depending on the chemical-triggered excision of a function-disrupting intein has been shown to modulate the activity of, for example, Cas9. However, the reported intein-Cas9 is leaky and exhibits some genome editing activity even in the absence of the chemical.

[0005] In view of the problems encountered in the art, there is a need for providing a genome engineering/editing system which is capable of being chemically controllable and/or inducible.

SUMMARY

[0006] In one aspect, the present disclosure refers to an endonuclease-based gene editing construct comprising the following formula (II):

$$B - L^1 - C_1 - L^2 - C_2 - L^3 - \mathbf{X} - L^4 - D - L^5 - E - L^7 - F - L^8 - G \quad \text{(II)},$$

wherein B is a localization sequence or derivatives thereof, or the binding tag; wherein both C1 and C2 are present or only C1 is present; wherein C1 and C2 are each independently selected from the group consisting of the localization sequence, derivatives thereof of the localization sequence, and a mutated hormone binding domain of the estrogen receptor (ERT2); wherein when C1 is one mutated hormone binding domain of the estrogen receptor (ERT2), C2 is another mutated hormone binding domain of the estrogen receptor (ERT2); wherein X is a CRISPR-associated endonuclease or a derivative thereof; wherein D is selected from the group consisting of a mutated hormone binding domain of the estrogen receptor (ERT2), the localization sequence and derivatives of the localization sequence; wherein E is absent or is selected from the group consisting of a mutated hormone binding domain of the estrogen receptor (ERT2) and a self-cleaving peptide; wherein F is absent or is selected from the group consisting of the self-cleaving peptide, the mutated hormone binding domain of the estrogen receptor (ERT2) and a selectable marker; wherein G is absent or is the selectable marker; wherein L1, L2, L3, L4, L5 , L7and L8 are linker sequences; wherein at least one of the linker sequences is present; wherein each of the linkers sequences is independently between 1 to 25 amino acids long; wherein each linker sequence independently comprises natural or unnatural or a mixture of natural and unnatural amino acids; wherein the linker sequences comprise the amino acids A, E, G, P, S and T; wherein, if any one or more of the linker sequences of L1 to L5 and L7 to L8 is absent, the neighbouring substituents are bound by a peptide bond; wherein L1 is selected from the group consisting of PR, TG, TGPGPGGS, TGPGPGGSAGDTTGPGTGPG, TGPGGS, TGPGGSAGDTTGPGGS and TGGGS; wherein L2 is selected from the group consisting of PRGGS,

GGSPRGGS, PR, GGSPRGGS and TPGGPRGGS; wherein L3 is selected from the group consisting of PG, SGSEGA, GASGSKTPG, SGSETPGTSESAGA, SGSETPGTGPGGA, SESATPESGA, GTSESATPESGGA, GGSGGSGA, GA, GGGS, TPESGA, SGSETPGTGA, SGSETPGTSEGA, PAG, PAGGGS, SGSETPGTPGGA, TPESGPGGA and GASGS; wherein L4 is GGGS; wherein L5 and L7 are independently PAG, SGS or PAGGGS; wherein L8 is selected from the linkers as disclosed in any of L1 to L5 and L7, wherein the construct has at least 90% sequence identity to SEQ ID NOs: 15 to 74.

[0007] In another aspect, the present disclosure refers to a nucleic acid sequence encoding the construct as disclosed herein.

[0008] In one example, there is disclosed a vector comprising the nucleic acid sequence as disclosed herein.

[0009] In another example, there is disclosed a host cell comprising the vector as disclosed herein.

[0010] In one aspect, the present disclosure refers to a kit comprising the construct, as disclosed herein, and tamoxifen, and/or a derivative thereof

[0011] In another aspect, the present disclosure refers to a method of editing a genome of a host cell using the construct of any one of the preceding claims, the method comprising transfecting, nucleofecting, or electroporating the host cell with the nucleic acid sequence as disclosed herein; and incubating the cell of step (a) with an inducing agent.

