



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

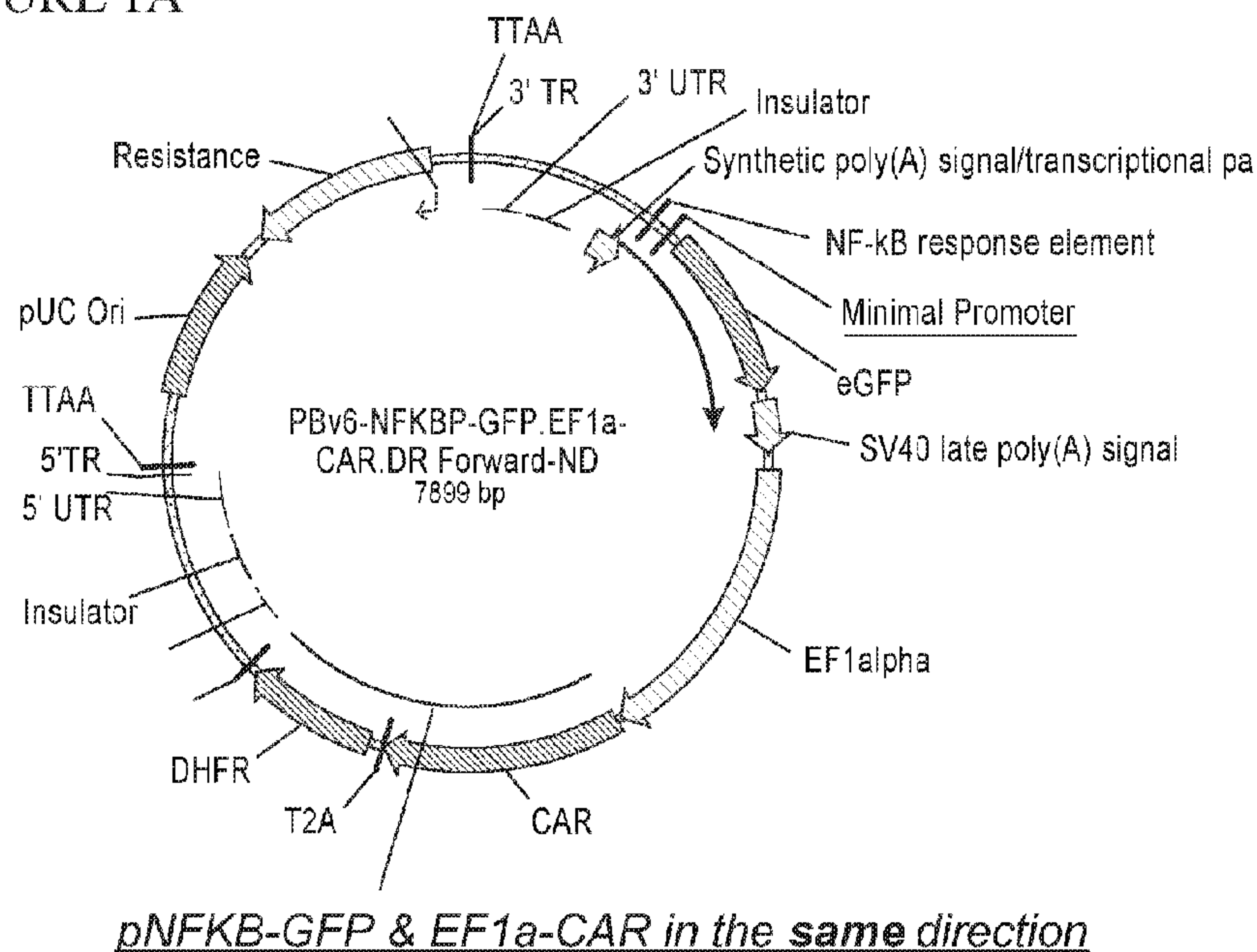
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2018/09/10
(87) Date publication PCT/PCT Publication Date: 2019/03/14
(85) Entrée phase nationale/National Entry: 2020/02/11
(86) N° demande PCT/PCT Application No.: US 2018/050288
(87) N° publication PCT/PCT Publication No.: 2019/051424
(30) Priorité/Priority: 2017/09/08 (US62/556,310)

(51) Cl.Int./Int.Cl. *C12N 15/63* (2006.01),
C12N 15/90 (2006.01)
(71) Demandeur/Applicant:
POSEIDA THERAPEUTICS, INC., US
(72) Inventeurs/Inventors:
OSTERAG, ERIC, US;
SHEDLOCK, DEVON, US
(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : COMPOSITIONS ET PROCEDES D'EXPRESSION GENIQUE CONDITIONNELLE MEDIEE PAR UN
RECEPTEUR DE LIGAND CHIMERIQUE (CLR)
(54) Title: COMPOSITIONS AND METHODS FOR CHIMERIC LIGAND RECEPTOR (CLR)-MEDIATED CONDITIONAL
GENE EXPRESSION

FIGURE 1A



(57) **Abrégé/Abstract:**

Disclosed are composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous reporter is expressed, and wherein the exogenous reporter, upon binding a ligand, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression. Methods for introducing compositions into cells and the use of the resultant cells in adoptive cell therapies are also provided.

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

CORRECTED VERSION

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
14 March 2019 (14.03.2019)



(10) International Publication Number
WO 2019/051424 A9

(51) International Patent Classification:

C12N 15/63 (2006.01) C12N 15/90 (2006.01)

(21) International Application Number:

PCT/US2018/050288

(22) International Filing Date:

10 September 2018 (10.09.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/556,310 08 September 2017 (08.09.2017) US

(71) Applicant: **POSEIDA THERAPEUTICS, INC.**
[US/US]; 4242 Campus Point Court, Suite 700, San Diego,
California 92121 (US).

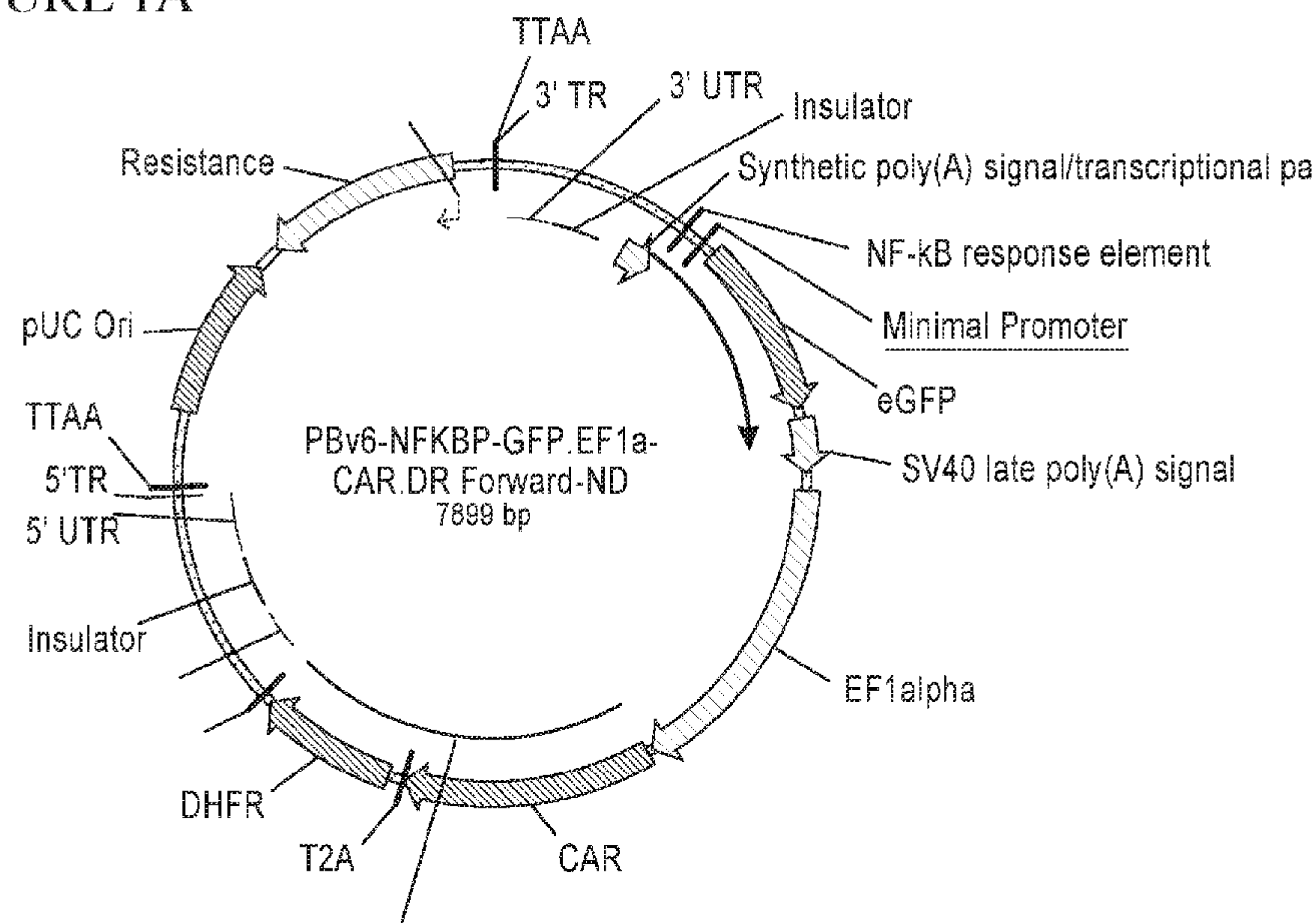
(72) Inventors: **OSTERAG, Eric**; 4242 Campus Point Court,
Suite 700, San Diego, California 92121 (US). **SHED-
LOCK, Devon**; 4242 Campus Point Court, Suite 700, San
Diego, California 92121 (US).

(74) Agent: **MILLER, Katherine J.** et al.; 1299 Pennsylvania
Avenue, NW, Suite 700, Washington, District of Columbia
2004 (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,

(54) Title: COMPOSITIONS AND METHODS FOR CHIMERIC LIGAND RECEPTOR (CLR)-MEDIATED CONDITIONAL GENE EXPRESSION

FIGURE 1A



pNFKB-GFP & EF1a-CAR in the same direction

(57) Abstract: Disclosed are composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous reporter is expressed, and wherein the exogenous reporter, upon binding a ligand, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression. Methods for introducing compositions into cells and the use of the resultant cells in adoptive cell therapies are also provided.

[Continued on next page]

WO 2019/051424 A9



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))*
— *with sequence listing part of description (Rule 5.2(a))*

(88) Date of publication of the international search report:

11 April 2019 (11.04.2019)

(48) Date of publication of this corrected version:

23 May 2019 (23.05.2019)

(15) Information about Correction:

see Notice of 23 May 2019 (23.05.2019)

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 162

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 162

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

COMPOSITIONS AND METHODS FOR CHIMERIC LIGAND RECEPTOR (CLR)- MEDIATED CONDITIONAL GENE EXPRESSION

RELATED APPLICATIONS

[01] This application claims the benefit of provisional application USSN 62/556,310, filed September 8, 2017, the contents of which are herein incorporated by reference in their entirety.

INCORPORATION OF SEQUENCE LISTING

[02] The contents of the text file named "POTH-027/001WO SeqListing.txt," which was created on September 10, 2018 and is 55,054 KB in size, are hereby incorporated by reference in their entirety.

FIELD OF THE DISCLOSURE

[03] The disclosure is directed to molecular biology, and more, specifically, to compositions and methods for use in a conditional gene expression system responsive to a chimeric ligand receptor (CLR)-mediated signal.

BACKGROUND

[04] There has been a long-felt but unmet need in the art for a method of controlling gene expression in genetically modified cells for the long-term delivery of therapeutic agents. The disclosure provides a solution by genetically modified cells that conditionally express genes upon activation of a cell-surface receptor.

SUMMARY

[05] The disclosure provides a composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous reporter is expressed, and wherein the exogenous reporter, upon binding a ligand, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression. In certain embodiments, the composition modifies gene expression by increasing gene

expression. In certain embodiments, the composition modifies gene expression by decreasing gene expression. In certain embodiments, the composition modifies gene expression by transiently modifying gene expression (e.g. for the duration of binding of the ligand to the exogenous receptor). In certain embodiments, the composition modifies gene expression acutely (e.g. the ligand reversibly binds to the exogenous receptor). In certain embodiments, the composition modifies gene expression chronically (e.g. the ligand irreversibly binds to the exogenous receptor).

[06] In certain embodiments of the compositions of the disclosure, the cell may be a prokaryotic cell. Prokaryotic cells of the disclosure include, but are not limited to, bacteria and archaea. For example, bacteria of the disclosure include, but are not limited to, *listeria monocytogenes*.

[07] In certain embodiments of the compositions of the disclosure, the cell may be a eukaryotic cell. Eukaryotic cells of the disclosure include, but are not limited to, yeast, plants, algae, insects, mammals, amphibians, birds, reptiles, marsupials, rodents, and humans. Preferred eukaryotic cells of the disclosure include, but are not limited to, human cells. Exemplary human cells of the disclosure include but are not limited to, immune cells (e.g. T cells), myeloid cells and bone marrow cells (e.g. hematopoietic stem cells (HSCs)).

[08] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises an endogenous receptor with respect to the genomic sequence of the cell. Exemplary receptors include, but are not limited to, intracellular receptors, cell-surface receptors, transmembrane receptors, ligand-gated ion channels, and G-protein coupled receptors.

[09] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the non-naturally occurring receptor is a synthetic, modified, recombinant, mutant or chimeric receptor. In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a T-cell receptor (TCR). In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a scaffold protein. In certain embodiments, including those wherein the non-naturally occurring receptor does not comprise a transmembrane domain, the non-naturally occurring receptor interacts with a second transmembrane, membrane-bound and/or an intracellular receptor that, following contact with the non-naturally occurring receptor, transduces an intracellular signal.

[010] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the non-naturally occurring receptor is a synthetic, modified, recombinant, mutant or chimeric receptor. In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a T-cell receptor (TCR). In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a scaffold protein. In certain embodiments, the non-naturally occurring receptor comprises a transmembrane domain. In certain embodiments, the non-naturally occurring receptor interacts with an intracellular receptor that transduces an intracellular signal. In certain embodiments, the non-naturally occurring receptor comprises an intracellular signalling domain. In certain embodiments, the non-naturally occurring receptor is a chimeric ligand receptor (CLR). In certain embodiments, the CLR is a chimeric antigen receptor.

