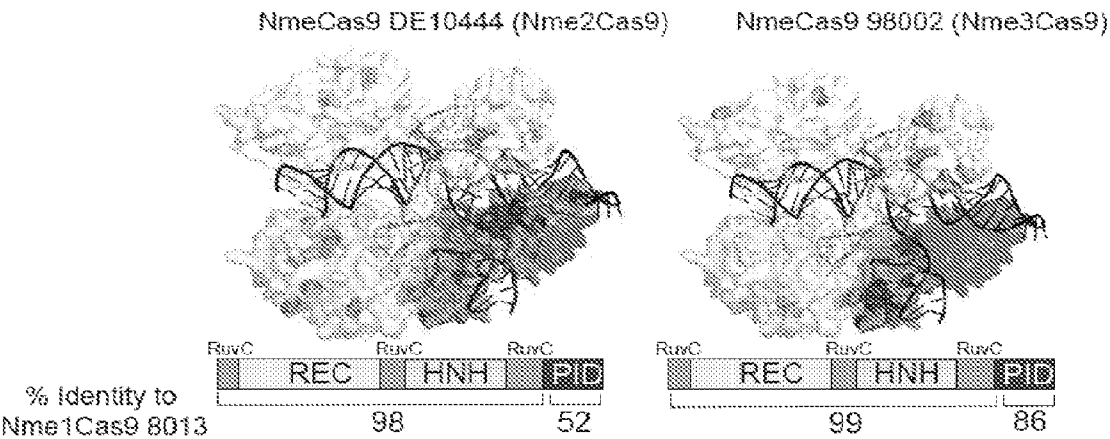


Fig. 7B

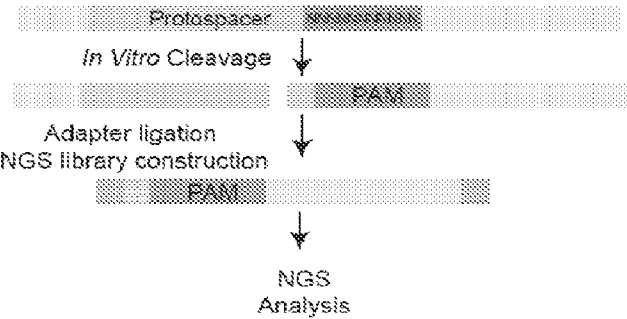


Fig. 7C

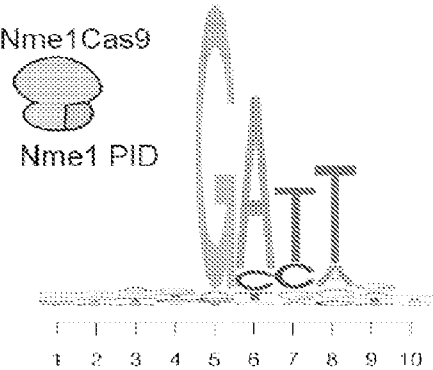


Fig. 7 (cont'd)

Fig. 7 D

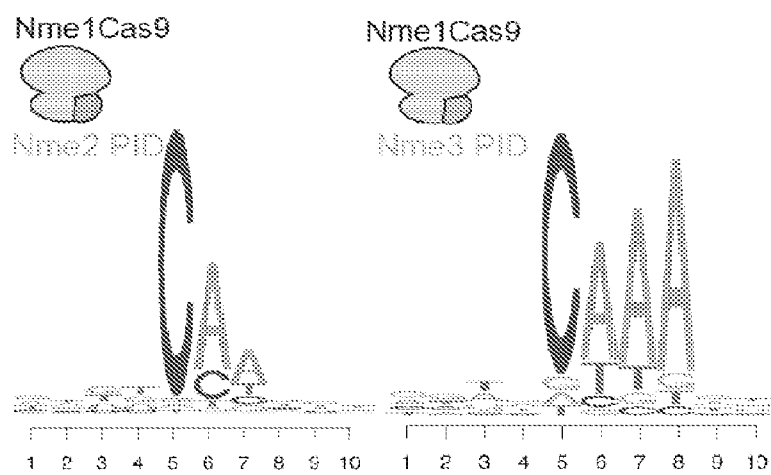


Fig. 7E

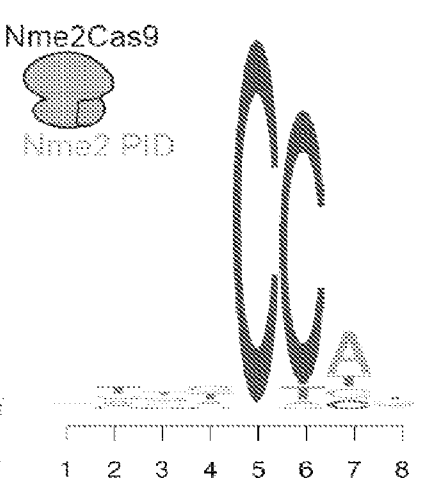


Fig. 8

Fig. 8A

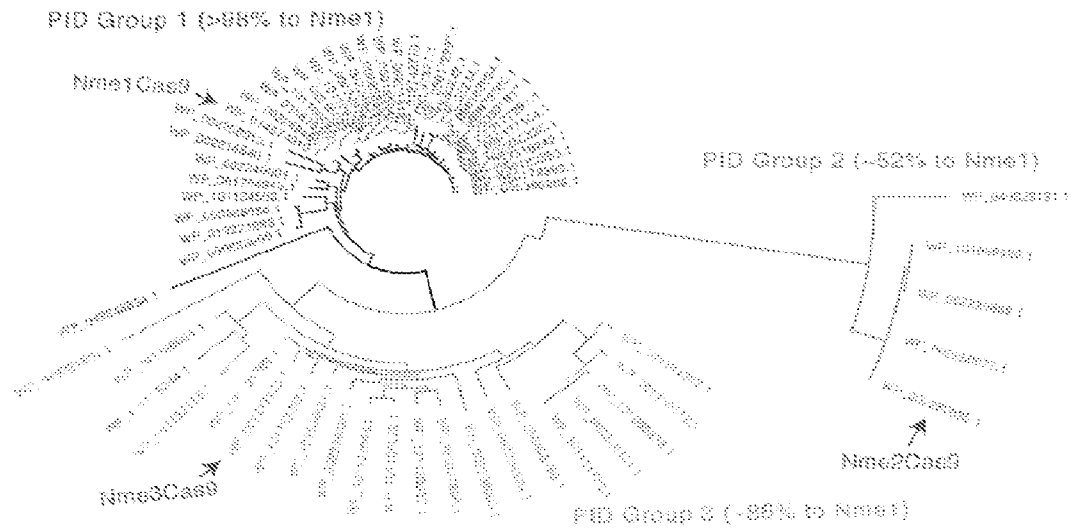


Fig. 8B

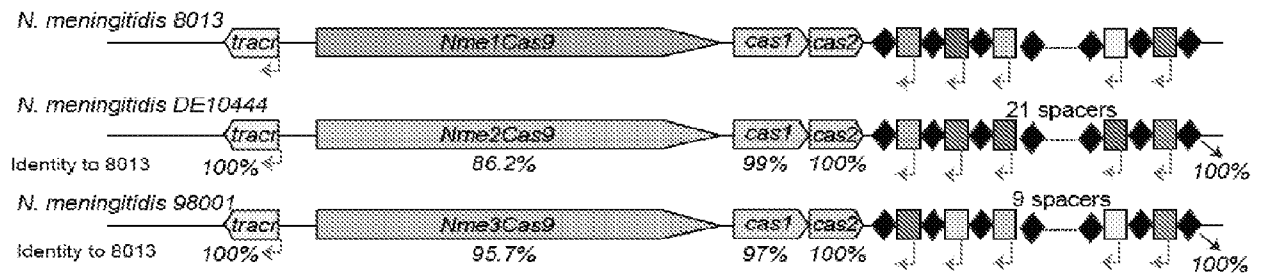


Fig. 8 (cont'd)

Fig. 8C

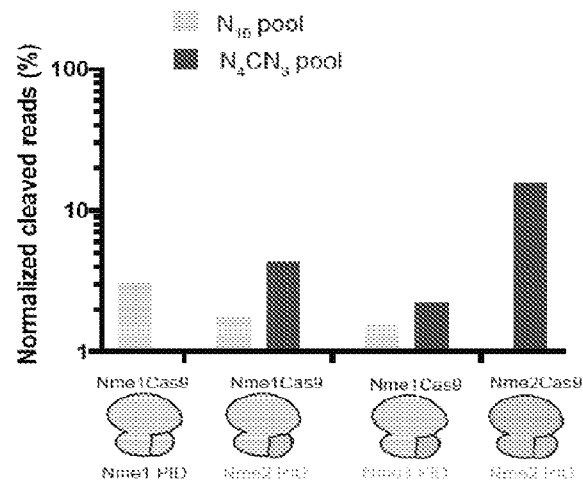


Fig. 8D

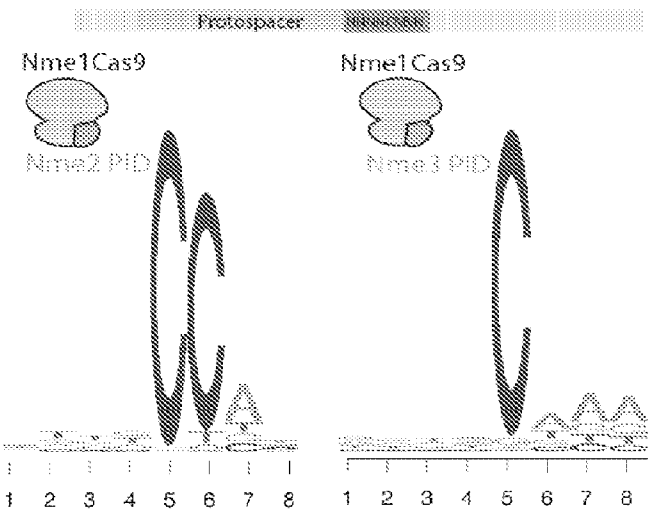


Fig. 9

Fig. 9A

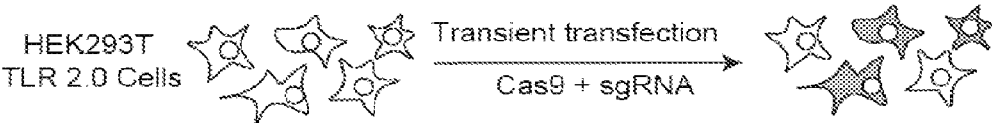


Fig. 9B

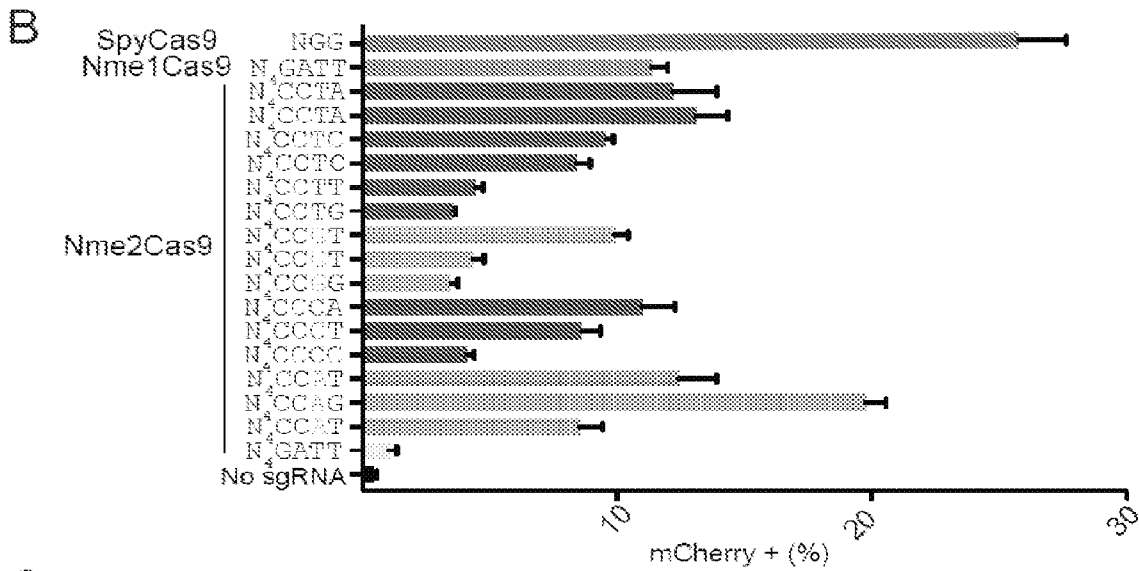


Fig. 9 (cont'd)

Figs. 9C-D

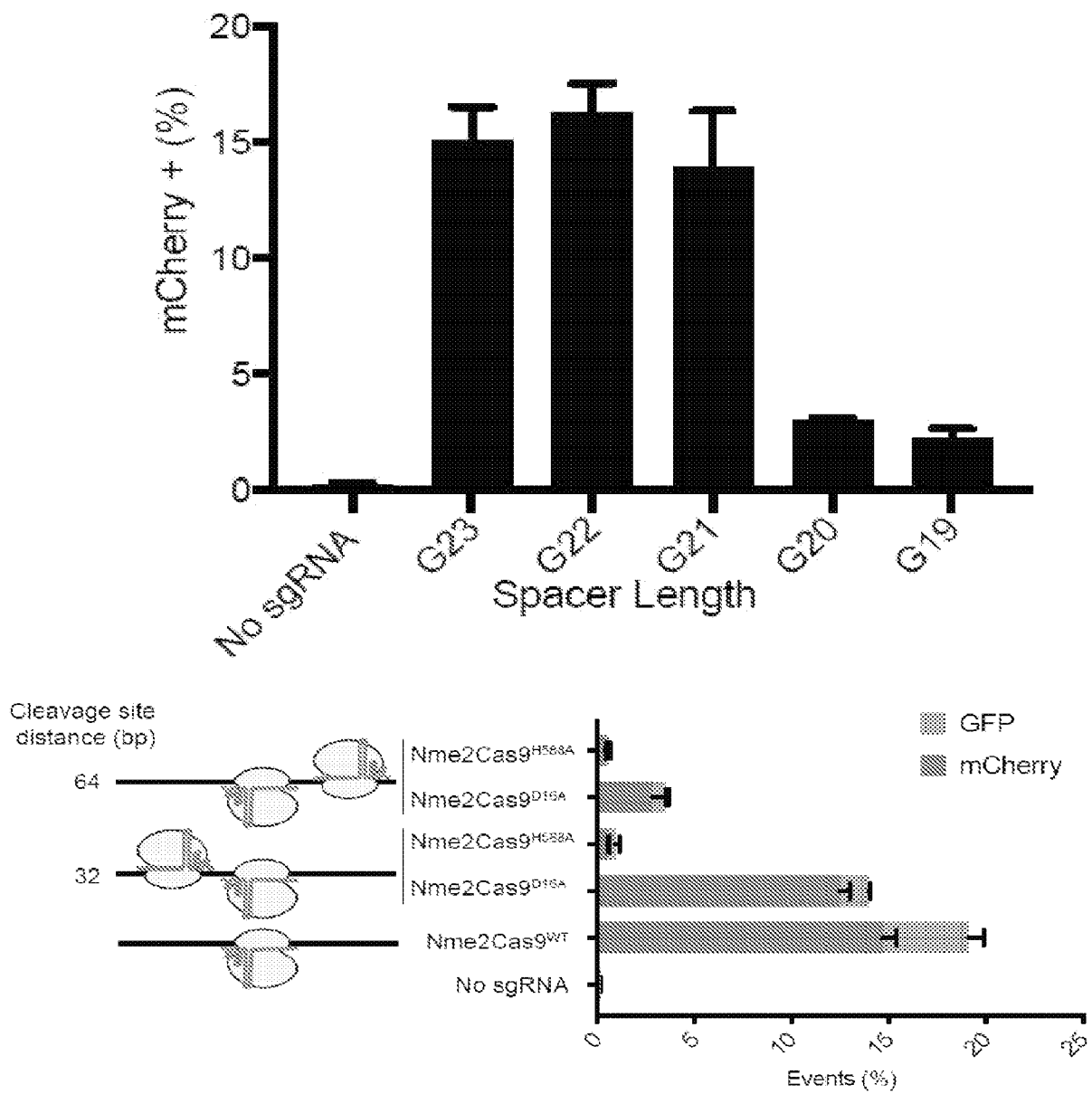


Fig. 10

Fig. 10A

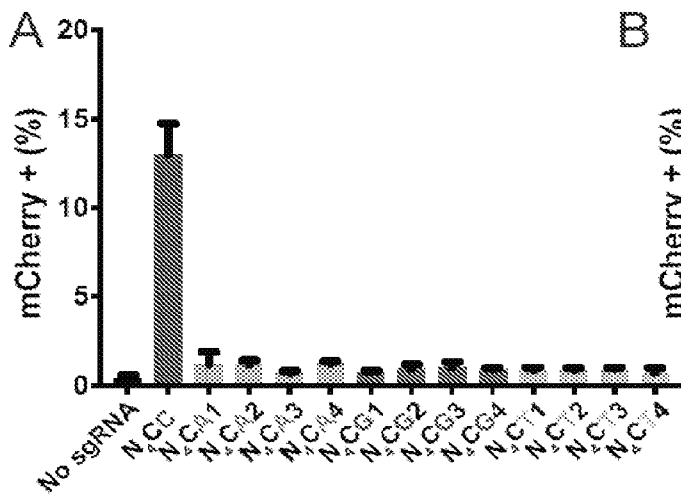


Fig. 10B

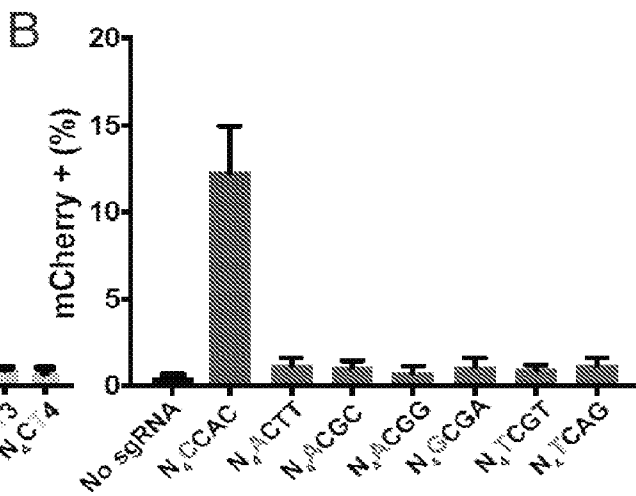


Fig. 10 (cont'd)