[0012] In yet another aspect, the present disclosure refers to a method of editing a genome of a host cell using the construct of any one of the preceding claims, wherein the host cell comprising the nucleic acid sequence as disclosed herein is incubated with an inducing agent.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

[0014] **Fig. 1 |** Building and testing a 4-HT-inducible Cas9. (a) Schematic of the ERT2-based strategy. A fusion of Cas9 (red) and ERT2 (crescent shape) was predicted to be sequestered in the cytoplasm. However, in the presence of an antagonist such as 4-HT (circle), the enzyme could enter the nucleus, form a complex with a sgRNA (brown and green), and generate a double-stranded break (triangles) in the cell's DNA (black). (b) Architectures of different Cas9–ERT2 fusions tested. Top,

FANCF_Off2-3_Adapter_REV	GTG CTC TTC CGA TCT ACC AGT TTG AGA CCT CTG ACC
FANCF_Off4_Adapter_FOR	GCG TTA TCG AGG TCG GCT CTG GGT ACA GTT CTG C
FANCF_Off4_Adapter_REV	GTG CTC TTC CGA TCT GCC ACA GAC GAA GAC ACA GA

[00124] **Table 1a** | List of #Cas9 variants constructed and tested. Amino acids for the different protein linkers are given in bold letters.

SEQ ID No.	No.	Details
15	17	NLS-**PR**-ERT2-**PG**-Cas9-ERT2
16	2	NLS-**TG**-ERT2-**SGSEGA**-Cas9-ERT2
17	9	NLS-**TG**-ERT2-**GASGSKTPG**-Cas9-ERT2
18	1	NLS-**TG**-ERT2-**SGSETPGTSESAGA**-Cas9-ERT2
19	5	NLS-**TG**-ERT2-**SGSETPGTGPGGA**-Cas9-ERT2
20	6	NLS-**TG**-ERT2-**SESATPESGA**-Cas9-ERT2
21	4	NLS-**TG**-ERT2-**GTSESATPESGGA**-Cas9-ERT2
22	3	NLS-**TG**-ERT2-**GGSGGSGA**-Cas9-ERT2
23	11	NLS-**TGPGPGGS**-ERT2-**GA**-Cas9-ERT2
24	14	NLS-**TGPGPGGSAGDTTGPGTGPG**-ERT2-**GA**-Cas9-ERT2
25	19	NLS-**TG**-ERT2-**GA**-Cas9-**GGGS**-ERT2
26	13	NLS-**TGGGS**-ERT2-**TPESGA**-Cas9-ERT2
27	15	NLS-**TGGGS**-ERT2-**SESATPESGA**-Cas9-ERT2
28	16	NLS-**TGGGS**-ERT2-**SGSEGA**-Cas9-ERT2
29	7	NLS-**TGGGS**-ERT2-**SGSETPGTGA**-Cas9-ERT2
30	12	NLS-**TGGGS**-ERT2-**SGSETPGTSEGA**-Cas9-ERT2
31	22	NLS-**TGGGS**-ERT2-**PRGGS**-ERT2-**TPESGA**-Cas9-ERT2
32	21	NLS-**TGGGS**-ERT2-**GGSPRGGS**-ERT2-**TPESGA**-Cas9-ERT2
33	23	NLS-**TG**-ERT2-**GA**-Cas9-ERT2-**PAG**-ERT2
34	24	NLS-**TG**-ERT2-**GA**-Cas9-ERT2-**PAGGGS**-ERT2
35	8	NLS-**TGGGS**-ERT2-**SGSETPGTPGGA**-Cas9-ERT2
36	27	NLS-**TGGGS**-ERT2-**PRGGS**-ERT2-**TPESGA**-Cas9-ERT2-**PAG**-ERT2
37	30	NLS-**TGGGS**-ERT2-**PR**-ERT2-**TPESGA**-Cas9-ERT2-**PAGGGS**-ERT2
38	25	NLS-**TG**-ERT2-**GGSPRGGS**-ERT2-**TPESGA**-Cas9-ERT2-**PAGGGS**-ERT2
39	28	NLS-**TGGGS**-ERT2-**TPGGPRGGS**-ERT2-**TPESGA**-Cas9-ERT2-**PAG**-ERT2

