(12) STANDARD PATENT (11) Application No. **AU 2017378482 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Enhanced hAT family transposon-mediated gene transfer and associated compositions, systems, and methods

(51) International Patent Classification(s)
C12N 9/12 (2006.01) ***C07K 14/705*** (2006.01)
A61K 35/545 (2015.01) ***C12N 5/074*** (2010.01)

(21) Application No: **2017378482** (22) Date of Filing: **2017.12.15**

(87) WIPO No: **WO18/112415**

(30) Priority Data

(31) Number	(32) Date	(33) Country
62/435,522	**2016.12.16**	**US**

(43) Publication Date: **2018.06.21**
(44) Accepted Journal Date: **2024.06.13**

(71) Applicant(s)
B-Mogen Biotechnologies, Inc.

(72) Inventor(s)
Largaespada, David;Moriarity, Branden;Webber, Beau;Otto, Neil;Kumar, Sandeep;Hogdal, Leah

(74) Agent / Attorney
Pizzeys Patent and Trade Mark Attorneys Pty Ltd, GPO Box 1374, Brisbane, QLD, 4001, AU

(56) Related Art
LI, XIANGHONG ET AL.: "A resurrected mammalian hAT transposable element and a closely related insect element are highly active in human cell culture", PNAS, vol. 110, 2013, pages E478 - E487

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 June 2018 (21.06.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/112415 A1

(51) International Patent Classification:
C12N 9/12 (2006.01) *C07K 14/705* (2006.01)
C12N 5/074 (2010.01) *A61K 35/545* (2015.01)

(21) International Application Number:
PCT/US2017/066829

(22) International Filing Date:
15 December 2017 (15.12.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/435,522 16 December 2016 (16.12.2016) US

(71) Applicant: **B-MOGEN BIOTECHNOLOGIES, INC.** [US/US]; 614 McKinley Place, NE, Minneapolis, Minnesota 55413 (US).

(72) Inventors: **LARGAESPADA, David**; 8285 Knollwood Drive, Mounds View, Minnesota 55112 (US). **MORIARITY, Branden**; 1520 Sherwood Road, Shoreview, Minnesota 55126 (US). **WEBBER, Beau**; 2607 130th Circle NW, Coon Rapids, Minnesota 55448 (US). **OTTO, Neil**; 20344 Babcock Avenue, Lester Prairie, Minnesota 55354 (US). **KUMAR, Sandeep**; 1122 Como Avenue SE, Apt. 105, Minneapolis, Minnesota 55414 (US). **HOGDAL, Leah**; 500 N 2nd Street, Apt. 112, Minneapolis, Minnesota 55401 (US).

(74) Agent: **HUANG, Wei** et al.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, California 94304-1050 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: ENHANCED hAT FAMILY TRANSPOSON-MEDIATED GENE TRANSFER AND ASSOCIATED COMPOSITIONS, SYSTEMS, AND METHODS

(57) Abstract: This disclosure provides various TcBuster transposases and transposons, systems, and methods of use.

WO 2018/112415 A1

ENHANCED hAT FAMILY TRANSPOSON-MEDIATED GENE TRANSFER AND ASSOCIATED COMPOSITIONS, SYSTEMS, AND METHODS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/435,522, filed December 16, 2016, which application is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

[0002] Transposable genetic elements, also called transposons, are segments of DNA that can be mobilized from one genomic location to another within a single cell. Transposons can be divided into two major groups according to their mechanism of transposition: transposition can occur (1) via reverse transcription of an RNA intermediate for elements termed retrotransposons, and (2) via direct transposition of DNA flanked by terminal inverted repeats (TIRs) for DNA transposons. Active transposons encode one or more proteins that are required for transposition. The natural active DNA transposons harbor a transposase enzyme gene.

[0003] DNA transposons in the hAT family are widespread in plants and animals. A number of active hAT transposon systems have been identified and found to be functional, including but not limited to, the Hermes transposon, Ac transposon, hobo transposon, and the Tol2 transposon. The hAT family is composed of two families that have been classified as the AC subfamily and the Buster subfamily, based on the primary sequence of their transposases. Members of the hAT family belong to Class II transposable elements. Class II mobile elements use a cut and paste mechanism of transposition. hAT elements share similar transposases, short terminal inverted repeats, and an eight base-pairs duplication of genomic target.

SUMMARY OF THE INVENTION

[0004] One aspect of the present disclosure provides a mutant TcBuster transposase, comprising an amino acid sequence at least 70% identical to full-length SEQ ID NO: 1 and having one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase has increased transposition efficiency in comparison to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1.

[0005] Another aspect of the present disclosure provides a mutant TcBuster transposase, comprising an amino acid sequence at least 70% identical to full-length SEQ ID NO: 1 and having one more amino acid substitutions in a DNA Binding and Oligomerization domain; an insertion domain; a Zn-BED domain; or a combination thereof. In some embodiments, the

mutant TcBuster transposase has increased transposition efficiency in comparison to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1.

[0006] Yet another aspect of the present disclosure provides a mutant TcBuster transposase comprising an amino acid sequence at least 70% identical to full-length SEQ ID NO: 1 and having one or more amino acid substitutions from **Table 1.**

[0007] In some embodiments, a mutant TcBuster transposase comprises one or more amino acid substitutions that increase a net charge at a neutral pH within or in proximity to a catalytic domain in comparison to SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1, and the one or more amino acids are located in proximity to D223, D289, or E589, when numbered in accordance to SEQ ID NO: 1. In some embodiments, the proximity is a distance of about 80, 75, 70, 60, 50, 40, 30, 20, 10, or 5 amino acids. In some embodiments, the proximity is a distance of about 70 to 80 amino acids.

[0008] In some embodiments, the amino acid sequence of the mutant TcBuster transposase is at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identical to full-length SEQ ID NO: 1.

[0009] In some embodiments, the one or more amino acid substitutions comprise a substitution to a lysine or an arginine. In some embodiments, the one or more amino acid substitutions comprise a substitution of an aspartic acid or a glutamic acid to a neutral amino acid, a lysine or an arginine. In some embodiments, the mutant TcBuster transposase comprises one or more amino acid substitutions from **Table 4**. In some embodiments, the mutant TcBuster transposase comprises one or more amino acid substitutions from **Table 2**. In some embodiments, the mutant TcBuster transposase comprises one or more amino acid substitutions from **Table 3**. In some embodiments, the mutant TcBuster transposase comprises amino acid substitutions V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitutions K573E and E578L, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitution I452K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitution A358K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitution V297K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitution N85S, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitutions I452F,

V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitutions A358K, V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the mutant TcBuster transposase comprises amino acid substitutions V377T, E469K, D189A, K573E and E578L, when numbered in accordance with SEQ ID NO: 1.

[0010] In some embodiments, the transposition efficiency is measured by an assay that comprises introducing the mutant TcBuster transposase and a TcBuster transposon containing a reporter cargo cassette into a population of cells, and detecting transposition of the reporter cargo cassette in genome of the population of cells.

[0011] Yet another aspect of the present disclosure provides a fusion transposase comprising a TcBuster transposase sequence and a DNA sequence specific binding domain. In some embodiments, the TcBuster transposase sequence has at least 70% identity to full-length SEQ ID NO: 1.

[0012] In some embodiments, the DNA sequence specific binding domain comprises a TALE domain, zinc finger domain, AAV Rep DNA-binding domain, or any combination thereof. In some embodiments, the DNA sequence specific binding domain comprises a TALE domain.

[0013] In some embodiments, the TcBuster transposase sequence has at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to full-length SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions in a DNA Binding and Oligomerization domain; an insertion domain; a Zn-BED domain; or a combination thereof. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions from **Table 1**. In some embodiments, the TcBuster transposase sequence has increased transposition efficiency in comparison to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions that increase a net charge at a neutral pH within or in proximity to a catalytic domain in comparison to SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1, and the one or more amino acid substitutions are located in proximity to D223, D289, or E589, when numbered in accordance to SEQ ID NO: 1. In some embodiments, the proximity is a distance of about 80, 75, 70, 60, 50, 40, 30, 20, 10, or 5 amino acids. In some embodiments, the proximity is a distance of about 70 to 80 amino acids. In some embodiments, the TcBuster

transposase sequence comprises one or more amino acid substitutions from **Table 2**. In some embodiments, the TcBuster transposase sequence comprises one or more amino acid substitutions from **Table 3**. In some embodiments, the TcBuster transposase sequence comprises amino acid substitutions V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitutions K573E and E578L, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitution I452K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitution A358K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitution V297K, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitution N85S, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitutions I452F, V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitutions A358K, V377T, E469K, and D189A, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence comprises amino acid substitutions V377T, E469K, D189A, K573E and E578L, when numbered in accordance with SEQ ID NO: 1. In some embodiments, the TcBuster transposase sequence has 100% identity to full-length SEQ ID NO: 1.

[0014] In some embodiments of a fusion transposase, the TcBuster transposase sequence and the DNA sequence specific binding domain are separated by a linker. In some embodiments, the linker comprises at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, or at least 50 amino acids. In some embodiments, the linker comprises SEQ ID NO: 9.

[0015] Yet another aspect of the present disclosure provides a polynucleotide that codes for the mutant TcBuster transposase as described herein.

[0016] Yet another aspect of the present disclosure provides a polynucleotide that codes for the fusion transposase as described herein.

[0017] In some embodiments, the polynucleotide comprises DNA that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the polynucleotide comprises messenger RNA (mRNA) that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the mRNA is chemically modified. In some embodiments, the polynucleotide comprises nucleic acid sequence encoding for a transposon recognizable by

the mutant TcBuster transposase or the fusion transposase. In some embodiments, the polynucleotide is present in a DNA vector. In some embodiments, the DNA vector comprises a mini-circle plasmid.

[0018] Yet another aspect of the present disclosure provides a cell producing the mutant TcBuster transposase or fusion transposase as described herein. Yet another aspect of the present disclosure provides a cell containing the polynucleotide as described herein.

[0019] Yet another aspect of the present disclosure provides a method comprising: introducing into a cell the mutant TcBuster transposase as described herein and a transposon recognizable by the mutant TcBuster transposase.

[0020] Yet another aspect of the present disclosure provides a method comprising: introducing into a cell the fusion transposase as described herein and a transposon recognizable by the fusion transposase.

[0021] In some embodiments of a method, the introducing comprises contacting the cell with a polynucleotide encoding the mutant TcBuster transposase or the fusion transposase. In some embodiments, the polynucleotide comprises DNA that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the polynucleotide comprises messenger RNA (mRNA) that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the mRNA is chemically modified.

[0022] In some embodiments of a method, the introducing comprises contacting the cell with a DNA vector that contains the transposon. In some embodiments, the DNA vector comprises a mini-circle plasmid. In some embodiments, the introducing comprises contacting the cell with a plasmid vector that contains both the transposon and the polynucleotide encoding the mutant TcBuster transposase or the fusion transposase. In some embodiments, the introducing comprises contacting the cell with the mutant TcBuster transposase or the fusion transposase as a purified protein.

[0023] In some embodiments of a method, the transposon comprises a cargo cassette positioned between two inverted repeats. In some embodiments, a left inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 3. In some embodiments, a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 3. In some embodiments, a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 4. In some embodiments, a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 4. In some embodiments, a left inverted repeat of the two inverted repeats

comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 5. In some embodiments, a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 5. In some embodiments, a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 6. In some embodiments, a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 6. In some embodiments, the cargo cassette comprises a promoter selected from the group consisting of: CMV, EFS, MND, EF1α, CAGCs, PGK, UBC, U6, H1, and Cumate. In some embodiments, the cargo cassette comprises a CMV promoter. In some embodiments, the cargo cassette is present in a forward direction. In some embodiments, the cargo cassette is present in a reverse direction. In some embodiments, the cargo cassette comprises a transgene. In some embodiments, the transgene codes for a protein selected from the group consisting of: a cellular receptor, an immunological checkpoint protein, a cytokine, and any combination thereof. In some embodiments, the transgene codes for a cellular receptor selected from the group consisting of: a T cell receptor (TCR), a B cell receptor (BCR), a chimeric antigen receptor (CAR), or any combination thereof. In some embodiments, the introducing comprises transfecting the cell with the aid of electroporation, microinjection, calcium phosphate precipitation, cationic polymers, dendrimers, liposome, microprojectile bombardment, fugene, direct sonic loading, cell squeezing, optical transfection, protoplast fusion, impalefection, magnetofection, nucleofection, or any combination thereof. In some embodiments, the introducing comprises electroporating the cell.

[0024] In some embodiments of a method, the cell is a primary cell isolated from a subject. In some embodiments, the subject is a human. In some embodiments, the subject is a patient with a disease. In some embodiments, the subject has been diagnosed with cancer or tumor. In some embodiments, the cell is isolated from blood of the subject. In some embodiments, the cell comprises a primary immune cell. In some embodiments, the cell comprises a primary leukocyte. In some embodiments, the cell comprises a primary T cell. In some embodiments, the primary T cell comprises a gamma delta T cell, a helper T cell, a memory T cell, a natural killer T cell, an effector T cell, or any combination thereof. In some embodiments, the primary immune cell comprises a CD3+ cell. In some embodiments, the cell comprises a stem cell. In some embodiments, the stem cell is selected from the group consisting of: embryonic stem cell, hematopoietic stem cell, epidermal stem cell, epithelial stem cell, bronchoalveolar stem cell, mammary stem cell, mesenchymal stem cell, intestine stem cell, endothelial stem cell, neural

stem cell, olfactory adult stem cell, neural crest stem cell, testicular cell, and any combination thereof. In some embodiments, the stem cell comprises induced pluripotent stem cell.

[0025] Yet another aspect of the present disclosure provides a method of treatment, comprising: (a) introducing into a cell a transposon and the mutant TcBuster transposase or the fusion transposase as described herein, which recognize the transposon, thereby generating a genetically modified cell; (b) administering the genetically modified cell to a patient in need of the treatment. In some embodiments, the genetically modified cell comprises a transgene introduced by the transposon. In some embodiments, the patient has been diagnosed with cancer or tumor. In some embodiments, the administering comprises transfusing the genetically modified cell into blood vessels of the patient.

[0026] Yet another aspect of the present disclosure provides a system for genome editing, comprising: the mutant TcBuster transposase or fusion transposase as described herein, and a transposon recognizable by the mutant TcBuster transposase or the fusion transposase.

[0027] Yet another aspect of the present disclosure provides a system for genome editing, comprising: the polynucleotide encoding a mutant TcBuster transposase or fusion transposase as described herein, and a transposon recognizable by the mutant TcBuster transposase or the fusion transposase.

[0028] In some embodiments of a system, the polynucleotide comprises DNA that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the polynucleotide comprises messenger RNA (mRNA) that encodes the mutant TcBuster transposase or the fusion transposase. In some embodiments, the mRNA is chemically modified. In some embodiments, the transposon is present in a DNA vector. In some embodiments, the DNA vector comprises a mini-circle plasmid. In some embodiments, the polynucleotide and the transposon are present in a same plasmid. In some embodiments, the transposon comprises a cargo cassette positioned between two inverted repeats. In some embodiments, a left inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 3. In some embodiments, a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 3. In some embodiments, a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 4. In some embodiments, a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 4. In some embodiments, a left inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 5. In some

embodiments, a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 5. In some embodiments, a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 6. In some embodiments, a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 6. In some embodiments, the cargo cassette comprises a promoter selected from the group consisting of: CMV, EFS, MND, EF1α, CAGCs, PGK, UBC, U6, H1, and Cumate. In some embodiments, the cargo cassette comprises a CMV promoter. In some embodiments, the cargo cassette comprises a transgene. In some embodiments, the transgene codes for a protein selected from the group consisting of: a cellular receptor, an immunological checkpoint protein, a cytokine, and any combination thereof. In some embodiments, the transgene codes for a cellular receptor selected from the group consisting of: a T cell receptor (TCR), a B cell receptor (BCR), a chimeric antigen receptor (CAR), or any combination thereof. In some embodiments, the cargo cassette is present in a forward direction. In some embodiments, the cargo cassette is present in a reverse direction.

INCORPORATION BY REFERENCE

[0029] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent that a term incorporated by reference conflicts with a term defined herein, this specification controls.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0031] **Fig. 1** shows the transposition efficiency of several exemplary TcBuster transposon vector constructs, as measured by percent of mCherry positive cells in cells that were transfected with wild-type (WT) TcBuster transposase and the exemplary TcBuster transposons.

[0032] **Fig. 2** shows nucleotide sequence comparison of exemplary TcBuster IR/DR sequence 1 and sequence 2.

[0033] **Fig. 3A** shows representative bright-field and fluorescent images of HEK-293T cells 2 weeks after transfection with exemplary TcBuster transposon Tn-8 (containing puro-mCherry

cassette; illustrated in **Fig. 1**) and WT TcBuster transposase or V596A mutant transposase (containing V596A substitution). The transfected cells were plated in 6-well plate with 1 μg/mL puromycin 2 days posttransfection, and were fixed and stained 2 weeks posttransfection with crystal violet for colony quantification. **Fig. 3B** shows representative pictures of the transfected cell colonies in 6-well plate 2 weeks posttransfection. **Fig. 3C** is a graph showing the quantification of colonies per each transfection condition 2 weeks posttransfection.

[0034] **Fig. 4** depicts the amino acid sequence alignment of TcBuster transposase versus a number of transposases in AC subfamily, with only regions of amino acid conservation being shown.

[0035] **Fig. 5** depicts the amino acid sequence alignment of TcBuster transposase versus a number of other transposase members in Buster subfamily. Certain exemplary amino acid substitutions are indicated above the protein sequences, along with the percentage shown on top of the alignment is the percentage of other Buster subfamily members that contain the amino acid that is contemplated being substituted into the TcBuster sequence, and the percentage shown below is the percentage of other Buster subfamily members that contain the canonical TcBuster amino acid at that position.

[0036] **Fig. 6** shows a vector map of an exemplary expression vector pcDNA-DEST40 that was used to test TcBuster transposase mutants.

[0037] **Fig. 7** is a graph quantifying the transposition efficiency of exemplary TcBuster transposase mutants, as measured by percent of mCherry positive cells in HEK-293T cells that were transfected with TcBuster transposon Tn-8 (illustrated in **Fig. 1**) with the exemplary transposase mutants.

[0038] **Fig. 8** depicts one exemplary fusion transposase that contains a DNA sequence specific binding domain and a TcBuster transposase sequence joined by an optional linker.

[0039] **Fig. 9** is a graph quantifying the transposition efficiency of exemplary TcBuster transposases containing different tags as measured by percent of mCherry positive cells in HEK-293T cells that were transfected with TcBuster transposon Tn-8 (illustrated in **Fig. 1**) with the exemplary transposases containing the tags.

[0040] **Fig. 10A** is a graph quantifying the transposition efficiency of exemplary TcBuster transposition systems in human CD3+ T cells as measured by percent of GFP positive cells. **Fig. 10B** is a graph quantifying viability of the transfected T cells 2 and 7 days post-transfection by flow cytometry. Data is relative to pulse control.

[0041] **Fig. 11** shows amino acid sequence of wild-type TcBuster transposase with certain amino acids annotated (SEQ ID NO: 1).

[0042] **Fig. 12** shows amino acid sequence of mutant TcBuster transposase containing amino acid substitutions D189A/V377T/E469K (SEQ ID NO: 78).

[0043] **Fig. 13** shows amino acid sequence of mutant TcBuster transposase containing amino acid substitutions D189A/V377T/E469K/I452K(SEQ ID NO: 79).

[0044] **Fig. 14** shows amino acid sequence of mutant TcBuster transposase containing amino acid substitutions D189A/V377T/E469K/N85S (SEQ ID NO: 80).

[0045] **Fig. 15** shows amino acid sequence of mutant TcBuster transposase containing amino acid substitutions D189A/V377T/E469K/A358K (SEQ ID NO: 81).

[0046] **Fig. 16** shows amino acid sequence of mutant TcBuster transposase containing amino acid substitutions D189A/V377T/E469K/K573E/E578L (SEQ ID NO: 13).

DETAILED DESCRIPTION OF THE INVENTION

[0047] **Overview**

[0048] DNA transposons can translocate via a non-replicative, 'cut-and-paste' mechanism. This requires recognition of the two terminal inverted repeats by a catalytic enzyme, i.e. transposase, which can cleave its target and consequently release the DNA transposon from its donor template. Upon excision, the DNA transposons may subsequently integrate into the acceptor DNA that is cleaved by the same transposase. In some of their natural configurations, DNA transposons are flanked by two inverted repeats and may contain a gene encoding a transposase that catalyzes transposition.

[0049] For genome editing applications with DNA transposons, it is desirable to design a transposon to develop a binary system based on two distinct plasmids whereby the transposase is physically separated from the transposon DNA containing the gene of interest flanked by the inverted repeats. Co-delivery of the transposon and transposase plasmids into the target cells enables transposition via a conventional cut-and-paste mechanism.

[0050] TcBuster is a member of the hAT family of DNA transposons. Other members of the family include Sleeping Beauty and PiggBac. Discussed herein are various devices, systems and methods relating to synergistic approaches to enhance gene transfer into human hematopoietic and immune system cells using hAT family transposon components. The present disclosure relates to improved hAT transposases, transposon vector sequences, transposase delivery methods, and transposon delivery methods. In one implementation, the present study identified specific, universal sites for making hyperactive hAT transposases. In another implementation, methods for making minimally sized hAT transposon vector inverted terminal repeats (ITRs) that conserve genomic space are described. In another implementation, improved methods to

deliver hAT family transposases as chemically modified in vitro transcribed mRNAs are described. In another implementation, methods to deliver hAT family transposon vectors as "miniature" circles of DNA are described, in which virtually all prokaryotic sequences have been removed by a recombination method. In another implementation, methods to fuse DNA sequence specific binding domains using transcription activator-like (TAL) domains fused to the hAT transposases are described. These improvements, individually or in combination, can yield unexpectedly high levels of gene transfer to the cell types in question and improvements in the delivery of transposon vectors to sequences of interest.

[0051] Mutant TcBuster Transposase

[0052] One aspect of the present disclosure provides a mutant TcBuster transposase. A mutant TcBuster transposase may comprise one or more amino acid substitutions in comparison to a wild-type TcBuster transposase (SEQ ID NO: 1).