[011] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the CLR is a chimeric antigen receptor. In certain embodiments, the chimeric ligand receptor comprises (a) an ectodomain comprising a ligand recognition region, wherein the ligand recognition region comprises at least scaffold protein; (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain. In certain embodiments, the ectodomain of (a) further comprises a signal peptide. In certain embodiments, the ectodomain of (a) further comprises a hinge between the ligand recognition region and the transmembrane domain. In certain embodiments, the signal peptide comprises a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR signal peptide. In certain embodiments, the signal peptide comprises a sequence encoding a human CD8 α signal peptide. In certain embodiments, the signal peptide comprises an amino acid sequence comprising MALPVTALLPLALLHAARP (SEQ ID NO:17000). In certain embodiments, the signal peptide is encoded by a nucleic acid sequence comprising atggcactgccagtcaccgccctgctgctgcctctggctctgctgctgcacgcagctagacca (SEQ ID NO:17001). In certain embodiments, the transmembrane domain comprises a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR transmembrane domain. In certain embodiments, the transmembrane domain comprises a sequence encoding a human CD8 α transmembrane domain. In certain embodiments, the transmembrane domain comprises an amino acid sequence comprising IYIWAPLAGTCGVLLLSLVITLYC (SEQ ID NO: 17002). In certain embodiments, the

transmembrane domain is encoded by a nucleic acid sequence comprising
atctacatttgggcaccactggccgggacctgtggagtgctgctgctgagcctggtcacacactgtactgc (SEQ ID NO: 17003). In certain embodiments, the endodomain comprises a human CD3 ζ endodomain. In certain embodiments, the at least one costimulatory domain comprises a human 4-1BB, CD28, CD3 ζ , CD40, ICOS, MyD88, OX-40 intracellular segment, or any combination thereof. In certain embodiments, the at least one costimulatory domain comprises a human CD3 ζ and/or a 4-1BB costimulatory domain. In certain embodiments, the CD3 ζ costimulatory domain comprises an amino acid sequence comprising
RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPPEMGGKPRRKNPQ
EGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALP
PR (SEQ ID NO: 17004). In certain embodiments, the CD3 ζ costimulatory domain is encoded by a nucleic acid sequence comprising
cgcgtgaagttagtcgatcagcagatgccccagcttacaacaggacagaaccagctgtataacgagctgaatctgggcccgcga
gaggaatatgacgtgctggataagcggagaggacgcgaccccgaaatgggaggcaagcccaggcgcaaaaaccctcaggaagg
cctgtataacgagctgcagaaggacaaaatggcagaagcctattctgagatcggcctgaagggggagcgacggagaggcaaagg
gcacgatgggctgtaccagggactgagcaccgcccacaaaggacacctatgatgctctgcataatgcaggcactgcctccaagg
(SEQ ID NO: 17005). In certain embodiments, the 4-1BB costimulatory domain comprises an amino acid sequence comprising
KRGRKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL (SEQ ID NO: 17006). In certain embodiments, the 4-1BB costimulatory domain is encoded by a nucleic acid sequence comprising
aagagaggcaggaagaaaactgctgtatatatttcaaacagcccttcattgcgccccgtgcagactaccaggaggaagacgggtgctcc
tgctgattccctgaggaagaggaaggcgggtgtgagctg (SEQ ID NO: 17007). In certain embodiments, the 4-1BB costimulatory domain is located between the transmembrane domain and the CD3 ζ costimulatory domain. In certain embodiments, the hinge comprises a sequence derived from a human CD8 α , IgG4, and/or CD4 sequence. In certain embodiments, the hinge comprises a sequence derived from a human CD8 α sequence. In certain embodiments, the hinge comprises an amino acid sequence comprising
TTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD (SEQ ID NO: 17008). In certain embodiments, the hinge is encoded by a nucleic acid sequence comprising
actaccacaccagcacctagaccaccaactccagctccaaccatcgcgagtcagcccctgagctgagacctgaggcctgcaggcc
agctgcaggaggagctgtgcacaccaggggctggacttcgctgcgac (SEQ ID NO: 17028). In certain embodiments, the hinge is encoded by a nucleic acid sequence comprising

ACCACAACCCCTGCCCCCAGACCTCCCACACCCGCCCCCTACCATCGCGAGTCAGC
CCCTGAGTCTGAGACCTGAGGCCTGCAGGCCAGCTGCAGGAGGAGCTGTGCACA
CCAGGGGGCCTGGACTTCGCCTGCGAC (SEQ ID NO: 17009). In certain embodiments,
the at least one protein scaffold specifically binds the ligand.

[012] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the CLR is a chimeric antigen receptor. In certain embodiments, the chimeric ligand receptor comprises (a) an ectodomain comprising a ligand recognition region, wherein the ligand recognition region comprises at least scaffold protein; (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain. In certain embodiments, the at least one protein scaffold comprises an antibody, an antibody fragment, a single domain antibody, a single chain antibody, an antibody mimetic, or a Centyrin. In certain embodiments, the ligand recognition region comprises one or more of an antibody, an antibody fragment, a single domain antibody, a single chain antibody, an antibody mimetic, and a Centyrin. In certain embodiments, the single domain antibody comprises or consists of a VHH. In certain embodiments, the antibody mimetic comprises or consists of an affibody, an affilin, an affimer, an affitin, an alphabody, an anticalin, an avimer, a DARPin, a Fynomer, a Kunitz domain peptide or a monobody. In certain embodiments, the Centyrin comprises or consists of a consensus sequence of at least one fibronectin type III (FN3) domain.

[013] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the CLR is a chimeric antigen receptor. In certain embodiments, the chimeric ligand receptor comprises (a) an ectodomain comprising a ligand recognition region, wherein the ligand recognition region comprises at least scaffold protein; (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain. In certain embodiments, the Centyrin comprises or consists of a consensus sequence of at least one fibronectin type III (FN3) domain. In certain embodiments, the at least one fibronectin type III (FN3) domain is derived from a human protein. In certain embodiments, the human protein is Tenascin-C. In certain embodiments, the consensus sequence comprises
LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYDL
TGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17010). In certain
embodiments, the consensus sequence comprises

MLPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGGEAINLTVPGSERSYD
LTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17011). In certain

embodiments, the consensus sequence is modified at one or more positions within (a) a A-B loop comprising or consisting of the amino acid residues TEDS at positions 13-16 of the consensus sequence; (b) a B-C loop comprising or consisting of the amino acid residues TAPDAAF at positions 22-28 of the consensus sequence; (c) a C-D loop comprising or consisting of the amino acid residues SEKVGGE at positions 38-43 of the consensus sequence; (d) a D-E loop comprising or consisting of the amino acid residues GSER at positions 51-54 of the consensus sequence; (e) a E-F loop comprising or consisting of the amino acid residues GLKPG at positions 60-64 of the consensus sequence; (f) a F-G loop comprising or consisting of the amino acid residues KGGHRSN at positions 75-81 of the consensus sequence; or (g) any combination of (a)-(f). In certain embodiments, the Centyrin comprises a consensus sequence of at least 5 fibronectin type III (FN3) domains. In certain embodiments, the Centyrin comprises a consensus sequence of at least 10 fibronectin type III (FN3) domains. In certain embodiments, the Centyrin comprises a consensus sequence of at least 15 fibronectin type III (FN3) domains. In certain embodiments, the scaffold binds an antigen with at least one affinity selected from a K_D of less than or equal to $10^{-9}M$, less than or equal to $10^{-10}M$, less than or equal to $10^{-11}M$, less than or equal to $10^{-12}M$, less than or equal to $10^{-13}M$, less than or equal to $10^{-14}M$, and less than or equal to $10^{-15}M$. In certain embodiments, the K_D is determined by surface plasmon resonance. In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the CLR is a chimeric antigen receptor. In certain embodiments, the chimeric ligand receptor comprises (a) an ectodomain comprising a ligand recognition region, wherein the ligand recognition region comprises at least a VHH antibody; (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain. In certain embodiments, the VHH is camelid. Alternatively, or in addition, in certain embodiments, the VHH is humanized. In certain embodiments, the sequence comprises two heavy chain variable regions of an antibody, wherein the complementarity-determining regions (CDRs) of the VHH are human sequences.

[014] In certain embodiments of the compositions of the disclosure, the sequence encoding the constitutive promoter of (b) comprises a sequence encoding an EF1 α promoter. In certain embodiments of the compositions of the disclosure, the sequence encoding the constitutive promoter of (b) comprises a sequence encoding a CMV promoter, a U6 promoter, a SV40

promoter, a PGK1 promoter, a Ubc promoter, a human beta actin promoter, a CAG promoter, or an EF1 α promoter.

[015] In certain embodiments of the compositions of the disclosure, the sequence encoding the inducible promoter of (a) comprises a sequence encoding an NF κ B promoter. In certain embodiments of the compositions of the disclosure, the sequence encoding the inducible promoter of (a) comprises a sequence encoding an interferon (IFN) promoter or a sequence encoding an interleukin-2 promoter. In certain embodiments of the compositions of the disclosure, the sequence encoding the inducible promoter of (a) comprises a sequence encoding a nuclear receptor subfamily 4 group A member 1 (NR4A1; also known as NUR77) promoter or a sequence encoding a NR4A1 promoter. In certain embodiments of the compositions of the disclosure, the sequence encoding the inducible promoter of (a) comprises a sequence encoding a T-cell surface glycoprotein CD5 (CD5) promoter or a sequence encoding a CD5 promoter. In certain embodiments, the interferon (IFN) promoter is an IFN γ promoter. In certain embodiments of the compositions of the disclosure, the inducible promoter is isolated or derived from the promoter of a cytokine or a chemokine. In certain embodiments, the cytokine or chemokine comprises IL2, IL3, IL4, IL5, IL6, IL10, IL12, IL13, IL17A/F, IL21, IL22, IL23, transforming growth factor beta (TGF β), colony stimulating factor 2 (GM-CSF), interferon gamma (IFN γ), Tumor necrosis factor (TNF α), LT α , perforin, Granzyme C (Gzmc), Granzyme B (Gzmb), C-C motif chemokine ligand 5 (CCL5), C-C motif chemokine ligand 4 (CCL4), C-C motif chemokine ligand 3 (CCL3), X-C motif chemokine ligand 1 (XCL1) and LIF interleukin 6 family cytokine (Lif).

[016] In certain embodiments of the compositions of the disclosure, including those wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding a NR4A1 promoter or a sequence encoding a NR4A1 promoter, the NR4A1 promoter is activated by T-cell Receptor (TCR) stimulation in T cells and by B-cell Receptor (BCR) stimulation in B cells, therefore, inducing expression of any sequence under control of the NR4A1 promoter upon activation of a T-cell or B-cell of the disclosure through a TCR or BCR, respectively.

[017] In certain embodiments of the compositions of the disclosure, including those wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding a CD5 promoter or a sequence encoding a CD5 promoter, the CD5 promoter is activated by T-cell Receptor (TCR) stimulation in T cells, therefore, inducing expression of any sequence

under control of the CD5 promoter upon activation of a T-cell of the disclosure through a TCR.

[018] In certain embodiments of the compositions of the disclosure, the inducible promoter is isolated or derived from the promoter of a gene comprising a surface protein involved in cell differentiation, activation, exhaustion and function. In certain embodiments, the gene comprises CD69, CD71, CTLA4, PD-1, TIGIT, LAG3, TIM-3, GITR, MHCII, COX-2, FASL and 4-1BB.

[019] In certain embodiments of the compositions of the disclosure, the inducible promoter is isolated or derived from the promoter of a gene involved in CD metabolism and differentiation. In certain embodiments of the compositions of the disclosure, the inducible promoter is isolated or derived from the promoter of Nr4a1, Nr4a3, Tnfrsf9 (4-1BB), Sema7a, Zfp36l2, Gadd45b, Dusp5, Dusp6 and Neto2.

[020] In certain embodiments of the compositions of the disclosure, the transgene comprises a sequence that is endogenous with respect to the genomic sequence of the cell.

[021] In certain embodiments of the compositions of the disclosure, the transgene comprises a sequence that is exogenous with respect to the genomic sequence of the cell. In certain embodiments, the exogenous sequence is a sequence variant of an endogenous sequence within the genome of the cell. In certain embodiments, the exogenous sequence is a wild type sequence of gene that is entirely or partially absent in the cell, and wherein the gene is entirely present in the genome of a healthy cell. In certain embodiments, the exogenous sequence is a synthetic, modified, recombinant, chimeric or non-naturally occurring sequence with respect to the genome of the cell. In certain embodiments, the transgene encodes a secreted protein. In certain embodiments, the secreted protein is produced and/or secreted from the cell at a level that is therapeutically effective to treat a disease or disorder in a subject in need thereof.

[022] In certain embodiments of the compositions of the disclosure, a first transposon comprises the inducible transgene construct of (a) and a second transposon comprises the receptor construct of (b). In certain embodiments of the compositions of the disclosure, a first vector comprises the first transposon and a second vector comprises the second transposon. In certain embodiments of the compositions of the disclosure, a vector comprises the first transposon and the second transposon. In certain embodiments, the first transposon and the second transposon are oriented in the same direction. In certain embodiments, the first

transposon and the second transposon are oriented in opposite directions. In certain embodiments, the vector is a plasmid. In certain embodiments, the vector is a nanoplasmid.

[023] In certain embodiments of the compositions of the disclosure, the vector is a viral vector. Viral vectors of the disclosure may comprise a sequence isolated or derived from a retrovirus, a lentivirus, an adenovirus, an adeno-associated virus or any combination thereof. The viral vector may comprise a sequence isolated or derived from an adeno-associated virus (AAV). The viral vector may comprise a recombinant AAV (rAAV). Exemplary adeno-associated viruses and recombinant adeno-associated viruses of the disclosure comprise two or more inverted terminal repeat (ITR) sequences located in cis next to a sequence encoding a construct of the disclosure. Exemplary adeno-associated viruses and recombinant adeno-associated viruses of the disclosure include, but are not limited to all serotypes (e.g. AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, and AAV9). Exemplary adeno-associated viruses and recombinant adeno-associated viruses of the disclosure include, but are not limited to, self-complementary AAV (scAAV) and AAV hybrids containing the genome of one serotype and the capsid of another serotype (e.g. AAV2/5, AAV-DJ and AAV-DJ8). Exemplary adeno-associated viruses and recombinant adeno-associated viruses of the disclosure include, but are not limited to, rAAV-LK03 and AAVs with the NP-59 and NP-84 capsid variants.

[024] In certain embodiments of the compositions of the disclosure, the vector is a nanoparticle. Exemplary nanoparticle vectors of the disclosure include, but are not limited to, nucleic acids (e.g. RNA, DNA, synthetic nucleotides, modified nucleotides or any combination thereof), amino acids (L-amino acids, D-amino acids, synthetic amino acids, modified amino acids, or any combination thereof), polymers (e.g. polymersomes), micelles, lipids (e.g. liposomes), organic molecules (e.g. carbon atoms, sheets, fibers, tubes), inorganic molecules (e.g. calcium phosphate or gold) or any combination thereof. A nanoparticle vector may be passively or actively transported across a cell membrane.

[025] In certain embodiments of the compositions of the disclosure, first transposon or the second transposon is a piggyBac transposon. In certain embodiments, the first transposon and the second transposon is a piggyBac transposon. In certain embodiments, the composition further comprises a plasmid or a nanoplasmid comprising a sequence encoding a transposase enzyme. In certain embodiments, the sequence encoding a transposase enzyme is an mRNA sequence. In certain embodiments, the transposase is a piggyBac transposase. In certain embodiments, the piggyBac transposase comprises an amino acid sequence comprising SEQ

ID NO: 1. In certain embodiments, the piggyBac transposase is a hyperactive variant and wherein the hyperactive variant comprises an amino acid substitution at one or more of positions 30, 165, 282 and 538 of SEQ ID NO: 1. In certain embodiments, the amino acid substitution at position 30 of SEQ ID NO: 1 is a substitution of a valine (V) for an isoleucine (I) (I30V). In certain embodiments, the amino acid substitution at position 165 of SEQ ID NO: 1 is a substitution of a serine (S) for a glycine (G) (G165S). In certain embodiments, the amino acid substitution at position 282 of SEQ ID NO: 1 is a substitution of a valine (V) for a methionine (M) (M282V). In certain embodiments, the amino acid substitution at position 538 of SEQ ID NO: 1 is a substitution of a lysine (K) for an asparagine (N) (N538K). In certain embodiments, the transposase is a Super piggyBac (SPB) transposase. In certain embodiments, the Super piggyBac (SPB) transposase comprises an amino acid sequence comprising SEQ ID NO: 2.