Claims

1. An endonuclease-based gene editing construct comprising the following formula (II):

$$B - L^1 - C_1 - L^2 - C_2 - L^3 - X - L^4 - D - L^5 - E - L^7 - F - L^8 - G$$

(II);

wherein B is a localization sequence or derivatives thereof, or the binding tag;

wherein both C_1 and C_2 are present or only C_1 is present; wherein C_1 and C_2 are each independently selected from the group consisting of the localization sequence, derivatives thereof of the localization sequence, and a mutated hormone binding domain of the estrogen receptor (ERT2); wherein when C_1 is one mutated hormone binding domain of the estrogen receptor (ERT2), C_2 is another mutated hormone binding domain of the estrogen receptor (ERT2);

wherein X is a CRISPR-associated endonuclease or a derivative thereof;

wherein D is selected from the group consisting of a mutated hormone binding domain of the estrogen receptor (ERT2), the localization sequence and derivatives of the localization sequence;

wherein E is absent or is selected from the group consisting of a mutated hormone binding domain of the estrogen receptor (ERT2) and a self-cleaving peptide;

wherein F is absent or is selected from the group consisting of the self-cleaving peptide, the mutated hormone binding domain of the estrogen receptor (ERT2) and a selectable marker;

wherein G is absent or is the selectable marker;

wherein L^1, L^2, L^3, L^4, L^5 , $L^7$and L^8 are linker sequences; wherein at least one of the linker sequences is present; wherein each of the linkers sequences is independently between 1 to 25 amino acids long; wherein each linker sequence independently comprises natural or unnatural or a mixture of natural and unnatural amino acids; wherein the linker sequences comprise the amino acids A, E, G, P, S and T;

wherein, if any one or more of the linker sequences of L^1 to L^5 and L^7 to L^8 is absent, the neighbouring substituents are bound by a peptide bond;

wherein L^1 is selected from the group consisting of PR, TG, TGPGPGGS, TGPGPGGSAGDTTGPGTGPG, TGPGGS, TGPGGSAGDTTGPGGS and TGGGS;

wherein L^2 is selected from the group consisting of PRGGS, GGSPRGGS, PR, GGSPRGGS and TPGGPRGGS;

wherein L^3 is selected from the group consisting of PG, SGSEGA, GASGSKTPG, SGSETPGTSESAGA, SGSETPGTGPGGA, SESATPESGA, GTSESATPESGGA, GGSGGSGA, GA, GGGS, TPESGA, SGSETPGTGA, SGSETPGTSEGA, PAG, PAGGGS, SGSETPGTPGGA, TPESGPGGA and GASGS;

wherein L^4 is GGGS;

wherein L^5 and L^7 are independently PAG, SGS or PAGGGS;

wherein L^8 is selected from the linkers as disclosed in any of L^1 to L^5 and L^7,

wherein the construct has at least 90% sequence identity to SEQ ID NOs: 15 to 74.

2. The construct of claim 1, wherein

a) the linker sequences comprise of the amino acids A, E, G, P, S and T;

b) the linker sequences consist of the amino acids A, E, G, P, S and T;

c) B is the localization sequence, C_1 is the mutated hormone binding domain of the estrogen receptor (ERT2) and C_2 is absent;

d) D is a localization sequence and E and F are each a mutated hormone binding domain of the estrogen receptor (ERT2);

e) A is absent, B is localization sequence, C_1 is the mutated hormone binding domain of the estrogen receptor (ERT2), C_2 is absent, X is a CRISPR-associated endonuclease or a derivative thereof, D is localization sequence, E is a mutated hormone binding domain of the estrogen receptor (ERT2) and F is absent;

f) B is a localization sequence, C_1 is the mutated hormone binding domain of the estrogen receptor (ERT2), C_2 is absent, X is a CRISPR-associated endonuclease or a derivative thereof, D is localization sequence and E and F are both each a mutated hormone binding domain of the estrogen receptor (ERT2); or

g) B is a localization sequence, C_1 and C_2 are each independently a mutated hormone binding domain of the estrogen receptor (ERT2), X is a CRISPR-associated endonuclease or a derivative thereof, D is localization sequence and E and F are both each a mutated hormone binding domain of the estrogen receptor (ERT2).

3. The construct of any of the preceding claims, wherein the CRISPR-associated endonuclease, or derivative thereof, is selected from the group consisting of a wild type CRISPR-associated protein 9 (Cas9), a mutated CRISPR-associated protein 9 (Cas9), wherein the mutated CRISPR-associated protein 9 (Cas9) is functional; a wild type Cpf1 (CRISPR from Prevotella and Francisella 1) protein, and a mutated Cpf1 protein, wherein the mutated Cpf1 protein is functional.