[0053] A mutant TcBuster transposase can comprise an amino acid sequence having at least 70% sequence identity to full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1). In some embodiments, a mutant TcBuster transposase can comprise an amino acid sequence having at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1). In some cases, a mutant TcBuster transposase can comprise an amino acid sequence having at least 98%, at least 98.5%, at least 99%, at least 99.1%, at least 99.2%, at least 99.3%, at least 99.4%, at least 99.5%, at least 99.6%, at least 99.7%, at least 99.8%, at least 99.9%, or at least 99.95% sequence identity to full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1).

[0054] A mutant TcBuster transposase can comprise an amino acid sequence having at least one amino acid different from full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1). In some embodiments, a mutant TcBuster transposase can comprise an amino acid sequence having at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or more amino acids different from full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1). In some cases, a mutant TcBuster transposase can comprise an amino acid sequence having at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, or at least 300 amino acid different from full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1). In some cases, a mutant TcBuster transposase can comprise an amino acid sequence having at most 3, at most 6, at most 12, at most 25, at most 35, at most 45, at most 55, at most 65, at most 75, at

most 85, at most 95, at most 150, at most 250, or at most 350 amino acid different from full length sequence of a wild-type TcBuster transposase (SEQ ID NO: 1).

[0055] As shown in **Fig. 4**, typically, a wild-type TcBuster transposase can be regarded as comprising, from N terminus to C terminus, a ZnF-BED domain (amino acids 76-98), a DNA Binding and Oligomerization domain (amino acids 112-213), a first Catalytic domain (amino acids 213-312), an Insertion domain (amino acids 312-543), and a second Catalytic domain (amino acids 583-620), as well as at least four inter-domain regions in between these annotated domains. Unless indicated otherwise, numerical references to amino acids, as used herein, are all in accordance to SEQ ID NO: 1. A mutant TcBuster transposase can comprise one or more amino acid substitutions in any one of these domains, or any combination thereof. In some cases, a mutant TcBuster transposase can comprise one or more amino acid substitutions in ZnF-BED domain, a DNA Binding and Oligomerization domain, a first Catalytic domain, an Insertion domain, or a combination thereof. A mutant TcBuster transposase can comprise one or more amino acid substitutions in at least one of the two catalytic domains.

[0056] An exemplary mutant TcBuster transposase can comprise one or more amino acid substitutions from **Table 1**. Sometimes, a mutant TcBuster transposase can comprise at least one of the amino acid substitutions from **Table 1**. A mutant TcBuster transposase can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, or more of the amino acid substitutions from **Table 1**.

Table 1

Amino Acid of Wild-type TcBuster Transposase (SEQ ID NO: 1)	Amino Acid Substitution
Q82	Q82E
N85	N85S
D99	D99A
D132	D132A
Q151	Q151S
Q151	Q151A
E153	E153K
E153	E153R
A154	A154P
Y155	Y155H
E159	E159A
T171	T171K
T171	T171R
K177	K177E
D183	D183K

D183	D183R
D189	D189A
T191	T191E
S193	S193K
S193	S193R
Y201	Y201A
F202	F202D
F202	F202K
C203	C203I
C203	C203V
Q221	Q221T
M222	M222L
I223	I223Q
E224	E224G
S225	S225W
D227	D227A
R239	R239H
E243	E243A
E247	E247K
P257	P257K
P257	P257R
Q258	Q258T
E263	E263A
E263	E263K
E263	E263R
E274	E274K
E274	E274R
S278	S278K
N281	N281E
L282	L282K
L282	L282R
K292	K292P
V297	V297K
K299	K299S
A303	A303T
H322	H322E
A332	A332S
A358	A358E
A358	A358K
A358	A358S
D376	D376A
V377	V377T
L380	L380N
I398	I398D
I398	I398S
I398	I398K

F400	F400L
V431	V431L
S447	S447E
N450	N450K
N450	N450R
I452	I452F
E469	E469K
K469	K469K
P510	P510D
P510	P510N
E517	E517R
R536	R536S
V553	V553S
P554	P554T
P559	P559D
P559	P559S
P559	P559K
K573	K573E
E578	E578L
K590	K590T
Y595	Y595L
V596	V596A
T598	T598I
K599	K599A
Q615	Q615A
T618	T618K
T618	T618K
T618	T618R
D622	D622K
D622	D622R
E5275	E5275K

[0057] An exemplary mutant TcBuster transposase comprises one or more amino acid substitutions, or combinations of substitutions, from **Table 2**. Sometimes, a mutant TcBuster transposase can comprise at least one of the amino acid substitutions, or combinations of substitutions, from **Table 2**. A mutant TcBuster transposase can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, or more of the amino acid substitutions, or combinations of substitutions, from **Table 2**.

Table 2

Amino Acid of Wild-type TcBuster Transposase (SEQ ID NO: 1)	Amino Acid Substitution
V377 and E469	V377T/E469K
V377, E469, and R536S	V377T/E469K/R536S
A332	A332S
V553 and P554	V553S/P554T
E517	E517R
K299	K299S
Q615 and T618	Q615A/T618K
S278	S278K
A303	A303T
P510	P510D
P510	P510N
N281	N281S
N281	N281E
K590	K590T
E5275	E5275K
Q258	Q258T
E247	E247K
S447	S447E
N85	N85S
V297	V297K
A358	A358K
I452	I452F
V377, E469, D189	V377T/E469K/D189A
K573, E578	K573E/E578L
I452, V377, E469, D189	I452F/V377T/E469K/D189A
A358, V377, E469, D189	A358K/V377T/E469K/D189A
K573, E578, V377, E469, D189	K573E/E578L/V377T/E469K/D189A
T171	T171R
D183	D183R
S193	S193R
P257	P257K
E263	E263R
L282	L282K
T618	T618K
D622	D622R
E153	E153K
N450	N450K
T171	T171K
D183	D183K
S193	S193K
P257	P257R

E263	E263K
L282	L282R
T618	T618R
D622	D622K
E153	E153R
N450	N450R
E247, E274, V297, A358	E247K/E274K/V297K/A358K

[0058] An exemplary mutant TcBuster transposase comprises one or more amino acid substitutions, or combinations of substitutions, from **Table 3**. Sometimes, a mutant TcBuster transposase can comprise at least one of the amino acid substitutions, or combinations of substitutions, from **Table 3**. A mutant TcBuster transposase can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, or more of the amino acid substitutions, or combinations of substitutions, from **Table 3**.

Table 3

Amino Acid of Wild-type TcBuster Transposase (SEQ ID NO: 1)	**Amino Acid Substitutions**
V377 and E469	V377T/E469K
V377, E469, and R536S	V377T/E469K/R536S
A332	A332S
V553 and P554	V553S/P554T
E517	E517R
K299	K299S
Q615 and T618	Q615A/T618K
S278	S278K
A303	A303T
P510	P510D
P510	P510N
N281	N281S
N281	N281E
K590	K590T
E5275	E5275K
Q258	Q258T
E247	E247K
S447	S447E
N85	N85S
V297	V297K
A358	A358K
I452	I452F
V377, E469, D189	V377T/E469K/D189A
K573, E578	K573E/E578L

[0059] Hyperactive mutant TcBuster transposase

[0060] Another aspect of the present disclosure is to provide a hyperactive mutant TcBuster transposase. A "hyperactive" mutant TcBuster transposase, as used herein, can refer to any mutant TcBuster transposase that has increased transposition efficiency as compared to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1.

[0061] In some embodiments, a hyperactive mutant TcBuster transposase may have increased transposition efficiency under certain situations as compared to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1. For example, the hyperactive mutant TcBuster transposase may have better transposition efficiency than the wild-type TcBuster transposase when being used to catalyze transposition of transposons having particular types of inverted repeat sequences. It is possible that with some other transposons having other types of inverted repeat sequences, the hyperactive mutant TcBuster transposase does not have increased transposition efficiency in comparison to the wild-type TcBuster transposase. In some other non-limiting examples, the hyperactive mutant TcBuster transposase may have increased transposition efficiency in comparison to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1, under certain transfection conditions. Without being limited, when compared to a wild-type TcBuster transposase, a hyperactive mutant TcBuster transposase may have better transposition efficiency when the temperature is higher than normal cell culture temperature; a hyperactive mutant TcBuster transposase may have better transposition efficiency in a relative acidic or basic aqueous medium; a hyperactive mutant TcBuster transposase may have better transposition efficiency when a particular type of transfection technique (e.g. electroporation) is performed.

[0062] Transposition efficiency can be measured by the percent of successful transposition events occurring in a population of host cells normalized by the amount of transposon and transposase introduced into the population of host cells. In many instances, when the transposition efficiency of two or more transposases is compared, the same transposon construct is paired with each of the two or more transposases for transfection of the host cells under same or similar transfection conditions. The amount of transposition events in the host cells can be examined by various approaches. For example, the transposon construct may be designed to contain a reporter gene positioned between the inverted repeats, and transfected cells positive for the reporter gene can be counted as the cells where successful transposition events occurs, which can give an estimate of the amount of the transposition events. Another non-limiting example includes sequencing of the host cell genome to examine the insertion of the cassette cargo of the transposon. In some embodiments, when the transposition efficiency of two or more different

transposons is compared, the same transposase can be paired with each of the different transposons for transfection of the host cells under same or similar transfection conditions. Similar approaches can be utilized for the measurement of transposition efficiency. Other methods known to one skilled in the art may also be implemented for the comparison of transposition efficiency.

[0063] Also provided herein are methods of obtaining a hyperactive mutant TcBuster transposase.

[0064] One exemplary method can comprise systemically mutating amino acids of TcBuster transposase to increase a net charge of the amino acid sequence. Sometimes, the method can comprise performing systematic alanine scanning to mutate aspartic acid (D) or glutamic acid (E), which are negatively charged at a neutral pH, to alanine residues. A method can comprise performing systemic mutation to lysing (K) or arginine (R) residues, which are positively charged at a neutral pH.

[0065] Without wishing to be bound by a particular theory, increase in a net charge of the amino acid sequence at a neutral pH may increase the transposition efficiency of the TcBuster transposase. Particularly, when the net charge is increased in proximity to a catalytic domain of the transposase, the transposition efficiency is expected to increase. It can be contemplated that positively charged amino acids can form points of contact with DNA target and allow the catalytic domains to act on the DNA target. It may also be contemplated that loss of these positively charged amino acids can decrease either excision or integration activity in transposases.

[0066] **Fig. 11** depicts the WT TcBuster transposase amino acid sequence, highlighting amino acids that may be points of contact with DNA. In **Fig. 11,** large bold lettering indicates catalytic triad amino acids; lettering with boxes indicates amino acids that when substituted to a positive charged amino acid increases transposition; italicized and lowercased lettering indicates positive charged amino acids that when substituted to a different amino acid decreases transposition; bold italicized and underlined indicates amino acids that when substituted to a positive charged amino acid increases transposition, and when substituted to a negative charged amino acid decreases transposition; underlined lettering indicates amino acids that could be positive charged amino acids based on protein sequence alignment to the Buster subfamily.

[0067] A mutant TcBuster transposase can comprise one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1. Sometimes, a mutant TcBuster transposase comprising one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1 can be hyperactive. Sometimes, the mutant

TcBuster transposase can comprise one or more substitutions to a positively charged amino acid, such as, but not limited to, lysine (K) or arginine (R). A mutant TcBuster transposase can comprise one or more substitutions of a negatively charged amino acid, such as, but not limited to, aspartic acid (D) or glutamic acid (E), with a neutral amino acid, or a positively charged amino acid.

[0068] One non-limiting example includes a mutant TcBuster transposase that comprises one or more amino acid substitutions that increase a net charge at a neutral pH within or in proximity to a catalytic domain in comparison to SEQ ID NO: 1. The catalytic domain can be the first catalytic domain or the second catalytic domain. The catalytic domain can also include both catalytic domains of the transposase.

[0069] An exemplary method of the present disclosure can comprise mutating amino acids that are predicted to be in close proximity to, or to make direct contact with, the DNA. These amino acids can be substituted amino acids identified as being conserved in other member(s) of the hAT family (e.g., other members of the Buster and/or Ac subfamilies). The amino acids predicted to be in close proximity to, or to make direct contact with, the DNA can be identified, for example, by reference to a crystal structure, predicted structures, mutational analysis, functional analysis, alignment with other members of the hAT family, or any other suitable method.

[0070] Without wishing to be bound by a particular theory, TcBuster transposase, like other members of the hAT transposase family, has a DDE motif, which may be the active site that catalyzes the movement of the transposon. It is contemplated that D223, D289, and E589make up the active site, which is a triad of acidic residues. The DDE motif may coordinate divalent metal ions and can be important in the catalytic reaction. In some embodiments, a mutant TcBuster transposase can comprise one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1, and the one or more amino acids are located in proximity to D223, D289, or E589, when numbered in accordance to SEQ ID NO: 1.

[0071] In certain embodiments, a mutant TcBuster transposase as provided herein does not comprise any disruption of the catalytic triad, i.e. D223, D289, or E589. A mutant TcBuster transposase may not comprise any amino acid substitution at D223, D289, or E589. A mutant TcBuster transposase may comprises amino acid substitution at D223, D289, or E589, but such substitution does not disrupt the catalytic activity contributed by the catalytic triad.

[0072] In some cases, the term "proximity" can refer to a measurement of a linear distance in the primary structure of the transposase. For instance, the distance between D223 and D289 in the primary structure of a wild-type TcBuster transposase is 66 amino acids. In certain

embodiments, the proximity can refer to a distance of about 70 to 80 amino acids. In many cases, the proximity can refer to a distance of about 80, 75, 70, 60, 50, 40, 30, 20, 10, or 5 amino acids.

[0073] In some cases, the term "proximity" can refer to a measurement of a spatial relationship in the secondary or tertiary structure of the transposase, i.e. when the transposase folds into its three dimensional configurations. Protein secondary structure can refer to three dimensional form of local segments of proteins. Common secondary structural elements include alpha helices, beta sheets, beta turns and omega loops. Secondary structure elements may form as an intermediate before the protein folds into its three dimensional tertiary structure. Protein tertiary structure can refer to the three dimensional shape of a protein. Protein tertiary structure may exhibit dynamic configurational change under physiological or other conditions. The tertiary structure will have a single polypeptide chain "backbone" with one or more protein secondary structures, the protein domains. Amino acid side chains may interact and bond in a number of ways. The interactions and bonds of side chains within a particular protein determine its tertiary structure. In many implementations, the proximity can refer to a distance of about 1Å, about 2Å, about 5Å, about 8Å, about 10Å, about 15Å, about 20Å, about 25Å, about 30Å, about 35Å, about 40Å, about 50Å, about 60Å, about 70Å, about 80Å, about 90Å, or about 100Å.

[0074] A neutral pH can be a pH value around 7. Sometimes, a neutral pH can be a pH value between 6.9 and 7.1, between 6.8 and 7.2, between 6.7 and 7.3, between 6.6 and 7.4, between 6.5 and 7.5, between 6.4 and 7.6, between 6.3 and 7.7, between 6.2-7.8, between 6.1-7.9, between 6.0-8.0, between 5-8, or in a range derived therefrom.

[0075] Non-limiting exemplary mutant TcBuster transposases that comprise one or more amino acid substitutions that increase a net charge at a neutral pH in comparison to SEQ ID NO: 1 include TcBuster transposases comprising at least one of the combinations of amino acid substitutions from **Table 4**. A mutant TcBuster transposase can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, or more of the amino acid substitutions from **Table 4**.

[0076] In some embodiments, a mutant TcBuster transposase can comprise one or more amino acid substitutions that increase a net charge at a non-neutral pH in comparison to SEQ ID NO: 1. In some cases, the net charge is increased within or in proximity to a catalytic domain at a non-neutral pH. In many cases, the net charge is increased in proximity to D223, D289, or E589, at a non-neutral pH. The non-neutral pH can be a pH value lower than 7, lower than 6.5, lower than 6, lower than 5.5, lower than 5, lower than 4.5, lower than 4, lower than 3.5, lower than 3, lower than 2.5, lower than 2, lower than 1.5, or lower than 1. The non-neutral pH can also be a pH

value higher than 7, higher than 7.5, higher than 8, higher than 8.5, higher than 9, higher than 9.5, or higher than 10.

Table 4

Amino Acid of Wild-type TcBuster Transposase (SEQ ID NO: 1)	Amino Acid Substitutions
E247	E247K
E274	E274K
V297	V297K
A358	A358K
S278	S278K
E247	E247R
E274	E274R
V297	V297R
A358	A358R
S278	S278R
T171	T171R
D183	D183R
S193	S193R
P257	P257K
E263	E263R
L282	L282K
T618	T618K
D622	D622R
E153	E153K
N450	N450K
T171	T171K
D183	D183K
S193	S193K
P257	P257R
E263	E263K
L282	L282R
T618	T618R
D622	D622K
E153	E153R
N450	N450R

[0077] In one exemplary embodiment, a method can comprise systemically mutating amino acids in the DNA Binding and Oligomerization domain. Without wishing to be bound by a particular theory, mutation in the DNA Binding and Oligomerization domain may increase the binding affinity to DNA target and promote oligomerization activity of the transposase, which consequentially may promote transposition efficiency. More specifically, the method can comprise systemically mutating amino acids one by one within or in proximity to the DNA

Binding and Oligomerization domain (e.g., amino acid 112 to 213). The method can also comprise mutating more than one amino acid within or in proximity to the DNA Binding and Oligomerization domain. The method can also comprise mutating one or more amino acids within or in proximity to the DNA Binding and Oligomerization domain, together with one or more amino acids outside the DNA Binding and Oligomerization domain.

[0078] In some embodiments, the method can comprise performing rational replacement of selective amino acid residues based on multiple sequence alignments of TcBuster with other hAT family transposases (Ac, Hermes, Hobo, Tag2, Tam3, Hermes, Restless and Tol2) or with other members of Buster subfamily (e.g., AeBuster1, AeBuster2, AeBuster3, BtBuster1, BtBuster2, CfBuster1, and CfBuster2). Without being bound by a certain theory, conservancy of certain amino acids among other hAT family transposases, especially among the active ones, may indicate their importance for the catalytic activity of the transposases. Therefore, replacement of unconserved amino acids in wild-type TcBuster sequence (SEQ ID NO: 1) with conserved amino acids among other hAT family may yield hyperactive mutant TcBuster transposase. The method may comprise obtaining sequences of TcBuster as well as other hAT family transposases; aligning the sequences and identifying the amino acids in TcBuster transposase with a different conserved counterpart among the other hAT family transposases; performing site-directed mutagenesis to produce mutant TcBuster transposase harboring the mutation(s).

[0079] A hyperactive mutant TcBuster transposase can comprise one or more amino acid substitutions based on alignment to other members of Buster subfamily or other members of hAT family. In many cases, the one or more amino acid substitutions can be substitutions of conserved amino acid for the unconserved amino acid in wild-type TcBuster sequence (SEQ ID NO: 1). Non-limiting examples of mutant TcBuster transposases include TcBuster transposases comprising at least one of the amino acid substitutions from **Table 5**. A mutant TcBuster transposase can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, or more of the amino acid substitutions from **Table 5**.

[0080] Another exemplary method can comprise systemically mutating acidic amino acids to basic amino acids and identifying hyperactive mutant transposase.

[0081] In some cases, mutant TcBuster transposase can comprise amino acid substitutions V377T, E469K, and D189A. A mutant TcBuster transposase can comprise amino acid substitutions K573E and E578L. A mutant TcBuster transposase can comprise amino acid substitution I452K. A mutant TcBuster transposase can comprise amino acid substitution

A358K. A mutant TcBuster transposase can comprise amino acid substitution V297K. A mutant TcBuster transposase can comprise amino acid substitution N85S. A mutant TcBuster transposase can comprise amino acid substitutions N85S, V377T, E469K, and D189A. A mutant TcBuster transposase can comprise amino acid substitutions I452F, V377T, E469K, and D189A. A mutant TcBuster transposase can comprise amino acid substitutions A358K, V377T, E469K, and D189A. A mutant TcBuster transposase can comprise amino acid substitutions V377T, E469K, D189A, K573E and E578L.

Table 5

Amino Acid of Wild-type TcBuster Transposase (SEQ ID NO: 1)	Amino Acid Substitution
Q151	Q151S
Q151	Q151A
A154	A154P
Q615	Q615A
V553	V553S
Y155	Y155H
Y201	Y201A
F202	F202D
F202	F202K
C203	C203I
C203	C203V
F400	F400L
I398	I398D
I398	I398S
I398	I398K
V431	V431L
P559	P559D
P559	P559S
P559	P559K
M222	M222L

[0082] Fusion Transposase

[0083] Another aspect of the present invention provides a fusion transposase. The fusion transposase can comprise a TcBuster transposase sequence and a DNA sequence specific binding domain.

[0084] The TcBuster transposase sequence of a fusion transposase can comprise an amino acid sequence of any of the mutant TcBuster transposases as described herein. The TcBuster

transposase sequence of a fusion transposase can also comprise an amino acid sequence of a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1.

[0085] A DNA sequence specific binding domain as described herein can refer to a protein domain that is adapted to bind to a DNA molecule at a sequence region ("target sequence") containing a specific sequence motif. For instance, an exemplary DNA sequence specific binding domain may selectively bind to a sequence motif TATA, while another exemplary DNA sequence specific binding domain may selectively bind to a different sequence motif ATGCNTAGAT (N denotes any one of A, T, G, and C).

[0086] A fusion transposase as provided herein may direct sequence specific insertion of the transposon. For instance, a DNA sequence specific binding domain may guide the fusion transposase to bind to a target sequence based on the binding specificity of the binding domain. Being bound to or restricted to a certain sequence region may spatially limit the interaction between the fusion transposase and the transposon, thereby limiting the catalyzed transposition to a sequence region in proximity to the target sequence. Depending on the size, three-dimensional configuration, and sequence binding affinity of the DNA binding domain, as well as the spatial relationship between the DNA binding domain and the TcBuster transposase sequence, and the flexibility of the connection between the two domains, the distance of the actual transposition site to the target sequence may vary. Proper design of the fusion transposase configuration can direct the transposition to a desirable target genomic region.

[0087] A target genomic region for transposition can be any particular genomic region, depending on application purposes. For instance, sometimes, it is desirable to avoid transcription start sites for the transposition, which may cause undesirable, or even harmful, change in expression level of certain important endogenous gene(s) of the cell. A fusion transposase may contain a DNA sequence specific binding domain that can target the transposition to a safe harbor of the host genome. Non-limiting examples of safe harbors can include HPRT, AAVS site (e.g. AAVS1, AAVS2, ETC.), CCR5, or Rosa26. Safe harbor sites can generally refer to sites for transgene insertion whose use exert little to none disrupting effects on genome integrity of the cell or cellular health and functions.