Fig. 10C

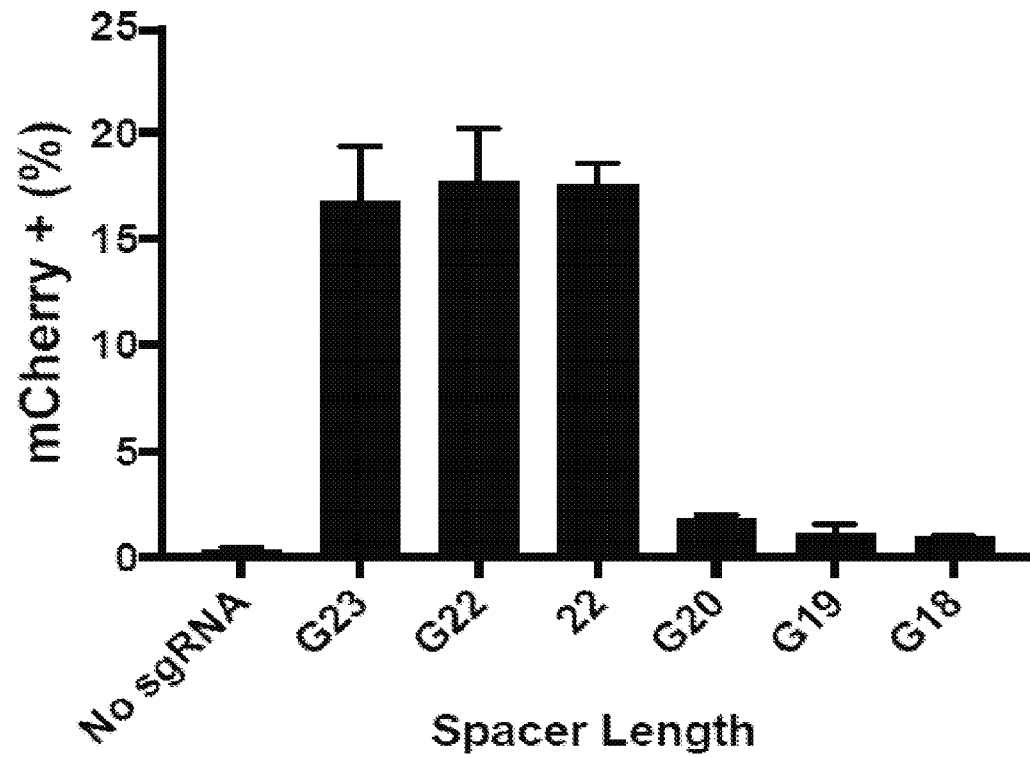


Fig. 10D

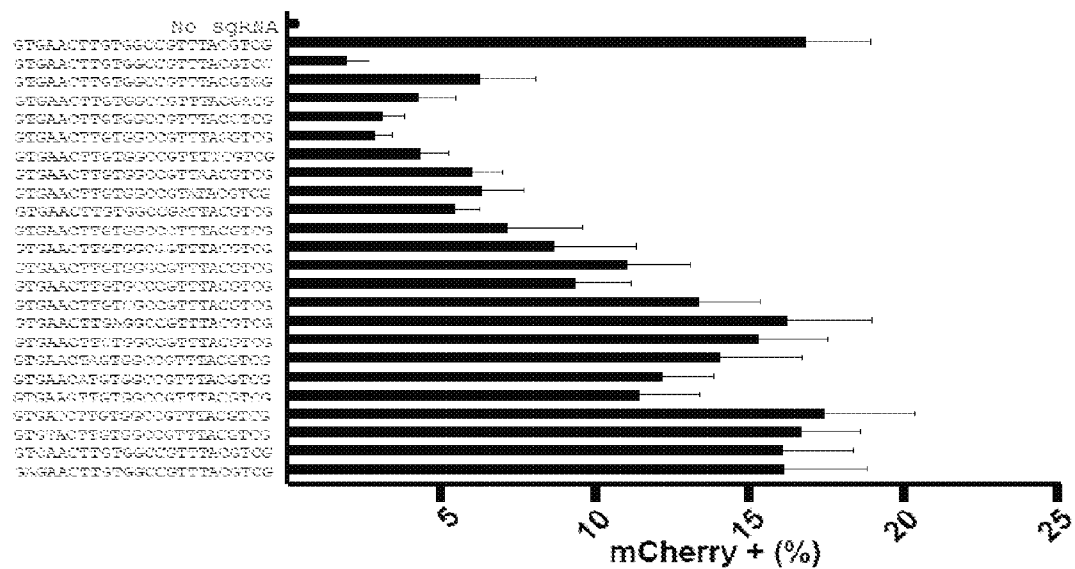


Fig. 11

Fig. 11A

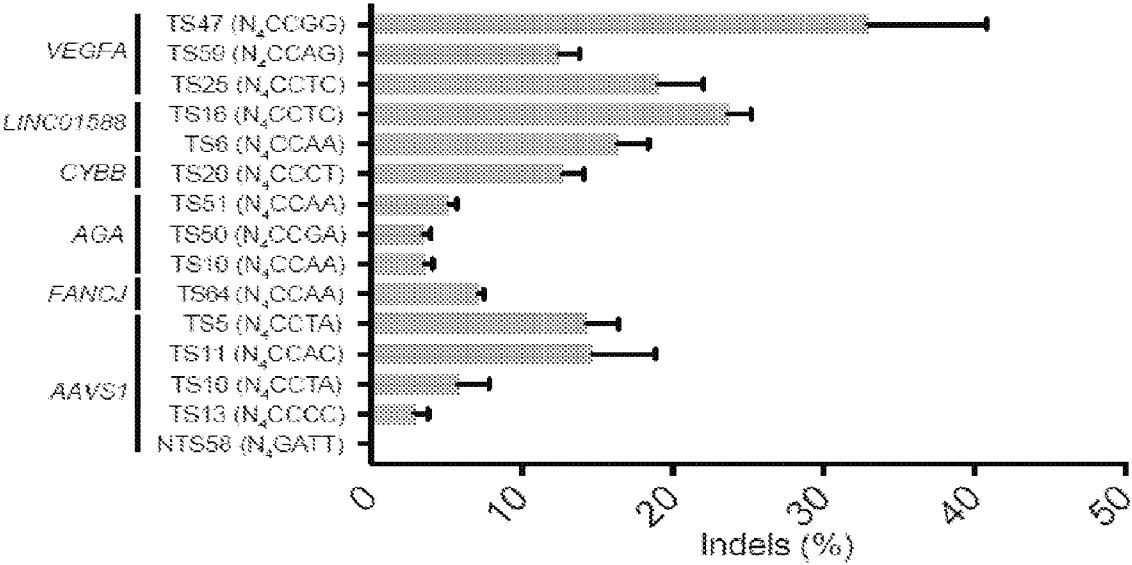


Fig. 11B

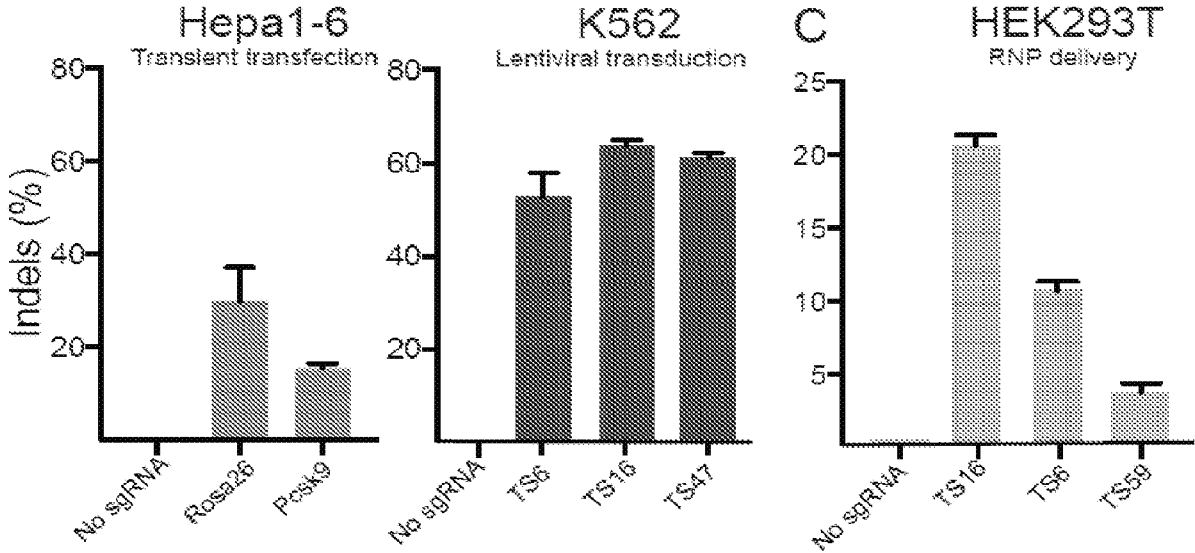


Fig. 12

Fig. 12A

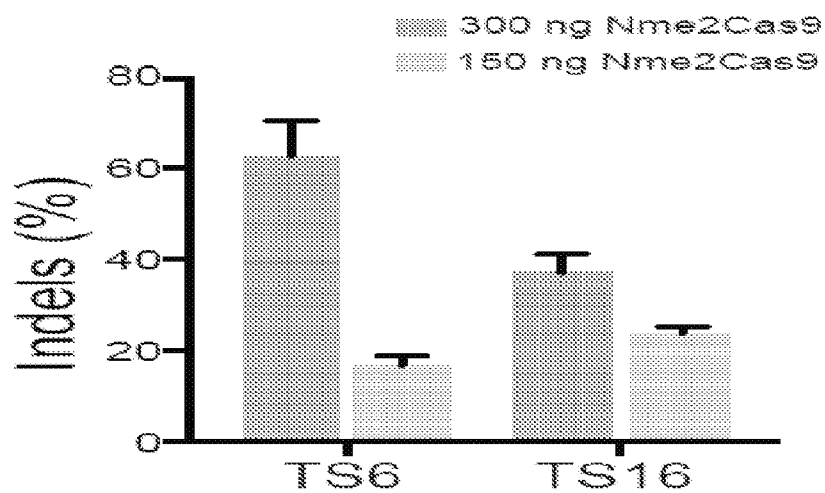


Fig. 12B

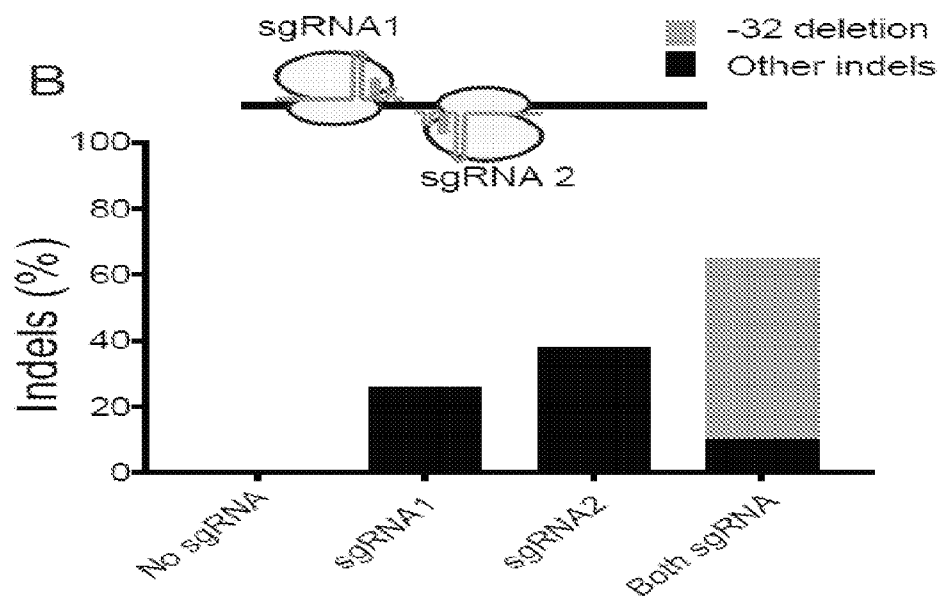


Fig 13

Fig. 13A

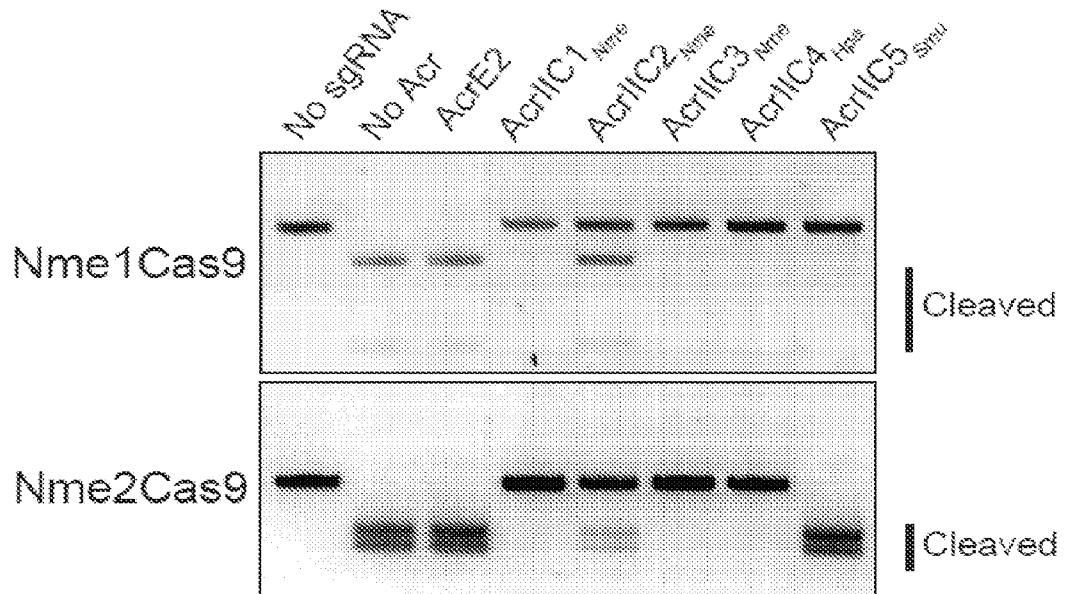


Fig. 13B

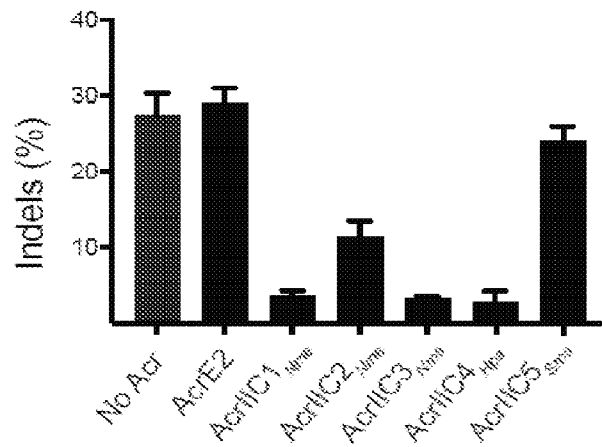


Fig. 13C

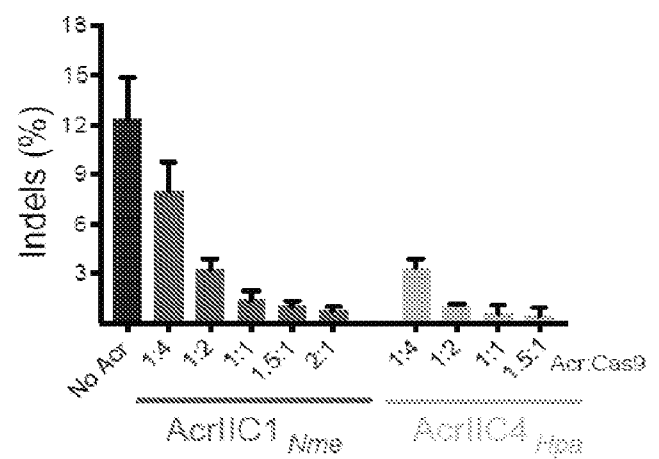


Fig. 14

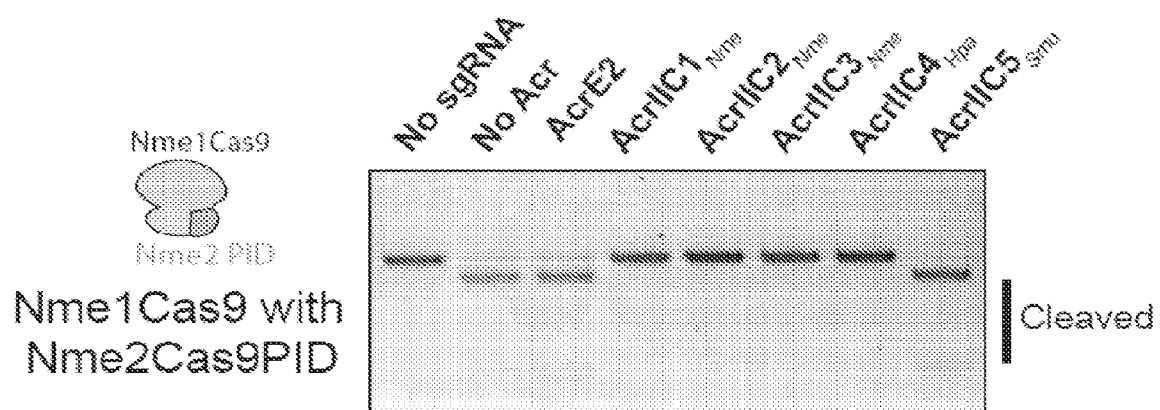


Fig. 15

Fig. 15A

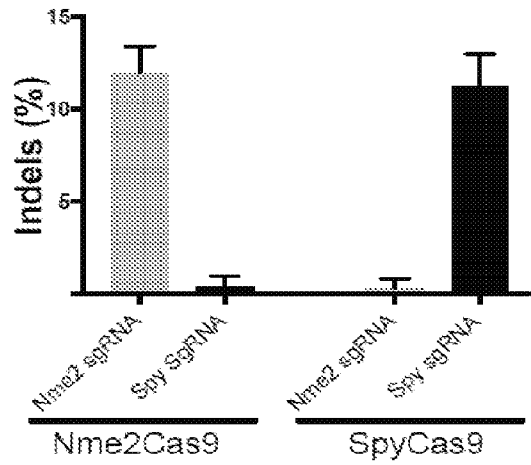


Fig. 15B

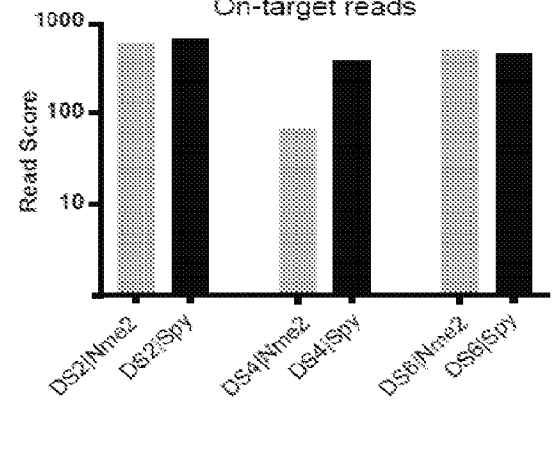


Fig. 15C

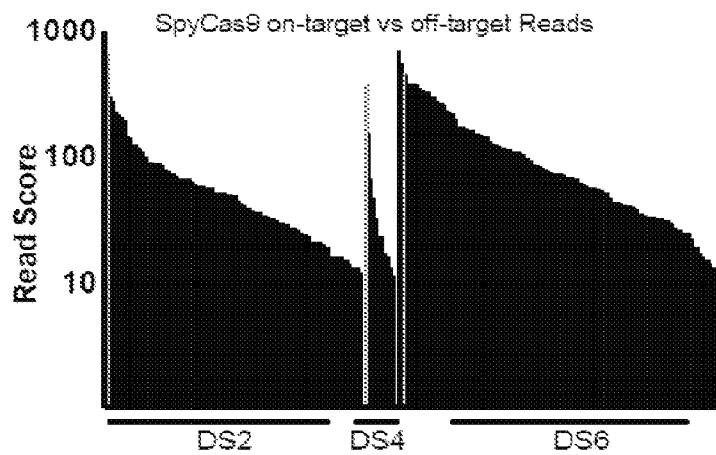


Fig. 15D

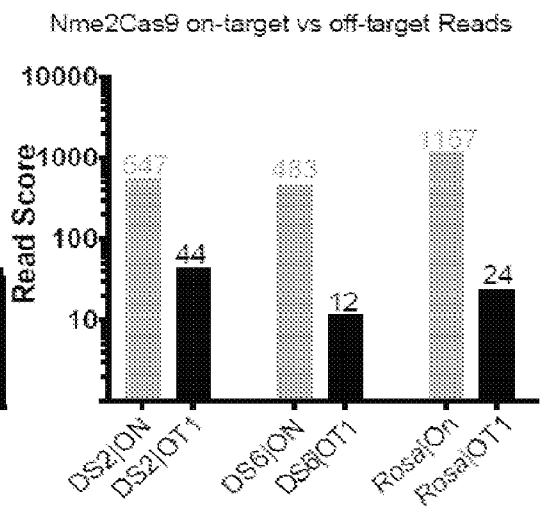


Fig. 15 (cont'd)

Fig. 15E

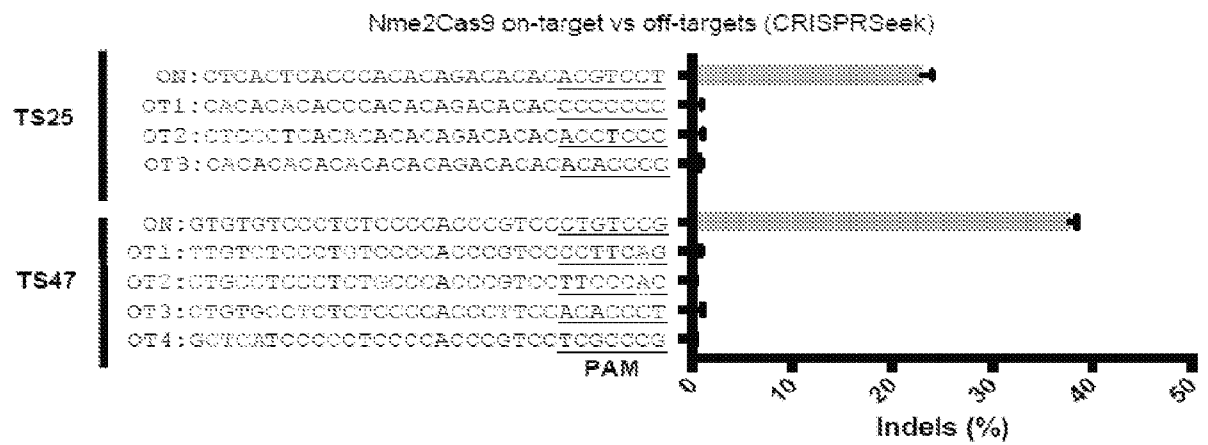


Fig. 16

Fig. 16A

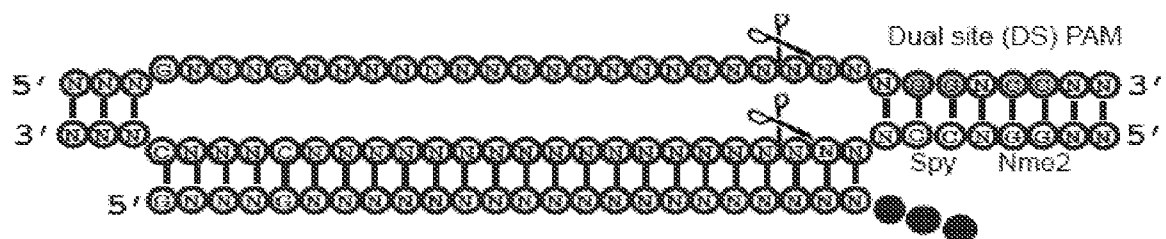


Fig. 16B

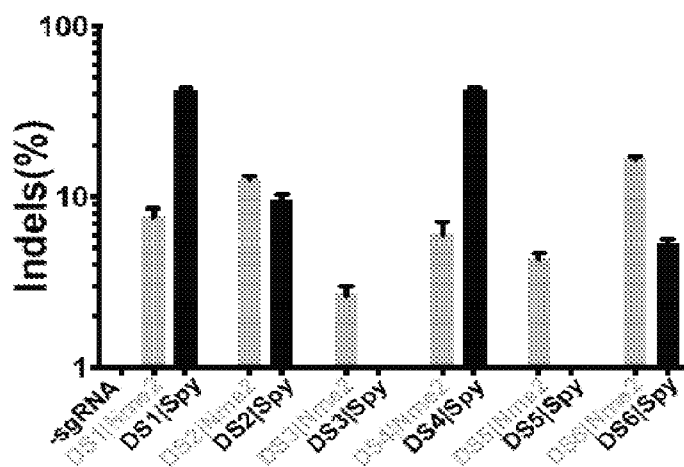


Fig. 16C

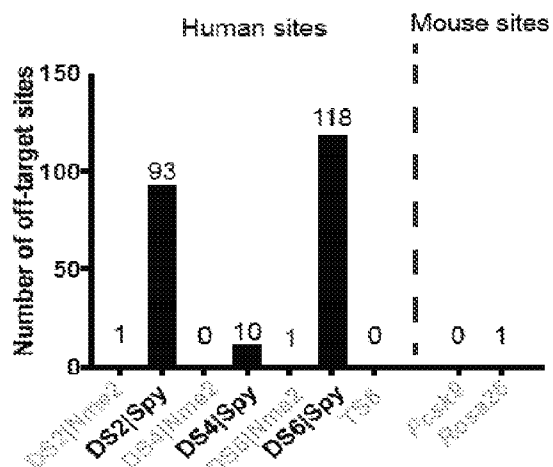


Fig. 16 (cont'd)