[026] In certain embodiments of the disclosure, the transposase enzyme is a piggyBac™ (PB) transposase enzyme. The piggyBac (PB) transposase enzyme may comprise or consist of an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MGSSLDDDEHI LSALLQSDDE LVGEDSDSEI SDHVSEDDVQ SDTEEFIDE VHEVQPTSSG
61 SEILDEQNVI EQPGSSLASN RILTLPORTI RGKNKHCWST SKSTRRSRVS ALNIVRSQRG
121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTGATFRD TNEDEIYAFF
181 GILVMTAVRK DNHMSTDDL FDRSLSMVYVS VMSRDRFDFL IRCLRMDDKS IRPTLRENDV
241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RMYIPNKPSK YGIKILMMCD
301 SGYKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWF TSIPLAKNLLQ
361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCFDGP LTLVSYKPKP AKMVYLLSSC
421 DEDASINEST GKPQMVMYYN QTKGGVDTLD QMCSVMTCR KTNRWPMALL YGMINIACIN
481 SFIIYSHNVS SKGEKVQSRK KEMRNLYMSL TSSFMRKRLE APTLKRYLRD NISNILPNEV
541 PGTSDDSTEE PVMKKRTYCT YCPSKIRKA NASCKKCKV ICREHNIDMC QSCF (SEQ ID NO:
17029) .

```

[027] In certain embodiments of the disclosure, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at one or more of positions 30, 165, 282, or 538 of the sequence:

```

1 MGSSLDDDEHI LSALLQSDDE LVGEDSDSEI SDHVSEDDVQ SDTEEFIDE VHEVQPTSSG
61 SEILDEQNVI EQPGSSLASN RILTLPORTI RGKNKHCWST SKSTRRSRVS ALNIVRSQRG
121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTGATFRD TNEDEIYAFF

```

181 GILVMTAVRK DNHMSTDDL F DRSLSMVYVS VMSRDREDFL IRCLRMDDKS IRPTLRENDV
 241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RMYIPNKPSK YGIKILMMCD
 301 SGYKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
 361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCFDGP LTLVSYKPKP AKMVYLLSSC
 421 DEDASINEST GKPQMVMYYN QTKGGVDTLD QMCSVMTCSP KTNRWEMALL YGMINIACIN
 481 SFIIYSHNVS SKGEKVQSRK KEMRNLYMSL TSSFMRKRLE APTLKRYLRD NISNILENEV
 541 PGTSDDSTEE PVMKKRTYCT YCPSKIRRKA NASCKKCKKV ICREHNIDMC QSCF (SEQ ID NO:
 17029) .

[028] In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at two or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 1. In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at three or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 1. In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at each of the following positions 30, 165, 282, and 538 of the sequence of SEQ ID NO: 1. In certain embodiments, the amino acid substitution at position 30 of the sequence of SEQ ID NO: 1 is a substitution of a valine (V) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 165 of the sequence of SEQ ID NO: 1 is a substitution of a serine (S) for a glycine (G). In certain embodiments, the amino acid substitution at position 282 of the sequence of SEQ ID NO: 1 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 538 of the sequence of SEQ ID NO: 1 is a substitution of a lysine (K) for an asparagine (N).

[029] In certain embodiments of the disclosure, the transposase enzyme is a Super piggyBac™ (SPB) transposase enzyme. In certain embodiments, the Super piggyBac™ (SPB) transposase enzymes of the disclosure may comprise or consist of the amino acid sequence of the sequence of SEQ ID NO: 1 wherein the amino acid substitution at position 30 is a substitution of a valine (V) for an isoleucine (I), the amino acid substitution at position 165 is a substitution of a serine (S) for a glycine (G), the amino acid substitution at position 282 is a substitution of a valine (V) for a methionine (M), and the amino acid substitution at position 538 is a substitution of a lysine (K) for an asparagine (N). In certain embodiments, the Super

piggyBac™ (SPB) transposase enzyme may comprise or consist of an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MGSSLDDEHI LSALLQSDDE LVGEDSDSEV SDHVSEDDVQ SDTEEAFFIDE VHEVQPTSSG
61  SEILDEQNVI EQPGSSLASN RILTLPORTI RGKNKHCWST SKSTRRSRVS ALNIVRSQRG
121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTSATFRD TNEDEIYAFF
181 GILVMTAVRK DNHMSTDDLF DRSLSMVYVS VMSRDRFDFL IRCLRMDDKS IRPTLRENDV
241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RYIIPNKPSK YGIKILMMCD
301 SGTKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCEFGP LTLVSYKPKP AKMVYLLSSC
421 DEDASINEST GKPQMVMYYN QTKGGVDTLQ QMCSVMTCSR KTNRWPMALL YGMINIACIN
481 SFIIYSHNVS SKGEKVQSRK KFMRLNLYMSL TSSFMRKRLE APTLKRYLRD NISNILPKEV
541 PGTSDDSTEE PVMKKRTYCT YCPSKIRPKA NASCKKCKKV ICREHNIDMC QSCF (SEQ ID NO:
17030) .

```

[030] In certain embodiments of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 3, 46, 82, 103, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 258, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 486, 503, 552, 570 and 591 of the sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 46, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 485, 503, 552 and 570. In certain embodiments, the amino acid substitution at position 3 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an asparagine (N) for a serine (S). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a serine (S) for an alanine (A). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 82 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a tryptophan (W) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position

119 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a proline (P) for an arginine (R). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an alanine (A) a cysteine (C). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a lysine (K) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a histidine (H) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an isoleucine (I) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a valine (V) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 185 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 187 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a glycine (G) for an alanine (A). In certain embodiments, the amino acid substitution at position 200 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a tryptophan (W) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 207 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a proline (P) for a valine (V). In certain embodiments, the amino acid substitution at position 209 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a phenylalanine (F) for a valine (V). In certain embodiments, the amino acid substitution at position 226 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a phenylalanine (F) for a methionine (M). In certain embodiments, the amino acid substitution at position 235 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an arginine (R) for a leucine (L). In certain embodiments, the amino acid substitution at position 240 of SEQ ID NO: 1 or SEQ ID NO: 1 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 241 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 243 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a lysine (K) for a proline (P). In certain embodiments, the amino acid substitution at position 258 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the amino acid substitution at

position 296 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a tryptophan (W) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a tyrosine (Y) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an alanine (A) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an isoleucine (I) for a proline (P). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a valine for a proline (P). In certain embodiments, the amino acid substitution at position 315 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a lysine (K) for an arginine (R). In certain embodiments, the amino acid substitution at position 319 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a glycine (G) for a threonine (T). In certain embodiments, the amino acid substitution at position 327 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an arginine (R) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 328 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a valine (V) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a glycine (G) for a cysteine (C). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 421 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a histidine (H) for the aspartic acid (D). In certain embodiments, the amino acid substitution at position 436 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an isoleucine (I) for a valine (V). In certain embodiments, the amino acid substitution at position 456 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a tyrosine (Y) for a methionine (M). In certain embodiments, the amino acid substitution at position 470 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid substitution at position 485 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a lysine (K) for a serine (S). In certain embodiments, the amino acid substitution at position 503 of SEQ ID

NO: 1 or SEQ ID NO: 2 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 503 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an isoleucine (I) for a methionine (M). In certain embodiments, the amino acid substitution at position 552 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a proline (P) for a glutamine (Q). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an arginine (R) for a glutamine (Q).

[031] In certain embodiments of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBacTM transposase enzyme may comprise or the Super piggyBacTM transposase enzyme may further comprise an amino acid substitution at one or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBacTM transposase enzyme may comprise or the Super piggyBacTM transposase enzyme may further comprise an amino acid substitution at two, three, four, five, six or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBacTM transposase enzyme may comprise or the Super piggyBacTM transposase enzyme may further comprise an amino acid substitution at positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position 194 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 372 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an alanine (A) for an arginine (R). In certain embodiments, the amino acid substitution at position 375 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an alanine (A) for a lysine (K). In certain embodiments, the amino acid substitution at position 450 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of an

asparagine (N) for an aspartic acid (D). In certain embodiments, the amino acid substitution at position 509 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a glycine (G) for a serine (S). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 1 or SEQ ID NO: 2 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 1. In certain embodiments, including those embodiments wherein the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 1, the piggyBac™ transposase enzyme may further comprise an amino acid substitution at positions 372, 375 and 450 of the sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 1, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 1, and a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 1. In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 1, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 1, a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 1 and a substitution of an asparagine (N) for an aspartic acid (D) at position 450 of SEQ ID NO: 1.

[032] In certain embodiments of the disclosure, the transposase enzyme is a Sleeping Beauty transposase enzyme (see, for example, US Patent No. 9,228,180, the contents of which are incorporated herein in their entirety). In certain embodiments, the Sleeping Beauty transposase is a hyperactive Sleeping Beauty (SB100X) transposase. In certain embodiments, the Sleeping Beauty transposase enzyme comprises an amino acid sequence at least 75% identical to:

MGKSKEISQDLRKKIVDLHKSGSSLGAISKRLKVPRSSVQTIVRKYKHHGTTQPSYRS
GRRRYLSPRDERTLVRKVQINPRTTAKDLVKMLEETGTKVSISTVKRVLYRHNKGR
SARKKPLLQNRHKKARLRFATAHGDKDRTFWRNVLWSDETKIELFGHNDHRYVWR
KKGEACKPKNTIPTVKHGGGSIMLWGCFAAGGTGALHKIDGIMRKENYVDILKQHL
KTSVRKLKLGRKWVFQMDNDPKHTSKVVAKWLKDNKVKVLEWPSQSPDLNPIENL
WAELKKRVRARRPTNLTQLHQLCQEEWAKIHPTYCGKLVEGYPKRLTQVKQFKGN
ATKY (SEQ ID NO:17031).

[033] In certain embodiments, including those wherein the Sleeping Beauty transposase is a hyperactive Sleeping Beauty (SB100X) transposase, the Sleeping Beauty transposase enzyme comprises an amino acid sequence at least 75% identical to:

MGKSKEISQDLRKRIVDLHKSGSSLGAISKRLAVPRSSVQTIVRKYKHHGTTQPSYRS
GRRRYLSPRDERTLVRKVQINPRTTAKDLVKMLEETGTKVSISTVKRVLYRHNKLGH
SARKKPLLQNRHKKARLRFATAHGDKDRTFWRNVLWSDETKIELFGHNDHRYVWR
KKGEACKPKNTIPTVKHGGGSIMLWGCFAAGGTGALHKIDGIMDAVQYVDILKQHL
KTSVRKLKLGRKWVFQHDNDPKHTSKVVAKWLKDNKVKVLEWPSQSPDLNPIENL
WAELKKRVRARRPTNLTQLHQLCQEEWAKIHPNYCGKLVEGYPKRLTQVKQFKGN
ATKY (SEQ ID NO: 17032).

[034] In certain embodiments of the compositions of the disclosure, the first transposon and/or the second transposon further comprises a selection gene. In certain embodiments, the selection gene comprises *neo*, DHFR (Dihydrofolate Reductase), TYMS (Thymidylate Synthetase), MGMT (O(6)-methylguanine-DNA methyltransferase), multidrug resistance gene (MDR1), ALDH1 (Aldehyde dehydrogenase 1 family, member A1), FRANCF, RAD51C (RAD51 Paralog C), GCS (glucosylceramide synthase), NKX2.2 (NK2 Homeobox 2) or any combination thereof. In certain embodiments, the selection gene comprises DHFR.

[035] In certain embodiments of the compositions of the disclosure, the first transposon and or the second transposon comprises an inducible caspase polypeptide comprising (a) a ligand binding region, (b) a linker, and (c) a truncated caspase 9 polypeptide, wherein the inducible caspase polypeptide does not comprise a non-human sequence. In certain embodiments, the non-human sequence is a restriction site. In certain embodiments, the ligand binding region inducible caspase polypeptide comprises a FK506 binding protein 12 (FKBP12) polypeptide. In certain embodiments, the amino acid sequence of the FK506 binding protein 12 (FKBP12) polypeptide comprises a modification at position 36 of the sequence. In certain embodiments, the modification is a substitution of valine (V) for phenylalanine (F) at position 36 (F36V). In certain embodiments, the FKBP12 polypeptide is encoded by an amino acid sequence comprising

GVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKVDSSRDRNKPFFKFMLGKQEVI
RGWEEGVAQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELLKLE (SEQ ID
NO: 17012). In certain embodiments, the FKBP12 polypeptide is encoded by a nucleic acid sequence comprising

GGGGTCCAGGTCGAGACTATTTACCAGGGGATGGGCGAACATTTCCAAAAAGG

GGCCAGACTTGCGTCGTGCATTACACCGGGATGCTGGAGGACGGGAAGAAAGTG
 GACAGCTCCAGGGATCGCAACAAGCCCTTCAAGTTCATGCTGGGAAAGCAGGAA
 GTGATCCGAGGATGGGAGGAAGGCGTGGCACAGATGTCAGTCGGCCAGCGGGCC
 AAAGTGACCATTAGCCCTGACTACGCTTATGGAGCAACAGGCCACCCAGGGATC
 ATTCCCCCTCATGCCACCCTGGTCTTCGAT GTGGAAGCTGCTGAAGCTGGAG (SEQ
 ID NO: 17013).

[036] In certain embodiments, the linker region of the inducible proapoptotic polypeptide is encoded by an amino acid comprising GGGGS (SEQ ID NO: 17014). In certain embodiments, the linker region of the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising GGAGGAGGAGGATCC (SEQ ID NO: 17015).

[037] In certain embodiments, the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid sequence that does not comprise an arginine (R) at position 87 of the sequence. In certain embodiments, the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid sequence that does not comprise an alanine (A) at position 282 the sequence. In certain embodiments, the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid comprising

GFGDVGALES LRGNADLAYILSMEPCGHCLIINN VNFCRESGLRTRTGSNIDCEKLRR
 RFSSLHFMVEVKGDLTAKKMVLALLELAQQDHGALDCCVVVILSHGCQASHLQFPG
 AVYGTDGCPVSVEKIVNIFNGTSCPSLGGKPKLFFIQACGGEQKDHGFEVASTSPEDE
 SPGSNPEPDATPFQEGLRTFDQLDAISSLPTPSDIFVSYSTFPGFVSWRDPKSGSWYVE
 TLDDIFEQWAHSEDLQSLLLRVANAVSVKGIYKQMPGCFNFLRKKLFFKTS (SEQ ID
 NO: 17016). In certain embodiments, the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising

TTTGGGGACGTGGGGGGCCCTGGAGTCTCTGCGAGGAAATGCCGATCTGGCTTAC
 ATCCTGAGCATGGAACCCTGCGGCCACTGTCTGATCATTAACAATGTGAACTTCT
 GCAGAGAAAGCGGACTGCGAACACGGACTGGCTCCAATATTGACTGTGAGAAGC
 TGCGGAGAAGGTTCTCTAGTCTGCACTTTATGGTCTGAAGTGAAAGGGGATCTGAC
 CGCCAAGAAAATGGTGCTGGCCCTGCTGGAGCTGGCTCAGCAGGACCATGGAGC
 TCTGGATTGCTGCGTGGTCGTGATCCTGTCCCACGGGTGCCAGGCTTCTCATCTG
 CAGTTCCCCGGAGCAGTGTACGGAACAGACGGCTGTCCTGTCAGCGTGGAGAAG
 ATCGTCAACATCTTCAACGGCACTTCTTGCCCTAGTCTGGGGGGGAAAGCCAAAAC
 TGTTCTTTATCCAGGCCTGTGGCGGGGAACAGAAAGATCACGGCTTCGAGGTGG

CCAGCACCCAGCCCTGAGGACGAATCACCAGGGAGCAACCCTGAACCAGATGCAA
CTCCATTCCAGGAGGGACTGAGGACCTTTGACCAGCTGGATGCTATCTCAAGCCT
GCCCCACTCCTAGTGACATTTTCGTGTCTTACAGTACCTTCCCAGGCTTTGTCTCAT
GGCGCGATCCCAAGTCAGGGAGCTGGTACGTGGAGACACTGGACGACATCTTTG
AACAGTGGGCCCATTTCAGAGGACCTGCAGAGCCTGCTGCTGCGAGTGGCAAACG
CTGTCTCTGTGAAGGGCATCTACAAACAGATGCCCGGGTGCTTCAATTTTCTGAG
AAAGAAACTGTTCTTTAAGACTTCC (SEQ ID NO: 17017).