[0088] A DNA sequence specific binding domain may be derived from, or be a variant of any DNA binding protein that has sequence-specificity. In many instances, a DNA sequence specific binding domain may comprise an amino acid sequence at least 40%, at least 50%, at least 60%, at least70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to a naturally occurring sequence-specific DNA binding protein. A DNA sequence specific binding domain may comprise an amino acid sequence at least 70% identical to a

naturally occurring sequence-specific DNA binding protein. Non-limiting examples of a naturally occurring sequence-specific DNA binding protein include, but not limited to, transcription factors from various origins, specific-sequence nucleases, and viral replication proteins. A naturally occurring sequence-specific DNA binding protein can also be any other protein having the specific binding capability from various origins. Selection and prediction of DNA binding proteins can be conducted by various approaches, including, but not limited to, using computational prediction databases available online, like DP-Bind (http://lcg.rit.albany.edu/dp-bind/) or DNABIND (http://dnabind.szialab.org/)

[0089] The term "transcription factor" can refer to a protein that controls the rate of transcription of genetic information from DNA to messenger DNA, by binding to a specific DNA sequence. A transcription factor that can be used in a fusion transposase described herein can be based on a prokaryotic transcription factor or a eukaryotic transcription factor, as long as it confers sequence specificity when binding to the target DNA molecule. Transcription factor prediction databases such as DBD (http://www.transcriptionfactor.org) may be used for selection of appropriate transcription factor for application of the disclosure herein.

[0090] A DNA sequence specific binding domain as used herein can comprise one or more DNA binding domain from a naturally occurring transcription factor. Non-limiting examples of DNA binding domains of transcription factors include DNA binding domains that belong to families like basic helix-loop-helix, basic-leucine zipper (bZIP), C-terminal effector domain of the bipartite response regulators, AP2/ERF/GCC box, helix-turn-helix, homeodomain proteins, lambda repressor-like, srf-like (serum response factor), paired box, winged helix, zinc fingers, multi-domain Cys2His2 (C2H2) zinc fingers, Zn2/Cys6, or Zn2/Cys8 nuclear receptor zinc finger.

[0091] A DNA sequence specific binding domain can be an artificially engineered amino acid sequence that binds to specific DNA sequences. Non-limiting examples of such artificially designed amino acid sequence include sequences created based on frameworks like transcription activator like effector nucleases (TALEs) DNA binding domain, zinc finger nucleases, adeno associated virus (AAV) Rep protein, and any other suitable DNA binding proteins as described herein.

[0092] Natural TALEs are proteins secreted by Xanthomonas bacteria to aid the infection of plant species. Natural TALEs can assist infections by binding to specific DNA sequences and activating the expression of host genes. In general, TALE proteins consist of a central repeat domain, which determines the DNA targeting specificity and can be rapidly synthesized de novo. TALEs have a modular DNA-binding domain (DBD) containing repetitive sequences of

residues. In some TALEs, each repeat region contains 34 amino acids. The term "TALE domain" as used herein can refer to the modular DBD of TALEs. A pair of residues at the 12th and 13th position of each repeat region can determine the nucleotide specificity and are referred to as the repeat variable diresidue (RVD). The last repeat region, termed the half-repeat, is typically truncated to 20 amino acids. Combining these repeat regions allows synthesizing sequence-specific synthetic TALEs. The C-terminus typically contains a nuclear localization signal (NLS), which directs a TALE to the nucleus, as well as a functional domain that modulates transcription, such as an acidic activation domain (AD). The endogenous NLS can be replaced by an organism-specific localization signal. For example, an NLS derived from the simian virus 40 large T-antigen can be used in mammalian cells. The RVDs HD, NG, NI, and NN target C, T, A, and G/A, respectively. A list of RVDs and their binding preferences under certain circumstances for nucleotides can be found in **Table 6**. Additional TALE RVDs can also be used for custom degenerate TALE-DNA interactions. For example, NA has high affinity for all four bases of DNA. Additionally, N*, where * is an RVD with a deletion in the 13th residue, can accommodate all letters of DNA including methylated cytosine. Also S* may have the ability to bind to any DNA nucleotide.

[0093] A number of online tools are available for designing TALEs to target a specific DNA sequence, for example TALE-NT (https://tale-nt.cac.cornell.edu/), Mojo hand (http://www.talendesign.org/). Commercially available kits may also assist in creating custom assembly of TALE repeat regions between the N and C-terminus of the protein. These methods can be used to assemble custom DBDs, which are then cloned into an expression vector containing a functional domain, e.g. TcBuster transposase sequence.

Table 6 RVD Binding Preference

	nucleotides			
RVD	**A**	**G**	**C**	**T**
NN	medium	medium		
NK		weak		
NI	medium			
NG				weak
HD			medium	
NS	weak	medium	weak	weak
NG				weak
N*			weak	weak
HN	weak	medium		
NT	weak	medium		
NP	weak		weak	medium

NH		medium		
SN		weak		
SH		weak		
NA	weak	strong	weak	weak
IG				weak
H*	poor	poor	weak	poor
ND			weak	
HI	medium			
HG				weak
NC				weak
NQ		weak		
SS		weak		
SN		weak		
S*	medium	medium	strong	medium
NV	weak	medium	poor	poor
HH	poor	poor	poor	poor
YG	poor	poor	poor	poor

[0094] TALEs can be synthesized de novo in the laboratory, for example, by combining digestion and ligation steps in a Golden Gate reaction with type II restriction enzymes. Alternatively, TALE can be assembled by a number of different approaches, including, but not limited to, Ligation-Independent Cloning (LIC), Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) assembly, and Iterative-Capped Assembly (ICA).

[0095] Zinc fingers (ZF) are ~30 amino acids that can bind to a limited combination of ~3 nucleotides. The C2H2 ZF domain may be the most common type of ZF and appears to be one of the most abundantly expressed proteins in eukaryotic cells. ZFs are small, functional and independently folded domains coordinated with zinc molecules in their structure. Amino acids in each ZF can have affinity towards specific nucleotides, causing each finger to selectively recognize 3--4 nucleotides of DNA. Multiple ZFs can be arranged into a tandem array and recognize a set of nucleotides on the DNA. By using a combination of different zinc fingers, a unique DNA sequence within the genome can be targeted. Different ZFPs of various lengths can be generated, which may allow for recognition of almost any desired DNA sequence out of the possible 64 triplet subsites.

[0096] Zinc fingers to be used in connection with the present disclosure can be created using established modular assembly fingers, such as a set of modular assembly finger domains developed by Barbas and colleagues, and also another set of modular assembly finger domains by ToolGen. Both set of domains cover all 3 bp GNN, most ANN, many CNN and some TNN triplets (where N can be any of the four nucleotides). Both have a different set of fingers, which allows for searching and coding different ZF modules as needed. A combinatorial selection-

based oligomerized pool engineering (OPEN) strategy can also be employed to minimize context-dependent effects of modular assembly involving the position of a finger in the protein and the sequence of neighboring fingers. OPEN ZF arrays are publicly available from the Zinc Finger Consortium Database.

[0097] AAV Rep DNA-binding domain is another DNA sequence specific binding domain that can be used in connection with the subject matter of the present disclosure. Viral cis -acting inverted terminal repeats (ITRs), and the trans-acting viral Rep proteins (Rep) are believed to be the factors mediating preferential integration of AAV into AAVS1 site of the host genome in the absence of a helper virus. AAV Rep protein can bind to specific DNA sequence in the AAVS1 site. Therefore, a site-specific DNA-binding domain can be fused together with a TcBuster transposase domain as described herein.

[0098] A fusion transposase as provided herein can comprise a TcBuster transposase sequence and a tag sequence. A tag sequence as provide herein can refer to any protein sequence that can be used as a detection tag of the fusion protein, such as, but not limited to, reporter proteins and affinity tags that can be recognized by antibodies. Reporter proteins include, but not limited to, fluorescent proteins (e.g. GFP, RFP, mCherry, YFP), β -galactosidase (β-gal), alkaline phosphatase (AP), chloramphenicol acetyl transferase (CAT), horseradish peroxidase (HRP). Non-limiting examples of affinity tags include polyhistidine (His tag), Glutathione S-Transferase (GST), Maltose Binding Protein (MBP), Calmodulin Binding Peptide (CBP), intein-chitin binding domain (intein-CBD), Streptavidin/Biotin-based tags, Epitope tags like FLAG, HA, c-myc, T7, Glu-Glu and many others.

[0099] A fusion transposase as provided herein can comprise a TcBuster transposase sequence and a DNA sequence specific binding domain or a tag sequence fused together without any intermediate sequence (e.g., "back-to-back"). In some cases, a fusion transposase as provided herein can comprise a TcBuster transposase sequence and a DNA sequence specific binding domain or a tag sequence joined by a linker sequence. **Fig. 8** is a schematic of an exemplary fusion transposase that comprises a DNA sequence specific binding domain and a TcBuster transposase sequence, joined by a linker. In an exemplary fusion transposase, a linker may serve primarily as a spacer between the first and second polypeptides. A linker can be a short amino acid sequence to separate multiple domains in a single polypeptide. A linker sequence can comprise linkers occurring in natural multi-domain proteins. In some instances, a linker sequence can comprise linkers artificially created. The choice of linker sequence may be based on the application of the fusion transposase. A linker sequence can comprise 3, 4, 5, 6, 7, 8, 9, 10, or more amino acids. In some embodiments, the linker sequence may comprise at least 3, at

least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, or at least 50 amino acids. In some embodiments, the linker sequence can comprise at most 4, at most 5, at most 6, at most 7, at most 8, at most 9, at most 10, at most 11, at most 12, at most 15, at most 20, at most 30, at most 40, at most 50, or at most 100 amino acids. In certain cases, it may be desirable to use flexible linker sequences, such as, but not limited to, stretches of Gly and Ser residues ("GS" linker) like (GGGGS)n (n=2-8), $(Gly)_8$, GSAGSAAGSGEF, $(GGGGS)_4$. Sometimes, it may be desirable to use rigid linker sequences, such as, but not limited to, (EAAAK)n (n=2-7), Pro-rich sequences like (XP)n, with X designating any amino acid.

[00100] In an exemplary fusion transposase provided herein, a TcBuster transposase sequence can be fused to the N-terminus of a DNA sequence specific binding domain or a tag sequence. Alternatively, a TcBuster transposase sequence can be fused to the C-terminus of a DNA sequence specific binding domain or a tag sequence. In some embodiments, a third domain sequence or more of other sequences can be present in between the TcBuster transposase and the DNA sequence specific binding domain or the tag sequence, depending on the application of the fusion transposase.

[00101] TcBuster Transposon

[00102] Another aspect of the present disclosure provides a TcBuster transposon that comprises a cassette cargo positioned between two inverted repeats. A TcBuster transposon can be recognized by a TcBuster transposase as described herein, e.g., a TcBuster transposase can recognize the TcBuster transposon and catalyze transposition of the TcBuster transposon into a DNA sequence.

[00103] The terms "inverted repeats", "terminal inverted repeats", "inverted terminal repeats", as used interchangeably herein, can refer to short sequence repeats flanking the transposase gene in a natural transposon or a cassette cargo in an artificially engineered transposon. The two inverted repeats are generally required for the mobilization of the transposon in the presence of a corresponding transposase. Inverted repeats as described herein may contain one or more direct repeat (DR) sequences. These sequences usually are embedded in the terminal inverted repeats (TIRs) of the elements. The term "cargo cassette" as used herein can refer to a nucleotide sequence other than a native nucleotide sequence between the inverted repeats that contains the TcBuster transposase gene. A cargo cassette can be artificially engineered.

[00104] A transposon described herein may contain a cargo cassette flanked by IR/DR sequences. In some embodiments, at least one of the repeats contains at least one direct repeat. As shown in Figs. 1 and 2, a transposon may contain a cargo cassette flanked by IRDR-L-Seq1 (SEQ ID NO: 3) and IRDR-R-Seq1 (SEQ ID NO: 4). In many cases, a left inverted repeat can

comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-L-Seq1 (SEQ ID NO: 3). Sometimes, a right inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-R-Seq1 (SEQ ID NO: 4). In other cases, a right inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-L-Seq1 (SEQ ID NO: 3). Sometimes, a left inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-R-Seq1 (SEQ ID NO: 4). The terms "left" and "right", as used herein, can refer to the 5' and 3' sides of the cargo cassette on the sense strand of the double strand transposon, respectively. It is also possible that a transposon may contain a cargo cassette flanked by IRDR-L-Seq2 (SEQ ID NO: 5) and IRDR-R-Seq2 (SEQ ID NO: 6). In many cases, a left inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-L-Seq2 (SEQ ID NO: 5). Sometimes, a right inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-R-Seq2 (SEQ ID NO: 6). In other cases, a right inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-L-Seq2 (SEQ ID NO: 5). Sometimes a left inverted repeat can comprise a sequence at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to IRDR-R-Seq2 (SEQ ID NO: 6). A transposon may contain a cargo cassette flanked by two inverted repeats that have different nucleotide sequences than the ones given in **Fig. 2**, or a combination of the various sequences known to one skilled in the art. At least one of the two inverted repeats of a transposon described herein may contain a sequence that is at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to any one of SEQ ID NOs: 3-6. At least one of inverted repeats of a transposon described herein may contain a sequence that is at least 80% identical to SEQ ID NO: 3 or 4. At least one of inverted repeats of a transposon described herein may contain a sequence that is at least 80% identical to SEQ ID NO: 5 or 6. The choice of inverted repeat sequences may vary depending on the expected transposition efficiency, the type of cell to be modified, the transposase to use, and many other factors.

[00105] In many implementations, minimally sized transposon vector inverted terminal repeats that conserve genomic space may be used. The ITRs of hAT family transposons diverge greatly with differences in right-hand and left-hand ITRs. In many cases, smaller ITRs consisting of just 100-200 nucleotides are as active as the longer native ITRs in hAT transposon vectors. These sequences may be consistently reduced while mediating hAT family transposition. These shorter ITRs can conserve genomic space within hAT transposon vectors.

[00106] The inverted repeats of a transposon provided herein can be about 50 to 2000 nucleotides, about 50 to 1000 nucleotides, about 50 to 800 nucleotides, about 50 to 600 nucleotides, about 50 to 500 nucleotides, about 50 to 400 nucleotides, about 50 to 350 nucleotides, about 50 to 300 nucleotides, about 50 to 250 nucleotides, about 50 to 200 nucleotides, about 50 to 180 nucleotides, about 50 to 160 nucleotides, about 50 to 140 nucleotides, about 50 to 120 nucleotides, about 50 to 110 nucleotides, about 50 to 100 nucleotides, about 50 to 90 nucleotides, about 50 to 80 nucleotides, about 50 to 70 nucleotides, about 50 to 60 nucleotides, about 75 to 750 nucleotides, about 75 to 450 nucleotides, about 75 to 325 nucleotides, about 75 to 250 nucleotides, about 75 to 150 nucleotides, about 75 to 95 nucleotides, about 100 to 500 nucleotides, about 100 to 400 nucleotides, about 100 to 350 nucleotides, about 100 to 300 nucleotides, about 100 to 250 nucleotides, about 100 to 220 nucleotides, about 100 to 200 nucleotides, or in any range derived therefrom.

[00107] In some cases, a cargo cassette can comprise a promoter, a transgene, or a combination thereof. In cargo cassettes comprising both a promoter and a transgene, the expression of the transgene can be directed by the promoter. A promoter can be any type of promoter available to one skilled in the art. Non-limiting examples of the promoters that can be used in a TcBuster transposon include EFS, CMV, MND, EF1α, CAGGs, PGK, UBC, U6, H1, and Cumate. The choice of a promoter to be used in a TcBuster transposition would depend on a number of factors, such as, but not limited to, the expression efficiency of the promoter, the type of cell to be genetically modified, and the desired transgene expression level.

[00108] A transgene in a TcBuster transposon can be any gene of interest and available to one skilled in the art. A transgene can be derived from, or a variant of, a gene in nature, or can be artificially designed. A transgene can be of the same species origin as the cell to be modified, or from different species. A transgene can be a prokaryotic gene, or a eukaryotic gene. Sometimes, a transgene can be a gene derived from a non-human animal, a plant, or a human being. A transgene can comprise introns. Alternatively, a transgene may have introns removed or not present.

[00109] In some embodiments, a transgene can code for a protein. Exemplary proteins include, but are not limited to, a cellular receptor, an immunological checkpoint protein, a cytokine, or any combination thereof. Sometimes, a cellular receptor as described herein can include, but not limited to a T cell receptor (TCR), a B cell receptor (BCR), a chimeric antigen receptor (CAR), or any combination thereof.

[00110] A cargo cassette as described herein may not contain a transgene coding for any type of protein product, but that is useful for other purposes. For instance, a cargo cassette may be used for creating frameshift in the insertion site, for example, when it is inserted in an exon of a gene in the host genome. This may lead to a truncation of the gene product or a null mutation. Sometimes, a cargo cassette may be used for replacing an endogenous genomic sequence with an exogenous nucleotide sequence, thereby modifying the host genome.

[00111] A transposon described herein may have a cargo cassette in either forward or reverse direction. In many cases, a cargo cassette has its own directionality. For instance, a cargo cassette containing a transgene would have a 5' to 3' coding sequence. A cargo cassette containing a promoter and a gene insertion would have promoter on the 5' site of the gene insertion. The term "forward direction", as used herein, can refer to the situation where a cargo cassette maintains its directionality on the sense strand of the double strand transposon. The term "reverse direction", as used herein, can refer to the situation where a cargo cassette maintains its directionality on the antisense strand of the double strand transposon.

[00112] Systems for Genome Editing and Methods of Use

[00113] Another aspect of the present disclosure provides a system for genome editing. A system can comprise a TcBuster transposase and a TcBuster transposon. A system can be used to edit a genome of a host cell, disrupting or modifying an endogenous genomic region of the host cell, inserting an exogenous gene into the host genome, replacing an endogenous nucleotide sequence with an exogenous nucleotide sequence or any combination thereof.

[00114] A system for genome editing can comprise a mutant TcBuster transposase or fusion transposase as described herein, and a transposon recognizable by the mutant TcBuster transposase or the fusion transposase. A mutant TcBuster transposase or the fusion transposase can be provided as a purified protein. Protein production and purification technologies are known to one skilled in the art. The purified protein can be kept in a different container than the transposon, or they can be kept in the same container.

[00115] In many cases, a system for genome editing can comprise a polynucleotide encoding a mutant TcBuster transposase or fusion transposase as described herein, and a transposon recognizable by the mutant TcBuster transposase or the fusion transposase. Sometimes, a

polynucleotide of the system can comprise DNA that encodes the mutant TcBuster transposase or the fusion transposase. Alternatively or additionally, a polynucleotide of the system can comprise messenger RNA (mRNA) that encodes the mutant TcBuster transposase or the fusion transposase. The mRNA can be produced by a number of approaches well known to one of ordinary skills in the art, such as, but not limited to, in vivo transcription and RNA purification, in vitro transcription, and de novo synthesis. In many cases, the mRNA can be chemically modified. The chemically modified mRNA may be resistant to degradation than unmodified or natural mRNAs or may degrade more quickly. In many cases, the chemical modification of the mRNA may render the mRNA being translated with more efficiency. Chemical modification of mRNAs can be performed with well-known technologies available to one skilled in the art, or by commercial vendors.

[00116] For many applications, safety dictates that the duration of hAT transposase expression be only long enough to mediate safe transposon delivery. Moreover, a pulse of hAT transposase expression that coincides with the height of transposon vector levels can achieve maximal gene delivery. The implementations are made using available technologies for the in vitro transcription of RNA molecules from DNA plasmid templates. The RNA molecules can be synthesized using a variety of methods for in vitro (e.g., cell free) transcription from a DNA copy. Methods to do this have been described and are commercially available. For example, the mMessage Machine in vitro transcription kit available through life technologies.

[00117] There are also a number of companies that can perform in vitro transcription on a fee for service basis. We have also found that that chemically modified RNAs for hAT expression work especially well for gene transfer. These chemically modified RNAs do not induce cellular immune responses and RNA generated using proprietary methods that also avoid the cellular immune response. These RNA preparations remove RNA dimers (Clean-Cap) and cellular reactivity (pseudouridine incorporation) produce better transient gene expression in human T cells without toxicity in our hands (data not shown). The RNA molecules can be introduced into cells using any of many described methods for RNA transfection, which is usually non-toxic to most cells. Methods to do this have been described and are commercially available. For example, the Amaxa nucleofector, Neon electroporator, and the Maxcyte platforms.

[00118] A transposon as described herein may be present in an expression vector. In many cases, the expression vector can be DNA plasmid. Sometimes, the expression vector can be a mini-circle vector. The term "mini-circle vector" as used herein can refer to small circular plasmid derivative that is free of most, if not all, prokaryotic vector parts (e.g., control sequences or non-functional sequences of prokaryotic origin). Under circumstances, the toxicity to the cells

created by transfection or electroporation can be mitigated by using the "mini-circles" as described herein.

[00119] A mini-circle vector can be prepared by well-known molecular cloning technologies available. First, a 'parental plasmid' (bacterial plasmid with insertion, such as transposon construct) in bacterial, such as E. coli, can be produced, which can be followed by induction of a site-specific recombinase. These steps can then be followed by the excision of prokaryotic vector parts via two recombinase-target sequences at both ends of the insert, as well as recovery of the resulting mini-circle vector. The purified mini-circle can be transferred into the recipient cell by transfection or lipofection and into a differentiated tissue by, for instance, jet injection. A mini-circle containing TcBuster transposon can have a size about 1.5kb, about 2 kb, about 2.2 kb, about 2.4 kb, about 2.6 kb, about 2.8 kb, about 3 kb, about 3.2 kb, about 3.4 kb, about 3.6 kb, about 3.8 kb, about 4 kb, about 4.2 kb, about 4.4 kb, about 4.6 kb, about 4.8 kb, about 5 kb, about 5.2 kb, about 5.4 kb, about 5.6 kb, about 5.8 kb, about 6 kb, about 6.5 kb, about 7 kb, about 8 kb, about 9 kb, about 10 kb, about 12 kb, about 25 kb, about 50 kb, or a value between any two of these numbers. Sometimes, a mini-circle containing TcBuster transposon as provided herein can have a size at most 2.1 kb, at most 3.1 kb, at most 4.1 kb, at most 4.5 kb, at most 5.1 kb, at most 5.5 kb, at most 6.5 kb, at most 7.5 kb, at most 8.5 kb, at most 9.5 kb, at most 11 kb, at most 13 kb, at most 15 kb, at most 30 kb, or at most 60 kb.