Fig. 16D

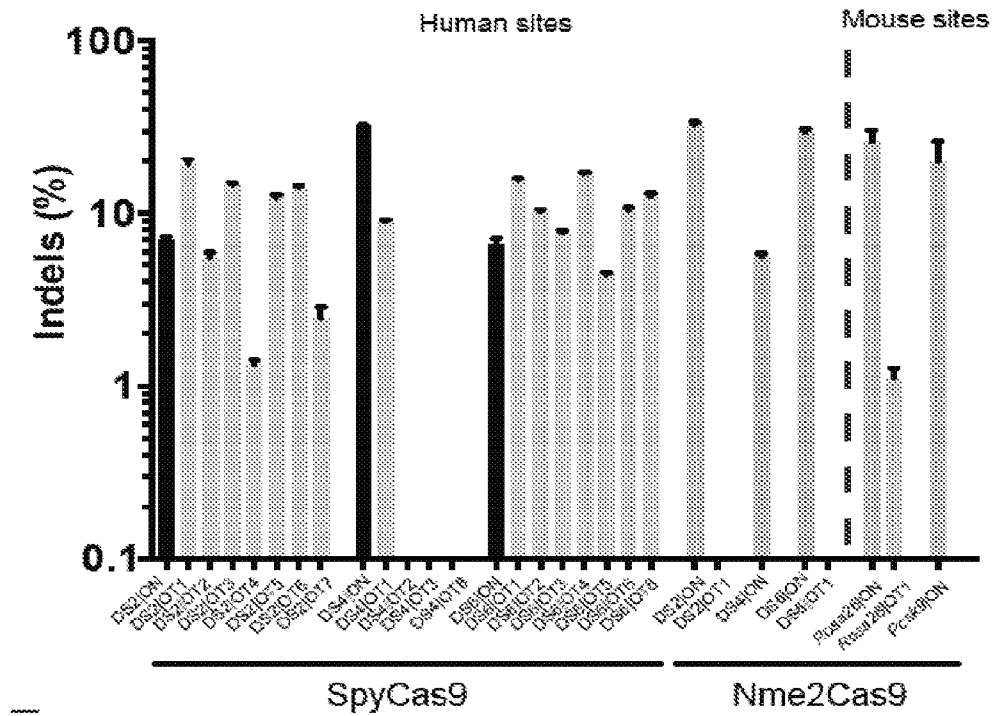


Fig. 16E

Rosa26|On: TGAGGACCGCCCTGGGCCTGGGAGAATCCCT
Rosa26|OT1: GAAGGACCACCCTAGGCCTGGGAGACTCCCT

Fig. 17

Fig. 17A

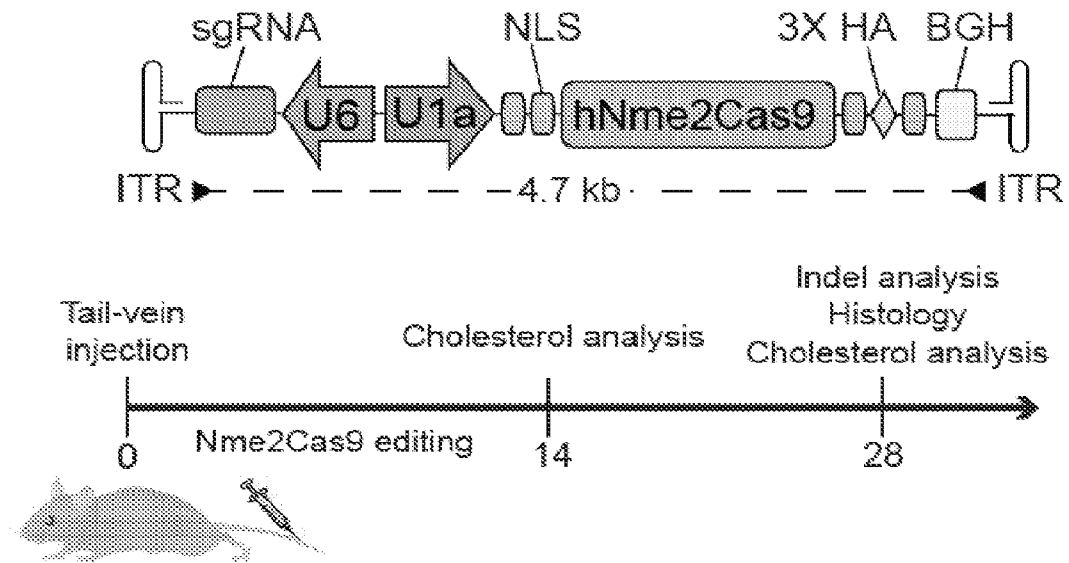


Fig. 17B

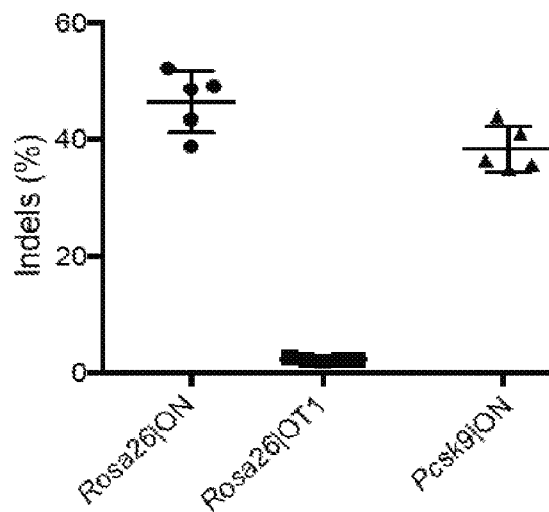


Fig. 17C

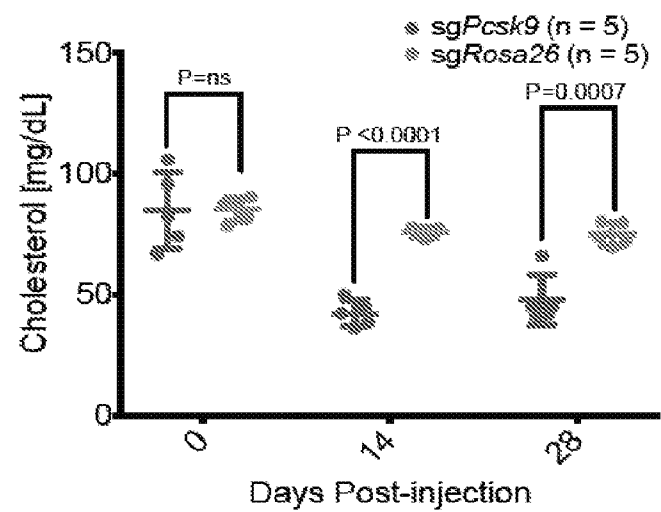


Fig. 18

Fig. 18A

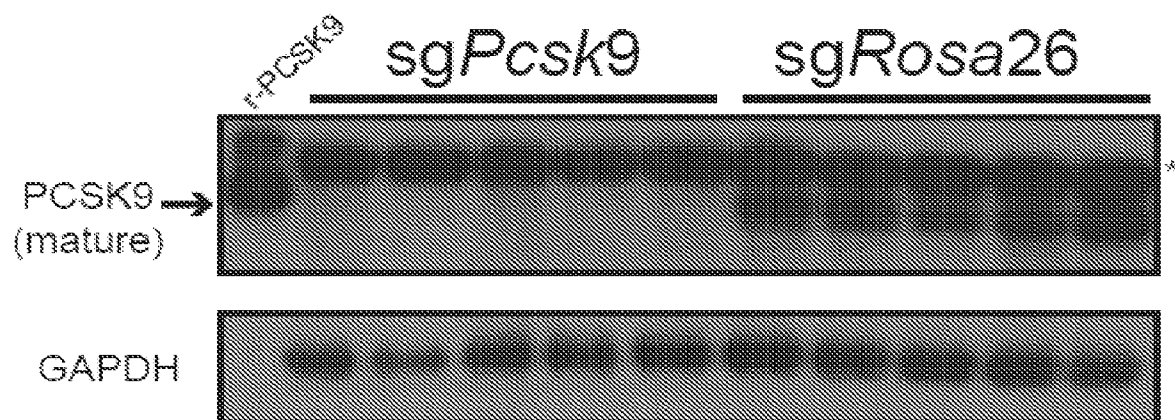


Fig. 18B

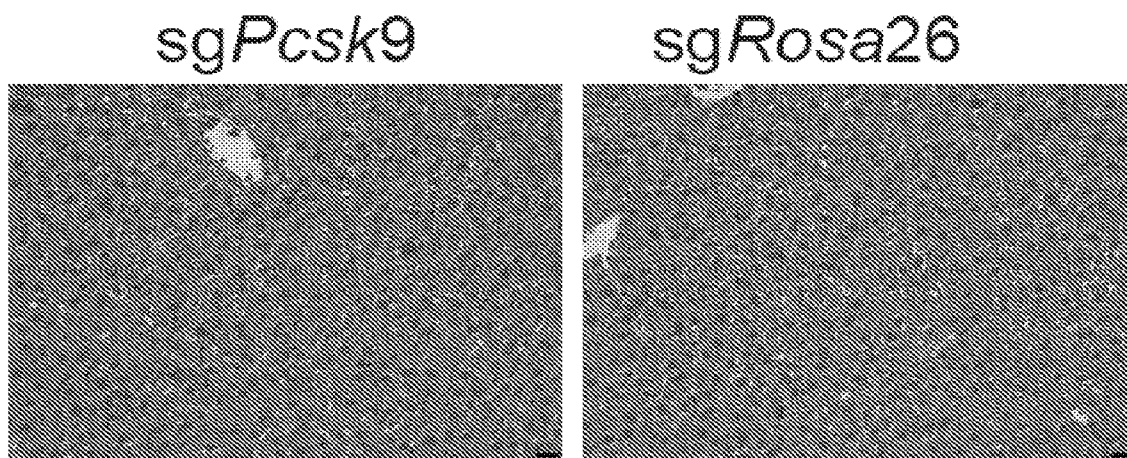


Fig. 19

Fig. 19A

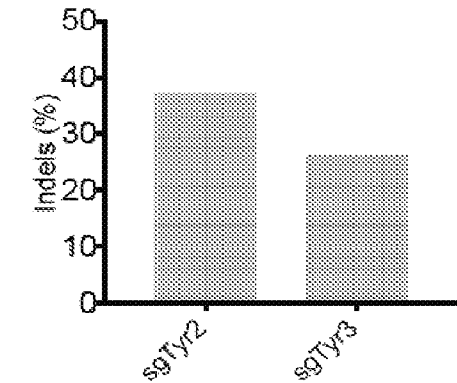


Fig. 19B

Mouse ID	AAV dosage	Indels (%)	Coat color
9-1	3 x 10 ⁹ GCs	100	Albino
9-2	3 x 10 ⁹ GCs	100	Albino
9-3	3 x 10 ⁹ GCs	84	mosaic and hypomorphic
9-4	3 x 10 ⁹ GCs	98	mosaic and hypomorphic
9-5	3 x 10 ⁹ GCs	100	mosaic and hypomorphic
8-1	3 x 10 ⁸ GCs	100	mosaic and hypomorphic
8-2	3 x 10 ⁸ GCs	96	mosaic and hypomorphic

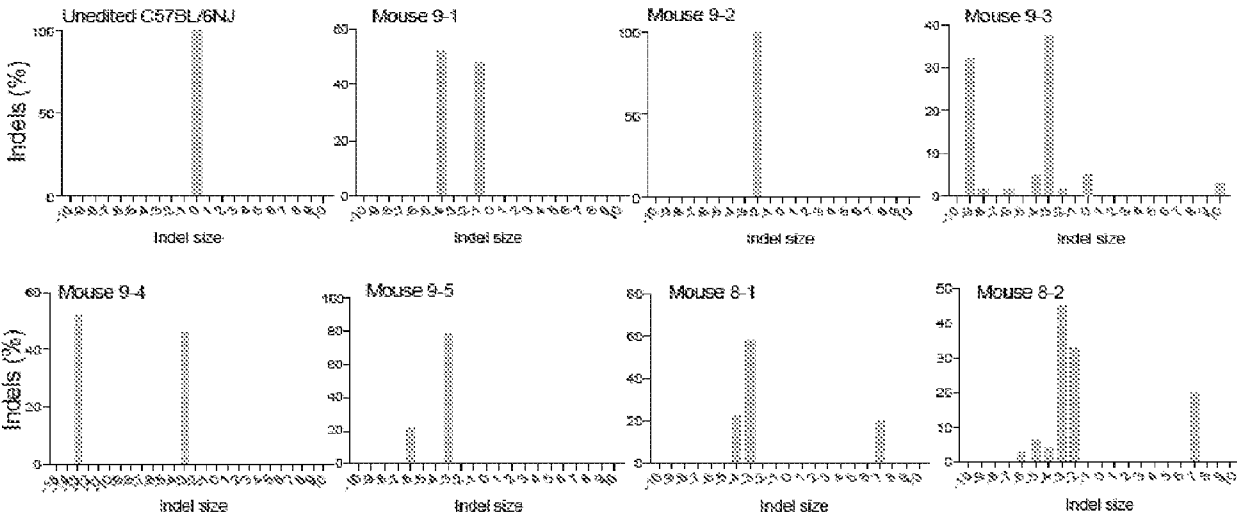
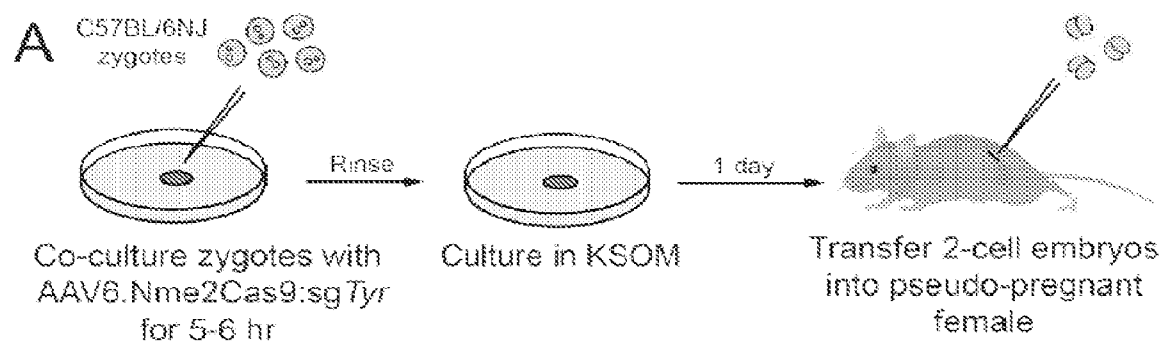
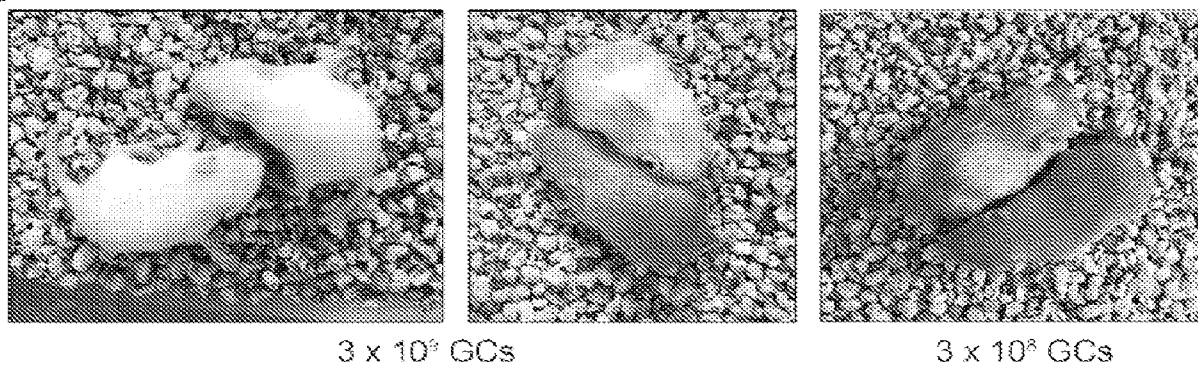


Fig. 19C

Fig. 20

Figs. 20A – C

**B****C**

AAV Dosage	Number of embryos transferred	Number of pups recovered	Number of <i>Tyr</i> -edited pups (albino or chinchilla)	<i>Tyr</i> editing frequency (%)
3 x 10 ⁶ GCs	50	5	5 (3)	100
3 x 10 ⁸ GCs	28	2*	2 (1)	100