[038] In certain embodiments, the inducible proapoptotic polypeptide is encoded by an amino acid sequence comprising

GVQVETISPGDGRTPFKRGQTCVVHYTGMLDGGKKVDSSRDRNKPFFKMLGKQEVI
RGWEEGVAQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELLKLEGGGGS
GFGDVGALESLRGNADLAYILSMEPCGHCLIINNVNFCRESGLRTRTGSNIDCEKLRR
RFSSLHFMVEVKGDLTAKKMVLALLELAQQDHGALDCCVVVILSHGCQASHLQFPG
AVYGTGDCPVSVEKIVNIFNGTSCPSLGGKPKLFFIQACGGEQKDHGFEVASTSPEDE
SPGSNPEPDATPFQEGLRTFDQLDAISSLPTPSDIFVSYSTFPGFVSWRDPKSGSWYVE
TLDDIFEQWAHSEDLQSLLLRVANAVSVKGIYKQMPGCFNFLRKKLFFKTS (SEQ ID

NO: 17018). In certain embodiments, the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising

Ggggtccaggtcgagactatlcaccaggggatggggaacattlccaaaagggggccagactlgeglcglgcattacaccgggat
gctggaggacgggaagaaagtgacagctccaggatcgcaacaagccctcaagttcatgctgggaaagcaggaagtatccga
ggatgggaggaaggcgtggcacagatgtcagtcggccagcgggccaactgaccattagccctgactacgcttatggagcaacag
gccaccagggatcatlccccctcatgccacctggtcttcgatgtggaactgctgaagctggaggaggaggaggatccggattg
gggacgtgggggcccctggagtcctcgcgaggaaalgccgatctggcttacatcctgagcatggaacctgcggccactgtctgatcal
laacaatgtgaacttctgcagagaaagcggactgcgaacacggactggctccaatattgactgtgagaagctgcggagaagggtctct
agtctgcactttatggtcgaagtgaaggggatctgaccgccaagaaaatgggtgctggccctgctggagctggctcagcaggacat
ggagctctggattgctgcgtggtcgtgatccgtcccacgggtgccaggcttctcatctgcagttccccggagcagtgtagcgaacag
acggctgtcctgtcagcgtggagaagatcgtaacatcttcaacggcacttcttgcctagctctggggggaagccaaaactgtcttta
tccaggcctgtggcggggaacagaaagatcacggcttcgaggtggccagcaccagccctgaggacgaatcaccaggagcaacc
ctgaaccagatgcaactccattccaggaggactgaggaccttgaccagctggatgctatctcaagcctgcccactcctagtgacatt
ttcgtgtcttacagtaccttcccaggcttctctcatggcgcatccaagtcaggagctggtacgtggagacactggacgacatctt
gaacagtgggcccattcagaggacctgcagagcctgctgctgcgagtggcaaacgctgtctctgtgaaggcatctacaacagatg
cccgggtgcttcaatttctgagaaagaaactgttctttaagacttcc (SEQ ID NO: 17019).

[039] In certain embodiments of the compositions of the disclosure, the first transposon and/or the second transposon comprises at least one self-cleaving peptide. In certain embodiments, the at least one self-cleaving peptide comprises a T2A peptide, a GSG-T2A peptide, an E2A peptide, a GSG-E2A peptide, an F2A peptide, a GSG-F2A peptide, a P2A peptide, or a GSG-P2A peptide. In certain embodiments, the at least one self-cleaving peptide comprises a T2A peptide. In certain embodiments, the T2A peptide comprises an amino acid sequence comprising EGRGSLLTCGDVEENPGP (SEQ ID NO: 17020). In certain embodiments, the GSG-T2A peptide comprises an amino acid sequence comprising GSGEGRGSLLTCGDVEENPGP (SEQ ID NO: 17021). In certain embodiments, the E2A peptide comprises an amino acid sequence comprising QCTNYALLKLAGDVESNPGP (SEQ ID NO: 17022). In certain embodiments, the GSG-E2A peptide comprises an amino acid sequence comprising GSGQCTNYALLKLAGDVESNPGP (SEQ ID NO: 17023). In certain embodiments, the F2A peptide comprises an amino acid sequence comprising VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 17024). In certain embodiments, the GSG-F2A peptide comprises an amino acid sequence comprising GSGVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 17025). In certain embodiments, the P2A peptide comprises an amino acid sequence comprising ATNFSLLKQAGDVEENPGP (SEQ ID NO: 17026). In certain embodiments, the GSG-P2A peptide comprises an amino acid sequence comprising GSGATNFSLLKQAGDVEENPGP (SEQ ID NO: 17027). In certain embodiments, the at least one self-cleaving peptide is positioned between (a) the selection gene and the inducible transgene construct or (b) the inducible transgene construct and the inducible caspase polypeptide. In certain embodiments, the at least one self-cleaving peptide is positioned between (a) the selection gene and the reporter construct or (b) the reporter construct and the inducible caspase polypeptide.

[040] The disclosure provides a cell comprising the composition of the disclosure.

[041] The disclosure provides a method of inducing conditional gene expression in a cell comprising (a) contacting the cell with a composition of the disclosure, under conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter and (b) contacting the exogenous receptor and a ligand that specifically binds thereto, to transduce an intracellular signal that targets the inducible promoter, thereby modifying gene expression. In certain embodiments, the cell is in vivo, ex vivo, in vitro or in situ. In certain embodiments, the cell is an immune cell. In certain embodiments, the immune cell is a T-cell, a Natural Killer (NK) cell, a

Natural Killer (NK)-like cell, a hematopoietic progenitor cell, a peripheral blood (PB) derived T cell or an umbilical cord blood (UCB) derived T-cell. In certain embodiments, the immune cell is a T-cell. In certain embodiments, the cell is autologous. In certain embodiments, the cell is allogeneic.

[042] The disclosure provides a method of treating a disease or disorder in a subject in need thereof, comprising administering to the subject a composition of the disclosure, under conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter, and administering a ligand to which the exogenous receptor selectively binds, wherein the binding of the ligand to the exogenous receptor transduces an intracellular signal to target the inducible promoter controlling the transgene, wherein the transgene is expressed, and wherein the product of the transgene is therapeutically-effective for treating the disease or disorder. In certain embodiments, the product of the transgene is a secreted protein. In certain embodiments, the secreted protein is a clotting factor. In certain embodiments, the clotting factor is factor IX. In certain embodiments, the disease or disorder is a clotting disorder.

[043] In certain embodiments of the methods of the disclosure, conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter comprise in vivo conditions. In certain embodiments, conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter comprise a temperature substantially similar to an internal temperature of a human body, a CO₂ level substantially similar to an internal CO₂ levels of a human body, an O₂ level substantially similar to an internal O₂ levels of a human body, an aqueous or saline environment with a level of electrolytes substantially similar to a level of electrolytes of an interior of a human body.

[044] In certain embodiments of the compositions and methods of the disclosure, the ligand to which the exogenous receptor specifically binds is non-naturally occurring. In certain embodiments, the ligand is a nucleic acid, an amino acid, a polymer, an organic small molecule, an inorganic small molecule, or a combination thereof. Exemplary ligands include, but are not limited to, synthetic, modified, recombinant, mutant, chimeric, endogenous or non-naturally occurring, proteins (soluble or membrane-bound), steroid hormones, gas particles, nucleic acids, growth factors, neurotransmitters, vitamins, and minerals.

[045] The disclosure provides a composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a ligand construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous ligand, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous ligand is expressed, and wherein the exogenous ligand, upon binding a receptor, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression. In certain embodiments, the ligand comprises a non-natural or synthetic sequence. In certain embodiments, the ligand comprises a fusion protein. In certain embodiments, the ligand is bound to the surface of the cell. In certain embodiments, the ligand comprises an intracellular domain. In certain embodiments, the intracellular domain transduces a signal in the cell expressing the ligand. In certain embodiments, the structure of the ligand is substantially similar to the structure of the receptor of the compositions of the disclosure. In certain embodiments, the signal transduced by the ligand and the signal transduced by the receptor comprise a bi-directional signal.

BRIEF DESCRIPTION OF THE DRAWINGS

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

[047] Figure 1A-B is a pair of schematic diagrams depicting NF-KB inducible vectors for expression in T-cells. Two T cell activation NF-KB inducible vectors were developed; one with the gene expression system (GES) in the forward orientation (A) and the other in the complementary direction (B), both preceding the constitutive EF1a promoter. These vectors also direct expression of a CAR molecule and a DHFR selection gene, separated by a T2A sequence. Both the conditional NF-KB inducible system and the EF1a directed genes are a part of a piggyBac transposon which can be permanently integrated into T cells using electroporation (EP). Once integrated into the genome, the T cells will constitutively express the CAR on the membrane surface and the DHFR within the cell, while expression of the NF-KB inducible gene, GFP, will be expressed to the highest level only upon T cell activation.

[048] Figure 2 is a pair of graphs depicting NF-KB inducible expression of GFP in activated T cells. T cells were nucleofected with a piggyBac vector expressing an anti-

BCMA CAR and a DHFR mutein gene under control of an EF1a promoter along with the absence (No GES control) or presence of an NF-KB inducible expression system driving GFP expression in either the forward (pNFkB-GFP forward) or reverse orientation (pNFkB-GFP reverse). Cells were cultured in the presence of methotrexate selection until the cells were almost completely resting (Day 19) and GFP expression was assessed at Day 5 and Day 19. At Day 5, all T cells are proliferating and highly stimulated, with cells harboring the NF-KB inducible expression cassette producing high levels of GFP due to strong NFkB activity. The No GES control cells did not express detectable levels of GFP. By Day 19, the GES T cells were almost fully resting and GFP expression was significantly lower than Day 5 (~1/8 MFI), since NFkB activity is lower. GFP expression is still observed at Day 19, which may be due to the long half-life of GFP protein (~30hr), or, basal level of NFkB activity through, for example, a TCR, a CAR, a cytokine receptor, or a growth factor receptor signal.

[049] Figure 3 is a series of graphs depicting anti-BCMA CAR-mediated activation of NF-KB inducible expression of GFP in presence of BCMA+ tumor cells. T cells were either unmodified (Mock T cells) or nucleofected with a piggyBac vector expressing an anti-BCMA CAR and a DHFR mutein gene under control of an EF1a promoter along with the absence (No GES control) or presence of an NF-KB inducible expression system driving GFP expression in either the forward (pNFkB-GFP forward) or reverse orientation (pNFkB-GFP reverse). All cells were cultured for 22 days, either with or without methotrexate selection (Mock T cells), until the cells were almost completely resting. Cells were then stimulated for 3 days in the absence (No stimulation) or presence of BCMA- (K562), BMCA+ (RPMI 8226), or positive control anti-CD3 anti-CD28 activation reagent (CD3/28 stimulation). GFP expression was undetectable under all conditions with the No GES control or Mock T cells. However, while pNFkB-GFP forward- and reverse-transposed cells exhibited little GFP expression over the No stimulation control when cultured with BCMA- K562 cells, they both demonstrated dramatic upregulation of gene expression either in the presence of BCMA+ tumor cells or under positive control conditions. Little difference in GFP expression was observed between the pNFkB-GFP forward- and reverse-transposed cells that were cocultured with BCMA+ tumor cells.

[050] Figure 4 is a series of graphs demonstrating that the Expression level of inducible gene can be regulated by number of response elements preceding the promoter T cells were nucleofected with a piggyBac vector encoding an anti-BCMA CARTyrin followed by a selection gene, both under control of a human EF1a promoter. Further, vectors either

additionally encoded the conditional NF-KB inducible gene expression system driving expression of a truncated CD19 protein (dCD19) and included a number of NFkB response elements (RE) varying from 0 - 5, no GES (No GES), or received an electroporation pulse but no piggyBac nucleic acid (Mock). Data are shown for only the GES in the reverse (opposite) direction/orientation. All cells were cultured for 18 days and included selection for piggyBac-modified T cells using methotrexate addition. Cells were then stimulated for 3 days using anti-CD3 anti-CD28 bead activation reagent and dCD19 surface expression was assessed by FACS at Days 0, 3 and 18, and data are shown as FACS histograms and MFI of target protein staining. Surface dCD19 expression was detected at low levels at Day 0 in all T cells transposed with vectors encoding the GES. At 3 days post-stimulation, dramatic upregulation of dCD19 expression was observed for all T cells expressing the GES, with a greater fold increase in surface expression in those with higher numbers of REs. Thus, surface dCD19 expression was directly proportional with the number of REs encoded in the GES. No dCD19 was detected on the surface of T cells that did not harbor the GES: No GES and Mock controls.

[051] Figure 5 is a schematic diagram showing the human coagulation pathway leading to blood clotting. Contact activation, for example by damaging an endothelium, activates an intrinsic clotting pathway. Tissue factors activate an extrinsic clotting pathway, for example following trauma. Both pathways converge onto the conversion of Prothrombin into Thrombin, which catalyzes the conversion of fibrinogen into fibrin. Polymerized fibrin together with platelets forms a clot. In the absence of Factor IX (circled), clotting is defective. Factor VIII (FVIII) deficiency leads to development of Hemophilia A. Factor IX (FIX) deficiency leads to development of Hemophilia B. Prior to the compositions and methods of the disclosure, the standard treatment for hemophilia B involved an infusion of recombinant FIX every 2 to 3 days, at an expense of approximately \$250,000 per year. In sharp contrast to this standard treatment option, T cells of the disclosure are maintained in humans for several decades.

[052] Figure 6 is a series of Fluorescence-Activated Cell Sorting (FACS plots) depicting FIX-secreting T cells. T cells encoding a human Factor IX transgene showed a T-cell phenotype in approximately 80% of cells. The 6 panels are described in order from left to right. (1) Forward scatter (FSC) on the x-axis versus side scatter (SSC) on the y-axis. The x-axis is from 0 to 250 thousand (abbreviated k) in increments of 50k, the y-axis is for 0 to 250k, in increments of 50k. (2) FSC on the x-axis versus the cell viability marker 7

aminoactinomycin D (7AAD). The x-axis is labeled from 0 to 250k in increments of 50k. The y-axis reads, from top to bottom, -10^3 , 0, 10^3 , 10^4 , 10^5 . (3) On the x-axis is shown anti-CD56-APC conjugated to a Cy7 dye (CDC56-APC-Cy7), units from 0 to 10^5 incrementing in powers of 10. On the y-axis is shown anti-CD3 conjugated to phycoerythrin (PE), units from 0 to 10^5 incrementing in powers of 10. (4) On the x-axis is shown anti-CD8 conjugated to fluorescein isothiocyanate (FITC), units from 0 to 10^5 incrementing in powers of 10. On the y-axis is shown anti-CD4 conjugated to Brilliant Violet 650 dye (BV650), units from 0 to 10^5 incrementing in powers of 10. (5) On the x-axis is shown an anti CD62L antibody conjugated to a Brilliant Violet 421 dye (BV421), units from 0 to 10^5 incrementing in powers of 10. On the y-axis is shown an anti-CD45RA antibody conjugated to PE and Cy7, units from 0 to 10^5 incrementing in powers of 10. This panel is boxed. (6) On the x-axis is shown an anti-CCR7 antibody conjugated to Brilliant Violet 786 (BV786), units from 0 to 10^5 incrementing in powers of 10. On the y-axis is shown anti-CD45RA conjugated to PE and Cy7, units from 0 to 10^5 incrementing in powers of 10.

[053] Figure 7A is a graph showing human Factor IX secretion during production of modified T cells of the disclosure. On the y-axis, Factor IX concentration in nanograms (ng) per milliliter (mL) from 0 to 80 in increments of 20. On the x-axis are shown 9 day and 12 day T cells.

[054] Figure 7B is a graph showing the clotting activity of the secreted Factor IX produced by the T cells. On the y-axis is shown percent Factor IX activity relative to human plasma, from 0 to 8 in increments of 2. On the x-axis are 9 and 12 day T cells.

[055] Figure 8 is a series of graphs demonstrating that the expression level of inducible gene can be regulated by number of response elements preceding the promoter in CD4 positive cells. Truncated CD19 (dCD19) expressing CAR-T cells were stimulated by BCMA+ H929 multiple myeloma cells at 2:1 CAR-T: H929 ratio. The expression of dCD19 was driven by the minimal promoter that enhanced by 0, 1, 2, 3, 4 or 5 repeats of the NF- κ B response element. The expression of BCMA CAR was driven by human elongation factor-1 α (EF-1 α) promoter, a constitutive promoter that is commonly used for gene expression in human T cells. Before tumor cell stimulation, the expression of CAR and dCD19 were both at basal levels compared to mock T cell control. The expression levels of CAR and dCD19 were both upregulated upon tumor stimulation (day 3) and then subsequently downregulated (day 9, 14) and eventually reached their respective basal levels when the cells resume a fully rested status again (day 20). However, CAR surface expression was equivalently up- or

down-regulated in all the CAR-T cell samples during cell activation and resting process, while the expression levels of dCD19 were directly proportional to the number of NF- κ B response elements (day 3, 9, 14). Data are shown as FACS histograms and MFI of target protein staining. Thus, surface dCD19 expression was directly proportional with the number of REs encoded in the GES. No dCD19 was detected on the surface of T cells that did not harbor the GES: No GES and Mock controls.