[00120] In certain embodiments, a system as described herein may contain a polynucleotide encoding a mutant TcBuster transposase or fusion transposase as described herein, and a transposon, which are present in a same expression vector, e.g. plasmid.

[00121] Yet another aspect of the present disclosure provides a method of genetic engineering. A method of genetic engineering can comprise introducing into a cell a TcBuster transposase and a transposon recognizable by the TcBuster transposase. A method of genetic engineering can also be performed in a cell-free environment. A method of genetic engineering in a cell-free environment can comprise combining a TcBuster transposase, a transposon recognizable by the transposase, and a target nucleic acid into a container, such as a well or tube.

[00122] A method described herein can comprises introducing into a cell a mutant TcBuster transposase provided herein and a transposon recognizable by the mutant TcBuster transposase. A method of genome editing can comprise: introducing into a cell a fusion transposase provided herein and a transposon recognizable by the fusion transposase.

[00123] The mutant TcBuster transposase or the fusion transposase can be introduced into the cell either as a protein or via a polynucleotide that encodes for the mutant TcBuster transposase

or the fusion transposase. The polynucleotide, as discussed above, can comprise a DNA or an mRNA that encodes the mutant TcBuster transposase or the fusion transposase.

[00124] In many instances, the TcBuster transposase or the fusion transposase can be transfected into a host cell as a protein, and the concentration of the protein can be at least 0.05nM, at least 0.1 nM, at least 0.2 nM, at least 0.5 nM, at least 1 nM, at least 2 nM, at least 5 nM, at least 10 nM, at least 50 nM, at least 100 nM, at least 200 nM, at least 500 nM, at least 1 μM, at least 2 μM, at least 5 μM, at least 7.5 μM, at least 10 μM, at least 15 μM, at least 20 μM, at least 25 μM, at least 50 μM, at least 100 μM, at least 200 μM, at least 500 μM, or at least 1 μM. Sometimes, the concentration of the protein can be around 1 μM to around 50 μM, around 2 μM to around 25 μM, around 5 μM to around 12.5 μM, or around 7.5 μM to around 10 μM.

[00125] In many cases, the TcBuster transposase or the fusion transposase can be transfected into a host cell through a polynucleotide, and the concentration of the polynucleotide can be at least about 5 ng/ml, 10 ng/ml, 20 ng/ml, 40 ng/ml, 50 ng/ml, 60 ng/ml, 80 ng/ml, 100 ng/ml, 120 ng/ml, 150 ng/ml, 180 ng/ml, 200 ng/ml, 220 ng/ml, 250 ng/ml, 280 ng/ml, 300 ng/ml, 500 ng/ml, 750 ng/ml, 1 μg /ml, 2 μg /ml, 3 μg /ml, 5 μg/ml, 50 μg/ml, 100 μg/ml, 150 μg/ml, 200 μg/ml, 250 μg/ml, 300 μg/ml, 350 μg/ml, 400 μg/ml, 450 μg/ml, 500 μg/ml, 550 μg/ml, 600 μg/ml, 650 μg/ml, 700 μg/ml, 750 μg/ml, or 800 μg/ml. Sometimes, the concentration of the polynucleotide can be between about 5-25 μg/ml, 25-50 μg/ml, 50-100 μg/ml, 100-150 μg/ml, 150-200 μg/ml, 200-250 μg/ml, 250-500 μg/ml, 5-800 μg/ml, 200-800 μg/ml, 250-800 μg/ml, 400-800 μg/ml, 500-800 μg/ml, or any range derivable therein. In many cases, the transposon is present in a separate expression vector than the transposase, and the concentration of the transposon can be at least about 5 ng/ml, 10 ng/ml, 20 ng/ml, 40 ng/ml, 50 ng/ml, 60 ng/ml, 80 ng/ml, 100 ng/ml, 120 ng/ml, 150 ng/ml, 180 ng/ml, 200 ng/ml, 220 ng/ml, 250 ng/ml, 280 ng/ml, 300 ng/ml, 500 ng/ml, 750 ng/ml, 1 μg /ml, 2 μg /ml, 3 μg /ml, 5 μg/ml, 50 μg/ml, 100 μg/ml, 150 μg/ml, 200 μg/ml, 250 μg/ml, 300 μg/ml, 350 μg/ml, 400 μg/ml, 450 μg/ml, 500 μg/ml, 550 μg/ml, 600 μg/ml, 650 μg/ml, 700 μg/ml, 750 μg/ml, or 800 μg/ml. Sometimes, the concentration of the transposon can be between about 5-25 μg/ml, 25-50 μg/ml, 50-100 μg/ml, 100-150 μg/ml, 150-200 μg/ml, 200-250 μg/ml, 250-500 μg/ml, 5-800 μg/ml, 200-800 μg/ml, 250-800 μg/ml, 400-800 μg/ml, 500-800 μg/ml, or any range derivable therein. It is possible the ratio of the transposon versus the polynucleotide coding for the transposase is at most 10000, at most 5000, at most 1000, at most 500, at most 200, at most 100, at most 50, at most 20, at most 10, at most 5, at most 2, at most 1, at most 0.1, at most 0.05, at most 0.01, at most 0.001, at most 0.0001, or any number in between any two thereof.

[00126] In some other cases, the transposon and the polynucleotide coding for the transposase are present in the same expression vector, and the concentration of the expression vector containing both transposon and the polynucleotide encoding transposase can be at least about 5 ng/ml, 10 ng/ml, 20 ng/ml, 40 ng/ml, 50 ng/ml, 60 ng/ml, 80 ng/ml, 100 ng/ml, 120 ng/ml, 150 ng/ml, 180 ng/ml, 200 ng/ml, 220 ng/ml, 250 ng/ml, 280 ng/ml, 300 ng/ml, 500 ng/ml, 750 ng/ml, 1 μg /ml, 2 μg /ml, 3 μg /ml, 5 μg/ml, 50 μg/ml, 100 μg/ml, 150 μg/ml, 200 μg/ml, 250 μg/ml, 300 μg/ml, 350 μg/ml, 400 μg/ml, 450 μg/ml, 500 μg/ml, 550 μg/ml, 600 μg/ml, 650 μg/ml, 700 μg/ml, 750 μg/ml, or 800 μg/ml. Sometimes, the concentration of the expression vector containing both transposon and the polynucleotide encoding transposase can be between about 5-25 μg/ml, 25-50 μg/ml, 50-100 μg/ml, 100-150 μg/ml, 150-200 μg/ml, 200-250 μg/ml, 250-500 μg/ml, 5-800 μg/ml, 200-800 μg/ml, 250-800 μg/ml, 400-800 μg/ml, 500-800 μg/ml, or any range derivable therein.

[00127] In some cases, the amount of polynucleic acids that may be introduced into the cell by electroporation may be varied to optimize transfection efficiency and/or cell viability. In some cases, less than about 100 pg of nucleic acid may be added to each cell sample (e.g., one or more cells being electroporated). In some cases, at least about 100 pg, at least about 200 pg, at least about 300 pg, at least about 400 pg, at least about 500 pg, at least about 600 pg, at least about 700 pg, at least about 800 pg, at least about 900 pg, at least about 1 microgram, at least about 1.5 μg, at least about 2 μg, at least about 2.5 μg, at least about 3 μg, at least about 3.5 μg, at least about 4 μg, at least about 4.5 μg, at least about 5 μg, at least about 5.5 μg, at least about 6 μg, at least about 6.5 μg, at least about 7 μg, at least about 7.5 μg, at least about 8 μg, at least about 8.5 μg, at least about 9 μg, at least about 9.5 μg, at least about 10 μg, at least about 11 μg, at least about 12 μg, at least about 13 μg, at least about 14 μg, at least about 15 μg, at least about 20 μg, at least about 25 μg, at least about 30 μg, at least about 35 μg, at least about 40 μg, at least about 45 μg, or at least about 50 μg, of nucleic acid may be added to each cell sample (e.g., one or more cells being electroporated). For example, 1 microgram of dsDNA may be added to each cell sample for electroporation. In some cases, the amount of polynucleic acids (e.g., dsDNA) required for optimal transfection efficiency and/or cell viability may be specific to the cell type.

[00128] The subject matter disclosed herein may find use in genome editing of a wide range of various types of host cells. In preferred embodiments, the host cells may be from eukaryotic organisms. In some embodiments, the cells may be from a mammal origin. In some embodiments, the cells may be from a human origin.

[00129] In general, the cells may be from an immortalized cell line or primary cells.

[00130] The terms "cell line" and "immortalized cell line", as used herein interchangeably, can refer to a population of cells from an organism which would normally not proliferate indefinitely but, due to mutation, may have evaded normal cellular senescence and instead can keep undergoing division. The subject matter provided herein may find use in a range of common established cell lines, including, but not limited to, human BC-1 cells, human BJAB cells, human IM-9 cells, human Jiyoye cells, human K-562 cells, human LCL cells, mouse MPC-11 cells, human Raji cells, human Ramos cells, mouse Ramos cells, human RPMI8226 cells, human RS4-11 cells, human SKW6.4 cells, human Dendritic cells, mouse P815 cells, mouse RBL-2H3 cells, human HL-60 cells, human NAMALWA cells, human Macrophage cells, mouse RAW 264.7 cells, human KG-1 cells, mouse M1 cells, human PBMC cells, mouse BW5147 (T200-A)5.2 cells, human CCRF-CEM cells, mouse EL4 cells, human Jurkat cells, human SCID.adh cells, human U-937 cells or any combination of cells thereof.

[00131] The term "primary cells" and its grammatical equivalents, as used herein, can refer to cells taken directly from an organism, typically living tissue of a multicellular organism, such as animals or plants. In many cases, primary cells may be established for growth in vitro. In some cases, primary cells may be just removed from the organism and have not been established for growth in vitro yet before the transfection. In some embodiments, the primary cells can also be expanded in vitro, i.e. primary cells may also include progeny cells that are generated from proliferation of the cells taken directly from an organism. In these cases, the progeny cells do not exhibit the indefinite proliferative property as cells in established cell lines. For instance, the host cells may be human primary T cells, while prior to the transfection, the T cells have been exposed to stimulatory factor(s) that may result in T cell proliferation and expansion of the cell population.

[00132] The cells to be genetically modified may be primary cells from tissues or organs, such as, but not limited to, brain, lung, liver, heart, spleen, pancreas, small intestine, large intestine, skeletal muscle, smooth muscle, skin, bones, adipose tissues, hairs, thyroid, trachea, gall bladder, kidney, ureter, bladder, aorta, vein, esophagus, diaphragm, stomach, rectum, adrenal glands, bronchi, ears, eyes, retina, genitals, hypothalamus, larynx, nose, tongue, spinal cord, or ureters, uterus, ovary, testis, and any combination thereof. In certain embodiments, the cells may include, but not limited to, hematocyte, trichocyte, keratinocyte, gonadotrope, corticotrope, thyrotrope, somatotrope, lactotroph, chromaffin cell, parafollicular cell, glomus cell, melanocyte, nevus cell, merkel cell, odontoblast, cementoblast, corneal keratocyte, retina muller cell, retinal pigment epithelium cell, neuron, glia, ependymocyte, pinealocyte, pneumocyte, clara cell, goblet cell, G cell, D cell, Enterochromaffin-like cell, gastric chief cell, parietal cell, foveolar

cell, K cell, D cell, I cell, paneth cell, enterocyte, microfold cell, hepatocyte, hepatic stellate cell, cholecystocyte, centroacinar cell, pancreatic stellate cell, pancreatic α cell, pancreatic β cell, pancreatic δ cell, pancreatic F cell, pancreatic ε cell, thyroid parathyroid, oxyphil cell, urothelial cell, osteoblast, osteocyte, chondroblast, chondrocyte, fibroblast, fibrocyte, myoblast, myocyte, myosatellite cell, tendon cell, cardiac muscle cell, lipoblast, adipocyte, interstitial cell of cajal, angioblast, endothelial cell, mesangial cell, juxtaglomerular cell, macula densa cell, stromal cell, interstitial cell, telocyte, simple epithelial cell, podocyte, kidney proximal tubule brush border cell, sertoli cell, leydig cell, granulosa cell, peg cell, germ cell, spermatozoon ovum, lymphocyte, myeloid cell, endothelial progenitor cell, endothelial stem cell, angioblast, mesoangioblast, pericyte mural cell, and any combination thereof. In many instances, the cell to be modified may be a stem cell, such as, but not limited to, embryonic stem cell, hematopoietic stem cell, epidermal stem cell, epithelial stem cell, bronchoalveolar stem cell, mammary stem cell, mesenchymal stem cell, intestine stem cell, endothelial stem cell, neural stem cell, olfactory adult stem cell, neural crest stem cell, testicular cell, and any combination thereof. Sometimes, the cell can be an induced pluripotent stem cell that is derived from any type of tissue.

[00133] In some embodiments, the cell to be genetically modified may be a mammalian cell. In some embodiments, the cell may be an immune cell. Non-limiting examples of the cell can include a B cell, a basophil, a dendritic cell, an eosinophil, a gamma delta T cell, a granulocyte, a helper T cell, a Langerhans cell, a lymphoid cell, an innate lymphoid cell (ILC), a macrophage, a mast cell, a megakaryocyte, a memory T cell, a monocyte, a myeloid cell, a natural killer T cell, a neutrophil, a precursor cell, a plasma cell, a progenitor cell, a regulatory T-cell, a T cell, a thymocyte, any differentiated or de-differentiated cell thereof, or any mixture or combination of cells thereof. In certain cases, the cell may be a T cell. In some embodiments, the cell may be a primary T cell. In certain cases, the cell may be an antigen-presenting cell (APC). In some embodiments, the cell may be a primary APC. The APCs in connection with the present disclosure may be a dendritic cell, macrophage, B cell, other non-professional APCs, or any combination thereof.

[00134] In some embodiments, the cell may be an ILC (innate lymphoid cell), and the ILC can be a group 1 ILC, a group 2 ILC, or a group 3 ILC. Group 1 ILCs may generally be described as cells controlled by the T-bet transcription factor, secreting type-1 cytokines such as IFN-gamma and TNF-alpha in response to intracellular pathogens. Group 2 ILCs may generally be described as cells relying on the GATA-3 and ROR-alpha transcription factors, producing type-2 cytokines in response to extracellular parasite infections. Group 3 ILCs may generally be described as cells controlled by the ROR-gamma t transcription factor, and produce IL-17 and/or IL-22.

[00135] In some embodiments, the cell may be a cell that is positive or negative for a given factor. In some embodiments, a cell may be a CD3+ cell, CD3- cell, a CD5+ cell, CD5- cell, a CD7+ cell, CD7- cell, a CD14+ cell, CD14- cell, CD8+ cell, a CD8- cell, a CD103+ cell, CD103- cell, CD11b+ cell, CD11b- cell, a BDCA1+ cell, a BDCA1- cell, an L-selectin+ cell, an L-selectin- cell, a CD25+, a CD25- cell, a CD27+, a CD27- cell, a CD28+ cell, CD28- cell, a CD44+ cell, a CD44- cell, a CD56+ cell, a CD56- cell, a CD57+ cell, a CD57- cell, a CD62L+ cell, a CD62L- cell, a CD69+ cell, a CD69- cell, a CD45RO+ cell, a CD45RO- cell, a CD127+ cell, a CD127- cell, a CD132+ cell, a CD132- cell, an IL-7+ cell, an IL-7- cell, an IL-15+ cell, an IL-15- cell, a lectin-like receptor G1positive cell, a lectin-like receptor G1 negative cell, or an differentiated or de-differentiated cell thereof. The examples of factors expressed by cells is not intended to be limiting, and a person having skill in the art will appreciate that the cell may be positive or negative for any factor known in the art. In some embodiments, the cell may be positive for two or more factors. For example, the cell may be CD4+ and CD8+. In some embodiments, the cell may be negative for two or more factors. For example, the cell may be CD25-, CD44-, and CD69-. In some embodiments, the cell may be positive for one or more factors, and negative for one or more factors. For example, a cell may be CD4+ and CD8-.

[00136] It should be understood that cells used in any of the methods disclosed herein may be a mixture (e.g., two or more different cells) of any of the cells disclosed herein. For example, a method of the present disclosure may comprise cells, and the cells are a mixture of CD4+ cells and CD8+ cells. In another example, a method of the present disclosure may comprise cells, and the cells are a mixture of CD4+ cells and naïve cells.

[00137] As provided herein, the transposase and the transposon can be introduced in to a cell through a number of approaches. The term "transfection" and its grammatical equivalents as used herein can generally refer to a process whereby nucleic acids are introduced into eukaryotic cells. The transfection methods that can be used in connection with the subject matter can include, but not limited to, electroporation, microinjection, calcium phosphate precipitation, cationic polymers, dendrimers, liposome, microprojectile bombardment, fugene, direct sonic loading, cell squeezing, optical transfection, protoplast fusion, impalefection, magnetofection, nucleofection, or any combination thereof. In many cases, the transposase and transposon described herein can be transfected into a host cell through electroporation. Sometimes, transfection can also be done through a variant of electroporation method, such as nucleofection (also known as NucleofectorTM technology). The term "electroporation" and its grammatical equivalents as used herein can refer to a process whereby an electrical field is applied to cells in order to increase the permeability of the cell membrane, allowing chemicals, drugs, or DNA to

be introduced into the cell. During electroporation, the electric filed is often provided in the form of "pulses" of very brief time periods, e.g. 5 milliseconds, 10 milliseconds, and 50 milliseconds. As understood by those skilled in the art, electroporation temporarily opens up pores in a cell's outer membrane by use of pulsed rotating electric fields. Methods and apparatus used for electroporation in vitro and in vivo are also well known. Various electric parameters can be selected dependent on the cell type being electroporated and physical characteristics of the molecules that are to be taken up by the cell, such as pulse intensity, pulse length, number of pulses).

[00138] Applications

[00139] The subject matter, e.g., the compositions (e.g., mutant TcBuster transposases, fusion transposases, TcBuster transposons), systems and methods, provided herein may find use in a wide range of applications relating to genome editing, in various aspects of modern life.

[00140] Under certain circumstances, advantages of the subject matter described herein may include, but not limited to, reduced costs, regulatory consideration, lower immunogenicity and less complexity. In some cases, a significant advantage of the present disclosure is the high transposition efficiency. Another advantage of the present disclosure, in many cases, is that the transposition system provided herein can be "tunable", e.g., transposition can be designed to target select genomic region rather than random insertion.

[00141] One non-limiting example is related to create genetically modified cells for research and clinical applications. For example, as discussed above, genetically modified T cells can be created using the subject matter provided herein, which may find use in helping people fighting against a variety of diseases, such as, but not limited to, cancer and infectious disease.

[00142] One particular example includes generation of genetically modified primary leukocytes using the methods provided herein, and administering the genetically modified primary leukocytes to a patient in need thereof. The generation of genetically modified primary leukocytes can include introducing into a leukocyte a transposon and a mutant TcBuster transposase or the fusion transposase as described herein, which can recognize the transposon, thereby generating a genetically modified leukocyte. In many cases, the transposon may comprise a transgene. The transgene can be a cellular receptor, an immunological checkpoint protein, a cytokine, and any combination thereof. Sometimes, a cellular receptor can include, but not limited to a T cell receptor (TCR), a B cell receptor (BCR), a chimeric antigen receptor (CAR), or any combination thereof. In some other cases, the transposon and the transposase are designed to delete or modify an endogenous gene, for instance, a cytokine, an immunological checkpoint protein, an oncogene, or any combination thereof. The genetic modification of the

primary leukocytes can be designed to facilitate immunity against an infectious pathogen or cancer cells that render the patient in diseased state.

[00143] Another non-limiting example is related to create genetically modified organisms for agriculture, food production, medicine, and pharmaceutics. The species that can be genetically modified span a wide range, including, but not limited to, plants and animals. The genetically modified organisms, such as genetically modified crops or livestock, may be modified in a certain aspect of their physiological properties. Examples in food crops include resistance to certain pests, diseases, or environmental conditions, reduction of spoilage, or resistance to chemical treatments (e.g. resistance to a herbicide), or improving the nutrient profile of the crop. Examples in non-food crops include production of pharmaceutical agents, biofuels, and other industrially useful goods, as well as for bioremediation. Examples in livestock include resistance to certain parasites, production of certain nutrition elements, increase in growth rate, and increase in milk production.

[00144] The term "about" and its grammatical equivalents in relation to a reference numerical value and its grammatical equivalents as used herein can include a range of values plus or minus 10% from that value. For example, the amount "about 10" includes amounts from 9 to 11. The term "about" in relation to a reference numerical value can also include a range of values plus or minus 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% from that value.

EXAMPLES

[00145] The examples below further illustrate the described embodiments without limiting the scope of this disclosure.

[00146] **EXAMPLE 1. MATERIALS AND METHODS**

[00147] This example describes several methods utilized for generation and evaluation of exemplary mutant TcBuster transposases.

[00148] *Site Directed Mutagenesis for TcBuster mutant preparation*

[00149] Putative hyperactive TcBuster (TcB) transposase mutants were identified by nucleotide sequence and amino acid alignment of hAT and buster subfamilies. The Q5 site-directed mutagenesis kit (New England BioLabs) was used for all site-directed mutagenesis. Following PCR mutagenesis, PCR products were purified with GeneJET PCR purification kit (Thermo Fisher Scientific). A 20uL ligation reaction of purified PCR products was performed using T4 DNA ligase (New England BioLabs). 5uL of ligation reaction was used for transformation in DH10Beta cells. Direct colony sequencing through Sequetech was used to confirm the presence

of desired mutations. DNA for confirmed mutations was prepped using ZymoPURE plasmid miniprep kits (Zymo Research).