Fig. 21

Figs 21A-E

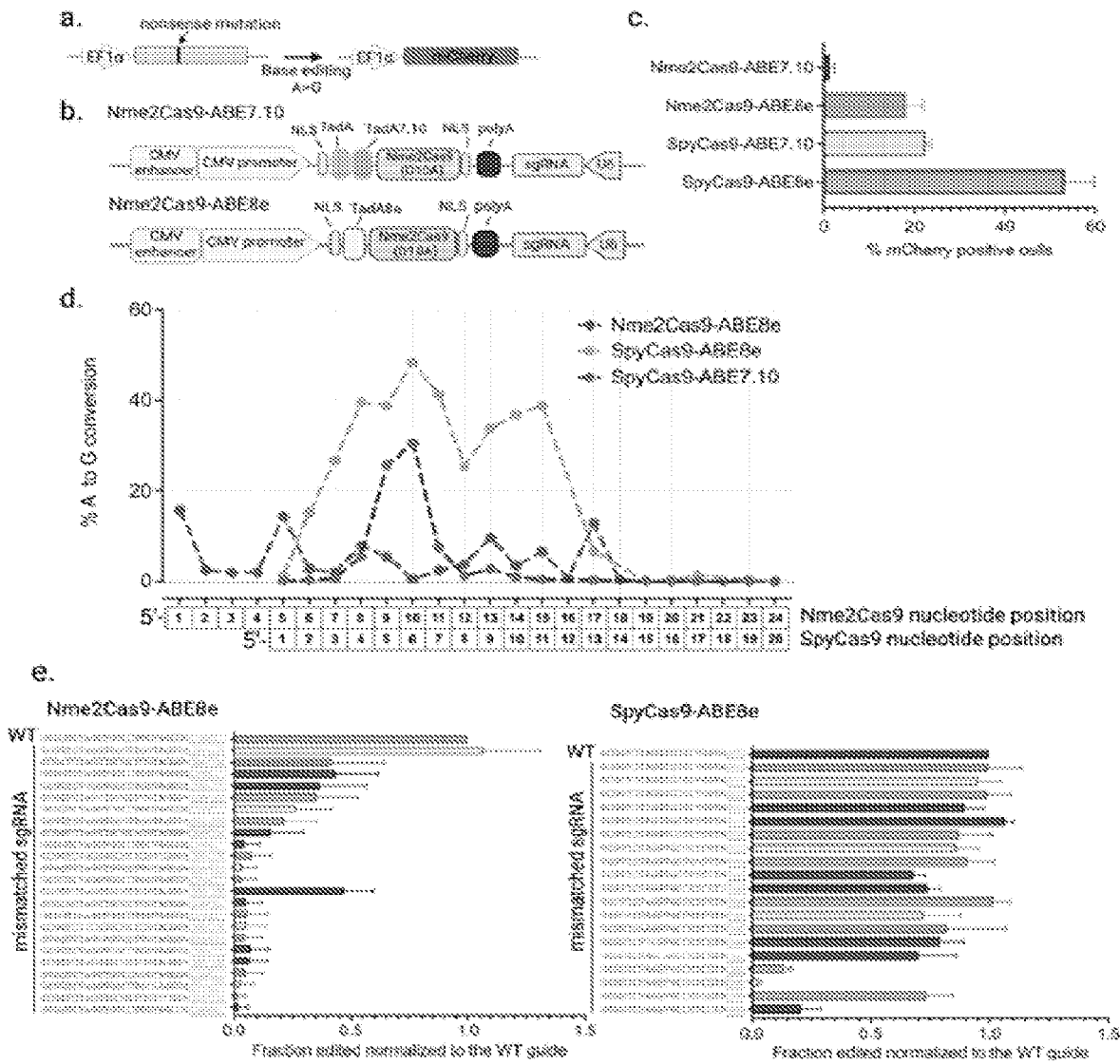


Fig. 22

Figs 22 A -C

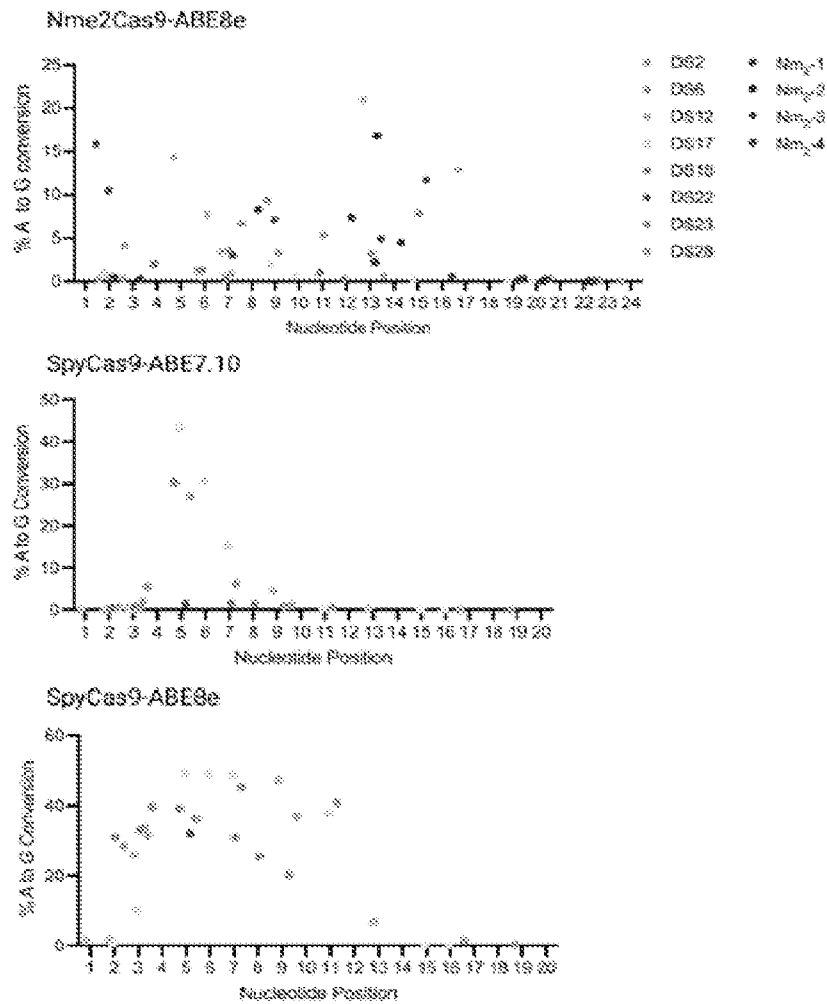


Fig. 23

Figs 23 A - D

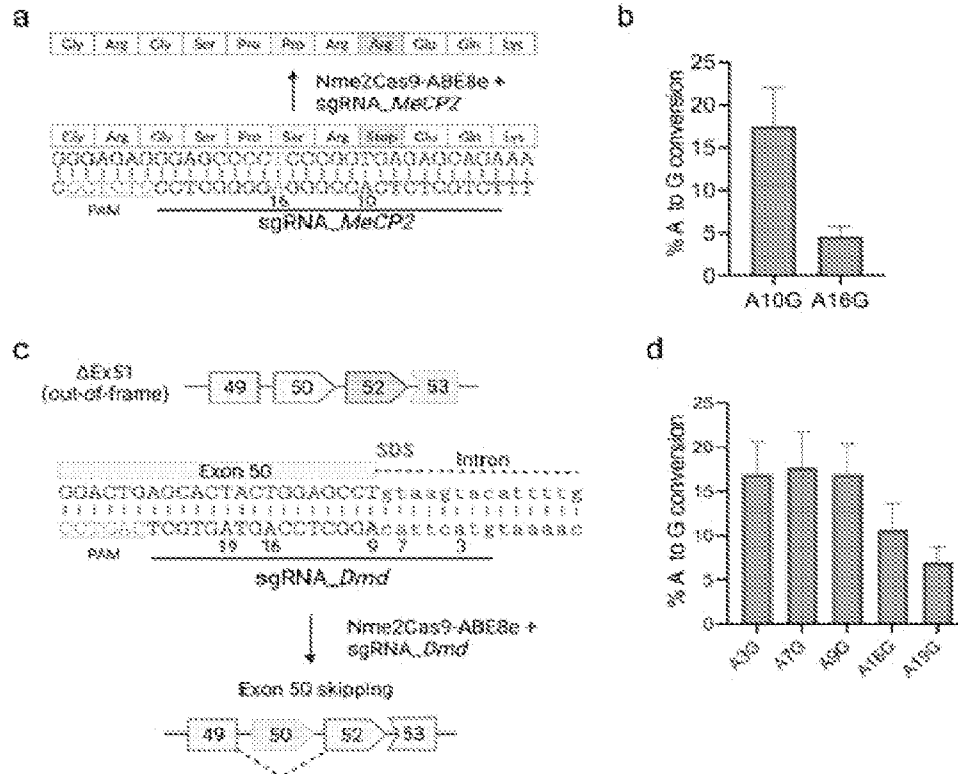


Fig. 24

Figs. 24A-C

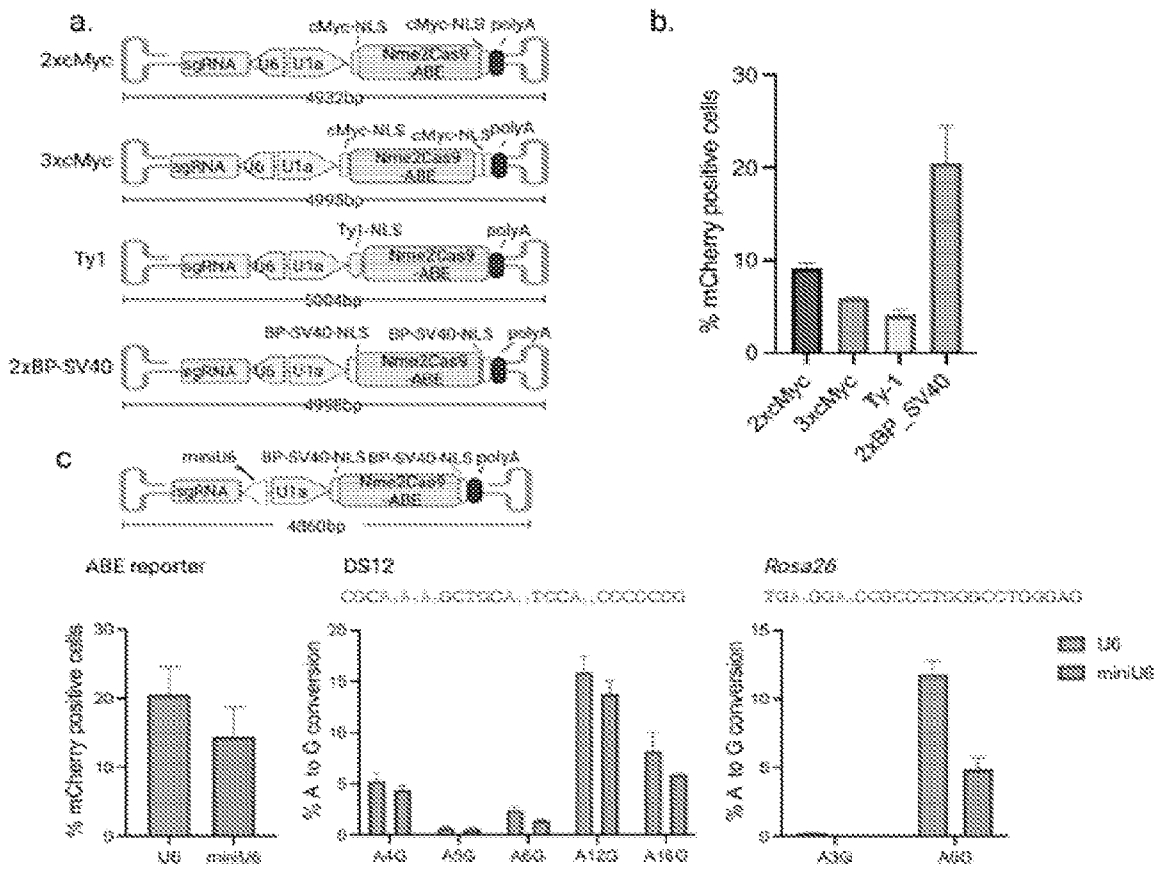


Fig. 25

Figs. 25A-C

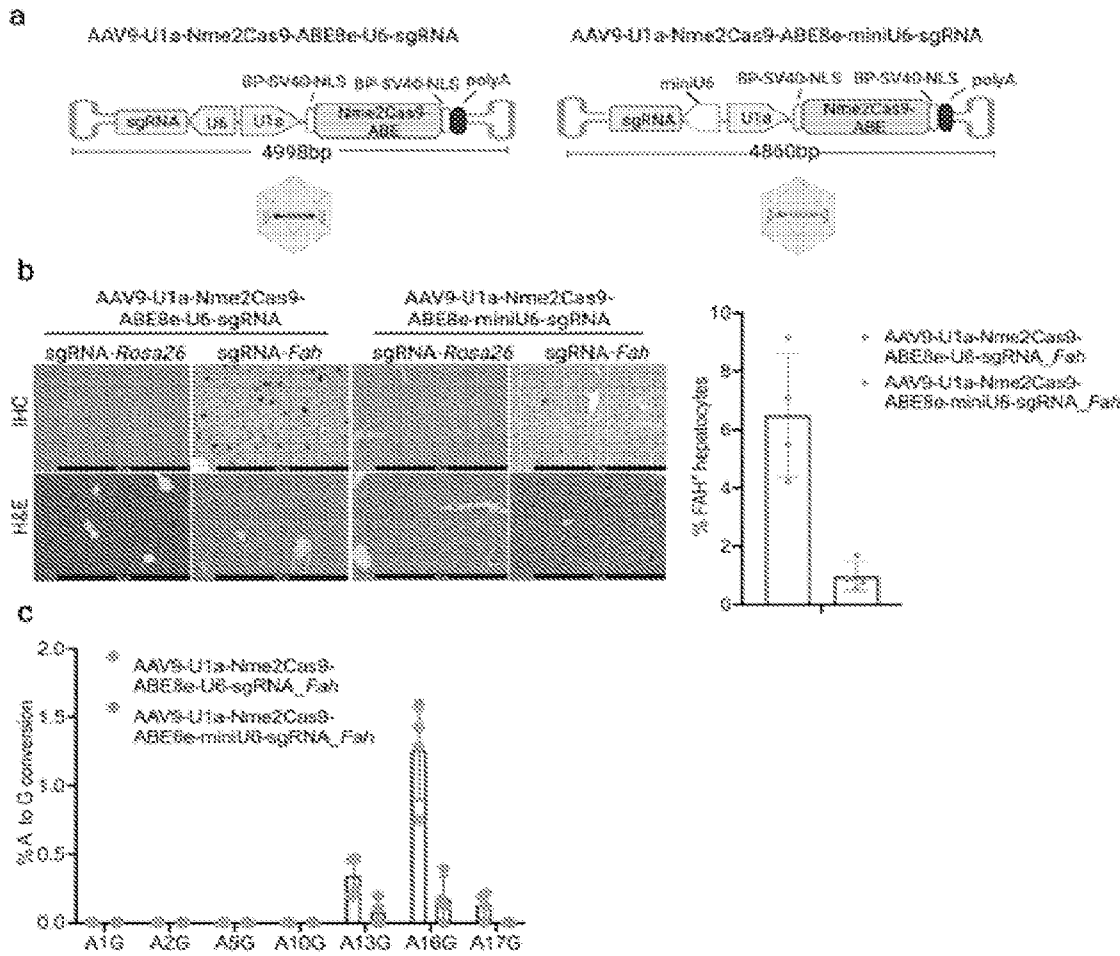


Fig. 26

Figs. 26A-I

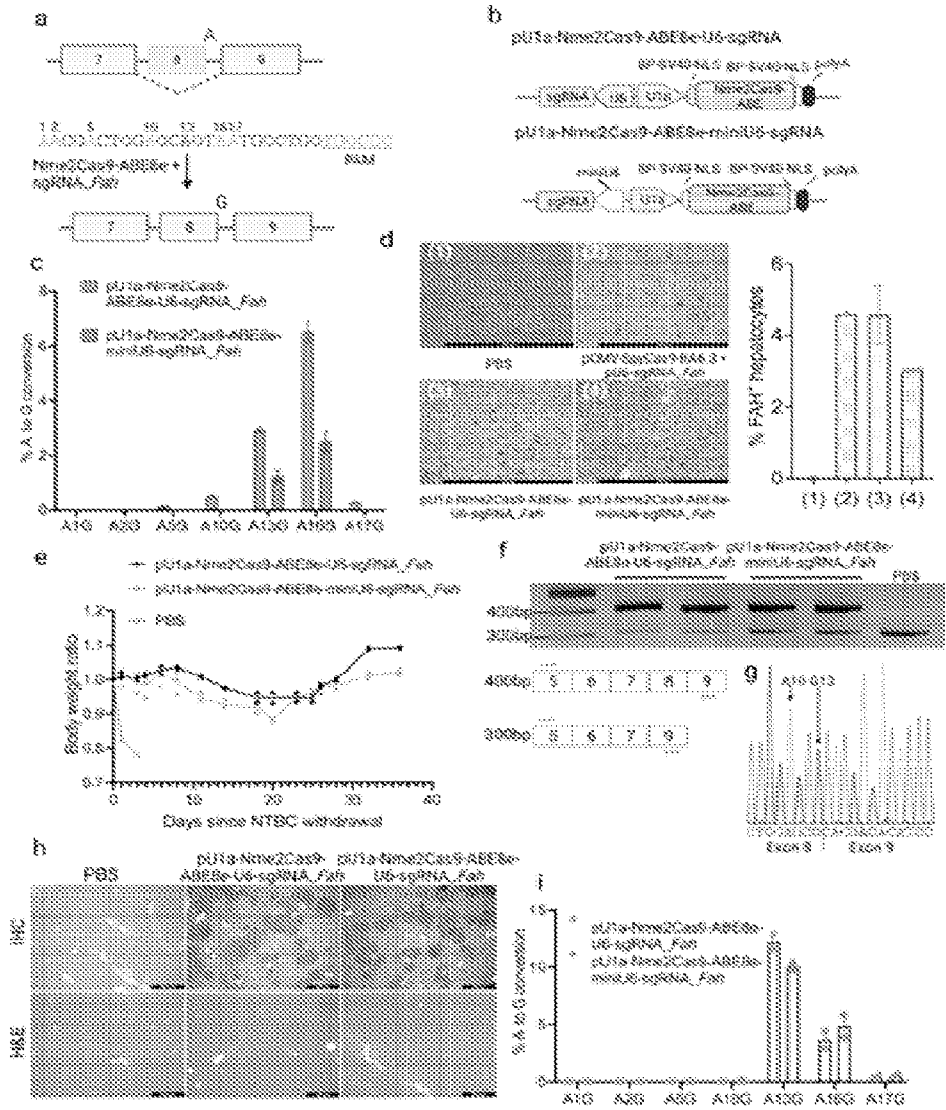


Fig. 27

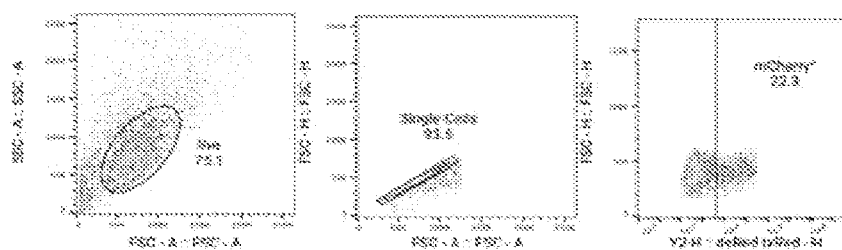


Fig. 28

Figs. 28A-B

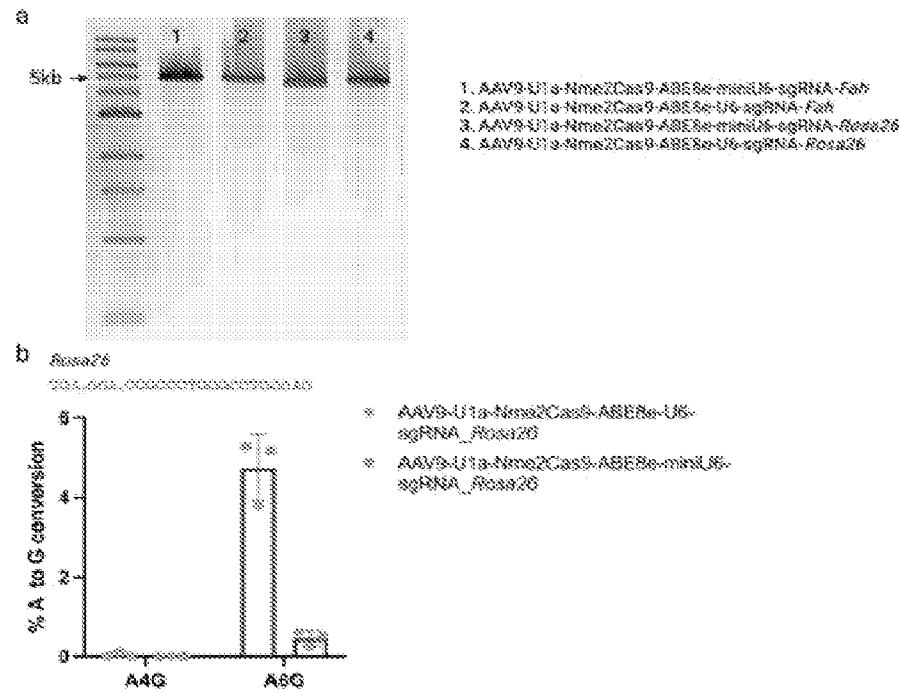


Fig. 29

Figs. 29A-B

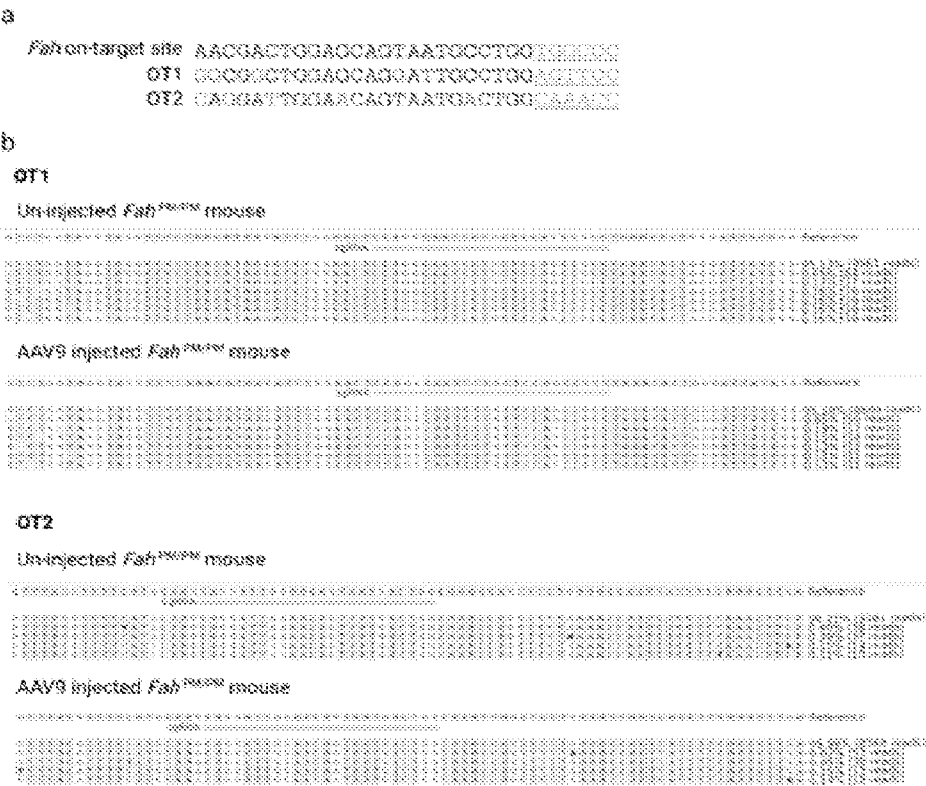
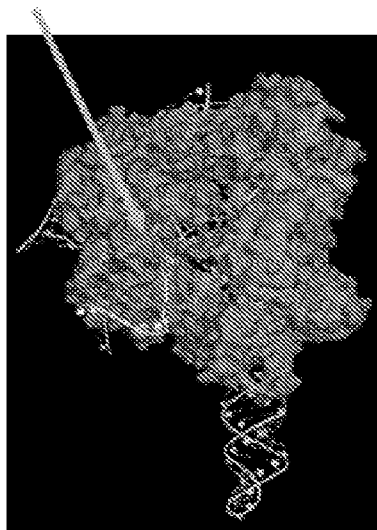