[056] Figure 9 is a series of graphs demonstrating that the expression level of inducible gene can be regulated by number of response elements preceding the promoter in CD8 positive cells. Truncated CD19 (dCD19) expressing CAR-T cells were stimulated by BCMA+ H929 multiple myeloma cells at 2:1 CAR-T: H929 ratio. The expression of dCD19 was driven by the minimal promoter that enhanced by 0, 1, 2, 3, 4 or 5 repeats of the NF- κ B response element. The expression of BCMA CAR was driven by human elongation factor-1 α (EF-1 α) promoter, a constitutive promoter that is commonly used for gene expression in human T cells. Before tumor cell stimulation, the expression of CAR and dCD19 were both at basal levels compared to mock T cell control. The expression levels of CAR and dCD19 were both upregulated upon tumor stimulation (day 3) and then subsequently downregulated (day 9, 14) and eventually reached their respective basal levels when the cells resume a fully rested status again (day 20). However, CAR surface expression was equivalently up- or down-regulated in all the CAR-T cell samples during cell activation and resting process, while the expression levels of dCD19 were directly proportional to the number of NF- κ B response elements (day 3, 9, 14). Data are shown as FACS histograms and MFI of target protein staining. Thus, surface dCD19 expression was directly proportional with the number of REs encoded in the GES. No dCD19 was detected on the surface of T cells that did not harbor the GES: No GES and Mock controls.

[057] Figure 10 is a bar graph depicting the knock out efficiency of targeting various checkpoint signaling proteins that could be used to armor T-cells. Cas-CLOVER was used to knockout the checkpoint receptors, PD-1, TGFBR2, LAG-3, TIM-3 and CTLA-4 in resting primary human pan T cells. Percent knock-out is shown on the y-axis. Gene editing resulted in 30-70% loss of protein expression at the cell surface as measured by flow cytometry.

[058] Figure 11 is a series of schematic diagrams of wildtype, null and switch receptors and their effects on intracellular signaling, either inhibitory or stimulatory, in primary T-cells. Binding of the wildtype inhibitory receptor expressed endogenously on a T-cell with its endogenous ligand results in transmission of an inhibitory signal which, in part, reduces T-

cell effector function. However, mutation (Mutated null) or deletion (Truncated null) of the intracellular domain (ICD) of a checkpoint receptor protein, such as PD1 (top panel) or TGFBR2 (bottom panel), reduces or eliminates its signaling capability when cognate ligand(s) is bound. Thus, expression of engineered mutated or truncated null receptors on the surface of modified T cells results in a competition with endogenously-expressed wildtype receptors for binding of the free endogenous ligand(s), effectively reducing or eliminating delivery of inhibitory signals by endogenously-expressed wildtype receptors. Specifically, any binding by a mutated or null receptor sequesters the endogenous ligand(s) from binding the wildtype receptor and results in dilution of the overall level of checkpoint signaling effectively delivered to the modified T-cell, thereby reducing or blocking checkpoint inhibition and functional exhaustion of the modified T cells. A switch receptor is created by replacement of the wildtype ICD with an ICD from either a co-stimulatory molecule (such as CD3z, CD28, 4-1BB) or a different inhibitory molecule (such as CTLA4, PD1, Lag3). In the former case, binding of the endogenous ligand(s) by the modified switch receptor results in the delivery of a positive signal to the T-cells, thereby helping to enhance stimulation of the modified T cell and potentially enhance target tumor cell killing. In the latter case, binding of the endogenous ligand(s) by the modified switch receptor results in the delivery of a negative signal to the T-cells, thereby eliminating stimulation of the modified T cell and potentially reducing target tumor cell killing. The signal peptide (purple arrow), extracellular domain (ECD) (bright green), transmembrane domain (yellow), intracellular signaling domain (ICD)(orange), and replacement ICD (green) are displayed in the receptor diagrams. “*” indicates a mutated ICD. “+” indicates the presence of a checkpoint signal. “-” indicates the absence of a checkpoint signal.

[059] Figure 12 is a schematic diagram showing an example of the design of null receptors with specific alterations that may help to increase expression of the receptor on the surface of modified T cells. Examples are shown for PD1 and TGFBR2 null receptors and the signal peptide domain (SP), transmembrane domain (TM) and extracellular domain (ECD) of truncated null receptors for PD1 (top panel) and TGFBR2 (bottom panel) are displayed. The first of the top four molecules is the wildtype PD-1 receptor, which encodes the wildtype PD-1 SP and TM. For the PD1 null receptor, replacement of PD1 wildtype SP or TM domain (green; light green) with the SP or TM domain of a human T cell CD8a receptor (red) is depicted. The second molecule encodes the CD8a SP along with the native PD-1 TM, the third encodes the wildtype PD-1 SP and the alternative CD8a TM, and the fourth encodes

both the alternative CD8a SP and TM. Similarly, for the null receptor of TGF β RII, replacement of the wildtype TGFBR II SP (pink) with a SP domain of a human T cell CD8a receptor (red). The names of the constructs and the amino acid lengths (aa) of each construct protein is listed on the left of the diagram.

[060] Figure 13 is a series of histograms depicting the expression of the PD1 and TGFBR II null Receptors on the surface of modified primary human T cells as determined by flow cytometry. Each of the six truncated null constructs from Figure 12 were expressed on the surface of primary human T cells. T cells were stained with either anti-PD1 (top; blue histograms) or anti-TGF β R II (bottom; blue histograms), or isotype control or secondary only (gray histograms). Cells staining positive for PD-1 or TGF β R II expression were gated (frequency shown above gate) and mean fluorescence intensity (MFI) value is displayed above each positive histogram. The names of the null receptor constructs are depicted above each plot. Both null receptor gene strategies, replacement of the wildtype SP with the alternative CD8a were successfully expressed. 02.8aSP-PD-1 and 02.8aSP-TGF β R II resulted in the highest level of expression at the T-cell surface. 02.8aSP-PD-1 null receptor exhibited an MFI of 43,680, which is 177-fold higher than endogenous T cell PD-1 expression and 2.8-fold higher than the wildtype PD-1 null receptor. 02.8aSP-TGF β R II null receptor exhibited an MFI of 13,809, which is 102-fold higher than endogenous T cell TGF β R II expression and 1.8-fold higher than the wildtype TGF β R II null receptor. Replacement of wildtype SP with the alternative CD8a SP for both PD1 and TGRBR II results in enhanced surface expression of the null or Switch receptor, which may help to maximize reduction or blockage of checkpoint inhibition upon binding and sequestration of the endogenous ligand(s).

[061] Figure 14 is a schematic depiction of the Csy4-T2A-Clo051-G4Slinker-dCas9 construct map (Embodiment 2).

[062] Figure 15 is a schematic depiction of the pRT1-Clo051-dCas9 Double NLS construct map (Embodiment 1).

[063] Figure 16 is a pair of graphs comparing the efficacy of knocking out expression of either B2M on the surface of Pan T-cells (left) or the α -chain of the T-cell Receptor on the surface of Jurkat cells (right) for either Embodiment 1 (pRT1-Clo051-dCas9 Double NLS, as shown in Figure 15) or Embodiment 2 (Csy4-T2A-Clo051-G4Slinker-dCas9, as shown in Figure 14) of a Cas-Clover fusion protein of the disclosure. For the right-hand graph, the fusion protein is provided at either 10 μ g or 20 μ g, as indicated.

[064] Figure 17 is a photograph of a gel electrophoresis analysis of mRNA encoding each of Embodiment 1 (Lane 2; pRT1-Clo051-dCas9 Double NLS, as shown in Figure 15) or Embodiment 2 (Lane 3; Csy4-T2A-Clo051-G4Slinker-dCas9, as shown in Figure 14). In addition, a previous preparation (“old version”) of mRNA encoding Embodiment 2 is included (Lane 4) for comparison. As shown, all mRNA samples encoding the two different embodiments migrate as distinct bands within the gel, are of high quality, and are similar in size, as expected.

DETAILED DESCRIPTION

[065] The disclosure provides a composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous reporter is expressed, and wherein the exogenous reporter, upon binding a ligand, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression.

Exogenous Receptors

[066] Exogenous receptors of the disclosure may comprise a non-naturally occurring receptor. In certain embodiments, the non-naturally occurring receptor is a synthetic, modified, recombinant, mutant or chimeric receptor. In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a T-cell receptor (TCR). In certain embodiments, the non-naturally occurring receptor comprises one or more sequences isolated or derived from a scaffold protein. In certain embodiments, the non-naturally occurring receptor comprises a transmembrane domain. In certain embodiments, the non-naturally occurring receptor interacts with an intracellular receptor that transduces an intracellular signal. In certain embodiments, the non-naturally occurring receptor comprises an intracellular signaling domain. In certain embodiments, the non-naturally occurring receptor is a chimeric ligand receptor (CLR). In certain embodiments, the CLR is a chimeric antigen receptor.

[067] In certain embodiments of the compositions of the disclosure, the exogenous receptor of (b) comprises a non-naturally occurring receptor. In certain embodiments, the CLR is a chimeric antigen receptor. In certain embodiments, the chimeric ligand receptor comprises (a) an ectodomain comprising a ligand recognition region, wherein the ligand

recognition region comprises at least scaffold protein; (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain.

[068] The disclosure provides chimeric receptors comprising at least one Centyrin. Chimeric ligand/antigen receptors (CLRs/CARs) of the disclosure may comprise more than one Centyrin, referred to herein as a CARTyrin.

[069] The disclosure provides chimeric receptors comprising at least one VHH. Chimeric ligand/antigen receptors (CLRs/CARs) of the disclosure may comprise more than one VHH, referred to herein as a VCAR.

[070] Chimeric receptors of the disclosure may comprise a signal peptide of human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR. A hinge/spacer domain of the disclosure may comprise a hinge/spacer/stalk of human CD8 α , IgG4, and/or CD4. An intracellular domain or endodomain of the disclosure may comprise an intracellular signaling domain of human CD3 ζ and may further comprise human 4-1BB, CD28, CD40, ICOS, MyD88, OX-40 intracellular segment, or any combination thereof. Exemplary transmembrane domains include, but are not limited to a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR transmembrane domain.

[071] The disclosure provides genetically modified cells, such as T cells, NK cells, hematopoietic progenitor cells, peripheral blood (PB) derived T cells (including T cells from G-CSF-mobilized peripheral blood), umbilical cord blood (UCB) derived T cells rendered specific for one or more ligands or antigens by introducing to these cells a CLR/CAR, CARTyrin and/or VCAR of the disclosure. Cells of the disclosure may be modified by electrotransfer of a transposon of the disclosure and a plasmid or a nanoplasmid comprising a sequence encoding a transposase of the disclosure (preferably, the sequence encoding a transposase of the disclosure is an mRNA sequence).

[072] In some embodiments, the armored T-cell comprises a composition comprising (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor, such as a CLR or CAR, wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous receptor is expressed, and wherein the exogenous receptor, upon binding a ligand or antigen, transduces an intracellular signal that targets directly or indirectly the inducible promoter regulating expression of the inducible transgene (a) to modify gene expression.

Chimeric Receptors

[073] Chimeric antigen receptors (CARs) and/or chimeric ligand receptors (CLRs) of the disclosure may comprise (a) an ectodomain comprising an antigen/ligand recognition region, (b) a transmembrane domain, and (c) an endodomain comprising at least one costimulatory domain. In certain embodiments, the ectodomain may further comprise a signal peptide. Alternatively, or in addition, in certain embodiments, the ectodomain may further comprise a hinge between the antigen/ligand recognition region and the transmembrane domain. In certain embodiments of the CARs of the disclosure, the signal peptide may comprise a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR signal peptide. In certain embodiments of the CARs of the disclosure, the signal peptide may comprise a sequence encoding a human CD8 α signal peptide. In certain embodiments, the transmembrane domain may comprise a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR transmembrane domain. In certain embodiments of the CARs of the disclosure, the transmembrane domain may comprise a sequence encoding a human CD8 α transmembrane domain. In certain embodiments of the CARs/CLRs of the disclosure, the endodomain may comprise a human CD3 ζ endodomain.

[074] In certain embodiments of the CARs/CLRs of the disclosure, the at least one costimulatory domain may comprise a human 4-1BB, CD28, CD40, ICOS, MyD88, OX-40 intracellular segment, or any combination thereof. In certain embodiments of the CARs of the disclosure, the at least one costimulatory domain may comprise a CD28 and/or a 4-1BB costimulatory domain. In certain embodiments of the CARs of the disclosure, the hinge may comprise a sequence derived from a human CD8 α , IgG4, and/or CD4 sequence. In certain embodiments of the CARs/CLRs of the disclosure, the hinge may comprise a sequence derived from a human CD8 α sequence.

[075] The CD28 costimulatory domain may comprise an amino acid sequence comprising RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPPEMGGKPRRKNPQ EGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALP PR (SEQ ID NO: 17004) or a sequence having at least 70%, 80%, 90%, 95%, or 99% identity to the amino acid sequence comprising RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPPEMGGKPRRKNPQ EGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALP PR (SEQ ID NO: 17004). The CD28 costimulatory domain may be encoded by the nucleic

acid sequence comprising

cgcgtagagtttagtcgacagcagatgccccagcttacaacaggagacagaaccagctgtataacgagctgaatctgggcccgcga
gaggaatatgacgtgctggataagcggagaggacgcgaccccgaaatgggaggcaagcccaggcgcaaaaaccctcaggaagg
cctgtataacgagctgcagaaggacaaaatggcagaagcctattctgagatcggcatgaagggggagcgacggagaggcaaagg
gcacgatgggctgtaccagggactgagcaccgccacaaaggacacctatgatgctctgcatatgcaggcactgcctccaagg

(SEQ ID NO: 17005). The 4-1BB costimulatory domain may comprise an amino acid sequence comprising KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEEGGCEL (SEQ ID NO: 17006) or a sequence having at least 70%, 80%, 90%, 95%, or 99% identity to the amino acid sequence comprising

KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEEGGCEL (SEQ ID NO: 17006).

The 4-1BB costimulatory domain may be encoded by the nucleic acid sequence comprising aagagaggcaggaagaaaactgctgtataatttcaaacagcccttcacgcgccccgtgcagactaccagagggaagacgggtgctcc
tgtcgattccctgaggaagaggaaggcgggtgtgagctg (SEQ ID NO: 17007). The 4-1BB costimulatory domain may be located between the transmembrane domain and the CD28 costimulatory domain.

[076] In certain embodiments of the CARs/CLRs of the disclosure, the hinge may comprise a sequence derived from a human CD8 α , IgG4, and/or CD4 sequence. In certain embodiments of the CARs/CLRs of the disclosure, the hinge may comprise a sequence derived from a human CD8 α sequence. The hinge may comprise a human CD8 α amino acid sequence comprising TTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD (SEQ ID NO: 17008) or a sequence having at least 70%, 80%, 90%, 95%, or 99% identity to the amino acid sequence comprising

TTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD (SEQ ID NO: 17008).

The human CD8 α hinge amino acid sequence may be encoded by the nucleic acid sequence comprising

actaccacaccagcacctagaccaccaactccagctccaaccatcgcgagtcagccctgagtctgagacctgaggcctgcaggcc
agctgcaggaggagctgtgcacaccaggggctggacttcgcctgcgac (SEQ ID NO: 17028).

ScFv

[077] The disclosure provides single chain variable fragment (scFv) compositions and methods for use of these compositions to recognize and bind to a specific target protein. ScFv compositions comprise a heavy chain variable region and a light chain variable region of an antibody. ScFv compositions may be incorporated into an antigen/ligand recognition region of a CAR or CLR of the disclosure. An antigen/ligand recognition region of a CAR or CLR

of the disclosure may comprise an ScFv or an ScFv composition of the disclosure. In some embodiments, ScFvs comprise fusion proteins of the variable regions of the heavy (VH) and light (VL) chains of an immunoglobulin, wherein the VH and VL domains are connected with a linker. ScFvs retain the specificity of the original immunoglobulin, despite removal of the constant regions and the introduction of the linker. An exemplary linker comprises a sequence of GGGGSGGGGSGGGGS (SEQ ID NO: 17033).