[00150] *Measuring transection efficiency in HEK-293T cells*

[00151] HEK-293T cells were plated at 300,000 cells per well of a 6 well plate one day prior to transfection. Cells were transfected with 500ng transposon carrying mCherry-puromycin cassette and 62.5ng TcB transposase using TransIT X2 reagent per manufacturer's instructions (Mirus Bio). Two days post-transfection, cells were re-plated with puromycin (1ug/mL) at a density of 3,000 cells/well of a 6 well plate in triplicate in DMEM complete media, or re-plated without puromycin selection. Stable integration of the transgene was assessed by colony counting of puromycin treated cells (each cell that survived drug selection formed a colony) or flow cytometry. For colony counting, two weeks post-puromycin selection, DMEM complete + puromycin media was removed. Cells were washed with 1X PBS and cells were stained with 1x crystal violet solution for 10 minutes. Plates were washed twice with PBS and colonies counted.

[00152] For flow cytometry analysis, stable integration of the transgene was assessed by detection of mCherry fluorescence in cells grown without drug selection. Transfected cells were harvested at indicated time points post-transfection, washed 1X with PBS and resuspended in 200uL RDFII buffer for analysis. Cells were analyzed using Novocyte (Acea Biosciences) and mCherry expression was assessed using the PE-Texas red channel.

[00153] *Screening of TcB transposase mutants in HEK-293T cells*

[00154] HEK-293T cells were plated at 75,000 cells per well of a 24 well plate one day prior to transfection. Cells were transfected with 500ng transposon and 125ng transposase using TransIT X2 reagent in duplicate per manufacturer's instructions (Mirus Bio). Stable integration of the transgene was assessed by detection of mCherry fluorescence. Cells were harvested at 14 days post-transfection, washed 1X with PBS and resuspended in 200uL RDFII buffer. Cells were analyzed using Novocyte (Acea Biosciences) and mCherry expression was assessed using the PE-Texas red channel.

[00155] *Transfection of TcBuster transposon and transposase in CD3+ T-cells*

[00156] CD3+ T-cells were enriched and cryopreserved from leukopaks (StemCellTechnologies). CD3+ T-cells were thawed and activated using CD3/CD28 Dynabeads (ThermoFisher) for 2 days in X-Vivo -15 media supplemented with human serum and IL-2, IL-15 and IL-7 cytokines. Prior to transfection, CD3/CD28 beads were removed, cells washed and electroporated using Neon Transfection system (ThermoFisher) with TcBuster transposon (mini-circle carrying TcBuster and Sleeping beauty IR/DRs and GFP cargo) and TcBuster or Sleeping Beauty transposases in RNA form. As a viability control, cells were "pulse" electroporated

without DNA or RNA. Electroporated cells were expanded for 21 days post-transfection and viability stable integration of GFP cargo was assessed by flow cytometry. Viability was measured by SSC-A vs FSC-A and standardized to pulse only control, and GFP expression was assessed using FITC channel on days 2, 7, 14 and 21.

[00157] EXAMPLE 2. EXEMPLARY TRANSPOSON CONSTRUCTS

[00158] The aim of this study was to examine transposition efficiency of different exemplary TcBuster transposon constructs. Inventors compared 10 TcBuster (TcB) transposon (Tn) configurations (**Fig. 1A**) to test their transposition efficiency in mammalian cells. These 10 TcB Tns differed in the promoter used (EFS vs CMV), IR/DR sequence and direction of the transposon cargo. The transposons each contained an identical cassette coding for mCherry linked by 2A to a drug-resistance gene, puromycin, so that transfected cells could be identified by fluorescence and/or selection with puromycin. HEK-293T cells were transfected with one of the 10 TcB Tns and TcB wild-type transposase (ratio of 1 transposon: 1 transposase). Stable integration of the transgene was assessed by flow cytometry by detection of mCherry fluorescence for 10-30 days post-transfection (**Fig. 1B**).

[00159] It was found that, under experimental conditions, stable expression of the transgene mCherry was greatly enhanced using the CMV promoter compared to EFS. Transposition appeared to only occur when sequence 1 IR/DRs was used. It was also found that transcription of the cargo in the reverse direction promoted greater transposition activity compared to the forward direction.

[00160] TcB Tn-8 showed the greatest transposition efficiency among the test 10 Tns by flow cytometry. To confirm the transposition efficiency of TcB Tn-8, HEK-293T cells were transfected with TcB Tn-8 with WT transposase or V596A mutant transposase. Two days post-transfection, cells were re-plated with puromycin (1ug/mL) at a density of 3,000 cells/well of a 6 well plate in triplicate in DMEM complete media. After selection for two weeks, each cell that survived drug selection formed a colony, which was assessed for mCherry expression (**Fig. 3A**) and counted to confirm stable integration of the transgene (**Figs. 3B-C**). Transposition efficiency of TcB-Tn 8 was confirmed by expression of mCherry and puromycin resistant colonies in HEK-293T cells.

[00161] EXAMPLE 3. EXEMPLARY TRANSPOSASE MUTANTS

[00162] The aim of this study was to generate TcBuster transposase mutants and examine their transposition efficiency.

[00163] To this end, inventors have generated a consensus sequence by comparing cDNA and amino acid sequences of wild-type TcB transposase to other similar transposases. For the

comparison, sleeping beauty was resurrected by the alignment of 13 similar transposases and SPIN by the alignment of SPIN like transposases from 8 separate organisms. SPIN and TcBuster are a part of the abundant hAT family of transposases.

[00164] The hAT transposon family consists of two subfamilies: AC, such as has hobo, hermes, and Tol2, and the Buster subfamily, such as SPIN and TcBuster. Amino acid sequence of TcBuster was aligned to amino acid sequences of both AC and Buster subfamily members to identify key amino acids that are not conserved in TcBuster that may be targets of hyperactive substitutions. Alignment of TcBuster to the AC subfamily members Hermes, Hobo, Tag2, Tam3, Herves, Restless, and Tol2 allowed us to identify amino acids within areas of high conservation that could be substituted in TcBuster (**Fig. 4**). Further, sequence alignment of TcBuster to the Buster subfamily led to a larger number of candidate amino acids that may be substituted (**Fig. 5**). Candidate TcB transposase mutants were generated using oligonucleotides comprising site mutations as listed in **Table 7**. The mutants were then sequence verified, cloned into pCDNA-DEST40 expression vector (**Fig. 6**) and mini-prepped prior to transfection.

Table 7

Amino Acid Substitutions	**Oligo Name**	**Oligonucleotide Sequence (5' -3')**	**SEQ ID NO**
Q82E	TCB Q82E FWD	GATTTGCGAGgAGGTAGTCAAC	14
Q82E	TCB Q82E REV	ACACAAAGTCCGTTGGGC	15
A358E	TCBA358E FWD	CGCGTCTTCGaaTTGCTGTGTGAC	16
A358E	TCBA358E REV	CGCATTCAACGGCCGAGA	17
A358S	TCBA358S FWD	GCGCGTCTTCagTTTGCTGTGTGACG	18
A358S	TCBA358S REV	GCATTCAACGGCCGAGAC	19
A358K	TCBA358K FWD	GCGCGTCTTCaagTTGCTGTGTGACG	20
A358K	TCBA358K REV	GCATTCAACGGCCGAGAC	21
S447E	TCBS447E FWD	CAAGGTAAATgagCGCATTAACAGTATTAAATC	22
S447E	TCBS447E REV	AAGATTGTGCTATTCGGC	23
I452F	TCBI452F FWD	CATTAACAGTtTTAAATCAAAGTTGAAG	24
I452F	TCBI452F REV	CGGCTATTTACCTTGAAG	25
N281E	TCBN281E FWD	CATCCCATGGgaaCTGTGTTACC	26
N281E	TCBN281E	GAGTGCTTTTCGAAATAGG	27

	REV		
I223Q	TCBI223Q FWD	CGGTCTTGCAcagCTGCTTGTGTTTG	28
I223Q	TCBI223Q REV	GCAACATCTGTTGACTCG	29
P510D	TCBP510D FWD	GTATTTTCCAgatACGTGTAATAATATCTCCTG	30
P510D	TCBP510D REV	TCCAGAAAGGTGTTCTTAAG	31
P510N	TCBP510N FWD	GTATTTTCCAaatACGTGTAATAATATCTCC	32
P510N	TCBP510N REV	TCCAGAAAGGTGTTCTTAAG	33
E517R	TCBE517R FWD	CTCCTGGGTGcggAATCCTTTCAATG	34
E517R	TCBE517R REV	ATATTATTACACGTAGGTGG	35
K590T	TCBK590T FWD	GAAATTAGCAcACGAGCTGTC	36
K590T	TCBK590T REV	TGGAAATTCGTCCATCAG	37
N885S	TCBN885S FWD	GCAGGTAGTCagcAATTCCTCAC	38
N885S	TCBN885S REV	TCGCAAATCACACAAAGTC	39
S109D	TCBS109D FOR	TAAAGGCAAGgacGAATACTTCAAAAGAAAATGTAAC	40
S109D	TCBS109D REV	TAAAGGCAAGgacGAATACTTCAAAAGAAAATGTAAC	41
K135E	TCBK135E FWD	GGACGATAACgagAACCTCCTGA	42
K135E	TCBK135E REV	CTTACGTATCGCTCAAAAGTATG	43
D99A	TcB-D99A F	ACGCCATTTGgcaACAAAGCATC	44
D99A	TcB-D99A R	TTCAGTTTGGCCGGGTTA	45
D132A	TcB-D132A-F	ATACGTAAGGgcaGATAACAAGAACC	46
D132A	TcB-D132A-R	CGCTCAAAAGTATGCTTC	47
E159A	TcB-E159A-F	TACCATAGCGgcgAAGTTGATCAAG	48
E159A	TcB-E159A-R	TATGCCTCGCCCTGTTTA	49
D189A	TcB-D189A-F	CCCCCTGTCCgcaACGACTATTTC	50
D189A	TcB-D189A-R	ACGAGATCAACTTTGCTC	51
D227A	TcB-D227A-F	CGAGTCAACAgcaGTTGCCGGTC	52
D227A	TcB-D227A-R	TCCATCTGCAGCGTAAAC	53
E243A	TcB-E243A-F	GTACATACATgcaAGCTCTTTTG	54
E243A	TcB-E243A-R	CTAACAAACACAAGCAGG	55
V377T	TcB-V377T-F	TCATACCGAAacgAGGTGGCTGTC	56

V377T	TcB-V377T-R	AGAAGAAGATTTTTATGCAGG	57
S225W	TcB-S225W-F	GATGGACGAGtggACAGATGTTGC	58
S225W	TcB-S225W-R	TGCAGCGTAAACCCACAT	59
Y155F	TcB-Y155F-F	GGGCGAGGCAtttACCATAGCGG	60
Y155F	TcB-Y155F-R	TGTTTAGCTATTCTCAAACTGACGAGATAAG	61
D132A	TcB-D132A-F	ATACGTAAGGgcaGATAACAAGAACC	62
D132A	TcB-D132A-R	CGCTCAAAAGTATGCTTC	63
E159A	TcB-E159A-F	TACCATAGCGgcgAAGTTGATCAAG	64
E159A	TcB-E159A-R	TATGCCTCGCCCTGTTTA	65
D189A	TcB-D189A-F	CCCCCTGTCCgcaACGACTATTTC	66
D189A	TcB-D189A-R	ACGAGATCAACTTTGCTC	67
D227A	TcB-D227A-F	CGAGTCAACAgcaGTTGCCGGTC	68
D227A	TcB-D227A-R	TCCATCTGCAGCGTAAAC	69
E243A	TcB-E243A-F	GTACATACATgcaAGCTCTTTTG	70
E243A	TcB-E243A-R	CTAACAAACACAAGCAGG	71
V377T	TcB-V377T-F	TCATACCGAAacgAGGTGGCTGTC	72
V377T	TcB-V377T-R	AGAAGAAGATTTTTATGCAGG	73
S224W	TcB-S224W-F	GATGGACGAGtggACAGATGTTGC	74
S224W	TcB-S224W-R	TGCAGCGTAAACCCACAT	75
Y155F	TcB-Y155F-F	GGGCGAGGCAtttACCATAGCGG	76
Y155F	TcB-Y155F-R	TGTTTAGCTATTCTCAAACTGACGAGATAAG	77

[00165] To examine the transposition efficiency of the TcB transposase mutants, HEK-293T cells were transfected with TcB Tn-8 (mCherry-puromycin cassette) with WT transposase or V596A mutant transposase, or the candidate transposase mutants in duplicate. Cells were grown in DMEM complete (without drug selection) and mCherry expression was assessed by flow cytometry on Day 14 post-transfection. Over 20 TcB transposase mutants were identified that had transposition efficiency greater than the wild-type transposase (**Fig. 7**). It was discovered that among these examined mutants, one mutant transposase containing a combination of three amino acid substitutions, D189A, V377T, and E469K, led to a substantial increase in transposition activity, as compared to mutants containing respective single substitutions. Mutants with high transposition activity also included, among others, K573E/E578L, I452F, A358K, V297K, N85S, S447E, E247K, and Q258T.

[00166] Among these examined mutants, it was discovered that most of substitutions to a positively charged amino acid, such as Lysine (K) or Arginine (R), in proximity to one of the

catalytic triad amino acids (D234, D289, and E589) increased transposition. In addition, removal of a positive charge, or addition of a negative charge decreased transposition. These data suggests that amino acids close to the catalytic domain may help promote the transposition activity of TcB, in particular, when these amino acids are mutated to positively charged amino acids.

[00167] The amino acid sequence of the hyperactive TcBuster mutant D189A/V377T/E469K (SEQ ID NO: 78) is illustrated in **Fig. 12**. Further mutational analysis of this mutant will be performed. As illustrated in **Fig. 13**, the TcBuster mutant D189A/V377T/E469K/I452F (SEQ ID NO: 79) will be constructed. As illustrated in **Fig. 14**, the TcBuster mutant D189A/V377T/E469K/N85S (SEQ ID NO: 80) will be constructed. As illustrated in **Fig. 15**, the Tc Buster mutant D189A/V377T/E469K/S358K (SEQ ID NO: 81) will be constructed. As illustrated in **Fig. 16**, the Tc Buster mutant D189A/V377T/E469K/K573E/E578L (SEQ ID NO: 13) will be constructed. In each of **Figs. 12-16**, the domains of TcBuster are indicated as follows: ZnF-BED (lowercase lettering), DNA Binding/oligomization domain (bold lettering), catalytic domain (underlined lettering), and insertion domain (italicized lettering); the core D189A/V377T/E469K substitutions are indicated in larger, bold, italicized, and underlined letters; and the additional substitutions are indicated in large, bold letters. Each of these constructs will be tested as already described and are anticipated to show hyperactivity in comparison to the wild type TcBuster.

[00168] **EXAMPLE 4. EXEMPLARY FUSION TRANSPOSASE CONTAINING TAG**

[00169] The aim of this study was to generate and examine the transposition efficiency of fusion TcBuster transposases. As an example, protein tag, GST or PEST domain, was fused to N-terminus of TcBuster transposase to generate fusion TcBuster transposases. A flexible linker GGSGGSGGSGGSGTS (SEQ ID NO: 9), which was encoded by SEQ ID NO: 10, was used to separate the GST domain / PEST domain from TcBuster transposase. The presence of this flexibility linker may minimize non-specific interaction in the fusion protein, thus increasing its activity. The exemplary fusion transposases were transfected with TcB Tn- 8 as described above and transposition efficiency was measured by mCherry expression on Day 14 by flow cytometry. Transposition efficiency was not affected by tagging of GFP or PEST domain (**Fig. 9**), suggesting that fusing the transposase DNA binding domains to direct integration of TcBuster cargo to select genomic sites, such as safe harbor sites, could be a viable option for TcBuster allowing for a safer integration profile.

[00170] EXAMPLE 5. EXEMPLARY FUSION TRANSPOSASE COMPRISING TALE DOMAIN

[00171] The aim of this study is to generate a fusion TcBuster transposase comprising a TALE domain and to examine the transposition activity of the fusion transposase. A TALE sequence (SEQ ID NO: 11) is designed to target human AAVS1 (hAAVS1) site of human genome. The TALE sequence is thus fused to N-terminus of a wild-type TcBuster transposase (SEQ ID NO: 1) to generate a fusion transposase. A flexible linker Gly4Ser2, which is encoded by SEQ ID NO: 12, is used to separate the TALE domain and the TcBuster transposase sequence. The exemplary fusion transposase has an amino acid sequence SEQ ID NO: 8.

[00172] The exemplary fusion transposase will be transfected with a TcB Tn-8 as described above into Hela cells with the aid of electroporation. The TcB Tn-8 comprises a reporter gene mCherry. The transfection efficiency can be examined by flow cytometry 2 days post-transfection that counts mCherry-positive cells. Furthermore, next-generation sequencing will be performed to assess the mCherry gene insertion site in the genome. It is expected that the designed TALE sequence can mediate the target insertion of the mCherry gene at a genomic site near hAAVS1 site.

[00173] EXAMPLE 6. TRANSPOSITION EFFICIENCY IN PRIMARY HUMAN T-CELLS

[00174] The aim of this study was to develop TcBuster transposon system to engineer primary CD3+ T cells. To this end, inventors incorporated an exemplary TcBuster transposon carrying a GFP transgene into a mini-circle plasmid. Activated CD3+ T cells were electroporated with TcB mini-circle transposon and RNA transposases, such as WT TcBuster transposase, and select exemplary mutants as described in Example 2. The transgene expression was monitored for 21 days post-electroporation by flow cytometry.

[00175] It was found that transposition of the TcB transposon was improved nearly two folds using the exemplary mutants, V377T/E469K and V377T/E469K/D189A, 14 days post-transfection compared to the WT TcBuster transposase and V596A mutant transposase (**Fig. 10A**). Further, mean transposition efficiency with the hyperactive mutants V377T/E469K and V377T/E469K/D189A was two (mean = 20.2) and three (mean = 24.1) times more efficient compared to SB11 (mean = 8.4), respectively.

[00176] Next, the viability of CD3+ T cells was assessed two days post- electroporation with the mini-circle TcB transposon and RNA transposase. It was found that viability was moderately decreased when CD3+ T-cells were transfected with TcB mini-circle and RNA transposase; however, the cells quickly recovered viability by Day 7 (**Fig. 10B**). These experiments

demonstrate the competency of the TcBuster transposon system, according to some embodiments of the present disclosure, in cellular engineering of primary T cells.

[00177] **EXAMPLE 7. GENERATION OF CHIMERIC ANTIGEN RECEPTOR-MODIFIED T CELLS FOR TREATMENT OF CANCER PATIENT**

[00178] A mini-circle plasmid containing aforementioned TcB Tn-8 construct can be designed to harbor a chimeric antigen receptor (CAR) gene between the inverted repeats of the transposon. The CAR can be designed to have specificity for the B-cell antigen CD19, coupled with CD137 (a costimulatory receptor in T cells [4-1BB]) and CD3-zeta (a signal-transduction component of the T-cell antigen receptor) signaling domains.

[00179] Autologous T cells will be obtained from peripheral blood of a patient with cancer, for example, leukemia. The T cells can be isolated by lysing the red blood cells and depleting the monocytes by centrifugation through a PERCOLL™ gradient. CD3+ T cells can be isolated by flow cytometry using anti-CD3/anti-CD28-conjugated beads, such as DYNABEAD M-450 CD3/CD28T. The isolated T cells will be cultured under standard conditions according to GMP guidance.

[00180] Genetic modification of the primary T cells will be conducted using a mutant TcBuster transposase (SEQ ID NO: 13) comprising amino acid substitutions V377T, E469K, D189A, K573E and E578L and the TcBuster Tn-8 transposase comprising the CAR, as described above. The T cells will be electroporated in the presence of the mutant TcBuster transposase and the CAR-containing Tn-8 transposase. Following transfection, T cells will be treated with immunostimulatory reagents (such as anti-CD3 antibody and IL-2, IL-7, and IL-15) for activation and expansion. Validation of the transfection will be performed by next-generation sequencing 2 weeks post-transfection. The transfection efficiency and transgene load in the transfected T cells can be determined to assist the design of treatment regimen. Certain measure will also be taken to eliminate any safety concern if risky transgene insertion site is uncovered by the sequencing results.

[00181] Infusion of the chimeric antigen receptor modified T cells (CAR-T cells) back to the cancer patient will start after validation of transgene insertion and *in vitro* expansion of the CAR-T cells to a clinically desirable level.

[00182] The infusion dose will be determined by a number of factors, including, but not limited to, the stage of the cancer, the treatment history of the patient, and the CBC (complete blood cell count) and vital signs of the patient on the day of treatment. Infusion dose may be escalated or deescalated depending on the progression of the disease, the repulsion reaction of the patient, and many other medical factors. In the meantime, during the treatment regimen, quantitative

polymerase-chain-reaction (qPCR) analysis will be performed to detect chimeric antigen receptor T cells in blood and bone marrow. The qPCR analysis can be utilized to make medical decision regarding the dosing strategy and other treatment plans.