Fig. 30

Deaminase Target DNA



180
degree
turn

N-terminus

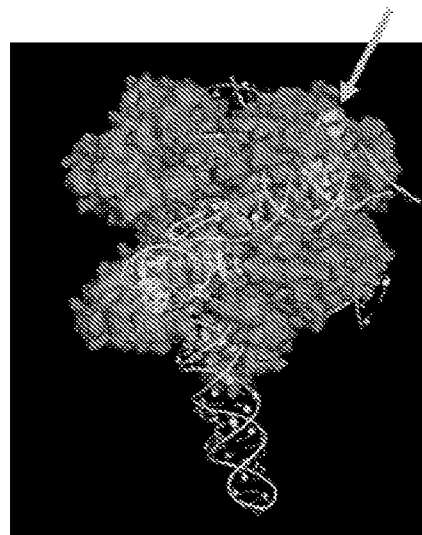


Fig. 31**Fig. 31A****Fig. 31B**

2xBPSV40-Nme2Cas9: 3369bp, 1122AA

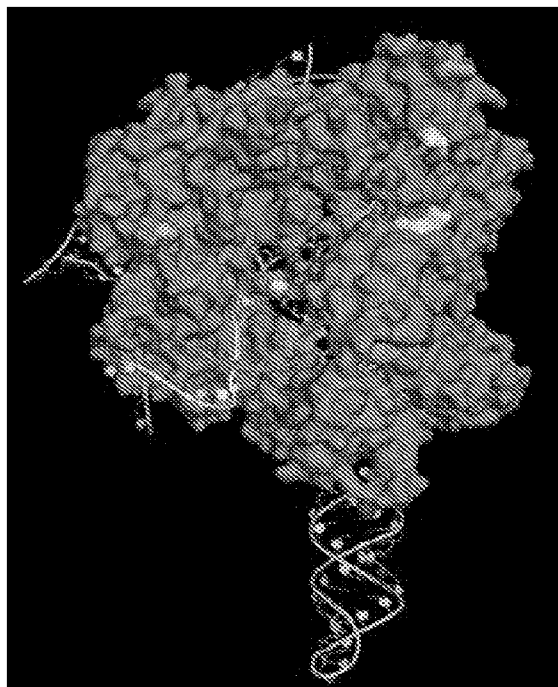
Fig. 31C

Fig. 32

Fig. 32A

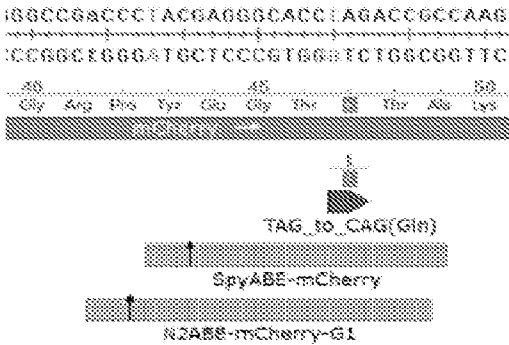


Fig. 32B

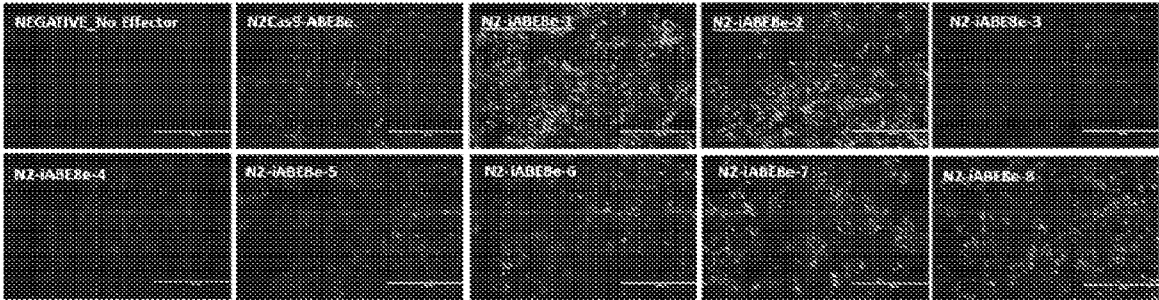


Fig. 33

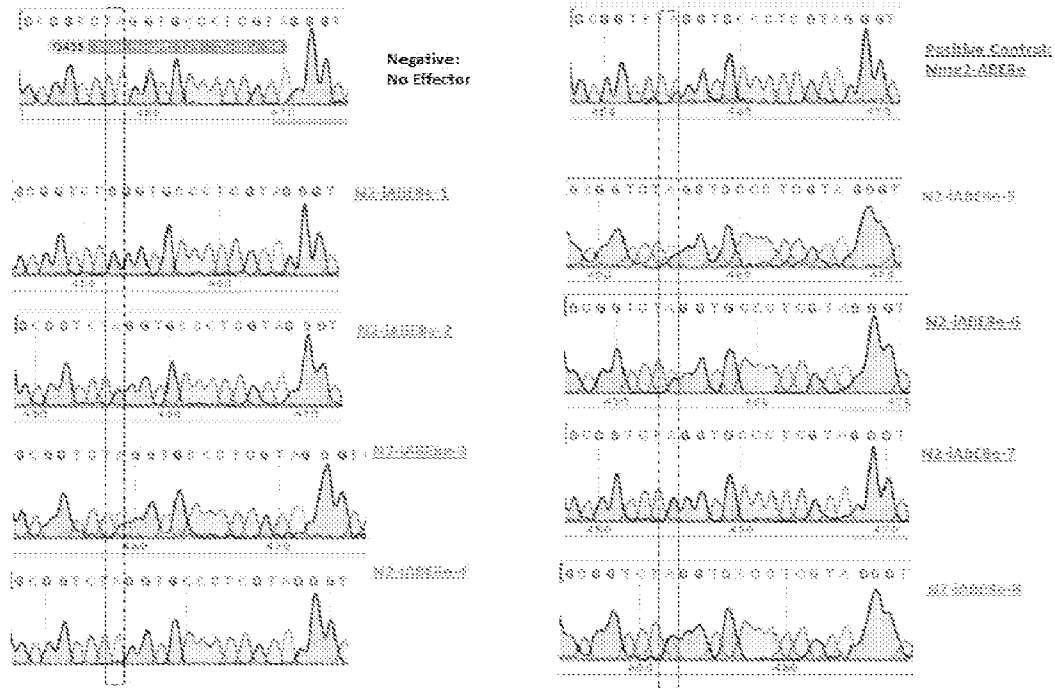


Fig. 34

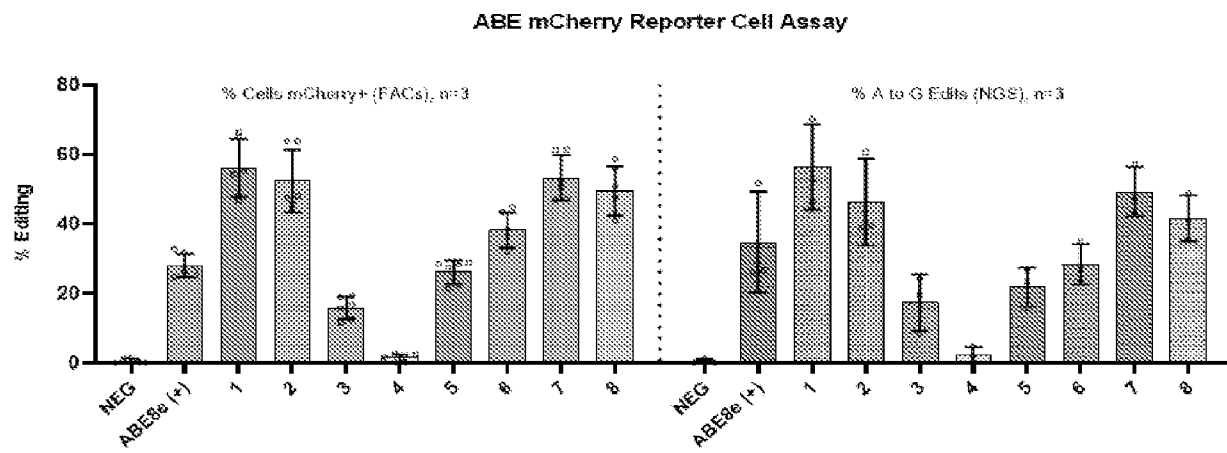


Fig. 34A

Fig. 34B

Fig. 35

Fig. 35A

LINC01588-DS12

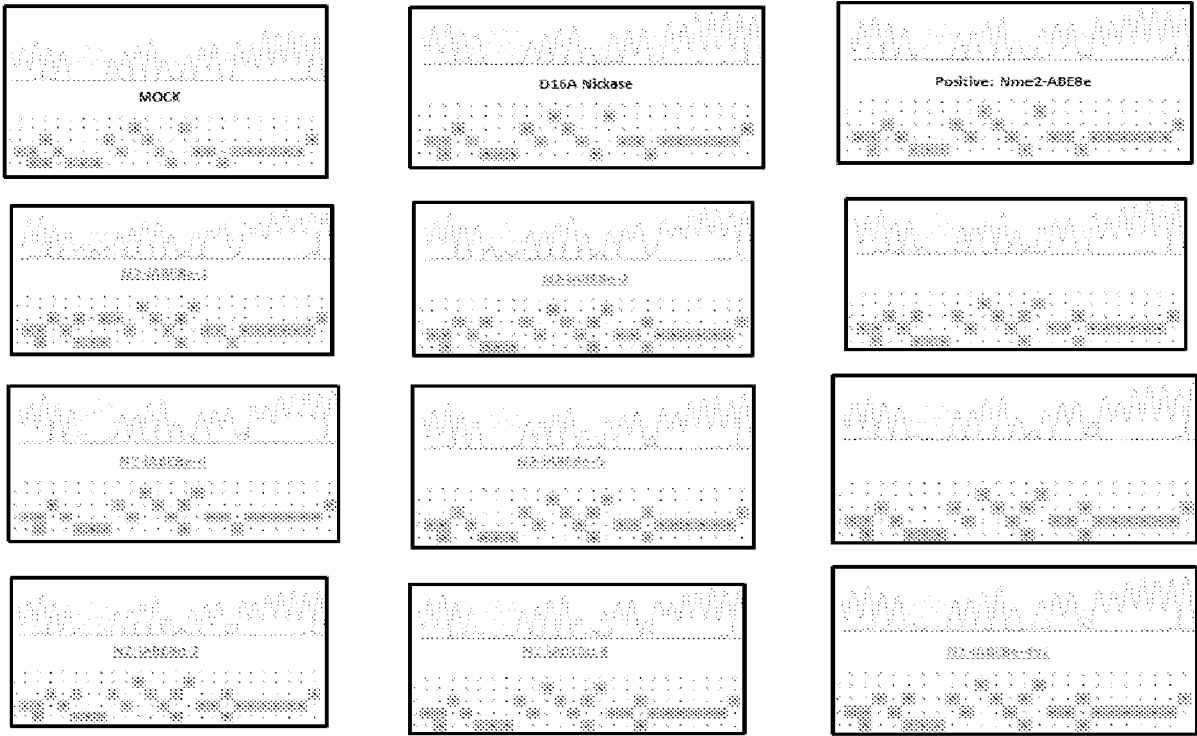


Fig. 35 (cont'd)

Fig. 35B

FANCF-DS28

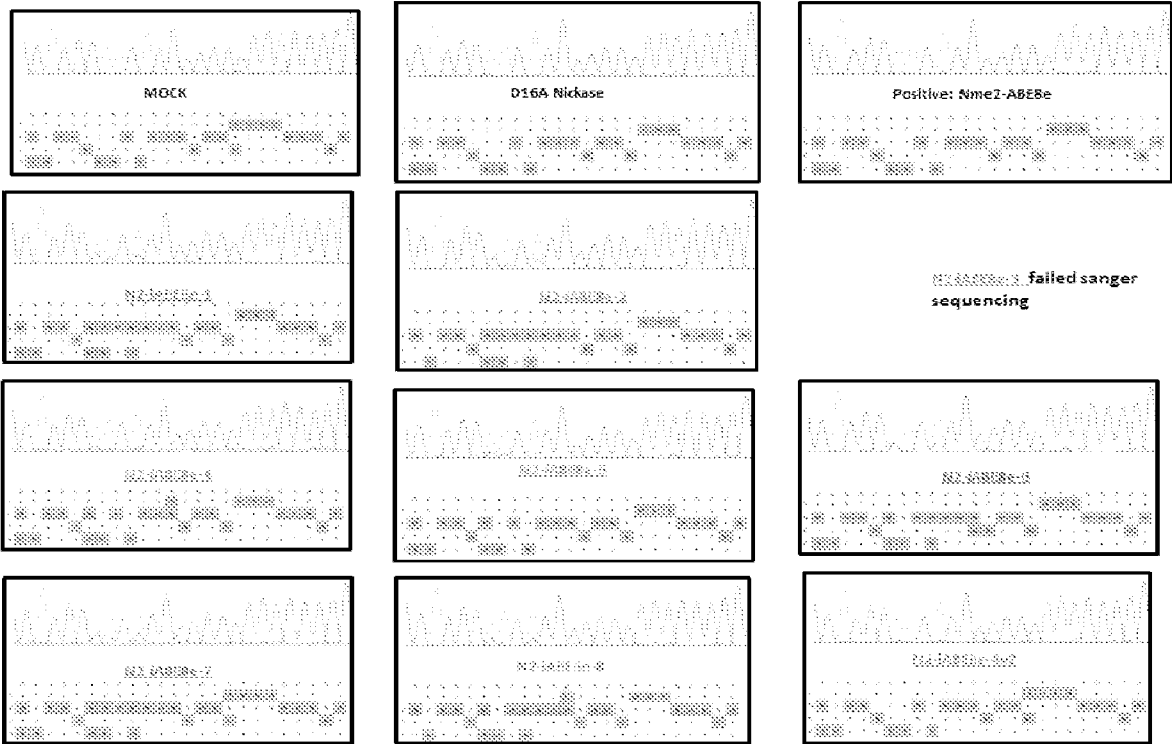


Fig. 35 (cont'd)

Fig. 35C

MECP2-G2

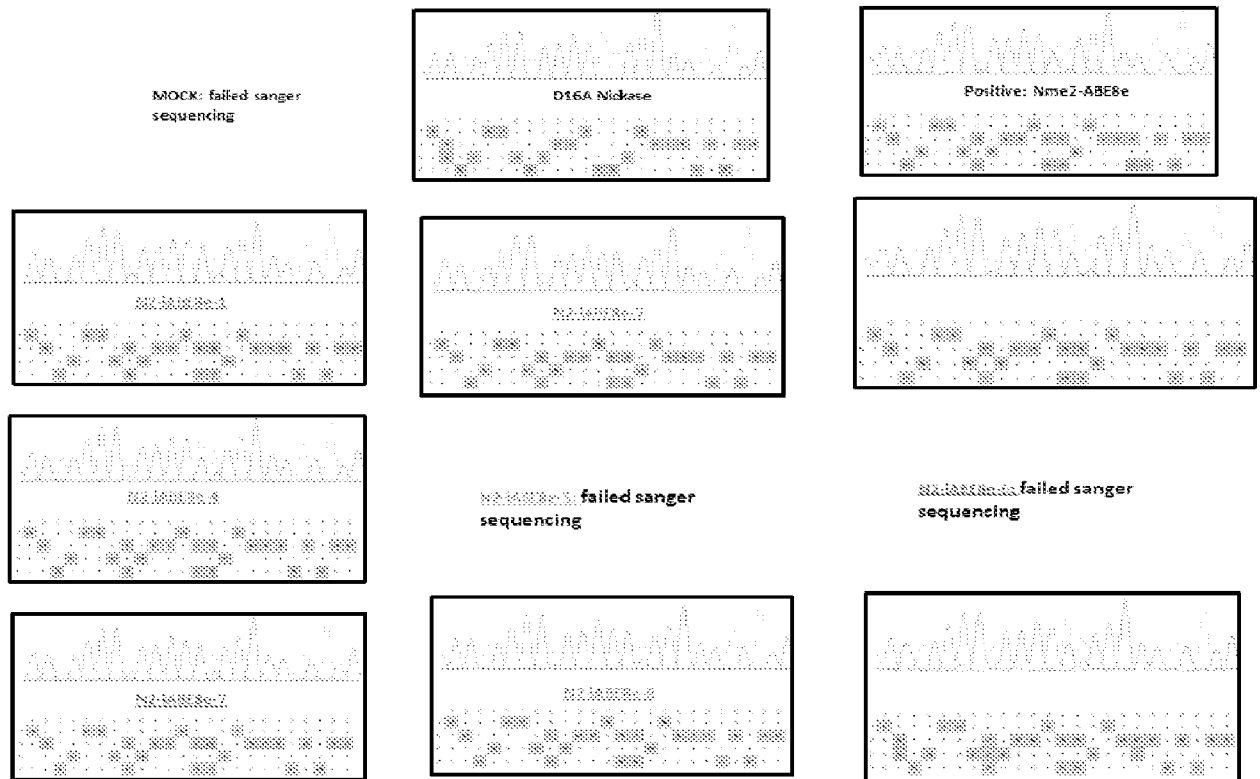


Fig. 36

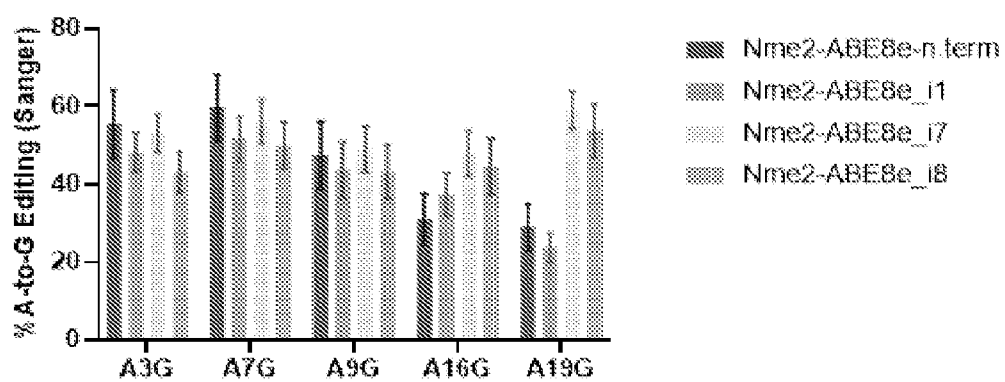


Fig. 37

Fig. 37A

Mouse CLN3 Exon 5

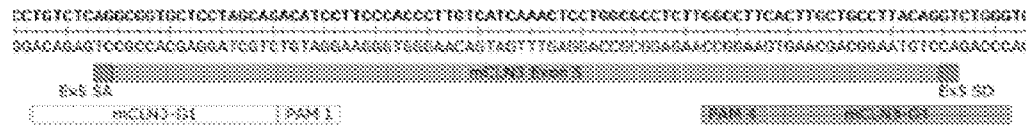


Fig. 37B

Human CLN3 Exon 5

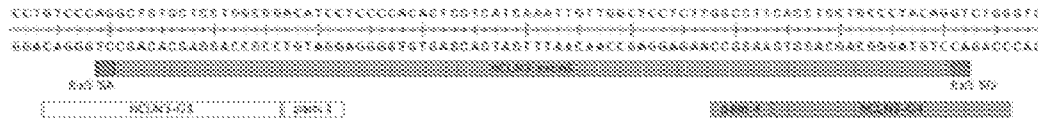


Fig. 38

Splice Donor
Guide-1

Mouse- TGTCTCAGGCGGTGCTCCTAGCAGACATCC

Human- TGTCCCAGGCTGTGCTCCTGGCGGACATCC

Splice AcceptorA
Guide 4

Mouse- CCCAGACCTGTAAGGCAGCAAGTGAAGGCC

Human- CCCAGACCTGTAGGGCAGCAGGTGAAGGCC

Fig. 39

Fig. 39A

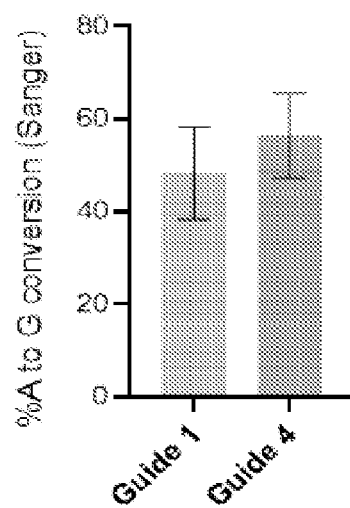


Fig. 39B

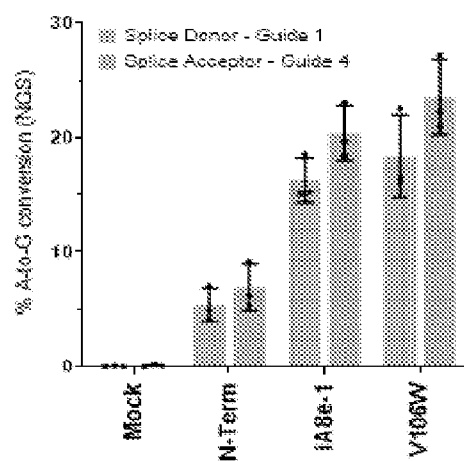


Fig. 40

Fig. 40A

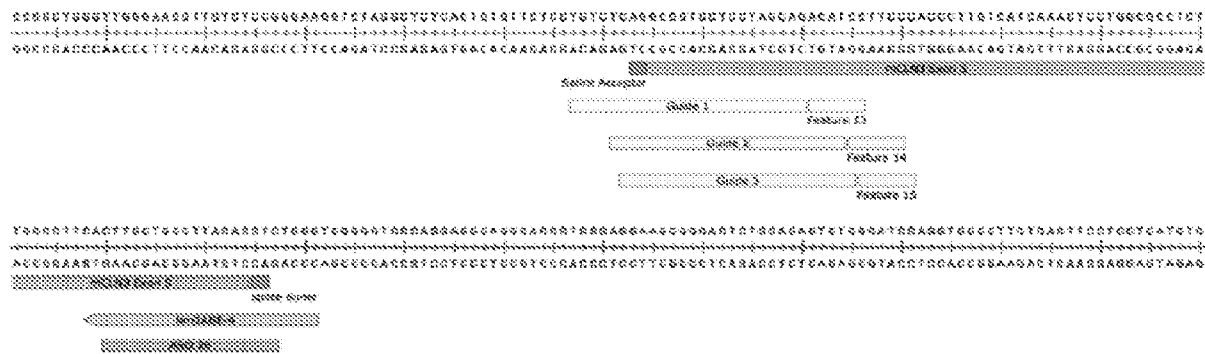


Fig. 40B

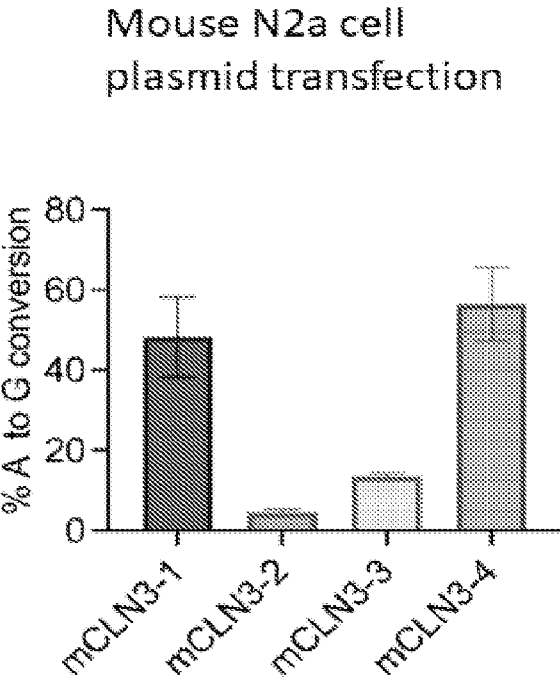


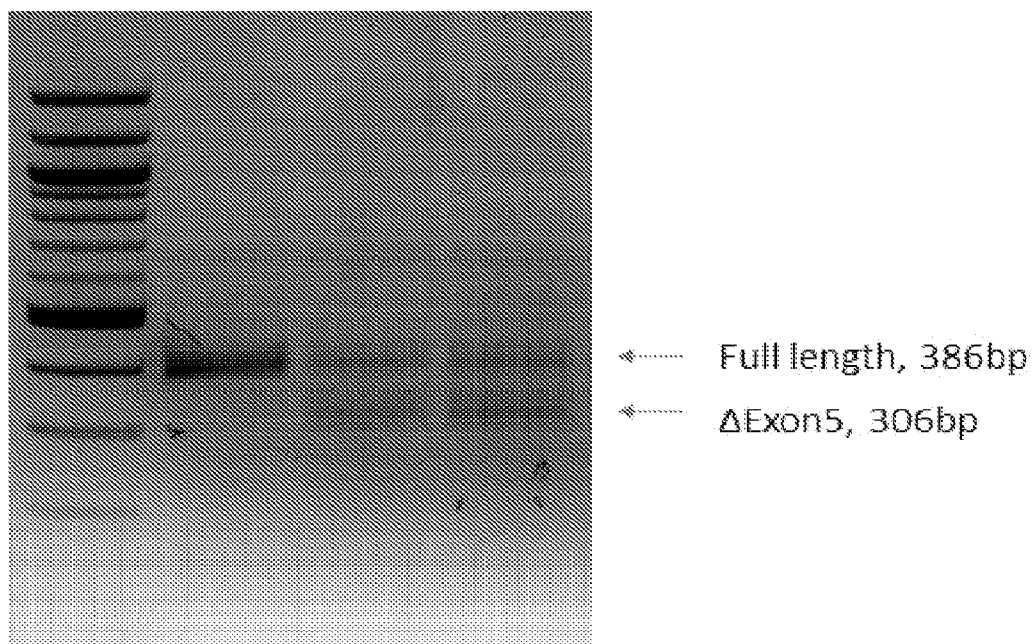
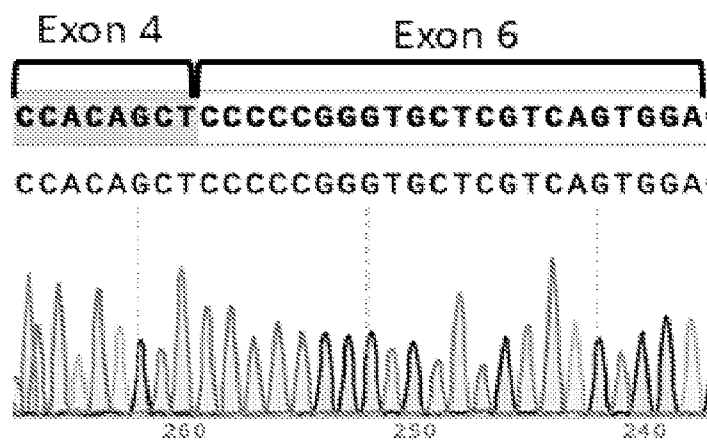
Fig. 40 (cont'd)**Fig. 40C****Fig. 40D**

Fig. 41

Fig. 41A

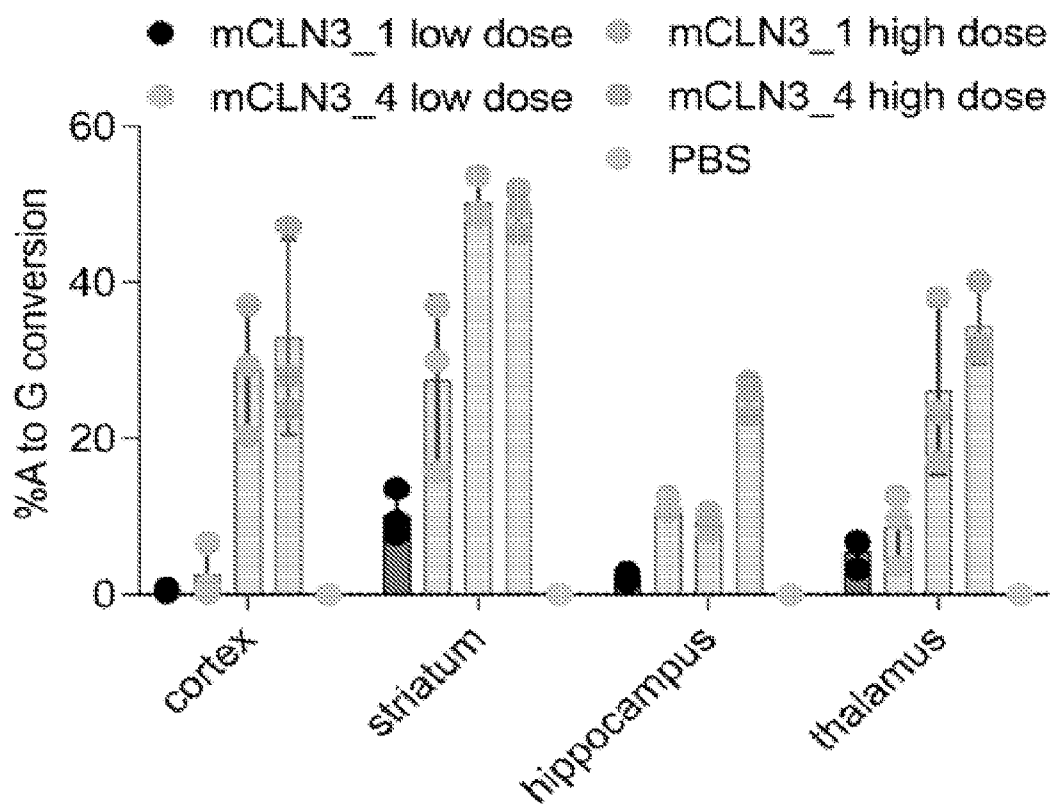


Fig. 41 (cont'd)