Centyrins

[078] Centyrins of the disclosure specifically bind to an antigen or a ligand of the disclosure. CARs and/or CLR s of the disclosure comprising one or more Centyrins that specifically bind an antigen may be used to direct the specificity of a cell, (e.g. a cytotoxic immune cell) towards a cell expressing the specific antigen. Alternatively or in addition, CLR s of the disclosure comprising a Centyrin that specifically binds a ligand antigen may transduce a signal intraceullularly to induce expression of a sequence under the control of an inducible promoter.

[079] Centyrins of the disclosure may comprise a protein scaffold, wherein the scaffold is capable of specifically binding an antigen or a ligand. Centyrins of the disclosure may comprise a protein scaffold comprising a consensus sequence of at least one fibronectin type III (FN3) domain, wherein the scaffold is capable of specifically binding an antigen or a ligand. The at least one fibronectin type III (FN3) domain may be derived from a human protein. The human protein may be Tenascin-C. The consensus sequence may comprise LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYDL TGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17010) or MLPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYD LTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17011).

[080] A Centyrin may comprise an amino sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% or any percentage in between of identity to the sequence of

LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYDL TGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17010) or MLPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYD LTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17011).

[081] A Centyrin may comprise an amino sequence having at least 74% identity to the sequence of

LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGGEAINLTVPGSERSYDL
TGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17010) or
MLPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGGEAINLTVPGSERSYD
LTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17011).

[082] The consensus sequence may be encoded by a nucleic acid sequence comprising
atgctgcctgcaccaaagaacctgggtggtgtctcatgtgacagaggatagtgccagactgtcatggactgtcccgacgcagccttcg
atagttttatcatcgtgtaccgggagaaacatcgaaaccggcgaggccattgtcctgacagtgccagggtccgaacgcctcttatgacctg
acagatctgaagcccggaactgagtactatgtgcagatcgccggcgctcaaaggaggcaatcagcttccctctgtccgcaatcttcac
caca (SEQ ID NO: 17034).

[083] The consensus sequence may be modified at one or more positions within (a) a A-B
loop comprising or consisting of the amino acid residues TEDS (SEQ ID NO: 17035) at
positions 13-16 of the consensus sequence; (b) a B-C loop comprising or consisting of the
amino acid residues TAPDAAF (SEQ ID NO: 17036) at positions 22-28 of the consensus
sequence; (c) a C-D loop comprising or consisting of the amino acid residues SEKVGGE (SEQ
ID NO: 17037) at positions 38-43 of the consensus sequence; (d) a D-E loop comprising or
consisting of the amino acid residues GSER (SEQ ID NO: 17038) at positions 51-54 of the
consensus sequence; (e) a E-F loop comprising or consisting of the amino acid residues
GLKPG (SEQ ID NO: 17039) at positions 60-64 of the consensus sequence; (f) a F-G loop
comprising or consisting of the amino acid residues KGGHRSN (SEQ ID NO: 17040) at
positions 75-81 of the consensus sequence; or (g) any combination of (a)-(f). Centyrins of the
disclosure may comprise a consensus sequence of at least 5 fibronectin type III (FN3)
domains, at least 10 fibronectin type III (FN3) domains or at least 15 fibronectin type III
(FN3) domains.

[084] The Centyrin may bind an antigen or a ligand with at least one affinity selected
from a K_D of less than or equal to $10^{-9}M$, less than or equal to $10^{-10}M$, less than or equal to
 $10^{-11}M$, less than or equal to $10^{-12}M$, less than or equal to $10^{-13}M$, less than or equal to
 $10^{-14}M$, and less than or equal to $10^{-15}M$. The K_D may be determined by surface plasmon
resonance.

Antibody Mimetic

[085] The term “antibody mimetic” is intended to describe an organic compound that
specifically binds a target sequence and has a structure distinct from a naturally-occurring
antibody. Antibody mimetics may comprise a protein, a nucleic acid, or a small molecule.
The target sequence to which an antibody mimetic of the disclosure specifically binds may be

an antigen. Antibody mimetics may provide superior properties over antibodies including, but not limited to, superior solubility, tissue penetration, stability towards heat and enzymes (e.g. resistance to enzymatic degradation), and lower production costs. Exemplary antibody mimetics include, but are not limited to, an affibody, an affilin, an affimer, an affitin, an alphabody, an anticalin, and avimer (also known as avidity multimer), a DARPin (Designed Ankyrin Repeat Protein), a Fynomer, a Kunitz domain peptide, and a monobody.

[086] Affibody molecules of the disclosure comprise a protein scaffold comprising or consisting of one or more alpha helix without any disulfide bridges. Preferably, affibody molecules of the disclosure comprise or consist of three alpha helices. For example, an affibody molecule of the disclosure may comprise an immunoglobulin binding domain. An affibody molecule of the disclosure may comprise the Z domain of protein A.

[087] Affilin molecules of the disclosure comprise a protein scaffold produced by modification of exposed amino acids of, for example, either gamma-B crystallin or ubiquitin. Affilin molecules functionally mimic an antibody's affinity to antigen, but do not structurally mimic an antibody. In any protein scaffold used to make an affilin, those amino acids that are accessible to solvent or possible binding partners in a properly-folded protein molecule are considered exposed amino acids. Any one or more of these exposed amino acids may be modified to specifically bind to a target sequence or antigen.

[088] Affimer molecules of the disclosure comprise a protein scaffold comprising a highly stable protein engineered to display peptide loops that provide a high affinity binding site for a specific target sequence. Exemplary affimer molecules of the disclosure comprise a protein scaffold based upon a cystatin protein or tertiary structure thereof. Exemplary affimer molecules of the disclosure may share a common tertiary structure of comprising an alpha-helix lying on top of an anti-parallel beta-sheet.

[089] Affitin molecules of the disclosure comprise an artificial protein scaffold, the structure of which may be derived, for example, from a DNA binding protein (e.g. the DNA binding protein Sac7d). Affitins of the disclosure selectively bind a target sequence, which may be the entirety or part of an antigen. Exemplary affitins of the disclosure are manufactured by randomizing one or more amino acid sequences on the binding surface of a DNA binding protein and subjecting the resultant protein to ribosome display and selection. Target sequences of affitins of the disclosure may be found, for example, in the genome or on the surface of a peptide, protein, virus, or bacteria. In certain embodiments of the disclosure,

an affitin molecule may be used as a specific inhibitor of an enzyme. Affitin molecules of the disclosure may include heat-resistant proteins or derivatives thereof.

[090] Alphabody molecules of the disclosure may also be referred to as Cell-Penetrating Alphabodies (CPAB). Alphabody molecules of the disclosure comprise small proteins (typically of less than 10 kDa) that bind to a variety of target sequences (including antigens). Alphabody molecules are capable of reaching and binding to intracellular target sequences. Structurally, alphabody molecules of the disclosure comprise an artificial sequence forming single chain alpha helix (similar to naturally occurring coiled-coil structures). Alphabody molecules of the disclosure may comprise a protein scaffold comprising one or more amino acids that are modified to specifically bind target proteins. Regardless of the binding specificity of the molecule, alphabody molecules of the disclosure maintain correct folding and thermostability.

[091] Anticalin molecules of the disclosure comprise artificial proteins that bind to target sequences or sites in either proteins or small molecules. Anticalin molecules of the disclosure may comprise an artificial protein derived from a human lipocalin. Anticalin molecules of the disclosure may be used in place of, for example, monoclonal antibodies or fragments thereof. Anticalin molecules may demonstrate superior tissue penetration and thermostability than monoclonal antibodies or fragments thereof. Exemplary anticalin molecules of the disclosure may comprise about 180 amino acids, having a mass of approximately 20 kDa. Structurally, anticalin molecules of the disclosure comprise a barrel structure comprising antiparallel beta-strands pairwise connected by loops and an attached alpha helix. In preferred embodiments, anticalin molecules of the disclosure comprise a barrel structure comprising eight antiparallel beta-strands pairwise connected by loops and an attached alpha helix.

[092] Avimer molecules of the disclosure comprise an artificial protein that specifically binds to a target sequence (which may also be an antigen). Avimers of the disclosure may recognize multiple binding sites within the same target or within distinct targets. When an avimer of the disclosure recognize more than one target, the avimer mimics function of a bi-specific antibody. The artificial protein avimer may comprise two or more peptide sequences of approximately 30-35 amino acids each. These peptides may be connected via one or more linker peptides. Amino acid sequences of one or more of the peptides of the avimer may be derived from an A domain of a membrane receptor. Avimers have a rigid structure that may

optionally comprise disulfide bonds and/or calcium. Avimers of the disclosure may demonstrate greater heat stability compared to an antibody.

[093] DARPin (Designed Ankyrin Repeat Proteins) of the disclosure comprise genetically-engineered, recombinant, or chimeric proteins having high specificity and high affinity for a target sequence. In certain embodiments, DARPins of the disclosure are derived from ankyrin proteins and, optionally, comprise at least three repeat motifs (also referred to as repetitive structural units) of the ankyrin protein. Ankyrin proteins mediate high-affinity protein-protein interactions. DARPins of the disclosure comprise a large target interaction surface.

[094] Fynomers of the disclosure comprise small binding proteins (about 7 kDa) derived from the human Fyn SH3 domain and engineered to bind to target sequences and molecules with equal affinity and equal specificity as an antibody.

[095] Kunitz domain peptides of the disclosure comprise a protein scaffold comprising a Kunitz domain. Kunitz domains comprise an active site for inhibiting protease activity. Structurally, Kunitz domains of the disclosure comprise a disulfide-rich alpha+beta fold. This structure is exemplified by the bovine pancreatic trypsin inhibitor. Kunitz domain peptides recognize specific protein structures and serve as competitive protease inhibitors. Kunitz domains of the disclosure may comprise Ecallantide (derived from a human lipoprotein-associated coagulation inhibitor (LACI)).

[096] Monobodies of the disclosure are small proteins (comprising about 94 amino acids and having a mass of about 10 kDa) comparable in size to a single chain antibody. These genetically engineered proteins specifically bind target sequences including antigens. Monobodies of the disclosure may specifically target one or more distinct proteins or target sequences. In preferred embodiments, monobodies of the disclosure comprise a protein scaffold mimicking the structure of human fibronectin, and more preferably, mimicking the structure of the tenth extracellular type III domain of fibronectin. The tenth extracellular type III domain of fibronectin, as well as a monobody mimetic thereof, contains seven beta sheets forming a barrel and three exposed loops on each side corresponding to the three complementarity determining regions (CDRs) of an antibody. In contrast to the structure of the variable domain of an antibody, a monobody lacks any binding site for metal ions as well as a central disulfide bond. Multispecific monobodies may be optimized by modifying the loops BC and FG. Monobodies of the disclosure may comprise an adnectin.

VHH

[097] In certain embodiments of the compositions and methods of the disclosure, a CAR or a CLR comprises a single domain antibody (SdAb). In certain embodiments, the SdAb is a VHH.

[098] The disclosure provides a CAR or a CLR comprising an antigen or ligand recognition region, respectively, that comprises at least one VHH (to produce a “VCAR” or “VCLR”). CARs and CLRs of the disclosure may comprise more than one VHH. For example, a bi-specific VCAR or VCLR may comprise two VHs. In some embodiments of the bi-specific VCAR or VCLR, each VHH specifically binds a distinct antigen.

[099] VHH proteins of the disclosure specifically bind an antigen or a ligand. CARs of the disclosure comprising one or more VHs that specifically bind an antigen may be used to direct the specificity of a cell, (e.g. a cytotoxic immune cell) towards a target cell expressing the specific antigen. CLRs of the disclosure comprising one or more VHs that specifically bind an antigen may transduce an intracellular signal upon binding a ligand of either VHH to activate expression of a sequence under the control of an inducible promoter.

[0100] Sequences encoding a VHH of the disclosure can be altered, added and/or deleted to reduce immunogenicity or reduce, enhance or modify binding, affinity, on-rate, off-rate, avidity, specificity, half-life, stability, solubility or any other suitable characteristic, as known in the art.

[0101] Optionally, VHH proteins can be engineered with retention of high affinity for the antigen or ligand and other favorable biological properties. To achieve this goal, the VHH proteins can be optionally prepared by a process of analysis of the parental sequences and various conceptual engineered products using three-dimensional models of the parental and engineered sequences. Three-dimensional models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate sequences and can measure possible immunogenicity (e.g., Immunofilter program of Xencor, Inc. of Monrovia, Calif.). Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate sequence, i.e., the analysis of residues that influence the ability of the candidate VHH protein to bind its antigen/ligand. In this way, residues can be selected and combined from the parent and reference sequences so that the desired characteristic, such as affinity for the target antigen(s)/ligand(s), is achieved. Alternatively, or in addition to, the above procedures, other suitable methods of engineering can be used.

VH

[0102] In certain embodiments of the compositions and methods of the disclosure, a CAR or a CLR comprises a single domain antibody (SdAb). In certain embodiments, the SdAb is a VH.

[0103] The disclosure provides CARs/CLRs comprising a single domain antibody (to produce a “VCAR” or a “VCLR”, respectively). In certain embodiments, the single domain antibody comprises a VH. In certain embodiments, the VH is isolated or derived from a human sequence. In certain embodiments, VH comprises a human CDR sequence and/or a human framework sequence and a non-human or humanized sequence (e.g. a rat Fc domain). In certain embodiments, the VH is a fully humanized VH. In certain embodiments, the VH is neither a naturally occurring antibody nor a fragment of a naturally occurring antibody. In certain embodiments, the VH is not a fragment of a monoclonal antibody. In certain embodiments, the VH is a UniDab™ antibody (TeneoBio).

[0104] In certain embodiments, the VH is fully engineered using the UniRat™ (TeneoBio) system and “NGS-based Discovery” to produce the VH. Using this method, the specific VH are not naturally-occurring and are generated using fully engineered systems. The VH are not derived from naturally-occurring monoclonal antibodies (mAbs) that were either isolated directly from the host (for example, a mouse, rat or human) or directly from a single clone of cells or cell line (hybridoma). These VHs were not subsequently cloned from said cell lines. Instead, VH sequences are fully-engineered using the UniRat™ system as transgenes that comprise human variable regions (VH domains) with a rat Fc domain, and are thus human/rat chimeras without a light chain and are unlike the standard mAb format. The native rat genes are knocked out and the only antibodies expressed in the rat are from transgenes with VH domains linked to a Rat Fc (UniAbs). These are the exclusive Abs expressed in the UniRat. Next generation sequencing (NGS) and bioinformatics are used to identify the full antigen-specific repertoire of the heavy-chain antibodies generated by UniRat™ after immunization. Then, a unique gene assembly method is used to convert the antibody repertoire sequence information into large collections of fully-human heavy-chain antibodies that can be screened *in vitro* for a variety of functions. In certain embodiments, fully humanized VH are generated by fusing the human VH domains with human Fcs *in vitro* (to generate a non-naturally occurring recombinant VH antibody). In certain embodiments, the VH are fully humanized, but they are expressed *in vivo* as human/rat chimera (human VH, rat

Fc) without a light chain. Fully humanized VHs are expressed *in vivo* as human/rat chimera (human VH, rat Fc) without a light chain are about 80kDa (vs 150 kDa).

[0105] VCARS/VCLRs of the disclosure may comprise at least one VH of the disclosure. In certain embodiments, the VH of the disclosure may be modified to remove an Fc domain or a portion thereof. In certain embodiments, a framework sequence of the VH of the disclosure may be modified to, for example, improve expression, decrease immunogenicity or to improve function.

Transposons/Transposases

[0106] Exemplary transposon/transposase systems of the disclosure include, but are not limited to, piggyBac transposons and transposases, Sleeping Beauty transposons and transposases, Helraiser transposons and transposases and Tol2 transposons and transposases.

[0107] The piggyBac transposase recognizes transposon-specific inverted terminal repeat sequences (ITRs) on the ends of the transposon, and moves the contents between the ITRs into TTAA chromosomal sites. The piggyBac transposon system has no payload limit for the genes of interest that can be included between the ITRs. In certain embodiments, and, in particular, those embodiments wherein the transposon is a piggyBac transposon, the transposase is a piggyBacTM or a Super piggyBacTM (SPB) transposase. In certain embodiments, and, in particular, those embodiments wherein the transposase is a Super piggyBacTM (SPB) transposase, the sequence encoding the transposase is an mRNA sequence.