Table 8 Amino Acid and Nucleotide Sequences

Sequence Description	**Amino Acid Sequence Or Nucleotide Sequence (SEQ ID NO)**
Wild-type TcBuster transposase	(accession number: ABF20545) MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTSRKKRKYDED YLNFGFTWTGDKDEPNGLCVICEQVVNNSSLNPAKLKRHLDTKHPTLKGKSEYFKRKC NELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGEAYTIAEKLIKPCTKDLTTCVF GEKFASKVDLVPLSDTTISRRIEDMSYFCEAVLVNRLKNAKCGFTLQMDESTDVAGLA ILLVFVRYIHESSFEEDMLFCKALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDG AKAMVGVIKGVIARIKKLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIK SRPLNARVFALLCDDLGSLHKNLLLHTEVRWLSRGKVLTRFWELRDEIRIFFNEREFA GKLNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKLWEECIT KNNTECFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPPTCNNISWVENPF NECGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPFWIKLMDEFPEISKRAVKELM PFVTTYLCEKSFSVYVATKTKYRNRLDAEDDMRLQLTTIHPDIDNLCNNKQAQKSH (SEQ ID NO: 1)
Wild-type TcBuster transposase	atgatgttgaattggctgaaaagtggaaagcttgaaagtcaatcacaggaacagagtt cctgctaccttgagaactctaactgcctgccaccaacgctcgattctacagatattat cggtgaagagaacaaagctggtaccacctctcgcaagaagcggaaatatgacgaggac tatctgaacttcggttttacatggactggcgacaaggatgagcccaacggactttgtg tgatttgcgagcaggtagtcaacaattcctcacttaacccggccaaactgaaacgcca tttggacacaaagcatccgacgcttaaaggcaagagcgaatacttcaaaagaaaatgt aacgagctcaatcaaaagaagcatacttttgagcgatacgtaagggacgataacaaga acctcctgaaagcttcttatctcgtcagtttgagaatagctaaacagggcgaggcata taccatagcggagaagttgatcaagccttgcaccaaggatctgacaacttgcgtattt ggagaaaaattcgcgagcaaagttgatctcgtcccccctgtccgacacgactatttcaa ggcgaatcgaagacatgagttacttctgtgaagccgtgctggtgaacaggttgaaaaa tgctaaatgtgggtttacgctgcagatggacgagtcaacagatgttgccggtcttgca atcctgcttgtgtttgttaggtacatacatgaaagctcttttgaggaggatatgttgt tctgcaaagcacttcccactcagacgacaggggaggagattttcaatcttctcaatgc ctatttcgaaaagcactccatcccatggaatctgtgttaccacatttgcacagacggt gccaaggcaatggtaggagttattaaaggagtcatagcgagaataaaaaaactcgtcc ctgatataaaagctagccactgttgcctgcatcgccacgctttggctgtaaagcgaat accgaatgcattgcacgaggtgctcaatgacgctgttaaaatgatcaacttcatcaag tctcggccgttgaatgcgcgctcttcgctttgctgtgtgacgatttggggagcctgc ataaaaatcttcttcttcataccgaagtgaggtggctgtctagaggaaaggtgctgac ccgattttgggaactgagagatgaaattagaattttcttcaacgaaagggaatttgcc

	gggaaattgaacgacaccagttggttgcaaaatttggcatatatagctgacatattca gttatctgaatgaagttaatctttccctgcaagggccgaatagcacaatcttcaaggt aaatagccgcattaacagtattaaatcaaagttgaagttgtgggaagagtgtataacg aaaaataacactgagtgttttgcgaacctcaacgatttttttggaaacttcaaacactg cgttggatccaaacctgaagtctaatattttggaacatctcaacggtcttaagaacac ctttctggagtattttccacctacgtgtaataatatctcctgggtggagaatcctttc aatgaatgcggtaacgtcgatacactcccaataaaagagagggaacaattgattgaca tacggactgatacgacattgaaatcttcattcgtgcctgatggtataggaccattctg gatcaaactgatggacgaatttccagaaattagcaaacgagctgtcaaagagctcatg ccatttgtaaccacttacctctgtgagaaatcattttccgtctatgtagccacaaaaa caaaatatcgaaatagacttgatgctgaagacgatatgcgactccaacttactactat ccatccagacattgacaacctttgtaacaacaagcaggctcagaaatcccactga (SEQ ID NO: 2)
IRDR-L-Seq1	Cagtgttcttcaacctttgccatccggcggaaccctttgtcgagatatttttttttat ggaaccttcatttagtaatacacccagatgagattttagggacagctgcgttgactt gttacgaacaaggtgagcccgtgctttggtctagccaagggcatggtaaagactatat tcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggctt gaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccg tatgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacg tagatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcgg tggcaatgccctgcctccggtgctcgccggagactgcgagatcatagatata (SEQ ID NO: 3)
IRDR-R-Seq1	>gatatcaagcttatcgataccgtcgacctcgagatttctgaacgattctaggttagg atcaaacaaaatacaatttatttttaaaactgtaagttaacttacctttgcttgtctaa accaaaaacaacaacaaaactacgaccacaagtacagttacatatttttgaaaattaa ggttaagtgcagtgtaagtcaactatgcgaatggataacatgtttcaacatgaaactc cgattgacgcatgtgcattctgaagagcggcgcggccgacgtctctcgaattgaagca atgactcgcggaaccccgaaagcctttgggtggaaccctagggttccgcggaacacag gttgaagaacactg (SEQ ID NO: 4)
IRDR-L-Seq2	Cctgcaggagtgttcttcaacctttgccatccggcggaaccctttgtcgagatatttt tttttatggaacccttcatttagtaatacacccagatgagattttagggacagctgcg ttgacttgttacgaacaaggtgagcccgtgctttggtaataaaaactctaaataagat ttaaatttgcatttatttaaacaaactttaaacaaaaagataaatattccaaataaaa taatatataaaataaaaaataaaaattaatgacttttttgcgcttgcttattattgca caaattatcaatatcgggatggatcgttgtttttt (SEQ ID NO: 5)
IRDR-R-Seq2	Gagccaattcagcatcatatttctgaacgattctaggttaggatcaaacaaaatacaa tttatttaaaactgtaagttaacttacctttgcttgtctaaacctaaaacaacaaca aaactacgaccacaagtacagttacatatttttgaaaattaaggttaagtgcagtgta agtcaactatgcgaatggataacatgtttcaacatgaaactccgattgacgcatgtgc

	attctgaagagcggcgcggccgacgtctctcgaattgaagcaatgactcgcggaaccc cgaaagcctttgggtggaaccctagggttccgcggaacacaggttgaagaacactg (SEQ ID NO: 6)
pcDNA-DEST 40	gacggatcgggagatctcccgatcccctatggtgcactctcagtacaatctgctctga tgccgcatagttaagccagtatctgctccctgcttgtgtgttggaggtcgctgagtag tgcgcgagcaaaatttaagctacaacaaggcaaggcttgaccgacaattgcatgaaga atctgcttagggttaggcgttttgcgctgcttcgcgatgtacgggccagatatacgcg ttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcat agcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgac cgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgcc aatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttg gcagtacatcaagtgtatcatatgccaagtacgcccccctattgacgtcaatgacggta aatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggca gtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatc aatgggcgtggatagcggtttgactcacggggatttccaagtctccaccccattgacg tcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaa ctccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagc agagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacga ctcactatagggagacccaagctggctagttaagctatcaacaagtttgtacaaaaaa gctgaacgagaaacgtaaaatgatataaatatcaatatattaaattagattttgcata aaaaacagactacataatactgtaaaacacaacatatccagtcactatggcggccgca ttaggcaccccaggctttacactttatgcttccggctcgtataatgtgtggattttga gttaggatccggcgagattttcaggagctaaggaagctaaaatggagaaaaaaatcac tggatataccaccgttgatatatcccaatggcatcgtaaagaacattttgaggcattt cagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttt taaagaccgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgc ccgcctgatgaatgctcatccggaattccgtatggcaatgaaagacggtgagctggtg atatgggatagtgttcacccttgttacaccgttttccatgagcaaactgaaacgtttt catcgctctggagtgaataccacgacgatttccggcagtttctacacatatattcgca agatgtggcgtgttacggtgaaaacctggcctatttccctaaagggtttattgagaat atgtttttcgtctcagccaatccctgggtgagtttcaccagttttgatttaaacgtgg ccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacgcaagg cgacaaggtgctgatgccgctggcgattcaggttcatcatgccgtctgtgatggcttc catgtcggcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcgggg cgtaaagatctggatccggcttactaaaagccagataacagtatgcgtatttgcgcgc tgatttttgcggtataagaatatatactgatatgtatacccgaagtatgtcaaaaaga ggtgtgctatgaagcagcgtattacagtgacagttgacagcgacagctatcagttgct caaggcatatatgatgtcaatatctccggtctggtaagcacaaccatgcagaatgaag cccgtcgtctgcgtgccgaacgctggaaagcggaaaatcaggaagggatggctgaggt cgcccggtttattgaaatgaacggctcttttgctgacgagaacagggactggtgaaat gcagtttaaggtttacacctataaaagagagagccgttatcgtctgtttgtggatgta

	cagagtgatattattgacacgcccgggcgacggatggtgatccccctggccagtgcac
	gtctgctgtcagataaagtctcccgtgaactttacccggtggtgcatatcggggatga
	aagctggcgcatgatgaccaccgatatggccagtgtgccggtctccgttatcggggaa
	gaagtggctgatctcagccaccgcgaaaatgacatcaaaaacgccattaacctgatgt
	tctggggaatataaatgtcaggctccgttatacacagccagtctgcaggtcgaccata
	gtgactggatatgttgtgttttacagtattatgtagtctgtttttttatgcaaaatcta
	atttaatatattgatatttatatcattttacgtttctcgttcagctttcttgtacaaa
	gtggttgatctagagggcccgcggttcgaaggtaagcctatccctaaccctctcctcg
	gtctcgattctacgcgtaccggtcatcatcaccatcaccattgagtttaaacccgctg
	atcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtg
	ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaa
	ttgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcagga
	cagcaaggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctct
	atggcttctgaggcggaaagaaccagctggggctctagggggtatccccacgcgccct
	gtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacact
	tgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttc
	gccggctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtg
	ctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggcc
	atcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagt
	ggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatt
	tataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaa
	atttaacgcgaattaattctgtggaatgtgtgtcagttagggtgtggaaagtccccag
	gctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccaggtg
	tggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattag
	tcagcaaccatagtcccgcccctaactccgcccatcccgcccctaactccgcccagtt
	ccgcccattctccgccccatggctgactaatttttttttatttatgcagaggccgaggc
	cgcctctgcctctgagctattccagaagtagtgaggaggcttttttggaggcctaggc
	ttttgcaaaaagctcccgggagcttgtatatccattttcggatctgatcaagagacag
	gatgaggatcgtttcgcatgattgaacaagatggattgcacgcaggttctccggccgc
	ttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgat
	gccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacc
	tgtccggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccac
	gacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcgggaagggactgg
	ctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccg
	agaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctac
	ctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaa
	gccggtcttgtcgatcaggatgatctggacgaagagcatcaggggctcgcgccagccg
	aactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgaccca
	tggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatc
	gactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctacccgtg
	atattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgctttacggtat

	cgccgctcccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctga
	gcgggactctggggttcgcgaaatgaccgaccaagcgacgcccaacctgccatcacga
	gatttcgattccaccgccgccttctatgaaaggttgggcttcggaatcgttttccggg
	acgccggctggatgatcctccagcgcggggatctcatgctggagttcttcgcccaccc
	caacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttc
	acaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatg
	tatcttatcatgtctgtataccgtcgacctctagctagagcttggcgtaatcatggtc
	atagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagcc
	ggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattg
	cgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatg
	aatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg
	ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa
	aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagc
	aaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccat
	aggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaa
	acccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctc
	tcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagc
	gtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgct
	ccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccgg
	taactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagcc
	actggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagt
	ggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaa
	gccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgct
	ggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctc
	aagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacg
	ttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaat
	taaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagtt
	accaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccat
	agttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggc
	cccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaa
	taaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctc
	catccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagt
	ttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggta
	tggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgtt
	gtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggcc
	gcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccat
	ccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtg
	tatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacat
	agcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaa
	ggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatc
	ttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaat

	gccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttt ttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttga atgtatttagaaaaataaacaaatagggqttccgcgcacatttccccgaaaagtgcca cctgacgtc (SEQ ID NO: 7)
Fusion Transposase containing wild-type TcBuster sequence and TALE DNA-binding domain targeting human AAVS1	atgctcgagatggatccctccgacgcttcgccggccgcgcaggtggatctacgcacgc tcggctacagtcagcagcagcaagagaagatcaaaccgaaggtgcgttcgacagtggc gcagcaccacgaggcactggtgggccatgggtttacacacgcgcacatcgttgcgctc agccaacacccggcagcgttagggaccgtcgctgtcacgtatcagcacataatcacgg cgttgccagaggcgacacacgaagacatcgttggcgtcggcaaacagtggtccggcgc acgcgccctggaggccttgttgactgatgctggtgagcttagaggacctcctttgcaa cttgatacaggccagcttctgaaaatcgccaagaggggtggggtcaccgcggtcgagg ccgtacacgcctggagaaatgcactgaccggggctcctcttaacCTGACCCCAGACCA GGTAGTCGCAATCGCGTCAAACGGAGGGGGAAAGCAAGCCCTGGAAACCGTGCAAAGG TTGTTGCCGGTCCTTTGTCAAGACCACGGCCTTACACCGGAGCAAGTCGTGGCCATTG CATCCCACGACGGTGGCAAACAGGCTCTTGAGACGGTTCAGAGACTTCTCCCAGTTCT CTGTCAAGCCCACGGGCTGACTCCCGATCAAGTTGTAGCGATTGCGTCGCATGACGGA GGGAAACAAGCATTGGAGACTGTCCAACGGCTCCTTCCCGTGTTGTGTCAAGCCCACG GTTTGACGCCTGCACAAGTGGTCGCCATCGCCTCCAATATTGGCGGTAAGCAGGCGCT GGAAACAGTACAGCGCCTGCTGCCTGTACTGTGCCAGGATCATGGACTGACGGCCAAG CTGGCCGGGGGCGCCCCCGCCGTGGGCGGGGGCCCCAAGGCCGCCGATAAATTCGCCG CCACCatgatgttgaattggctgaaaagtggaaagcttgaaagtcaatcacaggaaca gagttcctgctaccttgagaactctaactgcctgccaccaacgctcgattctacagat attatcggtgaagagaacaaagctggtaccacctctcgcaagaagcggaaatatgacg aggactatctgaacttcggttttacatggactggcgacaaggatgagcccaacggact ttgtgtgatttgcgagcaggtagtcaacaattcctcacttaacccggccaaactgaaa cgccatttggacacaaagcatccgacgcttaaaggcaagagcgaatacttcaaaagaa aatgtaacgagctcaatcaaaagaagcatacttttgagcgatacgtaagggacgataa caagaacctcctgaaagcttcttatctcgtcagtttgagaatagctaaacagggcgag gcatataccatagcggagaagttgatcaagccttgcaccaaggatctgacaacttgcg tatttggagaaaaattcgcgagcaaagttgatctcgtcccctgtccgacacgactat ttcaaggcgaatcgaagacatgagttacttctgtgaagccgtgctggtgaacaggttg aaaaatgctaaatgtgggtttacgctgcagatggacgagtcaacagatgttgccggtc ttgcaatcctgcttgtgtttgttaggtacatacatgaaagctcttttgaggaggatat gttgttctgcaaagcacttcccactcagacgacagggaggagattttcaatcttctc aatgcctatttcgaaaagcactccatcccatggaatctgtgttaccacatttgcacag acggtgccaaggcaatggtaggagttattaaaggagtcatagcgagaataaaaaaact cgtccctgatataaaagctagccactgttgcctgcatcgccacgctttggctgtaaag cgaataccgaatgcattgcacgaggtgctcaatgacgctgttaaaatgatcaacttca tcaagtctcggccgttgaatgcgcgcgtcttcgctttgctgtgtgacgatttggggag cctgcataaaaatcttcttcttcataccgaagtgaggtggctgtctagaggaaaggtg

	ctgacccgattttgggaactgagagatgaaattagaattttcttcaacgaaagggaat ttgccgggaaattgaacgacaccagttggttgcaaaatttggcatatatagctgacat attcagttatctgaatgaagttaatctttcctgcaagggccgaatagcacaatcttc aaggtaaatagccgcattaacagtattaaatcaaagttgaagttgtgggaagagtgta taacgaaaaataacactgagtgttttgcgaacctcaacgatttttggaaacttcaaa cactgcgttggatccaaacctgaagtctaatattttggaacatctcaacggtcttaag aacacctttctggagtattttccacctacgtgtaataatatctcctgggtggagaatc ctttcaatgaatgcggtaacgtcgatacactcccaataaaagagagggaacaattgat tgacatacggactgatacgacattgaaatcttcattcgtgcctgatggtataggacca ttctggatcaaactgatggacgaatttccagaaattagcaaacgagctgtcaaagagc tcatgccatttgtaaccacttacctctgtgagaaatcattttccgtctatgtagccac aaaaacaaaatatcgaaatagacttgatgctgaagacgatatgcgactccaacttact actatccatccagacattgacaacctttgtaacaacaagcaggctcagaaatcccact ga (SEQ ID NO: 8)
Flexible linker (Example 4)	GGSGGSGGSGGSGTS (SEQ ID NO: 9)
Flexible linker (Example 4)	GGAGGTAGTGGCGGTAGTGGGGGCTCCGGTGGGAGCGGCACCTCA (SEQ ID NO: 10)
TALE domain targeting hAAVS1 site (Example 5)	atgctcgagatggatccctccgacgcttcgccggccgcgcaggtggatctacgcacgc tcggctacagtcagcagcagcaagagaagatcaaaccgaaggtgcgttcgacagtggc gcagcaccacgaggcactggtgggccatgggtttacacacgcgcacatcgttgcgctc agccaacacccggcagcgttagggaccgtcgctgtcacgtatcagcacataatcacgg cgttgccagaggcgacacacgaagacatcgttggcgtcggcaaacagtggtccggcgc acgcgccctggaggccttgttgactgatgctggtgagcttagaggacctcctttgcaa cttgatacaggccagcttctgaaaatcgccaagaggggtggggtcaccgcggtcgagg ccgtacacgcctggagaaatgcactgaccggggctcctcttaacCTGACCCCAGACCA GGTAGTCGCAATCGCGTCAAACGGAGGGGGAAAGCAAGCCCTGGAAACCGTGCAAAGG TTGTTGCCGGTCCTTTGTCAAGACCACGGCCTTACACCGGAGCAAGTCGTGGCCATTG CATCCCACGACGGTGGCAAACAGGCTCTTGAGACGGTTCAGAGACTTCTCCCAGTTCT CTGTCAAGCCCACGGGCTGACTCCCGATCAAGTTGTAGCGATTGCGTCGCATGACGGA GGGAAACAAGCATTGGAGACTGTCCAACGGCTCCTTCCCGTGTTGTGTCAAGCCCACG GTTTGACGCCTGCACAAGTGGTCGCCATCGCCTCCAATATTGGCGGTAAGCAGGCGCT GGAAACAGTACAGCGCCTGCTGCCTGTACTGTGCCAGGATCATGGACTGAC (SEQ ID NO: 11)
Flexible linker (Example 5)	GGCCAAGCTGGCCGGGGGCGCCCCCGCCGTGGGCGGGGGCCCCAAGGCCGCCGATAAA TTCGCCGCCACC (SEQ ID NO: 12)
Mutant TcBuster transposase containing V377T,	MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTSRKKRKYDED YLNFGFTWTGDKDEPNGLCVICEQVVNNSSLNPAKLKRHLDTKHPTLKGKSEYFKRKC NELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGEAYTIAEKLIKPCTKDLTTCVF

E469K, D189A, K573E and E578L	GEKFASKVDLVPLSATTISRRIEDMSYFCEAVLVNRLKNAKCGFTLQMDESTDVAGLA ILLVFVRYIHESSFEEDMLFCKALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDG AKAMVGVIKGVIARIKKLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIK SRPLNARVFALLCDDLGSLHKNLLLHTETRWLSRGKVLTRFWELRDEIRIFFNEREFA GKLNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKLWEECIT KNNTKCFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPPTCNNISWVENPF NECGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPFWIKLMDEFPEISERAVKLLM PFVTTYLCEKSFSVYVATKTKYRNRLDAEDDMRLQLTTIHPDIDNLCNNKQAQKSH (SEQ ID NO: 13)

[00183] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A mutant TcBuster transposase comprising an amino acid sequence having at least 80% identical to full-length SEQ ID NO: 1 and at least one amino acid substitution selected from Q82E, N85S, D99A, D132A, Q151S, Q151A, E153K, E153R, A154P, Y155H, E159A, T171K, T171R, K177E, D183K, D183R, D189A, T191E, S193K, S193R, Y201A, F202D, F202K, C2031, C203V, Q221T, M222L, I223Q, E224G, S225W, D227A, R239H, E243A, E247K, P257K, P257R, Q258T, E263A, E263K, E263R, E274K, E274R, S278K, N281E, L282K, L282R, K292P, V297K, K2995, A303T, H322E, A332S, A358E, A358K, A358S, D376A, V377T, L380N, I398D, I398S, I398K, F400L, V431L, S447E, N450K, N450R, I452F, E469K, P510D, P510N, E517R, R536S, V553S, P554T, P559D, P559S, P559K, K573E, E578L, K590T, Y595L, T598I, K599A, Q615A, T618K, T618R, D622K, and D622R, or any combination thereof, when numbered in accordance with SEQ ID NO: 1, wherein the mutant TcBuster transposase has increased transposition efficiency in comparison to a wild-type TcBuster transposase having amino acid sequence SEQ ID NO: 1.

2. The mutant TcBuster transposase of claim 1, wherein the amino acid sequence of the mutant TcBuster transposase is at least 90%, at least 95%, at least 98%, or at least 99% identical to full-length SEQ ID NO: 1.

3. The mutant TcBuster transposase of claim 1 or claim 2, comprising amino acid substitution N85S.

4. The mutant TcBuster transposase of any one of claims 1-3, comprising amino acid substitution D99A.

5. The mutant TcBuster transposase of any one of claims 1 to 4, comprising amino acid substitution E247K.

6. The mutant TcBuster transposase of any one of claims 1 to 5, comprising amino acid substitution V377T.

7. The mutant TcBuster transposase of any one of claims 1 to 6, comprising amino acid substitution E469K.

8. The mutant TcBuster transposase of any one of claims 1 to 7, comprising at least two amino acid substitutions selected from N85S, D99A, E247K, V377T, and E469K.

9. The mutant TcBuster transposase of any one of claims 1 to 7, comprising at least three amino acid substitutions selected from N85S, D99A, E247K, V377T, and E469K.

10. The mutant TcBuster transposase of any one of claims 1 to 7, comprising at least four amino acid substitutions selected from N85S, D99A, E247K, V377T, and E469K.

11. The mutant TcBuster transposase of any one of claims 1 to 7, comprising at least five amino acid substitutions selected from N85S, D99A, E247K, V377T, and E469K.

12. The mutant TcBuster transposase of any one of claims 1 to 11, wherein the transposition efficiency is measured by an assay that comprises introducing the mutant TcBuster transposase and a TcBuster transposon containing a reporter cargo cassette into a population of cells, and detecting transposition of the reporter cargo cassette in genome of the population of cells.

13. A method of treatment, wherein the method comprises:

(a) introducing into a cell a transposon and the mutant TcBuster transposase of any one of claims 1-11, which recognizes the transposon, thereby generating a genetically modified cell; and

(b) administering the genetically modified cell to a patient in need of the treatment.

14. The method of claim 13, wherein the genetically modified cell comprises a transgene introduced by the transposon.

15. The method of claim 13 or claim 14, wherein the patient has been diagnosed with cancer or tumor.

16. The method of any one of claims 13 to 15, wherein the administering comprises transfusing the genetically modified cell into blood vessels of the patient.