Fig. 41B

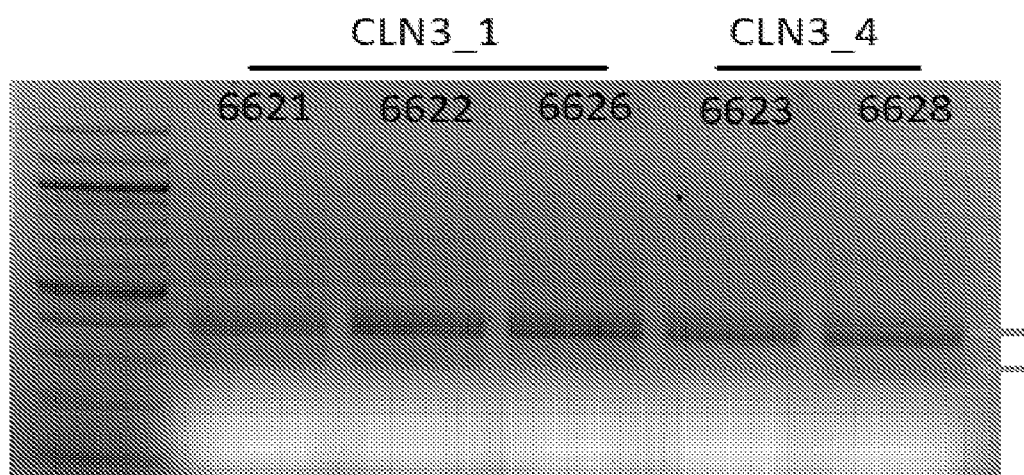


Fig. 41C

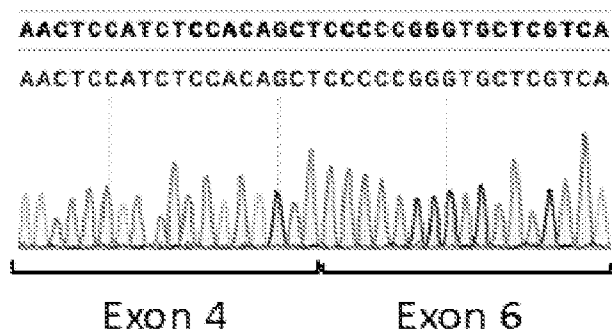


Fig. 42

Fig. 42A

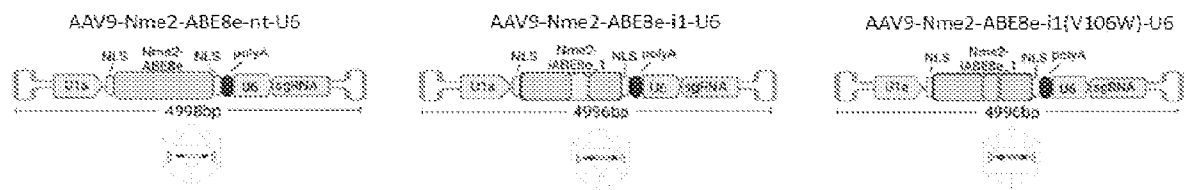


Fig. 42B

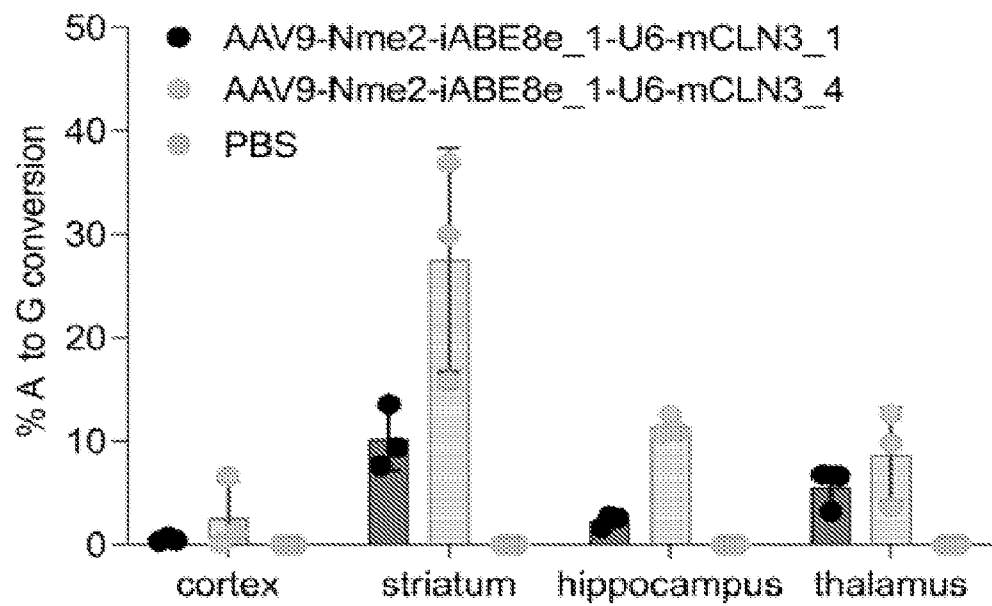


Fig. 42 (cont'd)

Fig. 42C

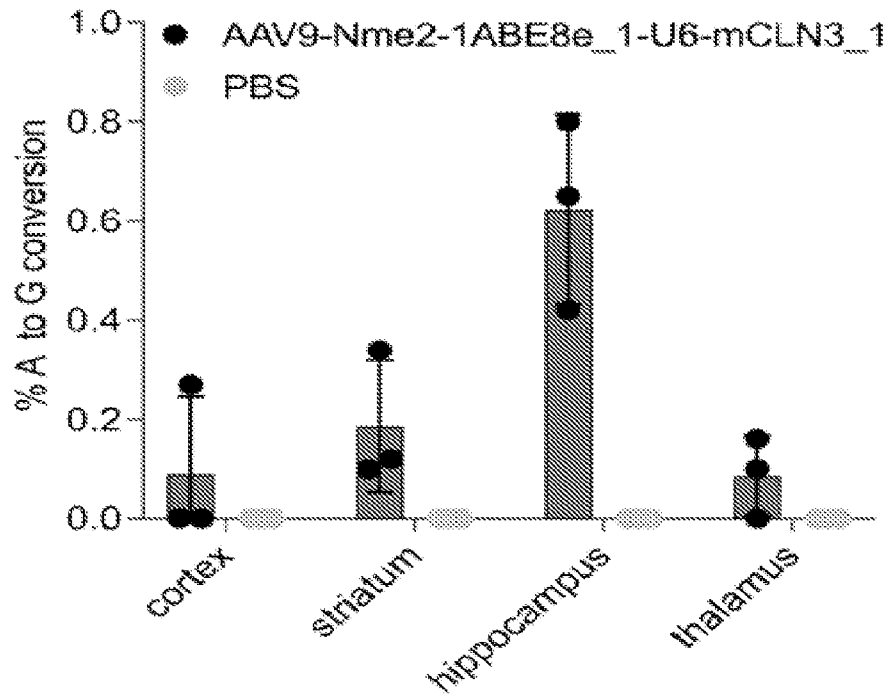


Fig. 42D

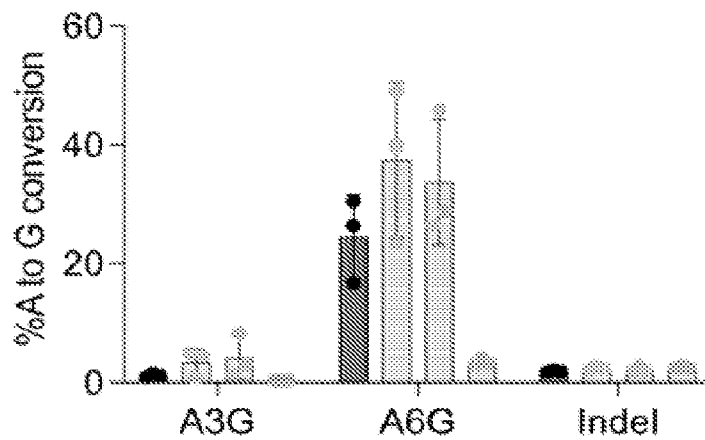


Fig. 42 (cont'd)

Fig. 42E

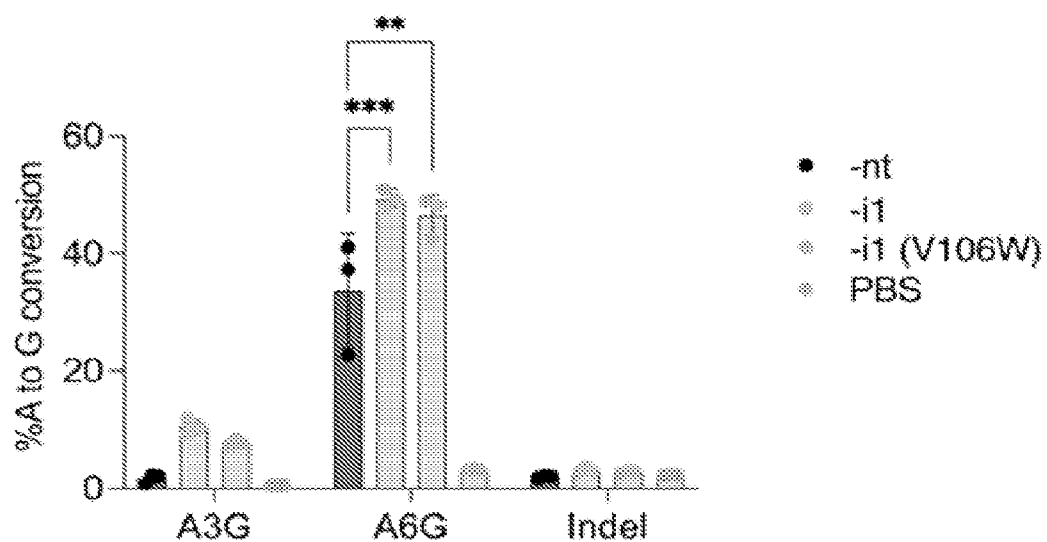


Fig. 43

Fig. 43A

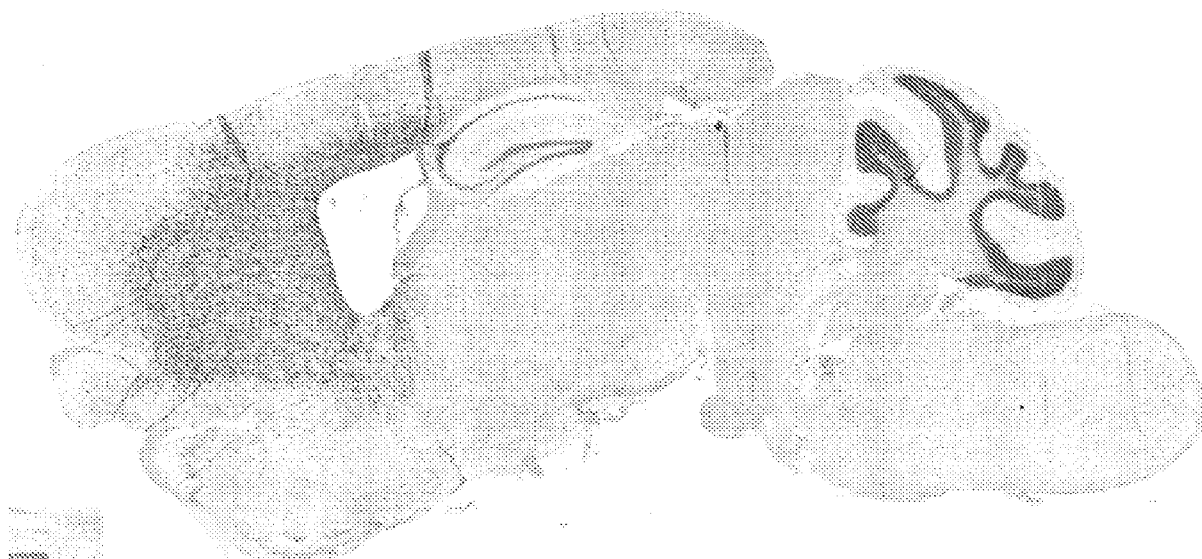


Fig. 43B



Fig. 43 (cont'd)

Fig. 43C

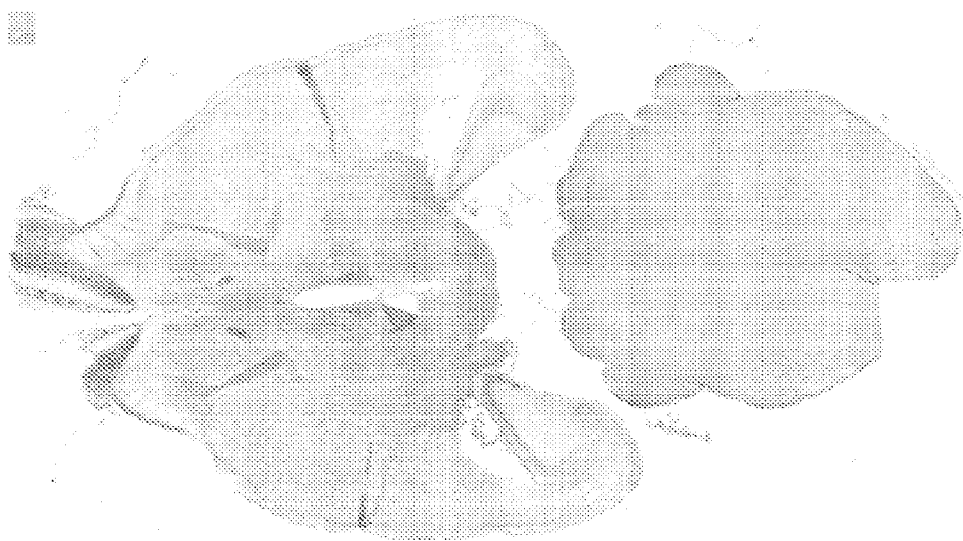


FIG. 44

Fig. 44A

Gly Arg Gly Ser Pro Pro Arg Arg Glu Gln Lys Pro Pro

Nme2-ABE Editing

Fig. 44B

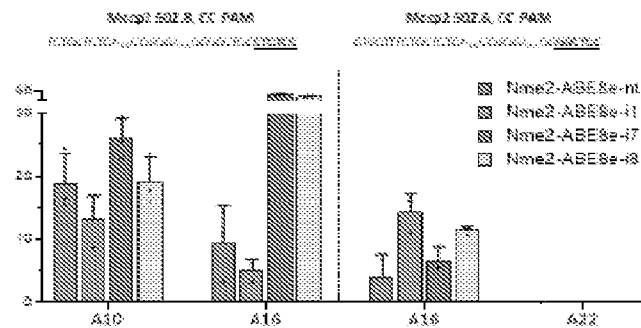


Fig. 45

Fig. 45A

Gln	Glu	Thr	Ala	Pro	Pro	Thr	Lys	Lys	Arg	Lys	Thr
Gln	Glu	Thr	Val	Leu	Pro	Ile	Lys	Lys	Cys	Lys	Thr
C A G G A G A C C G	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C	A A G A A G A A G	A A G A A G A A G	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C
G T G C T C T C G G	A T G A G G G G T G	A T G A G G G G T G	A T G A G G G G T G	A T G A G G G G T G	A T G A G G G G T G	A T G A G G G G T G	T T C T T C T T C	T T C T T C T T C	A C C C A C C C A C	A C C C A C C C A C	A C C C A C C C A C

Nme2-ABE editing

Fig. 45B

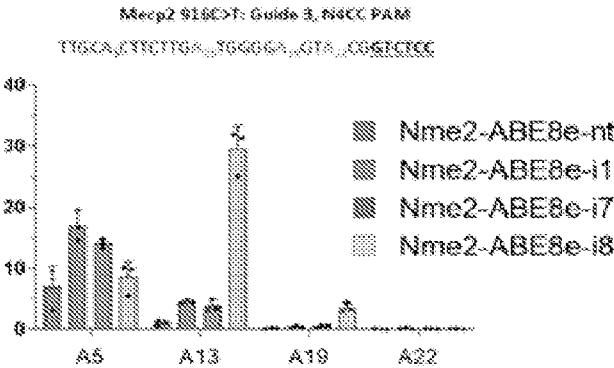


Fig. 46

1

Fig. 46A

Met	Val	Ile	Lys	Arg	Pro	Gly	Arg	Lys	Arg	Lys	Ala
Met	Val	Ile	Lys	Arg	Pro	Gly	Arg	Lys	STOP	Lys	Ala
A T G G T G A T C A A A C G C C C C G G C A G G A A G G A A A A G C T											
T A C C A C T A G T T T G C G G G G C C G T C C T T C A C T T T T C G A											

Nme2-ABE Editing

Fig. 46B

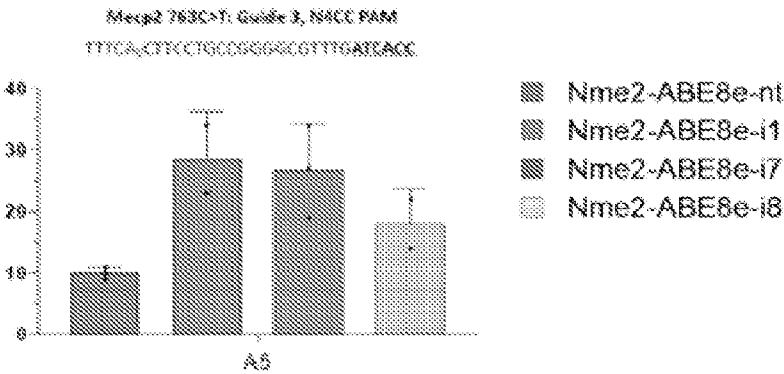


Fig. 47

Fig. 47A



Fig. 47B

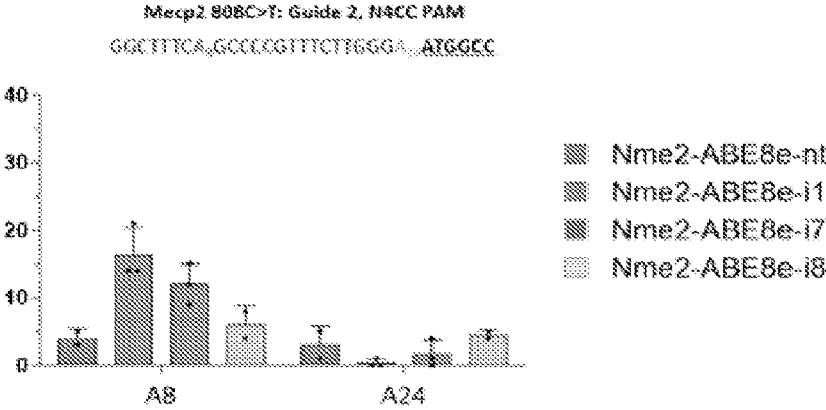


Fig. 48

Fig. 48A

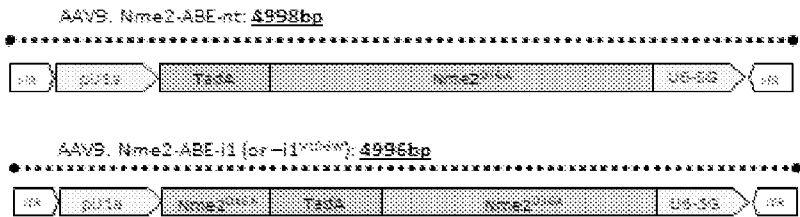


Fig. 48B

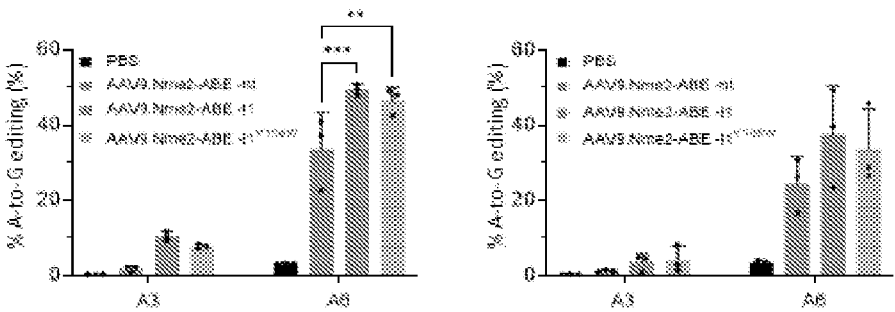


Fig. 48C

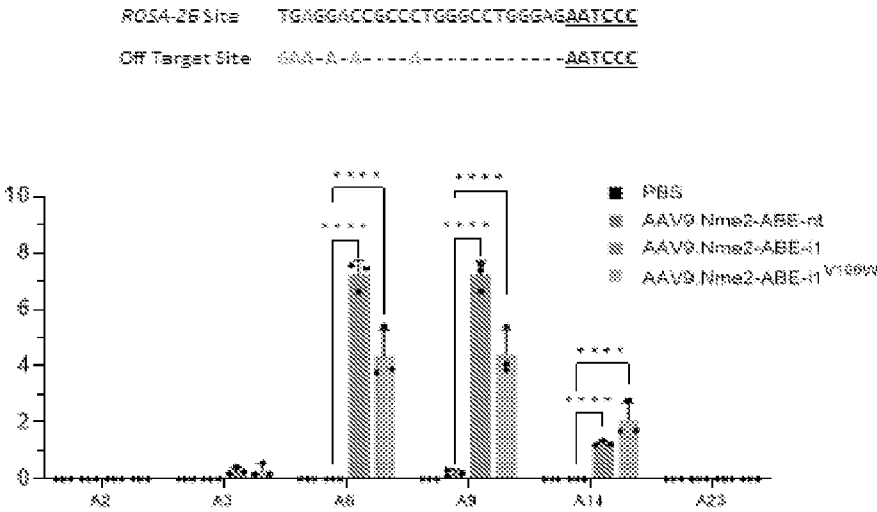


Fig. 49

Fig. 49A

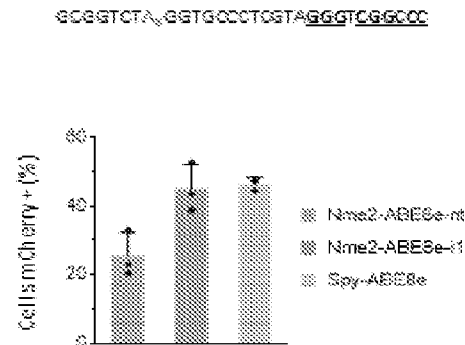


Fig. 49B

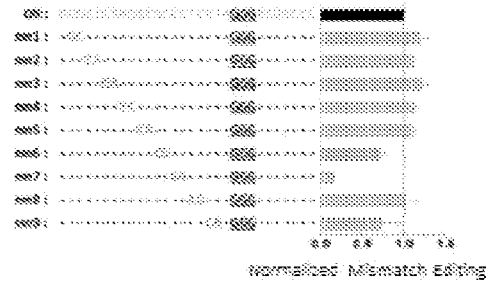


Fig. 49C

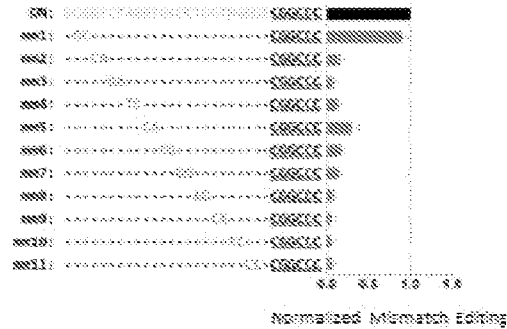


Fig. 49 (cont'd)