[0108] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBacTM (PB) transposase enzyme. The piggyBac (PB) transposase enzyme may comprise or consist of an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MGSSLDDEHI LSALLQSDDE LVGEDSDSEI SDHVSEDDVQ SDTEEAFFIDE VHEVQPTSSG
61 SEILDEQNVI EQPGSSLASN RILTLPQRTI RGKNKHCWST SKSTRRSRVS ALNIVRSQRG
121 PTRMCRNIDY PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTGATFRD TNEDEIYAFF
181 GILVMTAVRK DNHMSTDDL FDRSLSMVYVS VMSRDRFDFL IRCLRMDDKS IRPTLRENDV
241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF PMYI?NKPSK YGIKILMMCD
301 SGTKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCEFDGP LTLVSYKPKP AKMVYLLSSC
421 DEDASINEST GKPQVMYYN QTKGGVDTLD QMCSVMTCSR KTNRWPMALL YGMINIACIN
481 SFIIYSHNVS SKGEKVQSRK KFMRLYMSL TSSE?MRKRLE APTLKRYLRD NISNILPNEV
541 PGTSDDSTEE PVMKKRTYCT YCPSKIRKA NASCKKCKKV ICREHNIDMC QSCF (SEQ ID
NO: 14487) .

```

[0109] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at one or more of positions 30, 165, 282, or 538 of the sequence:

```

1  MGSSLDDEHI  LSALLQSDDE  LVGEDSDSEI  SDHVSEDDVQ  SDTEEAFFIDE  VHEVQPTSSG
61  SEILDEQNVI  EQPGSSLASN  RILTLPQRTI  RGKNKHCWST  SKSTRRSRVS  ALNIVRSQRG
121 PTRMCRNIYD  PLLCFKLFFT  DEIISEIVKW  TNAEISLKRR  ESMTGATFRD  TNEDEIYAFF
181 GILVMTAVRK  DNHMSTDDLF  DRSLSMVYVS  VMSRDRFDFL  IRCLRMDDKS  IRPTLRENDV
241 FTPVRKIWDL  FIHQCIQNYT  PGAHLTIDEQ  LLGFRGRCPF  RMYI?NKPSK  YGIKILMMCD
301 SGTKYMINGM  PYLGRGTQTN  GVPLGEYYVK  ELSKPVHGSC  RNITCDNWFT  SIPLAKNLLQ
361 EPYKLTIVGT  VRSNKREIPE  VLKNSRSRPV  GTSMFCEFDGP  LTLVSYKPKP  AKMVYLLSSC
421 DEDASINEST  GKPQVMYYN  QTKGGVDTLD  QMCSVMTCSR  KTNRWPMALL  YGMINIACIN
481 SFIIYSHNVS  SKGEKVQSRK  KFMRNLYMSL  TSSEMRKRLE  APTLKRYLRD  NISNILPNEV
541 PGTSDDSTEE  PVMKKRTYCT  YCPSKIRKA  NASCKKCKKV  ICREHNIDMC  QSCF (SEQ ID
NO: 14487) .

```

[0110] In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at two or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at three or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the transposase enzyme is a piggyBac™ (PB) transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at each of the following positions 30, 165, 282, and 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the amino acid substitution at position 30 of the sequence of SEQ ID NO: 14487 is a substitution of a valine (V) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 165 of the sequence of SEQ ID NO: 14487 is a substitution of a serine (S) for a glycine (G). In certain embodiments, the amino acid substitution at position 282 of the sequence of SEQ ID NO: 14487 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 538 of the sequence of SEQ ID NO: 14487 is a substitution of a lysine (K) for an asparagine (N).

[0111] In certain embodiments of the methods of the disclosure, the transposase enzyme is a Super piggyBac™ (SPB) transposase enzyme. In certain embodiments, the Super piggyBac™ (SPB) transposase enzymes of the disclosure may comprise or consist of the

amino acid sequence of the sequence of SEQ ID NO: 14487 wherein the amino acid substitution at position 30 is a substitution of a valine (V) for an isoleucine (I), the amino acid substitution at position 165 is a substitution of a serine (S) for a glycine (G), the amino acid substitution at position 282 is a substitution of a valine (V) for a methionine (M), and the amino acid substitution at position 538 is a substitution of a lysine (K) for an asparagine (N). In certain embodiments, the Super piggyBac™ (SPB) transposase enzyme may comprise or consist of an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MGSSLDDEHI  LSALLQSDDE  LVGEDSDSEV  SDHVSEDDVQ  SDTEEAFFIDE  VHEVQPTSSG
61  SEILDEQNVI  EQPGSSLASN  RILTLPQRTI  RGKNKHCWST  SKSTRRSRVS  ALNIVRSQRG
121 PTRMCRNIYD  PLLCFKLFFT  DEIISEIVKW  TNAEISLKRR  ESMTSATFRD  TNEDEIYAFF
181 GILVMTAVRK  DNHMSTDDL  FDRSLSMVYVS  VMSRDRFDFL  IRCLRMDDKS  IRPTLRENDV
241 FTPVRKIWDL  FIHQCIQNYT  PGAHLTIDEQ  LLGFRGRCPF  RVIYIPNKPSK  YGIKILMMCD
301 SGTKYMINGM  PYLGRGTQTN  GVPLGEYYVK  ELSKPVHGSC  RNITCDNWFT  SIPLAKNLLQ
361 EPYKLTIVGT  VRSNKREIPE  VLKNSRSRPV  GTSMFCFDGP  LTLVSYKPKP  AKMVYLLSSC
421 DEDASINEST  GKPQMVMYYN  QTKGGVDTLD  QMCSVMTCSR  KTNRWPMALL  YGMINIACIN
481 SFIIYSHNVS  SKGEKVQSRK  KFMRNLYMSL  TSSFMRKRLE  APTLKRYLRD  NISNILPKEV
541 PGTSDDSTEE  PVMKKRTYCT  YCPSKIRRKA  NASCKKCKKV  ICREHNIDMC  QSCF (SEQ ID NO:
14484) .

```

[0112] In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 3, 46, 82, 103, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 258, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 486, 503, 552, 570 and 591 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 46, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 485, 503, 552 and 570. In certain embodiments, the amino acid substitution at position 3 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an asparagine (N) for a serine (S). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an alanine (A). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 14487 or SEQ ID

NO: 14484 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 82 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position 119 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for an arginine (R). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) a cysteine (C). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a histidine (H) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 185 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 187 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for an alanine (A). In certain embodiments, the amino acid substitution at position 200 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 207 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a valine (V). In certain embodiments, the amino acid substitution at position 209 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a valine (V). In certain embodiments, the amino acid substitution at position 226 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a methionine (M). In certain embodiments, the amino acid substitution at position 235 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a leucine (L). In certain embodiments, the amino acid substitution at

position 240 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 241 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 243 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a proline (P). In certain embodiments, the amino acid substitution at position 258 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tyrosine (Y) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a proline (P). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine for a proline (P). In certain embodiments, the amino acid substitution at position 315 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for an arginine (R). In certain embodiments, the amino acid substitution at position 319 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a threonine (T). In certain embodiments, the amino acid substitution at position 327 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 328 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a cysteine (C). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 421 of

SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a histidine (H) for the aspartic acid (D). In certain embodiments, the amino acid substitution at position 436 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a valine (V). In certain embodiments, the amino acid substitution at position 456 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tyrosine (Y) for a methionine (M). In certain embodiments, the amino acid substitution at position 470 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid substitution at position 485 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a serine (S). In certain embodiments, the amino acid substitution at position 503 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 503 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a methionine (M). In certain embodiments, the amino acid substitution at position 552 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a glutamine (Q). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a glutamine (Q).

[0113] In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ transposase enzyme may comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ transposase enzyme may comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at two, three, four, five, six or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ transposase enzyme may

comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position 194 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 372 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for an arginine (R). In certain embodiments, the amino acid substitution at position 375 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for a lysine (K). In certain embodiments, the amino acid substitution at position 450 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an asparagine (N) for an aspartic acid (D). In certain embodiments, the amino acid substitution at position 509 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a serine (S). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487. In certain embodiments, including those embodiments wherein the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, the piggyBac™ transposase enzyme may further comprise an amino acid substitution at positions 372, 375 and 450 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 14487, and a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 14487. In certain embodiments, the piggyBac™ transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 14487, a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 14487 and a substitution of an asparagine (N) for an aspartic acid (D) at position 450 of SEQ ID NO: 14487.

[0114] The sleeping beauty transposon is transposed into the target genome by the Sleeping Beauty transposase that recognizes ITRs, and moves the contents between the ITRs into TA

chromosomal sites. In various embodiments, SB transposon-mediated gene transfer, or gene transfer using any of a number of similar transposons, may be used in the compositions and methods of the disclosure.

[0115] In certain embodiments, and, in particular, those embodiments wherein the transposon is a Sleeping Beauty transposon, the transposase is a Sleeping Beauty transposase or a hyperactive Sleeping Beauty transposase (SB100X).

[0116] In certain embodiments of the methods of the disclosure, the Sleeping Beauty transposase enzyme comprises an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MGKSKEISQD LRKKIVDLHK SGSSLGAISK RLKVPRSSVQ TIVRKYKHHG TTQPSYRSGR
61 RRVLSRDER TLVRKVQINP RTTAKDLVKM LEETGTKVSI STVKRVLYRH NLKGRSARKK
121 PLLQNRHKKA RLRFATAHGD KDRTEWRNVL WDETKEIEF GHNDHRYVWR KKGEACKPKN
181 TIPTVKHGGG SIMLWGCFAA GGTGALHKID GIMRKENYVD ILKQHLKTSV RKLKLGRKWV
241 FQMDNDPKHT SKVVAWLKD NKVKVLEWPS QSPDLNPIEN LWAELEKRVV ARRPTNLTQL
301 HQLCQEEWAK IHPTYCGKLV EGYPKRLTQV KQFKGNATKY (SEQ ID NO: 14485).

```

[0117] In certain embodiments of the methods of the disclosure, the hyperactive Sleeping Beauty (SB100X) transposase enzyme comprises an amino acid sequence at least 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MGKSKEISQD LRKRIVDLHK SGSSLGAISK RLAVPRSSVQ TIVRKYKHHG TTQPSYRSGR
61 RRVLSRDER TLVRKVQINP RTTAKDLVKM LEETGTKVSI STVKRVLYRH NLKGHSARKK
121 PLLQNRHKKA RLRFATAHGD KDRTEWRNVL WDETKEIEF GHNDHRYVWR KKGEACKPKN
181 TIPTVKHGGG SIMLWGCFAA GGTGALHKID GIMDAVQYVD ILKQHLKTSV RKLKLGRKWV
241 FQHDNDPKHT SKVVAWLKD NKVKVLEWPS QSPDLNPIEN LWAELEKRVV ARRPTNLTQL
301 HQLCQEEWAK IHPNYCGKLV EGYPKRLTQV KQFKGNATKY (SEQ ID NO: 14486).

```

[0118] The Helraiser transposon is transposed by the Helitron transposase. Helitron transposases mobilize the Helraiser transposon, an ancient element from the bat genome that was active about 30 to 36 million years ago. An exemplary Helraiser transposon of the disclosure includes Helibat1, which comprises a nucleic acid sequence comprising:

```

1 TCCTATATAA TAAAAGAGAA ACATGCAAAT TGACCATCCC TCCGCTACGC TCAAGCCACG
61 CCCACCAGCC AATCAGAAGT GACTATGCAA ATTAACCCAA CAAAGATGGC AGTTAAATTT
121 GCATACGCAG GTGTCAAGCG CCCAGGAGG CAACGGCGGC CGCGGGCTCC CAGGACCTTC
181 GCTGGCCCCG GGAGGCGAGG CCGGCCGCGC CTAGCCACAC CCGCGGGCTC CCGGGACCTT
241 CGCCAGCAGA GAGCAGAGCG GGAGAGCGGG CGGAGAGCGG GAGGTTTGGG GGAAGTTGGA
301 GAGCAGGAGG CCGCTGGACA TAGAGCAGAG CGAGAGAGAG GGTGGCTTGG AGGGCGTGGC
361 TCCCTCTGTC ACCCCAGCTT CCTCATCACA GCTGTGGAAA CTGACAGCAG GGAGGAGGAA
421 GTCCACCCCC CACAGAATCA GCCAGAATCA GCCGTTGGTC AGACAGCTCT CAGCGGCCTG

```

481 ACAGCCAGGA CTCTCATTCA CCTGCATCTC AGACCGTGAC AGTAGAGAGG TGGGACTATG
 541 TCTAAAGAAC AACTGTTGAT ACAACGTAGC TCTGCAGCCG AAAGATGCCG GCGTTATCGA
 601 CAGAAAATGT CTGCAGAGCA ACGTGCCTCT GATCTTGAAA GAAGGCGGCG CCTGCAACAG
 661 AATGTATCTG AAGASCAGCT ACTGGAAAAA CGTCCGCTCTG AAGCCGAAAA ACAGCGGCGT
 721 CATCGACAGA AAATGTCTAA AGACCAACGT GCCTTTGAAG TTGAAAGAAG GCGGTGGCGA
 781 CGACAGAATA TGTCTAGAGA ACAGTCATCA ACAAGTACTA CCAATACCGG TAGGAACTGC
 841 CTTCTCAGCA AAAATGGAGT ACATGAGGAT GCAATTCTCG AACATAGTTG TGGTGGAATG
 901 ACTGTTGAT GTGAATTTTG CCTATCACTA AATTTCTCTG ATGAAAAACC ATCCGATGGG
 961 AAATTTACTC GATGTTGTAG CAAAGGGAAA GTCTGTCCAA ATGATATACA TTTTCCAGAT
 1021 TACCCGGCAT ATTTAAAAAG ATTAATGACA AACGAAGATT CTGACAGTAA AAATTTCATG
 1081 GAAAATATTC GTTCCATAAA TAGTTCTTTT GCTTTTGCTT CCATGGGTGC AAATATTGCA
 1141 TCGCCATCAG GATATGGGCC ATACTGTTTT AGAATACACG GACAAGTTTA TCACCGTACT
 1201 GGAACCTTAC ATCCTTCGGA TGGTGTCTCT CGGAAGTTTG CTCAACTCTA TATTTTGGAT
 1261 ACAGCCGAAG CTACAAGTAA AAGATTAGCA ATGCCAGAAA ACCAGGGCTG CTCAGAAAGA
 1321 CTCATGATCA ACATCAACAA CCTCATGCAT GAAATAAATG AATTAACAAA ATCGTACAAG
 1381 ATGCTACATG AGGTAGAAAA GGAAGCCCAA TCTGAAGCAG CAGCAAAAGG TATTGCTCCC
 1441 ACAGAAGTAA CAATGGCGAT TAAATACGAT CGTAACAGTG ACCCAGGTAG ATATAATTCT
 1501 CCCCCTGTAA CCGAGGTTGC TGTCATATTC AGAAACGAAG ATGGAGAACC TCCTTTTGAA
 1561 AGGGACTTGC TCATTCATTG TAAACCAGAT CCAATAATC CAAATGCCAC TAAAATGAAA
 1621 CAAATCAGTA TCCTGTTTCC TACATTAGAT GCAATGACAT ATCCTATTCT TTTTCCACAT
 1681 GGTGAAAAAG GCTGGGGAAC AGATATTGCA TTAAGACTCA GAGACAACAG TGTAATCGAC
 1741 AATAATACTA GACAAAATGT AAGGACACGA GTCACACAAA TGCAGTATTA TGGATTTTCA
 1801 CTCTCTGTGC GGGACACGTT CAATCCTATT TTAAATGCAG GAAAATTAAC TCAACAGTTT
 1861 ATTGTGGATT CATATTCAAA AATGGAGGCC AATCGGATAA ATTTTCATCA AGCAAACCAA
 1921 TCTAAGTTGA GAGTTGAAAA ATATAGTGGT TTGATGGATT ATCTCAAATC TAGATCTGAA
 1981 AATGACAATG TGCCGATTGG TAAAATGATA ATACTTCCAT CATCTTTTGA GGGTAGTCCC
 2041 AGAAATATGC AGCAGCGATA TCAGGATGCT ATGGCAATTG TAACGAAGTA TGGCAAGCCC
 2101 GATTTATTCA TAACCATGAC ATGCAACCCC AAATGGGCAG ATATTACAAA CAATTTACAA
 2161 CGCTGGCAAA AAGTTGAAAA CAGACCTGAC TTGGTAGCCA GAGTTTTTAA TATTAAGCTG
 2221 AATGCTCTTT TAAATGATAT ATGTAAATTC CATTTATTTG GCAAAGTAAT AGCTAAAATT
 2281 CATGTCATTG AATTTCAGAA ACGCGGACTG CCTCACGCTC ACATATTATT GATATTAGAT
 2341 AGTGAGTCCA AATTACGTTT AGAAGATGAC ATTGACCGTA TAGTTAAGGC AGAAATTCCA
 2401 GATGAAGACC AGTGTCTCTG ACTTTTTTCAA ATTGTAAAAT CAAATATGGT ACATGGACCA
 2461 TGTGGAATAC AAAATCCAAA TAGTCCATGT ATGGAAAATG GAAAATGTTC AAAGGGATAT
 2521 CCAAAGAAT TTCAAAATGC GACCATTTGA AATATTGATG GATATCCCAA ATACAAACGA
 2581 AGATCTGGTA GCACCATGTC TATTGGAAAT AAAGTTGTCT ATAACACTTG GATTGTCCCT
 2641 TATAACCCGT ATTTGTGCCT TAAATATAAC TGTCATATAA ATGTTGAAGT CTGTGCATCA
 2701 ATTAAAAGTG TCAAAATATT ATTTAAATAC ATCTATAAAG GGCACGATTG TGCAAATATT
 2761 CAAATTTCTG AAAAAAATAT TATCAATCAT GACGAAGTAC AGGACTTCAT TGACTIONAGG
 2821 TATGTGAGCG CTCCTGAGGC TGTTTGGAGA CTTTTTGCAA TGCGAATGCA TGACCAATCT