17. A system for genome editing, wherein the system comprises the mutant TcBuster transposase of any one of claims 1-11, and a transposon recognizable by the mutant TcBuster transposase.

18. A system for genome editing, wherein the system comprises a polynucleotide encoding a mutant TcBuster transposase of any one of claims 1-11, and a transposon recognizable by the mutant TcBuster transposase

19. The system of claim 18, wherein the polynucleotide comprises DNA that encodes the mutant TcBuster transposase.

20. The system of claim 18 or claim 19, wherein the polynucleotide comprises messenger RNA (mRNA) that encodes the mutant TcBuster transposase.

21. The system of claim 20, wherein the mRNA is chemically modified.

22. The system of any one of claims 18 to 21, wherein the transposon is present in a DNA vector.

23. The system of claim 22, wherein the DNA vector comprises a mini-circle plasmid.

24. The system of any one of claims 18 to 23, wherein the polynucleotide and the transposon are present in the same plasmid.

25. The system of any one of claims 18 to 24, wherein the transposon comprises a cargo cassette positioned between two inverted repeats.

26. The system of claim 25, wherein a left inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 3.

27. The system of claim 25 or claim 26, wherein a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 3.

28. The system of any one of claims 25 to 27, wherein a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 4.

29. The system of any one of claims 25 to 28, wherein a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 4.

30. The system of claim 25, wherein a left inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 5.

31. The system of claim 25 or claim 30, wherein a left inverted repeat of the two inverted repeats comprises SEQ ID NO: 5.

32. The system of any one of claims 25, 30, or 31, wherein a right inverted repeat of the two inverted repeats comprises a sequence having at least 50%, at least 60%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 6.

33. The system of any one of claims 25, 30 or 31, wherein a right inverted repeat of the two inverted repeats comprises SEQ ID NO: 6.

34. The system of any one of claims 25 to 33, wherein the cargo cassette comprises a promoter selected from the group consisting of: CMV, EFS, MND, EF1α, CAGCs, PGK, UBC, U6, H1, and Cumate.

35. The system of any one of claims 25 to 33, wherein the cargo cassette comprises a CMV promoter.

36. The system of any one of claims 25 to 35, wherein the cargo cassette comprises a transgene.

37. The system of claim 36, wherein the transgene codes for a protein selected from the group consisting of: a cellular receptor, an immunological checkpoint protein, a cytokine, and any combination thereof.

38. The system of claim 36 or 37, wherein the transgene codes for a cellular receptor selected from the group consisting of: a T cell receptor (TCR), a B cell receptor (BCR), a chimeric antigen receptor (CAR), or any combination thereof.

39. The system of any one of claims 25 to 38, wherein the cargo cassette is present in a forward direction.

40. The system of any one of claims 25 to 38, wherein the cargo cassette is present in a reverse direction.

41. Use of the mutant TcBuster transposase of any one of claims 1-11 in the manufacture of a medicament for treating a patient in need thereof, wherein the treatment comprises:

(a) introducing into a cell a transposon and the mutant TcBuster transposase, which recognizes the transposon, thereby generating a genetically modified cell; and

(b) administering the genetically modified cell to the patient.

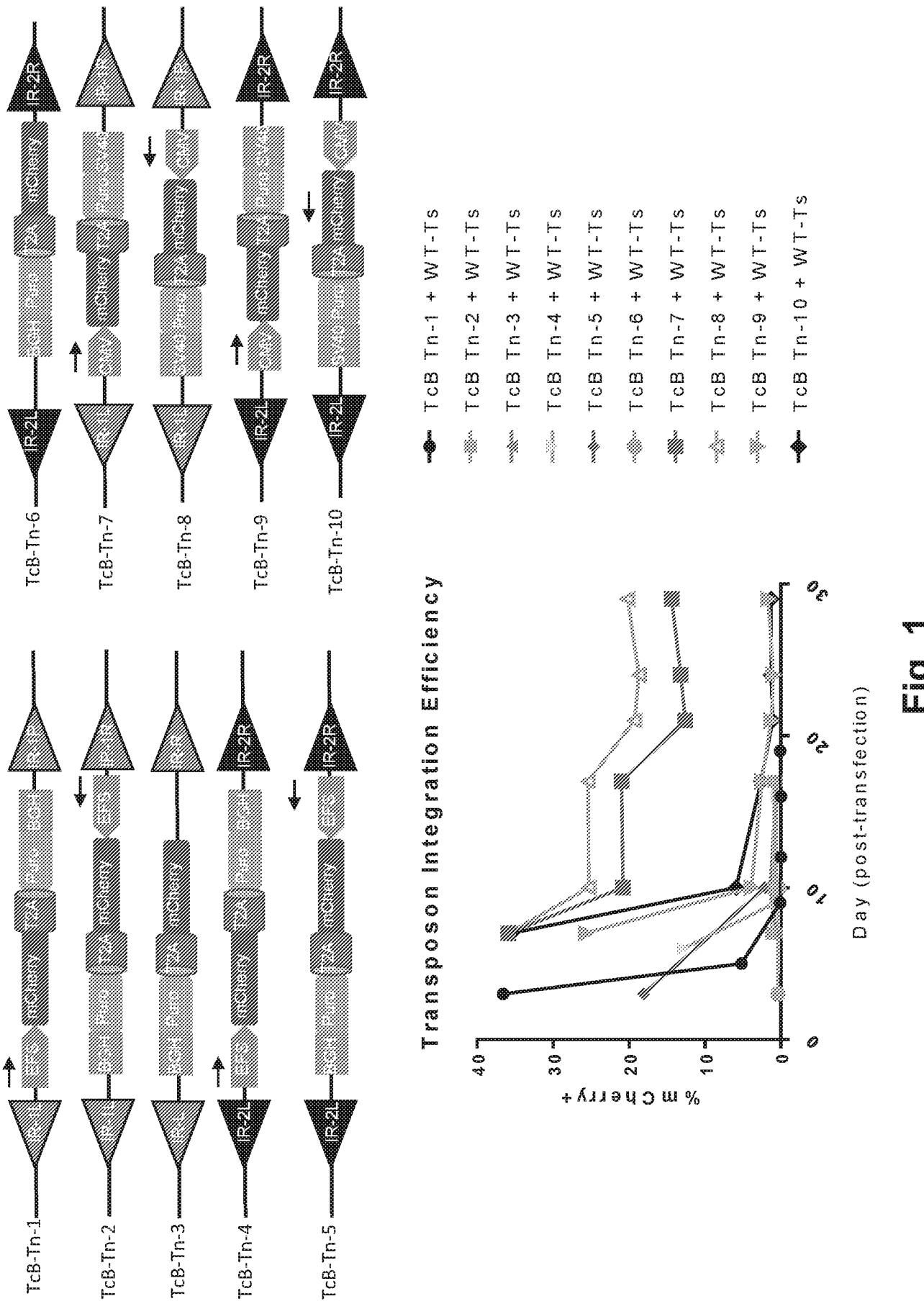
TcB-Tn-1
TcB-Tn-2
TcB-Tn-3
TcB-Tn-4
TcB-Tn-5
TcB-Tn-6
TcB-Tn-7
TcB-Tn-8
TcB-Tn-9
TcB-Tn-10
IR-1L
IR-1R
IR-2L
IR-2R
EFS
mCherry
T2A
PuroR
BsdR
GAV
Transposon Integration Efficiency
% mCherry+
Day (post-transfection)
TcB Tn-1 + WT-Ts
TcB Tn-2 + WT-Ts
TcB Tn-3 + WT-Ts
TcB Tn-4 + WT-Ts
TcB Tn-5 + WT-Ts
TcB Tn-6 + WT-Ts
TcB Tn-7 + WT-Ts
TcB Tn-8 + WT-Ts
TcB Tn-9 + WT-Ts
TcB Tn-10 + WT-Ts

Fig. 1

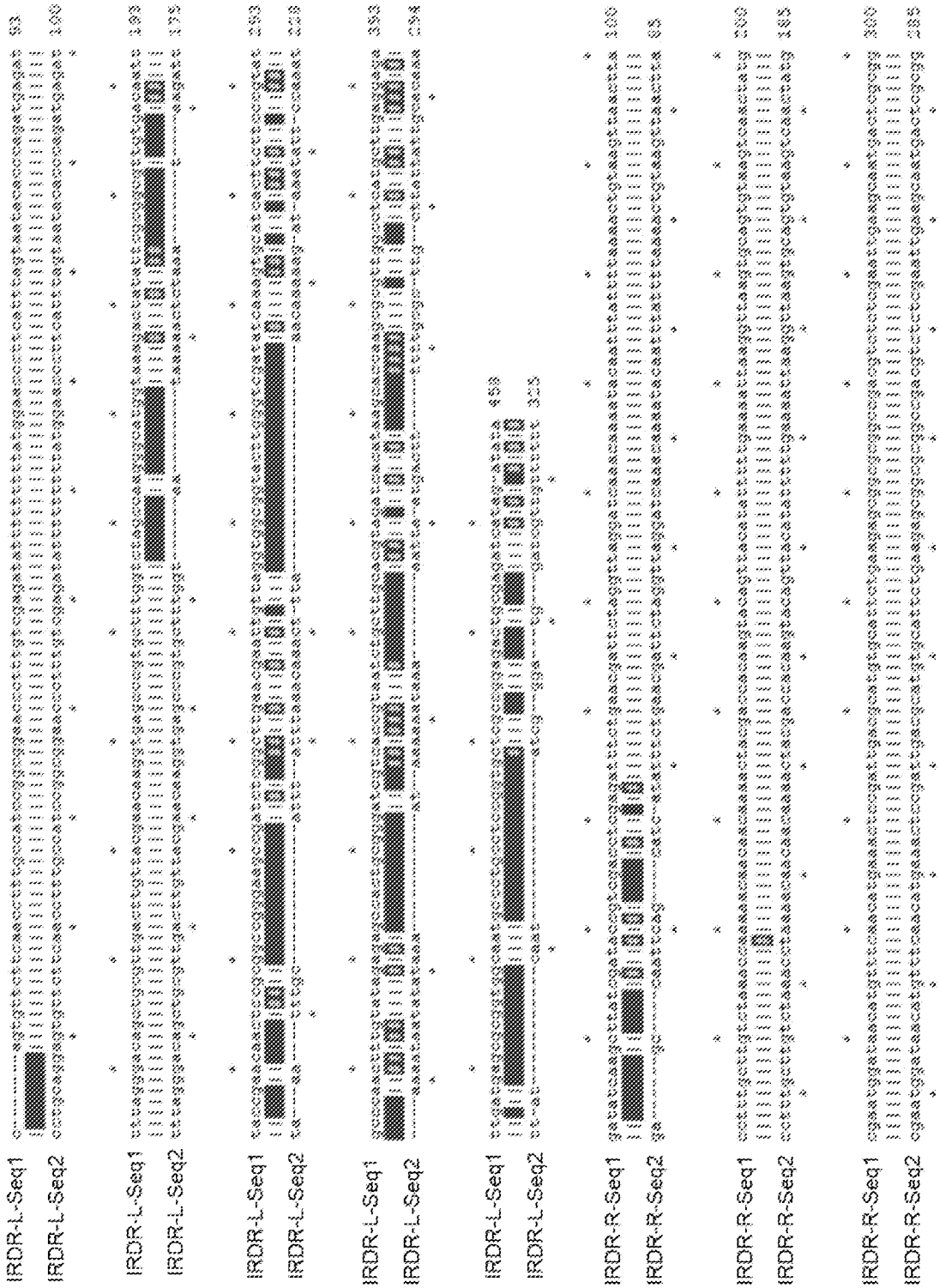
IRDR-L-Seq1
IRDR-L-Seq2
IRDR-L-Seq1
IRDR-L-Seq2
IRDR-L-Seq1
IRDR-L-Seq2
IRDR-L-Seq1
IRDR-L-Seq2
IRDR-L-Seq1
IRDR-L-Seq2
IRDR-R-Seq1
IRDR-R-Seq2
IRDR-R-Seq1
IRDR-R-Seq2
IRDR-R-Seq1
IRDR-R-Seq2
IRDR-R-Seq1
IRDR-R-Seq2

Fig. 2

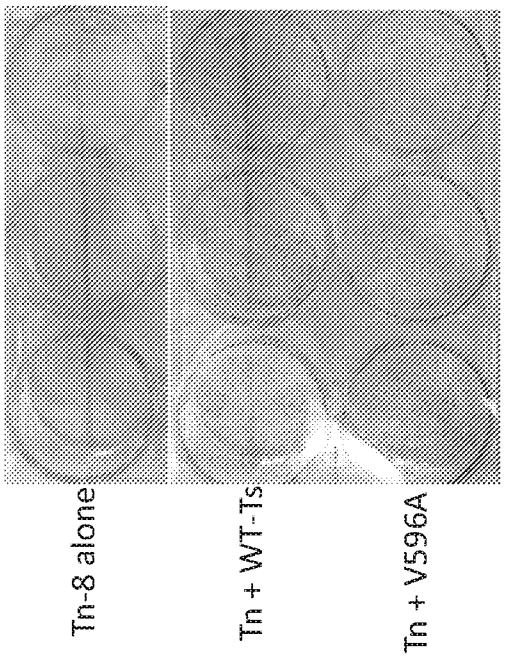

Fig. 3B

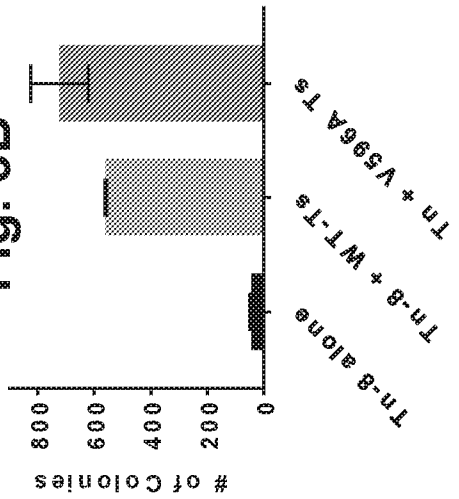

Fig. 3C

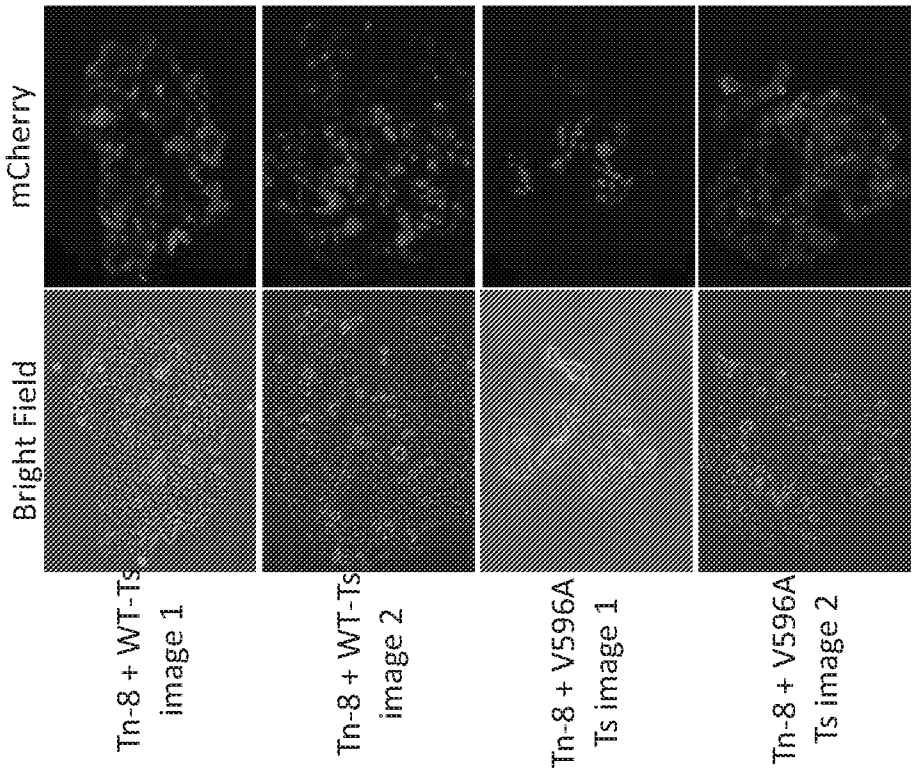

Fig. 3A

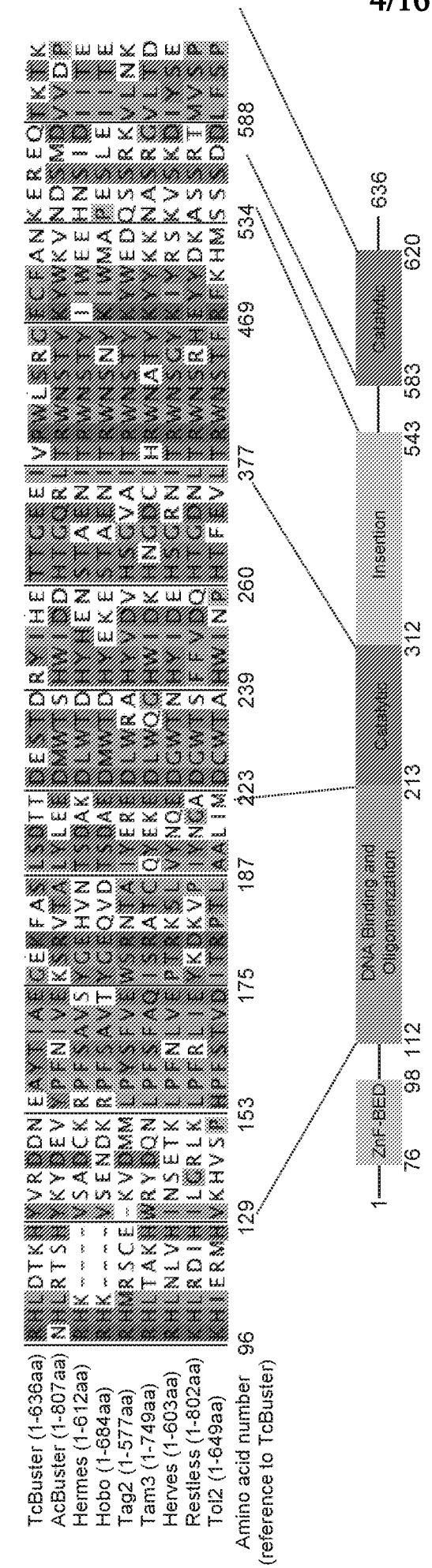
TcBuster (1-636aa)
AcBuster (1-807aa)
Hermes (1-612aa)
Hobo (1-684aa)
Tag2 (1-577aa)
Tam3 (1-749aa)
Herves (1-603aa)
Restless (1-802aa)
Tol2 (1-649aa)
Amino acid number (reference to TcBuster)
96
129
153
175
187
223
239
260
377
469
534
588
636
620
583
543
312
213
112
98
76
1
Zn-BED
DNA Binding and Oligomerization
Catalytic
Insertion
Catalytic

Fig. 4

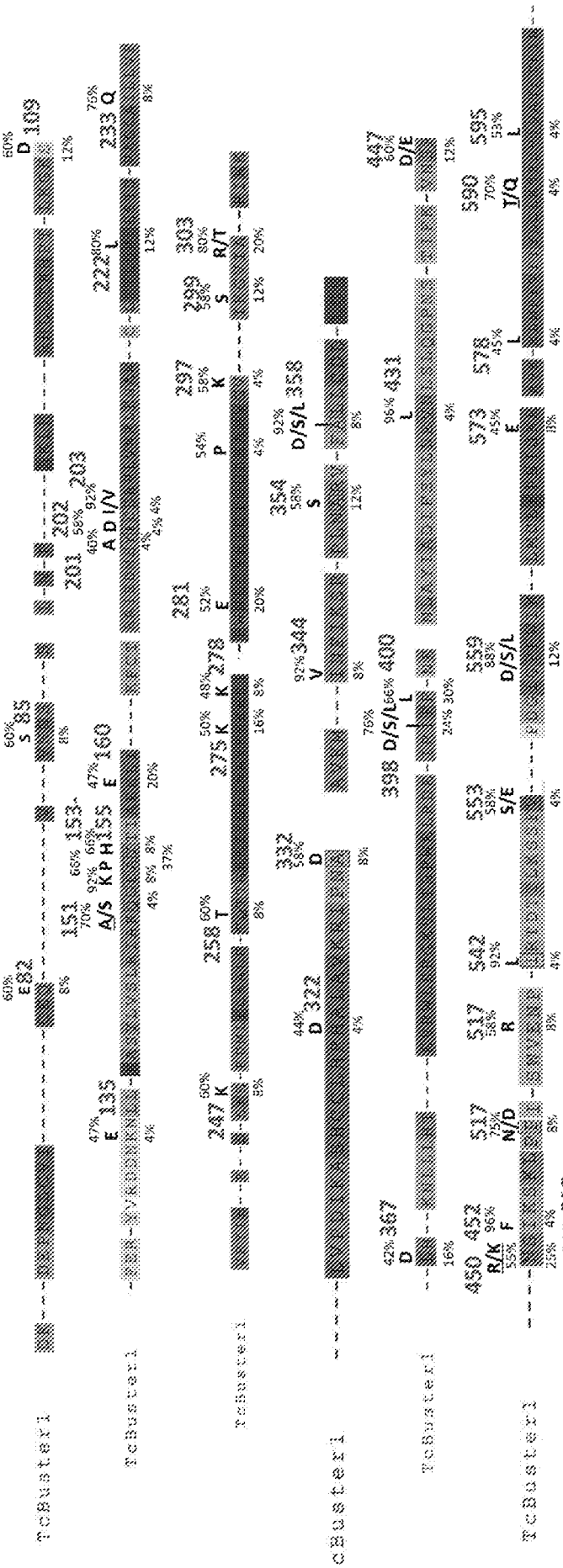

Fig. 5

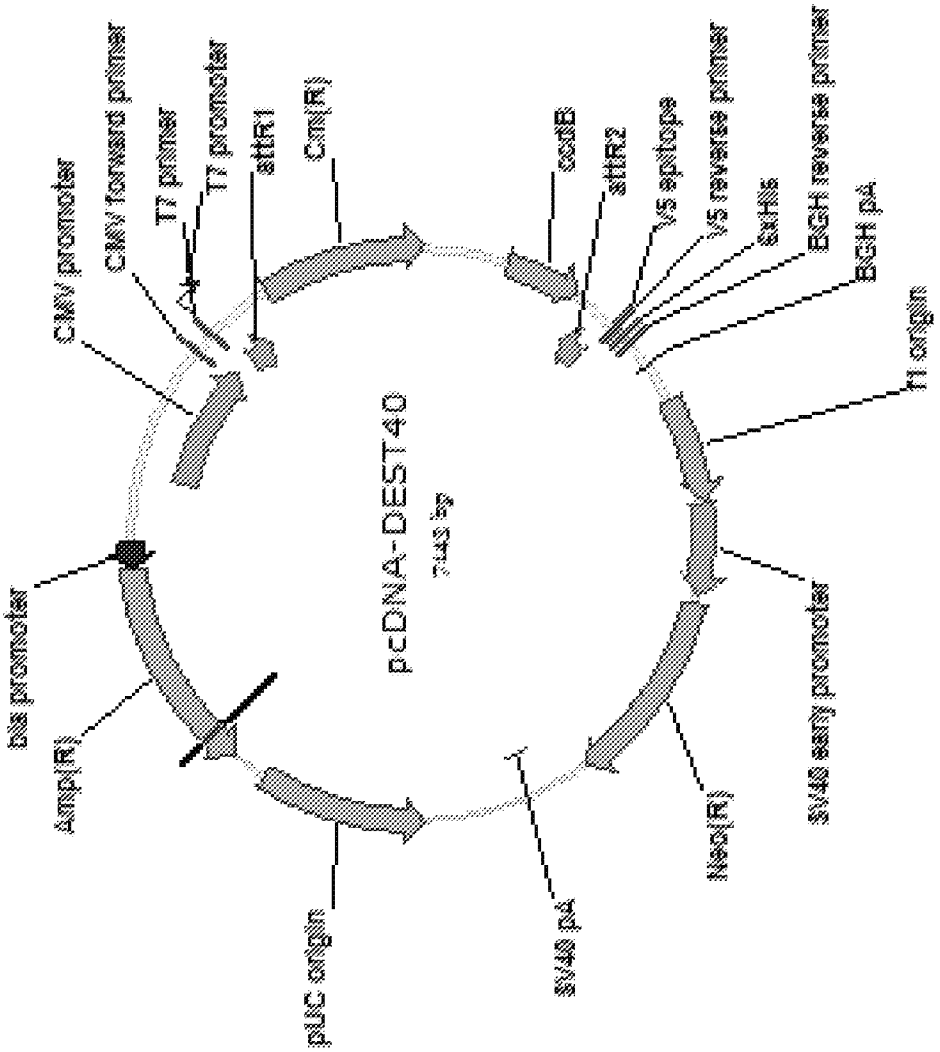
pcDNA-DEST40
7443 bp
CMV promoter
CMV forward primer
T7 primer
T7 promoter
attR1
Cm(R)
ccdB
attR2
V5 epitope
V5 reverse primer
6xHis
BGH reverse primer
BGH pA
f1 origin
SV40 early promoter
Neo(R)
SV40 pA
pUC origin
bla promoter
Amp(R)