Fig. 49D

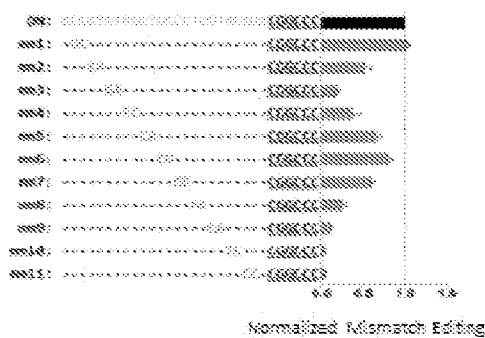


Fig. 49E

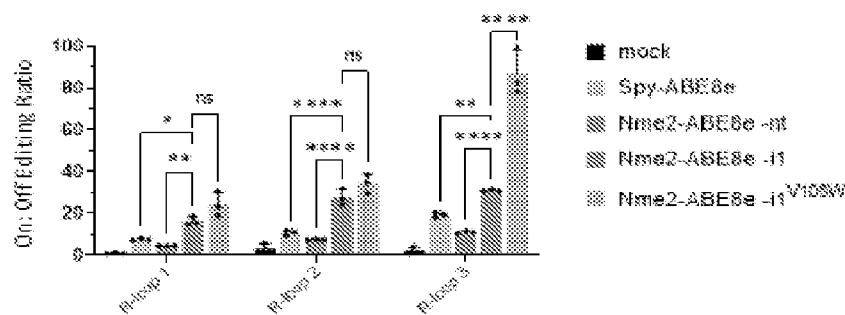


Fig. 50

Fig. 50A

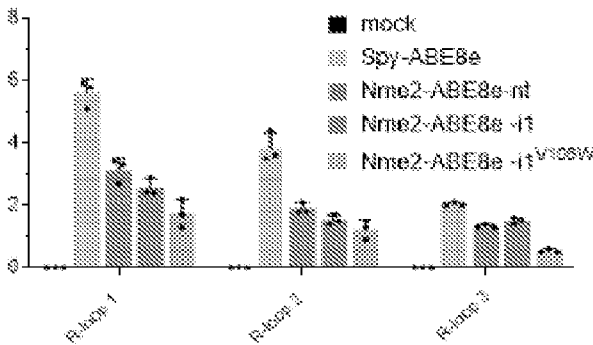


Fig. 50B

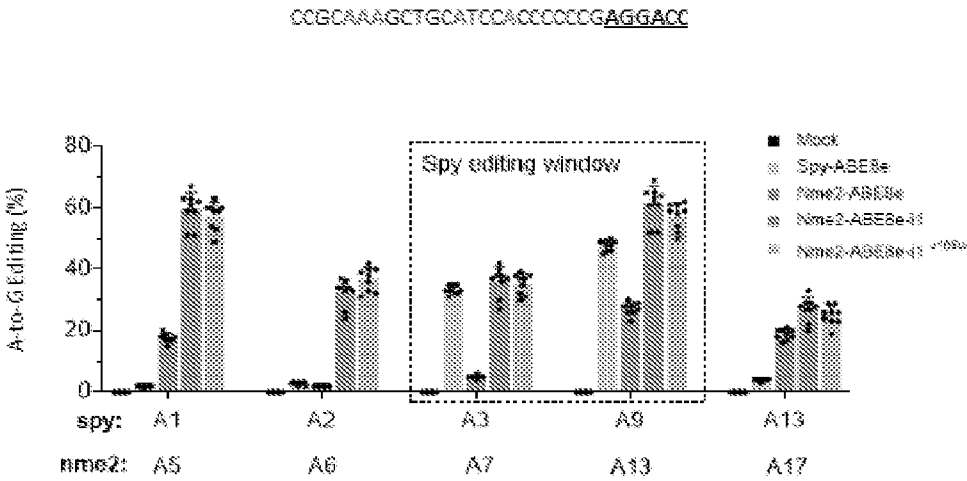


Fig. 51

Fig. 51A

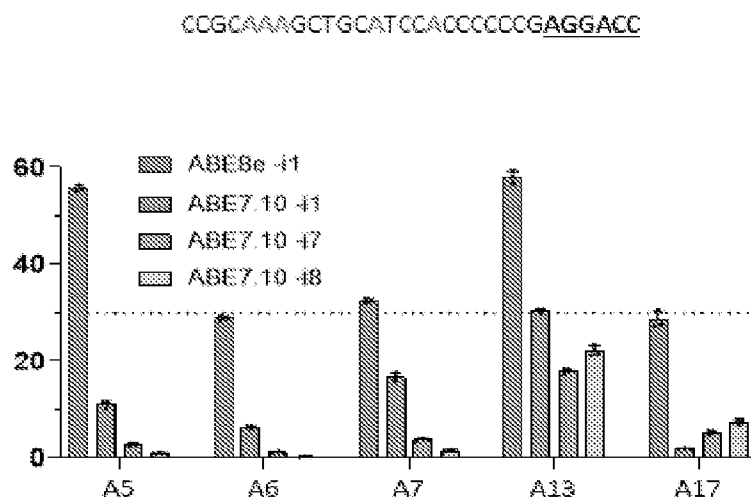


Fig. 51B

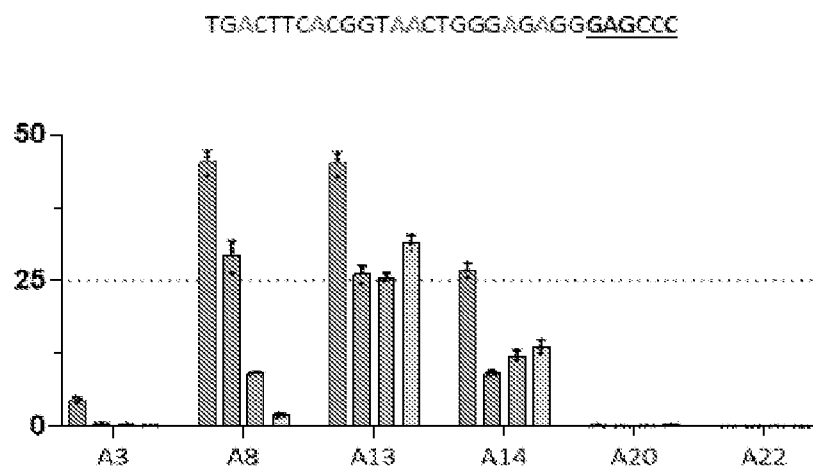


Fig. 51 (cont'd)

Fig.51C

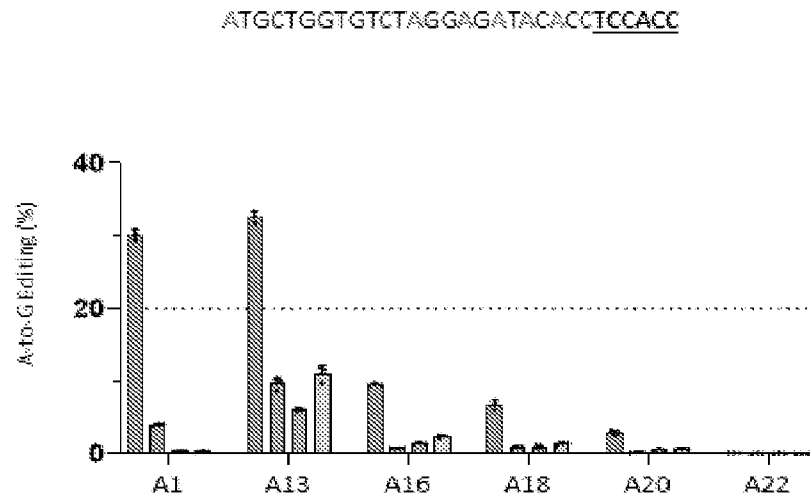


Fig. 52

Fig. 52A

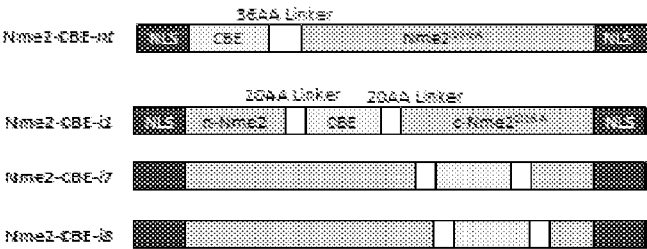


Fig. 52B

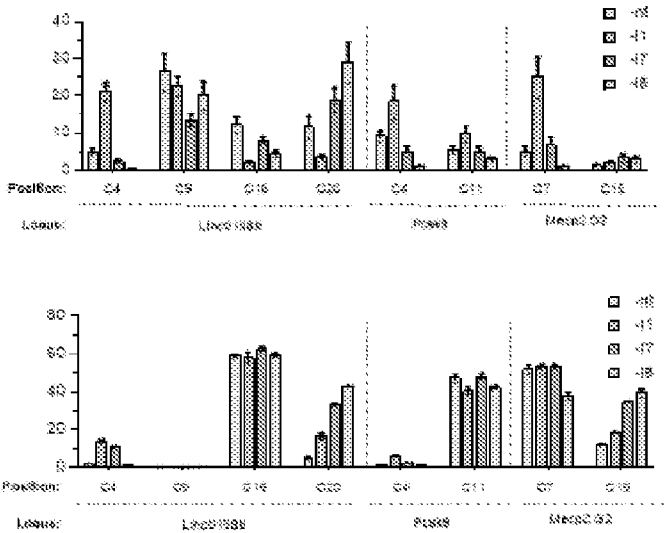


Fig. 52 (cont'd)

Fig. 52C

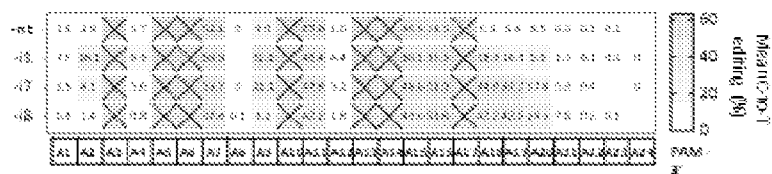


Fig. 53

Fig. 53A

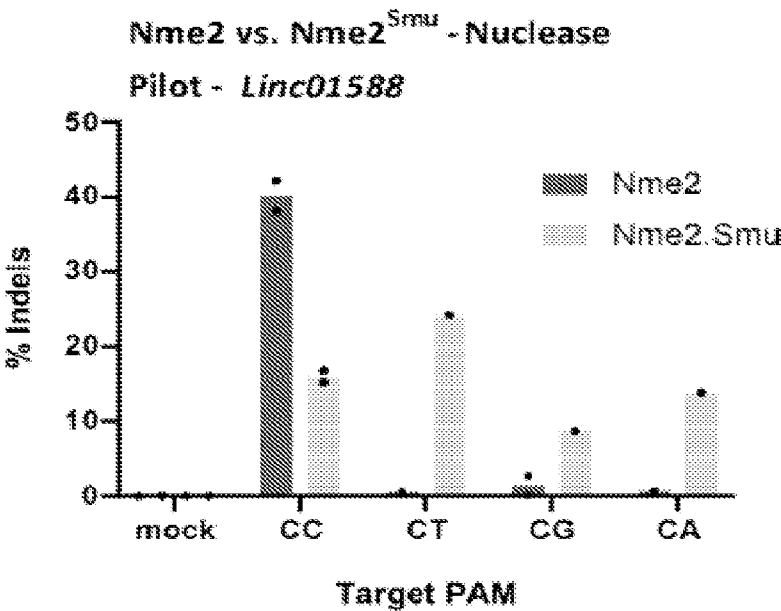


Fig. 53B

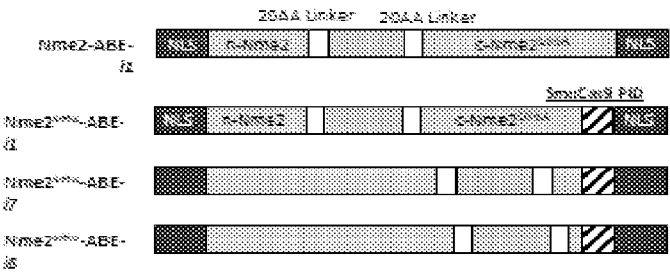


Fig. 53 (cont'd)

Fig. 53C

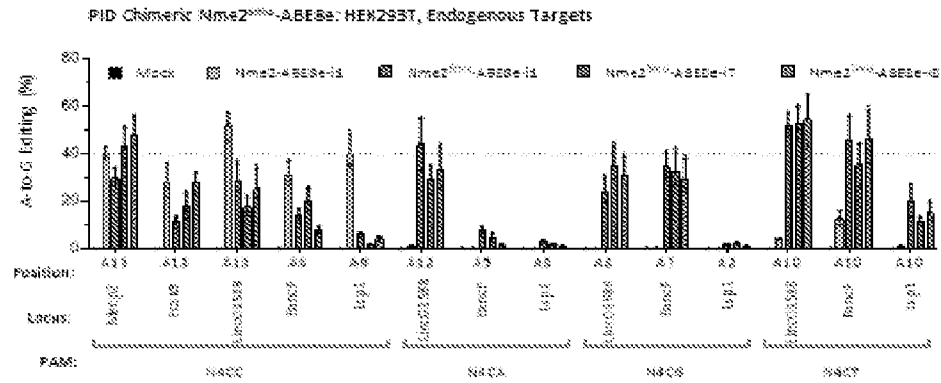


Fig. 53 (cont'd)

Fig. 53E

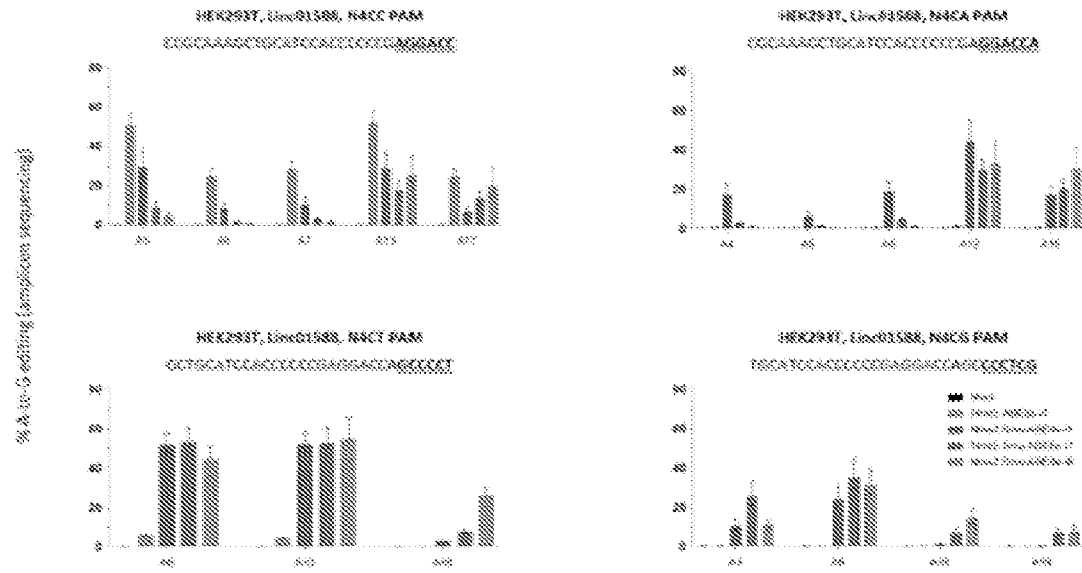


Fig. 53F

Fig. 54

Fig. 54A

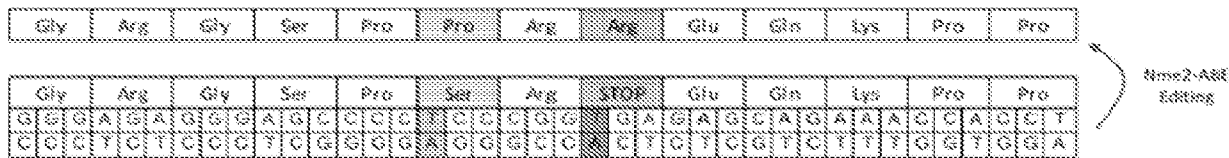


Fig. 54B

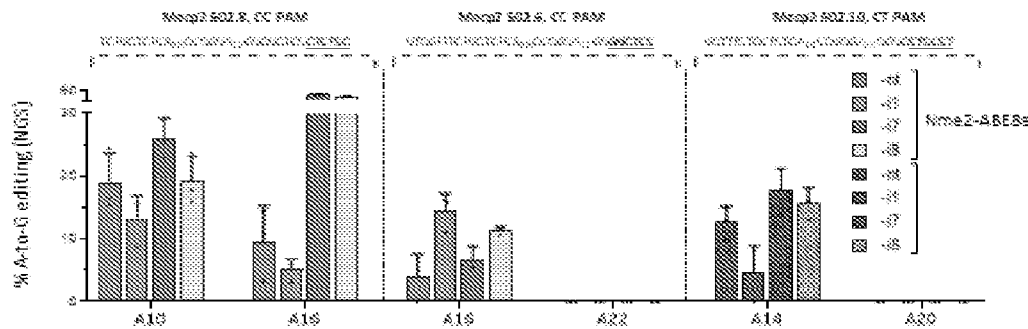


Fig. 55

Fig. 55A

Gln	Glu	Thr	Ala	Pro	Pro	Thr	Lys	Lys	Arg	Lys	Thr
C	A	G	G	A	G	A	C	C	G	T	A
G	T	C	C	T	C	T	G	G	C	A	T

Gln	Glu	Thr	Val	Leu	Pro	Ile	Lys	Lys	Cys	Lys	Thr
C	A	G	G	A	G	A	C	C	G	T	A
G	T	C	C	T	C	T	G	G	C	A	T

Meep2-288 Editing

Fig. 55B

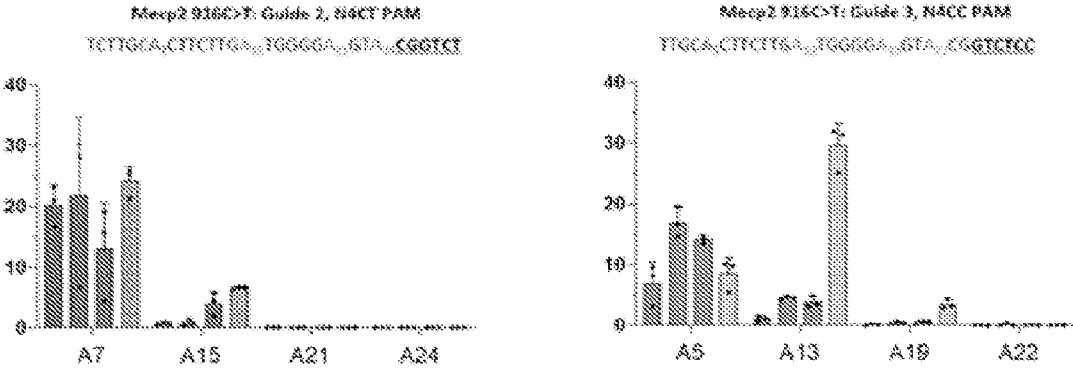


Fig. 56

Fig. 56A

Met	Val	Ile	Lys	Arg	Pro	Gly	Arg	Lys	Arg	Lys	As
Met	Val	Ile	Lys	Arg	Pro	Gly	Arg	Lys	STOP	Lys	As
A T G G T G A T C A A A C G C C C C G G C A G G A A G T G A A A A G C T											
T A C C A C T A G T T T G C G G G C C G T C C T T C A C T T T T C G A											

Splice site
editing

Fig. 56B

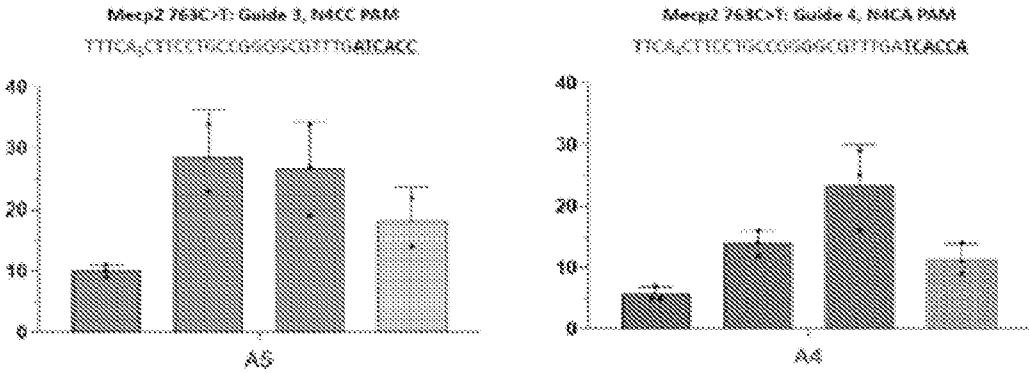


Fig. 57

Fig. 57A

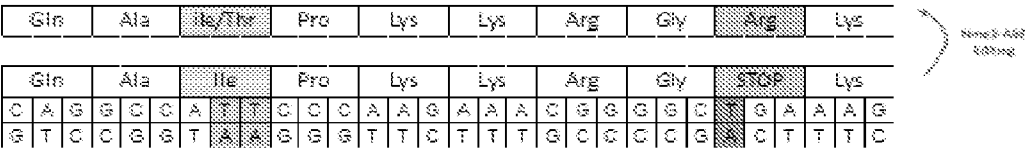


Fig. 57B

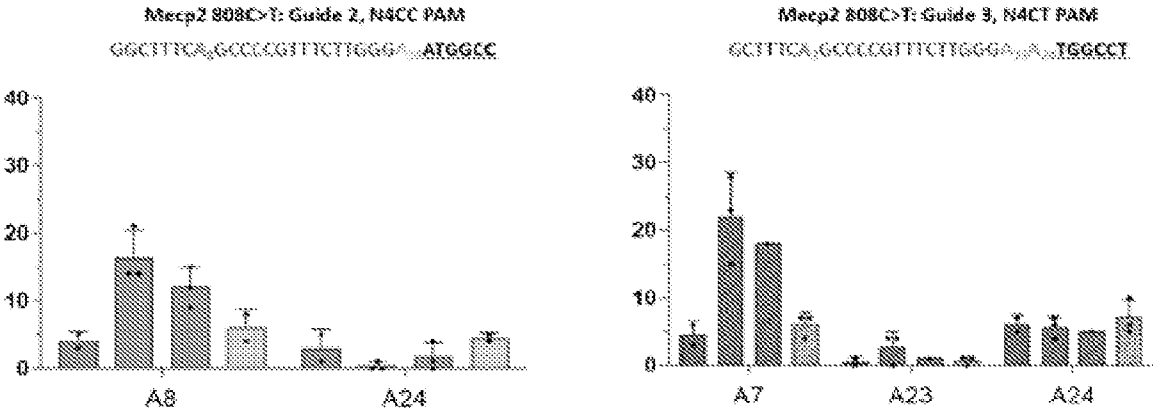


Fig. 58

Fig. 58A

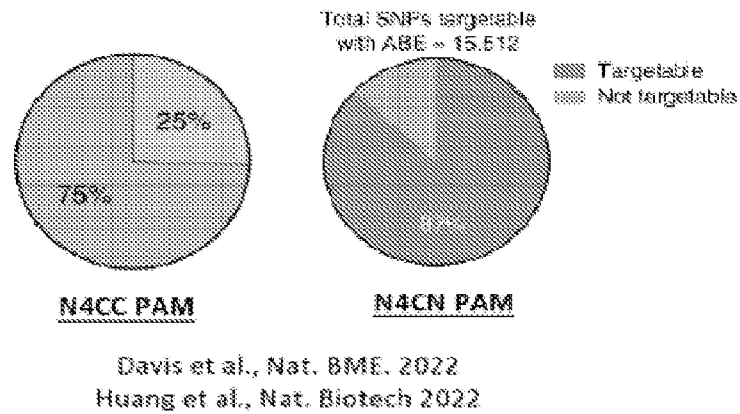


Fig. 58B

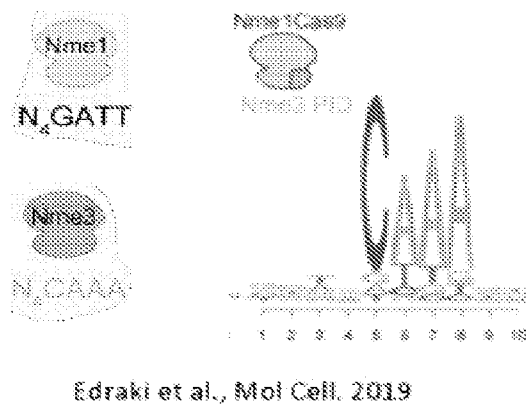
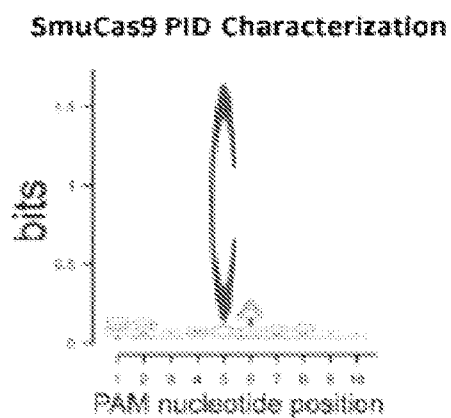