4. The construct of claim 3, wherein

a) the CRISPR-associated protein 9 (Cas9), or derivative thereof, is selected from the group consisting of *Streptococcus pyogenes, Streptococcus thermophiles, Listeria innocua, Staphylococcus aureus* and *Neisseria meningitidis;*

b) the CRISPR-associated protein 9 (Cas9), or derivative thereof, has at least 95% sequence identity to SEQ ID NO: 1;

c) the Cpf1 protein, or derivative thereof, is selected from the group consisting of *Acidaminococcus, Lachnospiraceae, Parcubacteria, Butyrivibrio proteoclasticus, Peregrinibacteria, Porphyromonas crevioricanis, Prevotella disiens, Moraxella bovoculi, Smithella, Leptospira inadai, Francisella novicida, Candidatus Methanoplasma termitum* and *Eubacterium eligens;* or

d) the Cpf1 protein, or derivative thereof, has at least 95% sequence identity to SEQ ID NO: 2 or 3.

5. The construct of any of the preceding claims, wherein

a) the localization sequence is selected from the group consisting of nuclear localization sequence, mitochondrial localization sequence and derivatives thereof, optionally wherein the at least one or more nuclear localization sequences (NLS) are selected from the group consisting of Simian Vacuolating Virus 40 (SV40) Large T-antigen, Nucleoplasmin, Importin α, EGL-13, c-MYC, TUS, AR, PLSCR1, PEP, TPX2, RB, TP53, N1N2, PB2, CBP80, SRY, hnRNP A1, HRP1, Borna Disease Virus p10, Ty1 Integrase, and the Chelsky consensus sequence;

b) at least one or more nuclear localization sequences (NLS) are monopartite or bipartite NLS; or

c) at least one or more nuclear localization sequences (NLS) are classical NLS (cNLS) or proline-tyrosine (PY)-NLS.

6. The construct of any one of the preceding claims, wherein the binding tag is selected from the group consisting of a V5 epitope tag, a FLAG tag, a tandem FLAG-tag, a triple FLAG tag (3xFLAG), a Human influenza hemagglutinin (HA) tag, a tandem HA tag, a triple HA tag (3xHA), a sextuple Histidine tag (6xHIS), biotin, c-MYC, a Glutathione-S-transferase (GST) tag, a Strep-tag, a Strep-tag II, a S-tag, a natural histidine affinity tag (HAT), a Calmodulin-binding peptide (CBP) tag, a Streptavidin-binding peptide (SBP) tag, a Chitin-binding domain, a Maltose-binding protein (MBP) and derivatives thereof; optionally wherein the

binding tag is located at either the N-terminus or the C-terminus of the construct, or at both ends of the construct.

7. The construct of claim 6, wherein the V5 epitope tag sequence is SEQ ID NO: 12 or a derivative thereof.

8. The construct of any one of claims 1 to 7, wherein the self-cleaving peptide is a 2A self-cleaving peptide or a derivative thereof; optionally, wherein the 2A self-cleaving peptide is SEQ ID NO: 13 or a derivative thereof.

9. The construct of any one of the preceding claims, wherein the construct has at least 90% sequence identity to SEQ ID NOs:36 to 40, 43 and 244 to 249.

10. The construct of any one of the preceding claims, wherein the construct has a sequence selected from the group consisting of SEQ ID NO. 36 to 40, 43 and 244 to 249.

11. The construct of any one of the preceding claims, wherein the construct has a sequence selected from the group consisting of SEQ ID NO: 37, SEQ ID NO: 74 and SEQ ID NO: 249.

12. A nucleic acid sequence encoding the construct of any one of the preceding claims.

13. A kit comprising the construct according to any of claims 1 to 9 and tamoxifen, and/or a derivative thereof.

14. A method of editing a genome of a host cell using the construct of any one of claims 1 to 9, the method comprising
(a) transfecting the host cell with the nucleic acid sequence as defined in claim 10; and
(b) incubating the host cell of step a. with an inducing agent.

15. A method of editing a genome of a host cell using the construct of any one of claims 1 to 9, wherein the host cell comprising the nucleic acid sequence as defined in claim 10 is incubated with an inducing agent.