Fig. 6

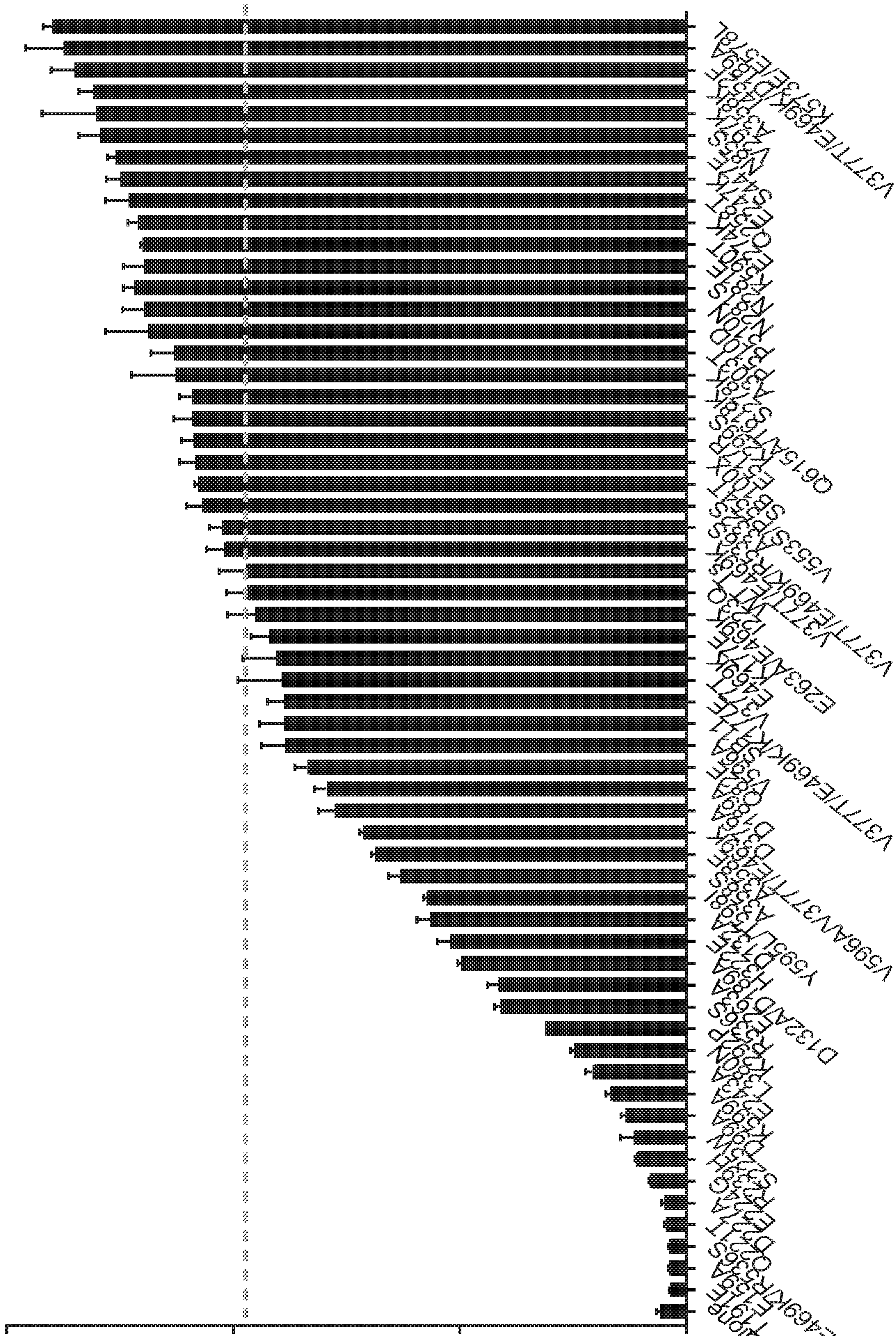

Fig. 7

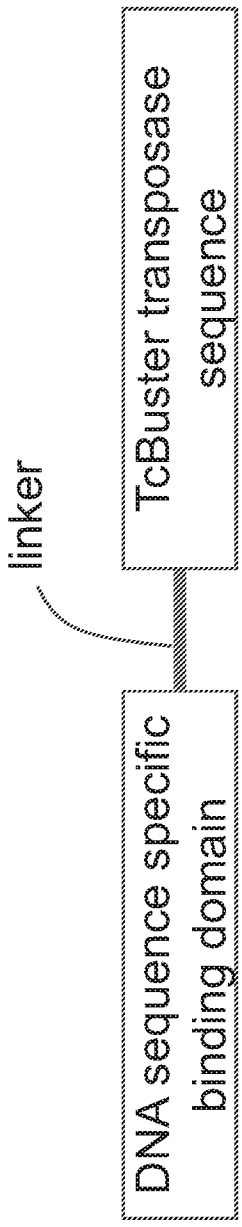
TcBuster transposase sequence
linker
DNA sequence specific binding domain

Fig. 8

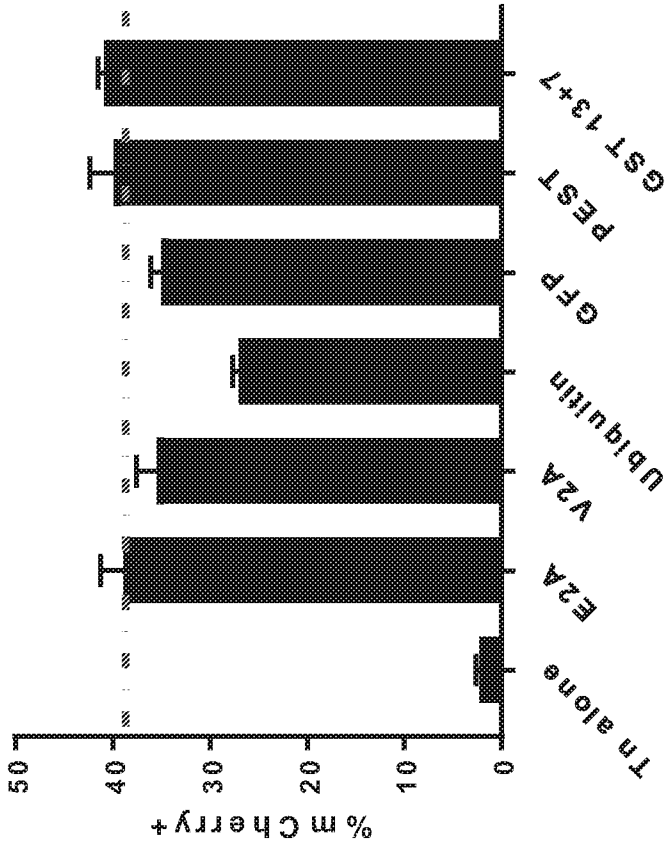
% mCherry+
0
10
20
30
40
50
Tn alone
E2A
V2A
Ubiquitin
GFP
PEST
GST 13+7

Fig. 9

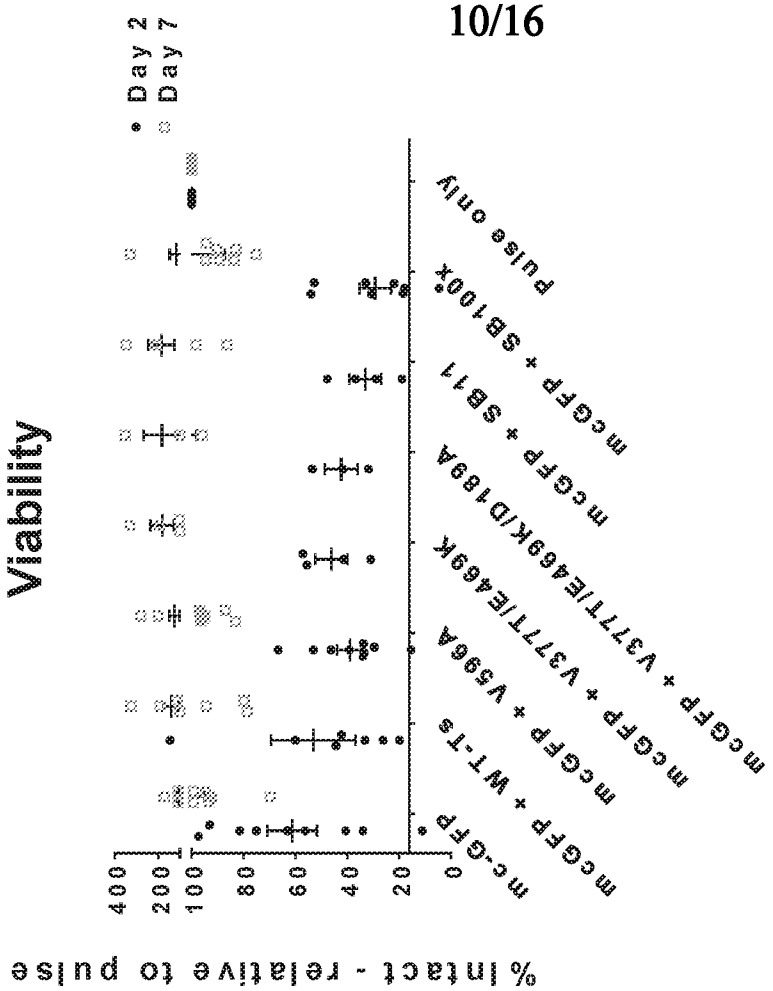

Fig. 10B

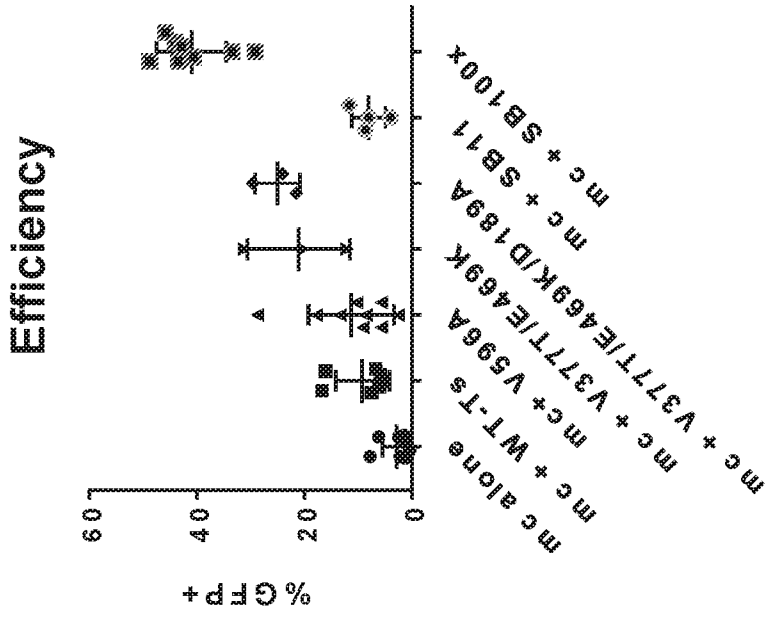

Fig. 10A

TcBuster
MMLNWLKSGKLESQSEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
RKKRKYDEDYLNFGFTWTGDKDEPNGLCVICEQVVNNSSLNPAKLKRHL
DTKHPTLKGKSEYFKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRI
AKQGEAYTIAEKLIKPCTKDLTICVFGEKFASKVDLVPLSDTTISRRIEDM
SYFCEAVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEE
DMLFCKALPTQTTGEEIFNLLNAYHEKHSPWNLCYHICTDGAKAMV
GVKGVIARIKKLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIK
SRPLNARVFALLCDDLGSLHKNLLLHTEVRWLSRGKVLTRFWELRDEIRI
FFNEREFAGKLNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRI
NSIKSKLKLWEECITKNNTECFANLNDFLETSNTALDPNLKSNILEHLNGLK
NTFLEYFPPTCNNISWVENPFNECGNVDTLPIKEREQLIDIRTDTTLKSSF
VPDGIGPFWIKLMDEFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTK
YRNRLDAEDDMRLQLTIHPDIDNLCNNKQAQKSH

Fig. 11

D189A/V377T/E469K
MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
RKKRKYDEDYLNFGFTWTGDKDEPNGIcviceqvvnnssInpaklkrhIDTKHPTL
KGKSEY**FKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGE**

AYTIAEKLIKPCTKDLTTCVFGEKFASKVDLVPLS*A*TTISRRIEDMSYFCE
AVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEDMLFCK
ALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDGAKAMVGVIKGVIARIK
KLVPDI*KASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIKSRPLNARVFA*

*LLCDDLGSLHKNLLLHTE*T*RWLSRGKVLTRFWELRDEIRIFFNEREFAGK*
LNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKLWE

*ECITKNNT*K*CFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPPT*
*CNNISWVENPFNECGNVDTLPIKEREQLIDIR*TDTTLKSSFVPDGIGPFWIK
LMDEFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTKYRNRLDAEDDMR
LQLTTIHPDIDNLCNNKQAQKSH

Fig. 12

D189A/V377T/E469K +I452F
MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
RKKRKYDEDYLNFGFTWTGDKDEPNGIcviceqvvmnssInpakIkrhIDTKHPTL
KGKSEYFKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGE
AYTIAEKLIKPCTKDLTTCVFGEKFASKVDLVPLSATTISRRIEDMSYFCE
AVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEDMLFCK
ALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDGAKAMVGVIKGVIARIK
KLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIKSRPLNARVFA
LLCDDLGSLHKNLLLHTETRWLSRGKVLTRFWELRDEIRIFFNEREFAGK
LNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSFKSKLKLW
EECITKNNTKCFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPP
TCNNISWVENPFNEQGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPFWI
KLMDEFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTKYRNRLDAEDDM
RLQLTTIHPDIDNLCNNKQAQKSH

Fig. 13

D189A/V377T/E469K +N85S

MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS

RKKRKYDEDYLNFGFTWTGDKDEPNGIcviceqvvSnssInpaklkrhlDTKHPT

LKGKSEY**FKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGE**

AYTIAEKLIKPCTKDLTTCVFGEKFASKVDLVPLSATTISRRIEDMSYFCE

AVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEDMLFCK

ALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDGAKAMVGVIKGVIARIK

KLVPDI*KASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIKSRPLNARVFA*

*LLCDDLGSLHKNLLLHTE*T*RWLSRGKVLTRFWELRDEIRIFFNEREFAGK*

LNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKLWE

*ECITKNNT*K*CFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPPT*

*CNNISWVENPFNECGNVDTLPIKEREQLIDIR*TDTTLKSSFVPDGIGPFWIK

LMDEFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTKYRNRLDAEDDMR

LQLTTIHPDIDNLCNNKQAQKSH

Fig. 14

D189A/V377T/E469K +A358K
MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
RKKRKYDEDYLNFGFTWTGDKDEPNGlovlceqvvnnsslnpaklkrhlDTKHPTL
KGKSEYFKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGE
AYTIAEKLIKPCTKDLTTCVFGEKFASKVDLVPLSATTISRRIEDMSYFCE
AVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEDMLFCK
ALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDGAKAMVGVIKGVIARIK
KLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIKSRPLNARVF
KLLCDDLGSLHKNILLHTETRWLSRGKVLTRFWELRDEIRIFFNEREFA
GKLNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKL
WEECITKNNTKCFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYF
PPTCNNISWVENPFNECGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPF
WIKLMDEFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTKYRNRLDAED
DMRLQLTTIHPDIDNLCNNKQAQKSH

Fig. 15

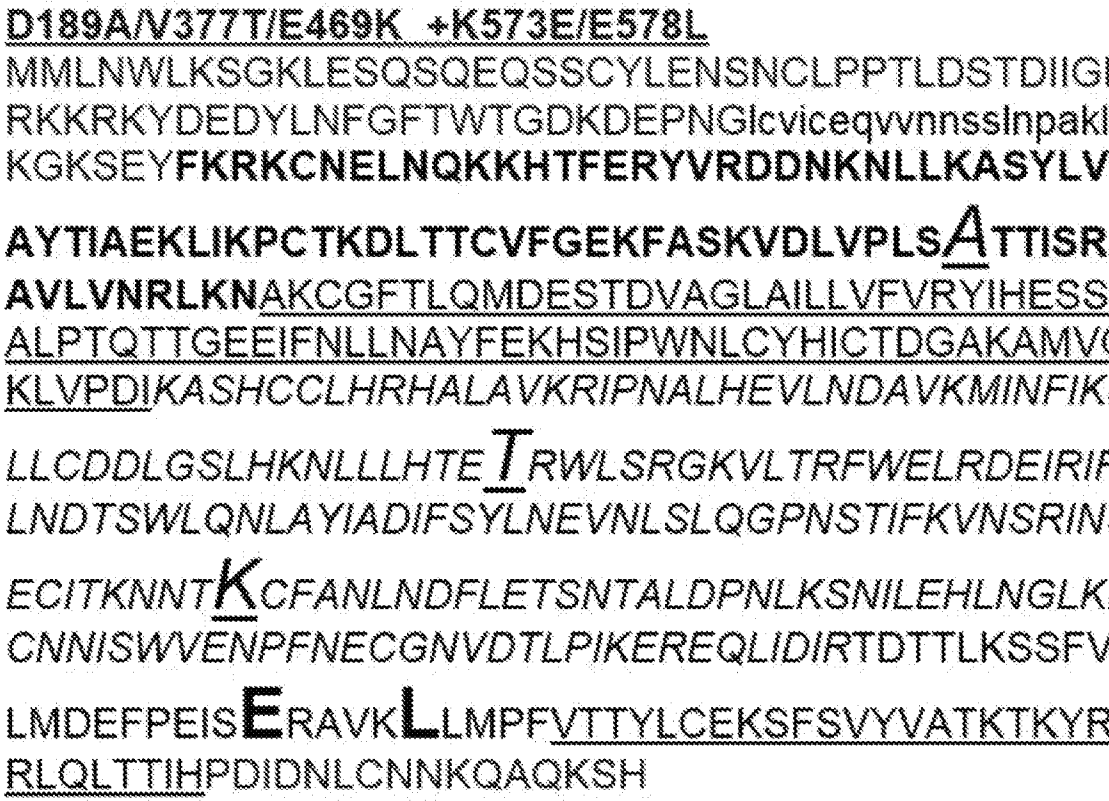
D189A/V377T/E469K +K573E/E578L
MMLNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
RKKRKYDEDYLNFGFTWTGDKDEPNGIcviceqvvnnssInpaklkrhIDTKHPTL
KGKSEYFKRKCNELNQKKHTFERYVRDDNKNLLKASYLVSLRIAKQGE
AYTIAEKLIKPCTKDLTTCVFGEKFASKVDLVPLSATTISRRIEDMSYFCE
AVLVNRLKNAKCGFTLQMDESTDVAGLAILLVFVRYIHESSFEEDMLFCK
ALPTQTTGEEIFNLLNAYFEKHSIPWNLCYHICTDGAKAMVGVIKGVIARIK
KLVPDIKASHCCLHRHALAVKRIPNALHEVLNDAVKMINFIKSRPLNARVFA
LLCDDLGSLHKNLLLHTETRWLSRGKVLTRFWELRDEIRIFFNEREFAGK
LNDTSWLQNLAYIADIFSYLNEVNLSLQGPNSTIFKVNSRINSIKSKLKLWE
ECITKNNTKCFANLNDFLETSNTALDPNLKSNILEHLNGLKNTFLEYFPPT
CNNISWVENPFNECGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPFWIK
LMDEFPEISERAVKLLMPFVTTYLCEKSFSVYVATKTKYRNRLDAEDDM
RLQLTTIHPDIDNLCNNKQAQKSH

Fig. 16