Fig. 58 (cont'd)

Fig. 58C



J. Lee et al., Sontheimer et al., mBio, 2018

J. Wei et al., Y. Wang, Biorxiv, 2022

Fig. 59

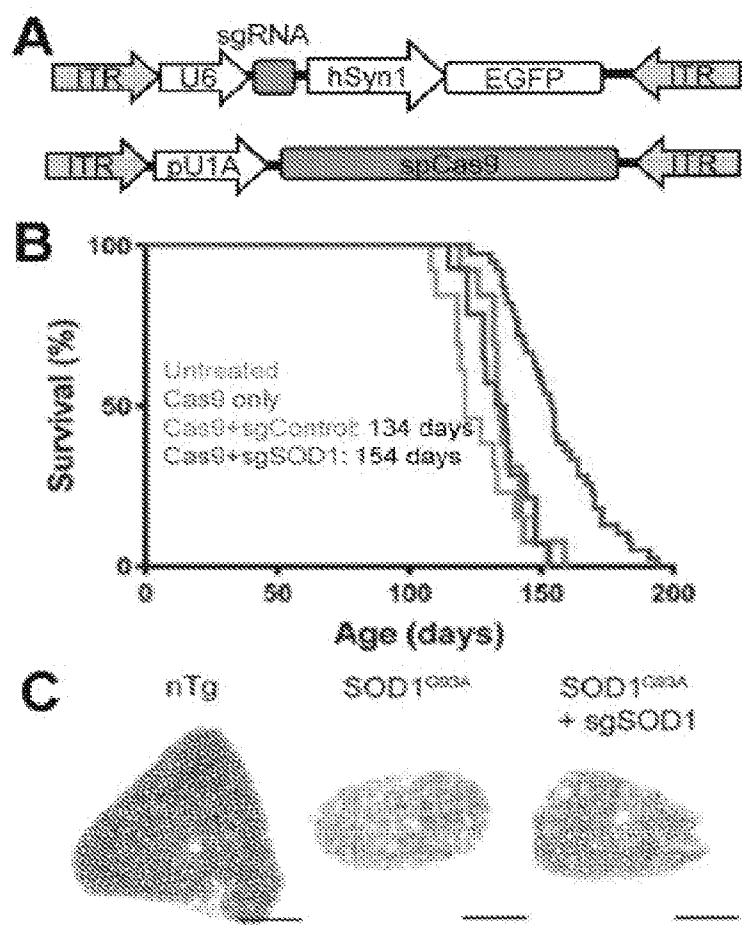


Fig. 60

Fig.s 60A and 60B

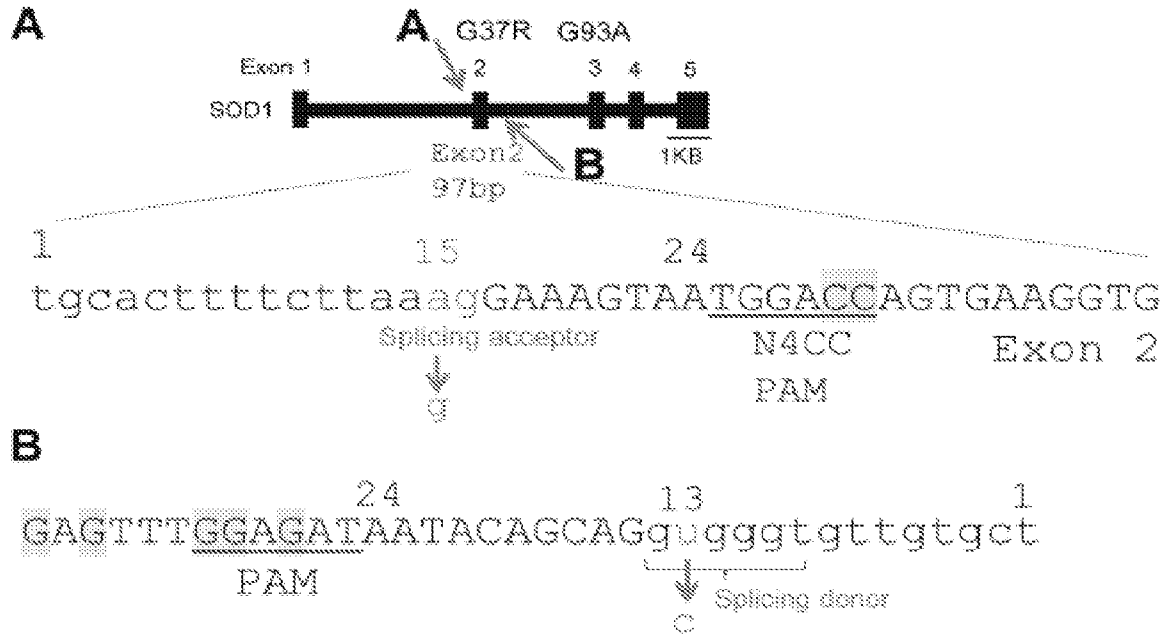


Fig. 61

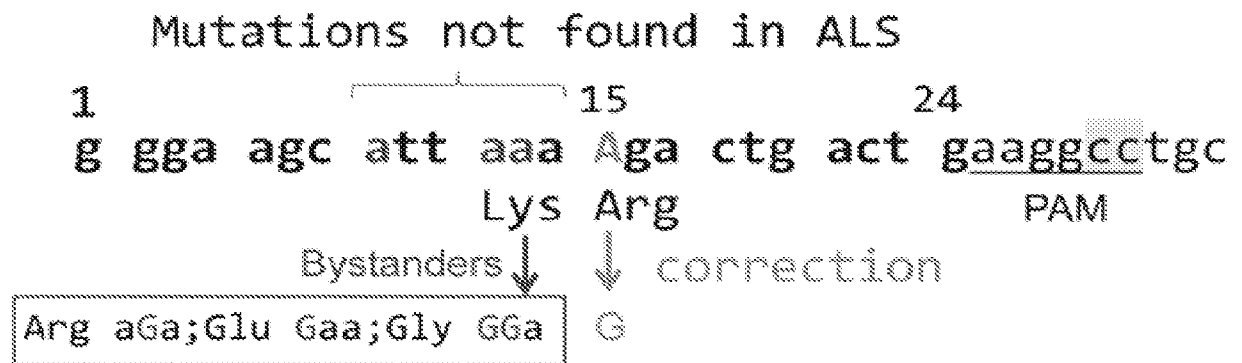


Fig. 62

Fig. 62A

KRTADGSEFESPKKKRKVEDMAAFKPNPINYILGLDIOIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVP
 KTQDSLAMARRLARSYRRLTRRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKL
 TPLEWSAVLLHLIKHRCGYLSQRKNEGETADKELGALLKGVANNAHALQTQDFRTPAELALNKFESGHI
 KNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLQHCTFEPA
 EFKAANKNTYTAERFIWLTYLNNLRILEQGSERPLTDTERATLMDEFYRKSKLTYAQARKLLGLEDTAFFKG
 LRYGKDNAAEASTLMEMKAYHAISRALKKEQLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRKDKRVQP
 EILEALLKHISFDKFFVQISLKALRRIVPLMEQGGKRYDEACAEIVGDHYGKKNTEEKIYLPPIPADEIRNPVVL
 ALSQARKVINGVVRRYGSPARIHETAKEVGGKSFKDKKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKS
 KDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLQSENONKONOTPY
 EYFNGKDNSREWQEFKARVETSRFPRSKKQRILLOKFDDEGFKECNLNDTRYVNRFLCQFVADHILLTOKG
 KRRVFASNGQITNLLRQFWGLRKVRAENDRHHALDAYVVACSTVAMQOKITRFVRYKEMNAFDGKTIDK
 ETGKVLHOKTHFPQPWEFFAQEV MIRVFGKCPDGKPEFEEADTPEKLRITLLAEKLSRPEAVHIEYVTPLEVS
 APNRKMSGAHKDTLRSKRFVKHNEKISVKRVWLTEUKLADLENMVNYKNGREIELYEALKARLEAYGG
 NAKQAFDPKDNPFYKKGQQLVKA VRVEKTQESGVLLNKKKNAYTIADNATMVRVDVYTCAQKNYLVPVY
 YVQYVAQQILPNRAVTSKGSEADWDLIDESFEFKFSLSRGDLVEMISNKGRIFGYYNGLDRANGSIGREHDL
 EELSKGKDGVRHVGVKTATAFNKYHVDPLGKEIHRCSSEFPRPTLKIKSKKEDKRTADGSEFESPKKKRKV

Fig. 62B

gtctaa

FIG. 63

Fig. 63A

PKKKRKVEDIKRPAATTEKAGQAKKTEKMAAFKPNPINYILGLDICIASVGVAMVEIDEEENPIRLIDLGV RVF
 ERAEVPKTGDSLAMARRLARSVRRLTRRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLFNTFWQLRAA
 ALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKEELGALLKGVANNAHALQTGDFRTPAELALNKFE
 KESGHIRNQRODYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGQLKEGIETLLMTQRPALSQDAVQKMLGH
 CTFEPAEPKAAKNTYTAERFIWLTCLNNURILEQQSERFLTQTERATLMDEPYRKSKLTYAQARKLLGLEQ
 TATFKQLRYGKDNAEASTLMEKAYHAISRALEKKEGLEKDKKSPLNLSSSELQDEIGTAFSLPKTDEDITGRL
 KDRVQPEILEALLKHISFDKFVOISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTTEKIYLPPIPADEI
 RNPVVLRAISOARKVINGVVRRYGSPARIHETAREVGKSPKDRKEIEKROEENRKDRKAAAKFREYFPN
 FVGEFPSKDILKRLLYEQQHOKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSPNNKVLVLGSENQNK
 GNQTPYEYFNGKDNREWQEFKARVETSRFPRSKKORILLQKFDGEGFKECNINDTRYVNRFLCOFVADHI
 LLTGKGRRVFASNGQITNLLRGFWGLRKVRAENDRIHALDAVVVACSTVAMQOKITRFVRYKEMNAFD
 GKTIDKETOKVLHQKTHFFQPWEEFAQEV MIRVFQKPDOKPEFEEADTPEKLRTLLAEKLSRPEAVHEYV
 TPLFVSRAPNRKMSQAHKDTLRSKRFPVKHNEKISVKRVWLTEIKLADLENMNYKNOREIELYEALKAR
 LEAYGGNAKQAFDFKDNPFYKKGGLVKAVRVEKTOESGVLLNKKNAYTIADNGDMVRVDVYTKACK
 NYLVPVYVWQVAQGHLPNRAVTSQKSEADWDLIDESFEFKFSLSRGDLVEMISNKGHIFGYVNGLDRANGS
 IGIREHIDLEKSKGKDOVIRVGVKTATAFNKYHYDMLGKEHRCSEHPTLKNKSKKKRPAATKKAGQAKK
 KKYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAAAAFAAKKKKLD*

Fig. 63B

87/95

FIG. 64

Fig. 64A

KRTADGSEFESPKKKRKVEDMAAFKPNPINYILGLAIGIASVQWAMVEIDEEENPIRLIDLQVRVPERAEVP
 KTGDSLAMARBLARSVRRLTERRAHRLRLRAPRLKKREGVLQAADFENGLIKSLPNTPWOLRAAALDRKL
 TPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLEGVANNAHALQTGDFRTPAELALNKFEEKESGHI
 RNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIELLMTQRPALSODAVQKMLGHCTFEPA
 EPKAAKNTYTAEERFIWLTUNNLRILEQGGSGGSGGSGGSGGSGGSEVEFSHEYWMRHALTLAKRAR
 DEREVPVGAVLVNNRVIGEGWNRAIGLIHDPTAHAEIMALRQGGLVMQNVRLIDATLVYVTFEPCVMCAG
 AMIHSRIGRVVFGVNRNSKRGAAAGSLMNVLNYPOMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKK
 AQSSINGSSGSETPGTSESATPESSGGSERPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRY
 GKDNABEASTLMEKAYHIAISRALEKEGLKDEKSPILNLSSELQDEIGTAFSLFKTDDEDITGRKLDREVQPEILE
 ALLKHISFDICFVQISLKALRRIVPLMFQGGKRYDEACAETYGDIYGGKNTTEKIYLPPIPADEIRNPVVLRLALS
 QARKVINGVVRRYGSPARIHETAREVOKSFKDRKEIEKROEENRKDRKAAAKFREYFPNFVGEFPSKDIL
 KLRLYEQQHGGKCLYSOKEINLVRLNEKGYVEIDHALFFSRTWDDSFNNKVLVLOSENQNKONQTPYEYFN
 GKDNSREWQEFKARVETSRFFRSKKQRILLQKFDDEGGFKECNLNDTRYVNRFLCQFVADHILLTOKGKRR
 VFASNGQITNLLRGFWGLRKVRAENDRIIHALDAVVVACSTVAMQKKITRFVRYKEMNAFDGKTIDKETG
 KVLHQKTHFPQPEFFAQEVMIRVFGIGDQKPEFEEADTPEKLRITLLAEKLSRPEAVHEYVTPLFVSRAP
 NRKMSGAIHKDTLRSKRFFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNA
 KQAFDPKDNPFYKKGGLVKAVRVEKTOESGVLLNKKKNAYTIADNATMVRVDVYTKAGKNYLVPVYV
 WQVAQGIPLPNRAVTSOKSEADWDLIDESFEFKPSLSRGDLVEMISNKORIFGYNGLDRANGSIOIPEHDL
 KSKGKDGVRHVGVKTATAFNKYHVDPLGKEIRCSSEPRPTLKIKSKKEDKRTADGSEFESPKKKRKV*

FIG. 64B

89/95

FIG. 65

Fig. 65A

KRTADGSEFESPKKKRKVEDMAAFKPNPINYILGLAIGIASVQWAMVEIDEEENPIRLIDLCVRYPERAEVP
 KTGDSLAMARBLARSVRRLTERRAHRLRLRAPRLKKREGVLQAADFDENGLIKSLPNTPWOLRAAALDRKL
 TPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEEKESGHI
 RNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIEILLMTQRPALSODAVQKMLGHCTFEPA
 EPKAAKNTYTAEFFIWUTKUNNLRILEQQSERPLTDTERATLMDEPYRKSKLTYYAQARKLLQLEDTAFFKG
 LRYGKDNAAEASTLMEMKAYHAISRALEKEGLKDKK5PLNLSSELQDEIGTAFSLFKTDEDITGRLLKDRVQP
 EILEALLKHISFDKFPVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVL
 ALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQENRDKREKAAAKFREYFPNFVGEPKS
 KDILKLRLYEQQHGGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSPNNKVLVLGSENQNKGNQTPY
 EYFNGKONSREWGEFKARVETSRFPRSKKORILLQKFDEDDGFKECNLDTRYVNRFLCOFVADHILLTGKG
 KRRVFASNGQITNLLRGFWGLRKVRAENDRIHALDAVVVACSTVAMQOKITRFVRYKEMNAFDGKTIDK
 EGGSGSGSGSGSGSGSGSGSGSEVEFSHEYWMRHATLAKRARDEREVPGAVLVLNRRVIGEGWNRAI
 GLHDPTAHAEIMALRQOGLVMQNYELLIDATLYVTTEPCVMCAGAMHISRIGRVVFQVENSIRGAAOSLM
 NVLNYPGMNHRVETTEGILADECAALLCDFYRMPROVFENAQKKAQSSIN**GSSGSETPGTSESATPESSTGK**
 VLHQKTHFPPOPWEFFAQEVMIRVFGKPDCKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNR
 KMSGAHKDTLRSARKFPVKHNEKISVKRVWLTEIKLADLENMNVYKNGREIELYEALKARLEAYOGNAKQ
 AFDPKDNFFYKKGGQQLVKAVRVEKTQESGVLLNKKNAYTIADNATMVRVDVYTKAGKNYLVPVYVWQ
 VAQCILPNRAVTSKSEADWDLIDESFEFKFSLSPGDLVEMISNKGRIFGYYNGIDRANOSIGREHDLKES
 KKGKGVHRVGVKTATAFENKYHVDPLCKEIHRCSSSEPRPTLKIKSKK**EDKRTADGSEFESPKKKRKV***

Fig. 65B

91/95

Fig. 66

Fig. 66A

KRTADGSEFESPKKKRKVEDMAAFKPNPINYILGLAIGIASVQWAMVEIDEEENPIRLIDQVVRVPERAEVP
 KTGDSLAMARRLARSVRRLTYRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWOLRAAALDRKL
 TPLEWSAVLLHLIKHRGYLSQRKNEGETADELGALEGVANNAHALQTGDFRTPAELALNKFEEKESGHI
 RNQRQDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGOLKEQIETLLMTQRPALSODAVQKMLGHCTFEPA
 EPKAAKNTYTAERFIWLTCLNNLRILEOGSERPLTDTERATLMDEPYRKSELTYYAQARKLLGLEDTAFFKG
 LRYGKDNAAEASTLMEMKAYHAISRALEKEGLKDKKSPNLSSSELQDEIGTAFSLFKTDEDITGRLLKDRVQP
 EILEALLKHISFDKFPVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVL
 ALSQARKVINGVVRRYGSPARIHETAREVQKSFKDRKEIEKRQEEENRKDREKAAAKFREYFPNFVQEPKS
 KDILKLRLYEQQHGGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPY
 EYPNGKDNSREWOEFKARVETSRFPRSKKORILLQKFDEDDGFKECNLNDTRYVNRFLCOPVADIHLLTGKQ
 KRRVFASNGQITNLLRGFWOLRKVRAENDRHHALDAVVVACSTVAMQOKITRFPVRKEMNAPDGKTIDK
 ETOKVLHQKTHFPQPWEFFAQEV MIRVFQKPDOKP**GSGSGSGSGSGSGSGSGSGSG**SEVEFSHEYWMRHA
 LTLAKRARDEREVPVGAVLVLNKVGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTFE
 PCVMCAGAMIHSRIGRVVFGVRNSKRGAAQSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQ
 VPNAQKKAQSSIN**GSSGSETPGTSESATP**ESSGEFEEADTPEKLRITLLAEKLSSRPEAVHEYVTTLPVSRAPN
 RKMSGAHKDTLRSKRFVKHNEKISVKRVWLTEKLADLENMVNYKNGREIELYEALKARLEAYGGNAK
 QAFDPKDNPFYKKGGLVKAVRVEKTQESGVLLNKKNAYTIADNATMVRVDVYTKAQKNYLVPVYVW
 QVAQGILPNRAVTSCKSEADWDLIDESFEFKFSLSRGDLVEMISNKGRIFOYYNQLDRAANGSIGIREHDLK
 SKGRDGVHRVGVKTATAPNKYHVDPLQKEIFRCSSEPRPITLKIKSK**EDKRTADGSEFESPKKKRKV***

Fig. 66B

93/95

FIG. 67

Fig. 67A

KRTADGSEFESPKKKRKVGGSGGGSGGGSGSEVEFSHEYWMRIHALTLAKRARDEREVPVGAVLVLNIRV
 IGEQWNRRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTPEPCVMCAGAMHSRIQRVVFGVRNSK
 RGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN**SGSSSGSSGSETP**
GTSESATPESSGGSSGGSSMAAFKPNFNYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPETG
 DSLAMARRLARSVRLTRRRRAHRLLRARFLLKREGVLQAADFDENOLIKSLPNTFWQLRAAALDRKUTPL
 EWSAVLLHLIKHROYLSPKNEGETADKELGALLKQVANNALHALQTGDFRTPAELALNKPEKESGHIRNQ
 RGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGOLKEGIEITLLMTQRPALSGDAVQKMLGHCTFEPAPFK
 AAKNTYTAERFIWLTCLNNLRILEQCSERPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRY
 GKDNABASTLMEMKAYHAISSRALEKEGLKDEKSPILNLSSSELQDEIGTAFSLFKTDEDITGRLLKDRVQPEILE
 ALLKHISFDKFVQISLKALRRIVPLMFQCKRYDEACAETYGDIHYGCKNTEEKIYLPPIPADEIRNPVVLRLALS
 QARKVINGVYRRYGGSPARIHETAREVCKSPKDRKEIEKROEENRKDRKAAAKFREYFPNFVCEPKSKDIL
 KLRLYEQQHGGKCLYSOKEINLVRLNEKGYVEIDHALFFSRTWDDSFNNKVLVLOSENQNKONQTPYEYFN
 GKDNSREWQEFKARVETSRFFRSKKQRILLQKFDDEGDFKECNLNDTRYVNRFLCQFVADHILLTOKGKRR
 VFASNGQITNLLRGFWGLRKVRAENDRIIHALDAVVVACSTVAMQCKITRFVRYKEMNAFDGKTIDKETG
 KVLHQKTHFPQPWEFFAOEVMIRVFGIGDOKPEFEEADTPEKLRITLLAEKLSRPEAVHEYVTPLFVSRAP
 NRKMSGAIHKDTLRSKRFPVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNA
 KQAFDPKDNPFYKKGQQLVKAVRVEKTOESGVLLNKKNAYTIADNATMVRVDVYTKAGKNYLVPVYV
 WQVAQGIPLPNRAVTSOKSEADWDLIDESFEFKFSLSRGDLVEMISNKRIFGYYNGLDRANGSIOIPEHDL
 KSKGKDCVHRVGVKTATAFNKYHVDPLGKEIRCSSEPRPTLKIKSKK**EDKRTADGSEFESPKKKRKV***

Fig. 67B

95/95