2881 CATGCAATCA CAAGATTAGC TATTCATTTG CCAAATGATC AGAATTTGTA TTTTCATACC
 2941 GATGATTTTG CTGAAGTTTT AGATAGGGCT AAAAGGCATA ACTCGACTTT GATGGCTTGG
 3001 TTCTTATTGA ATAGAGAAGA TTCTGATGCA CGTAATTATT ATTATTGGGA GATTCCACAG
 3061 CATTATGTGT TTAATAATTC TTTGTGGACA AAACGCCGAA AGGCTGGGAA TAAAGTATTA
 3121 GGTAGACTGT TCACTGTGAG CTTTAGAGAA CCAGAACGAT ATTACCTTAG ACTTTTGCTT
 3181 CTGCATGTAA AAGGTGCGAT AAGTTTTGAG GATCTGCGAA CTGTAGGAGG TGTAACCTAT
 3241 GATACATTTT ATGAAGCTGC TAAACACCGA GGATTATTAC TTGATGACAC TATCTGGAAA
 3301 GATACGATTG ACGATGCAAT CATCCTTAAT ATGCCCAAAC AACTACGGCA ACTTTTTGCA
 3361 TATATATGTG TGTTTGGATG TCCTTCTGCT GCAGACAAAT TATGGGATGA GAATAAATCT
 3421 CATTTTATTG AAGATTTCTG TTGGAAATTA CACCGAAGAG AAGGTGCCTG TGTGAACCTG
 3481 GAAATGCATG CCCTTAACGA AATTCAGGAG GTATTCACAT TGCATGGAAT GAAATGTTCA
 3541 CATTTCAAAC TTCCGGACTA TCCTTTATTA ATGAATGCAA ATACATGTGA TCAATTGTAC
 3601 GAGCAACAAC AGGCAGAGGT TTTGATAAAT TCTCTGAATG ATGAACAGTT GGCAGCCTTT
 3661 CAGACTATAA CTTCAGCCAT CGAAGATCAA ACTGTACACC CCAAATGCTT TTTCTTGGAT
 3721 GGTCCAGGTG GTAGTGGAAA AACATATCTG TATAAAGTTT TAACACATTA TATTAGAGGT
 3781 CGTGGTGGTA CTGTTTTACC CACAGCATCT ACAGGAATTG CTGCAAATTT ACTTCTTGGT
 3841 GGAAGAACCT TTCATTCCCA ATATAAATTA CCAATTCCAT TAAATGAAAC TTCAATTTCT
 3901 AGACTCGATA TAAAGAGTGA AGTTGCTAAA ACCATTAAAA AGGCCCAACT TCTCATTATT
 3961 GATGAATGCA CCATGGCATC CAGTCATGCT ATAAACGCCA TAGATAGATT ACTAAGAGAA
 4021 ATTATGAATT TGAATGTTGC ATTTGGTGGG AAAGTTCTCC TTCTCGGAGG GGATTTTCGA
 4081 CAATGTCTCA GTATTGTACC ACATGCTATG CGATCGGCCA TAGTACAAAC GAGTTTAAAG
 4141 TACTGTAATG TTTGGGGATG TTTCAGAAAG TTGTCTCTTA AAACAAATAT GAGATCAGAG
 4201 GATTCTGCTT ATAGTGAATG GTTAGTAAAA CTTGGAGATG GCAAACCTTG TAGCAGTTTTT
 4261 CATTTAGGAA TGGATATTAT TGAAATCCCC CATGAAATGA TTTGTAACGG ATCTATTATT
 4321 GAAGCTACCT TTGGAAATAG TATATCTATA GATAATATTA AAAATATATC TAAACGTGCA
 4381 ATTCTTTGTC CAAAAAATGA GCATGTTCAA AAATTAAATG AAGAAATTTT GGATATACTT
 4441 GATGGAGATT TTCACACATA TTTGAGTGAT GATTCCATTG ATTCAACAGA TGATGCTGAA
 4501 AAGGAAAATT TTCCCATCGA ATTTCTTAAT AGTATFACTC CTTCGGGAAT GCCGTGTCAT
 4561 AAATTAAAAT TGAAAGTGGG TGCAATCATC ATGCTATTGA GAAATCTTAA TAGTAAATGG
 4621 GGTCTTTGTA ATGGTACTAG ATTTATTATC AAAAGATTAC GACCTAACAT TATCGAAGCT
 4681 GAAGTATTAA CAGGATCTGC AGAGGGAGAG GTTGTCTGTA TTCCAAGAA TGAATTTGTCC
 4741 CCATCTGACA CTGGCCTCCC ATTTAAATTA ATTCTGAAGAC AGTTTCCCGT GATGCCAGCA
 4801 TTTGCGATGA CTATTAATAA ATCACAAGGA CAACTCTAG ACAGAGTAGG AATATTCCTA
 4861 CCTGAACCCG TTTTCGCACA TGGTCAGTTA TATGTTGCTT TCTCTCGAGT TCGAAGAGCA
 4921 TGTGACGTTA AAGTTAAAGT TGTAATAACT TCATCACAAG GGAAATTAGT CAAGCACTCT
 4981 GAAAGTGTTT TTAATCTTAA TGTGGTATAC AGGGAGATAT TAGAATAAGT TTAATCACTT
 5041 TATCAGTCAT TGTTTGCATC AATGTTGTTT TTATAFCATG TTTTGTGTGT TTTTATATCA
 5101 TGTCTTTGTT GTTGTATAT CATGTTGTTA TTGTTTATTT ATTAATAAAT TTATGTATTA
 5161 TTTTCATATA CATTTTACTC ATTTCTTTC ATCTCTCACA CTTCTATTAT AGAGAAAGGG
 5221 CAAATAGCAA TATTAAAATA TTTCTCTAA TTAATTCCTT TTCAATGTGC ACGAATTTCC

5281 TGCACCGGGC CACTAG (SEQ ID NO: 14652).

[0119] Unlike other transposases, the Helitron transposase does not contain an RNase-H like catalytic domain, but instead comprises a RepHel motif made up of a replication initiator domain (Rep) and a DNA helicase domain. The Rep domain is a nuclease domain of the HUH superfamily of nucleases.

[0120] An exemplary Helitron transposase of the disclosure comprises an amino acid sequence comprising:

```

1 MSKEQLLIQR SSAAERCRRY RQKMSAEQRA SDLERRRLQ QNVSEEQLLE KRRSEAEKQR
61 RHRQKMSKDQ RAFEVERRRW RRQNSREQS STSTNTGRN CLLSKNGVHE DAILEHSCGG
121 MTRVCEFLS LNFSDKPSD GKTRCCSKG KVCNDIHF DPAYLKRLM TNEDSDSKNF
181 MENIRSINSS FAFASMGANI ASPSGYGPYC FRIHGQVYHR TGTLHPSDGV SRKFAQLYIL
241 DTAEATSKRL AMPENQGCSE RLMINNNLM HEINELTKSY KMLHEVEKEA QSEAAKGIA
301 PTEVTMAIKY DRNSDPGRYN SPRVTEVAVI FRNEDGEPPF ERDLLIHCKP DPNNPNATKM
361 KQISILEPTL DAMTYPILFP HGEKGWGTDI ALRLRDNVI DNNTRQNVRT RVTQMYYGF
421 HLSVRDTFNP ILNAGKLTQQ FIVDSYSKME ANRINFIKAN QSKLRVEKYS GLMDYLKSR
481 ENDNVPIGKM IILPSSFEGS PRNMQORYQD AMAIVTKYK PDLEITMTCN PKWADITNNL
541 QRWQKVENRP DLVARVENIK LNALLNDICK FHLFGKVIK IHVIEFQKRG LPHAHILLIL
601 DSEKLRSED DIDRIVKAEI PDEDQCPRLF QIVKSNMVHG PCGIQNPNP CMENGKCSKG
661 YPKEFQNTI GNIDGYPKYK RRSSTMSIG NKVDNTWIV PYNPYLCLKY NCHINVEVCA
721 SIKSVKYLEF YIYKGHDCAN IQISEKNIIN HDEVQDFIDS RYVSAPEAVW RLFAMRMHDQ
781 SHAITRLAIH LPNDQNLIFH TDDFAEVLDR AKRHNTLMA WFLNREDSD ARNYYYWEIP
841 QHYVFNNSLW TKRRKGGNKV LGRLFTVSFR EPERYYLRL LLHVKGALSF EDLRTVGGVT
901 YDTFHEAAKH RGLLLDDTIW KDTIDDAIIL NMPKQLRQLF AYICVFGCPS AADKLWDENK
961 SHFIEDFCWK LHRREGACVN CEMHALNEIQ EVFTLHGMKC SHFKLPDYPL LMNANTCDQL
1021 YEQQQAEVLI NSLNDEQLAA FQTITSAIED QTVHPKCFEL DPGGSGKTY LYKVLTHYIR
1081 GRGGTVLPTA STGIAANLLL GGRTEHSQYK LPIPLNETSI SRLDIKSEVA KTIKKAQLLI
1141 IDECTMASSH AINAIDRLR EIMNLNVAFG GKVLGGLGDF RQCLSIVPHA MRSIVQTSI
1201 KYCNVWGCFR KLSLKTNMRS EDSAYSEWLV KLGDKLDSS FHLGMDIEI PHEMIGNSI
1261 IEATFGNSIS IDNIKNISKR AILCPKNEHV QKLNEEILDI LDGDEHTYLS DDSIDSTDDA
1321 EKENFPFIEFL NSITPSGMPC HKLKLKVGAI IMLLRNLNSK WGLCNGTRFI IKRLRPNIIE
1381 AEVLTGSAEG EVVLIPRIDL SPSDTGLPFK LIRRQFPVMP AFAMTINKSQ GQTLDRVGIF
1441 LPEPVEAHGQ LYVAESRVR ACVKVKVNV TSSQGKLVKH SESVFTLNVV YREILE (SEQ ID
NO: 14501).
```

[0121] In Helitron transpositions, a hairpin close to the 3' end of the transposon functions as a terminator. However, this hairpin can be bypassed by the transposase, resulting in the transduction of flanking sequences. In addition, Helitron transposition generates covalently closed circular intermediates. Furthermore, Helitron transpositions can lack target site

duplications. In the Helraiser sequence, the transposase is flanked by left and right terminal sequences termed LTS and RTS. These sequences terminate with a conserved 5'-TC/CTAG-3' motif. A 19 bp palindromic sequence with the potential to form the hairpin termination structure is located 11 nucleotides upstream of the RTS and consists of the sequence GTGCACGAATTTTCGTGCACCGGGCCACTAG (SEQ ID NO: 14500).

[0122] Tol2 transposons may be isolated or derived from the genome of the medaka fish, and may be similar to transposons of the hAT family. Exemplary Tol2 transposons of the disclosure are encoded by a sequence comprising about 4.7 kilobases and contain a gene encoding the Tol2 transposase, which contains four exons. An exemplary Tol2 transposase of the disclosure comprises an amino acid sequence comprising the following:

```

1 MEEVCDSSAA ASSTVQNQPQ DQEHWPYLR EEFSLSGVNK DSFKMKCVLC LPLNKEISAF
61 KSSPSNLRKH IERMHPNYLK NYSKLTAQKR KIGTSTHASS SKQLKVDSVF PVKHVSPVTV
121 NKAILRYIIQ GLHPFSTVDL PSFKELISTL QPGISVITRP TLRSKIAEAA LIMKQKVTA
181 MSEVEWIATT TDCWTARRKS FIGVTAHWIN PGSLEHSA LACKRLMGSH TFEVLASAMN
241 DIHSEYEIRD KVVCTTTDSG SNFMKAFRVF GVENNDIETE ARRCESDDTD SEGCGEGSDG
301 VEFQDASRVL DQDDGFEFQL PKHQKCACHL LNLVSSVDAQ KALSNEHYKK LYRSVFGKCQ
361 ALWNKSSRSA LAEAVESES RLQLLRPNQT RWNSTEMAVD RILQICKEAG EGALRNICTS
421 LEVPMFNPAE MLFLTEWANT MRPVAKVLDI LQAETNTQLG WLLPSVHQLS LKLQRLHHS
481 RYCDPLVDAL QQGIQTRFKH MFEDFEIIAA AILLPKERTS WTNDETIIKR GMDYIRVHLE
541 PLDHKKELAN SSSDDEDEFFA SLKPTTHEAS KELDGYLACV SDTRESLLTF PAICSLSIKT
601 NTPLPASAAC ERLFSTAGLL FSPKRRLDT NNFENQLLLK LNLRFYNFE (SEQ ID NO:
14502).
```

[0123] An exemplary Tol2 transposon of the disclosure, including inverted repeats, subterminal sequences and the Tol2 transposase, is encoded by a nucleic acid sequence comprising the following:

```

1 CAGAGGTGTA AAGTACTTGA GTAATTTTAC TTGATTACTG TACTTAAGTA TTATTTTGG
61 GGATTTTAC TTTACTTGAG TACAATTAAA AATCAATACT TTTACTTTTA CTTAATTACA
121 TTTTTTTAGA AAAAAAGTA CTTTTTACTC CTTACAATTT TATTTACAGT CAAAAAGTAC
181 TTATTTTFTG GAGATCACTT CATTCTATTT TCCCTTGCTA TTACCAAACC AATTGAATTG
241 CGCTGATGCC CAGTTTAATT TAAATGTTAT TTATTCTGCC TATGAAAATC GTTTTCACAT
301 TATATGAAAT TGGTCAGACA TGTTCAATTG TCCTTTGGAA GTGACGTCAT GTCACATCTA
361 TTACCACAAT GCACAGCACC TTGACCTGGA AATTAGGGAA ATTATAACAG TCAATCAGTG
421 GAAGAAAATG GAGGAAGTAT GTGATTCAFC AGCAGCTGCG AGCAGCACAG TCCAAAATCA
481 GCCACAGGAT CAAGAGCACC CGTGGCCGTA TCTTCGCGAA TTCTTTTCTT TAAGTGGTGT
541 AAATAAAGAT TCATTCAAGA TGAAATGTGT CCTCTGFTCT CCGCTTAATA AAGAAATATC
601 GGCCTTCAAA AGTTCGCCAT CAAACCTAAG GAAGCATATT GAGGTAAGTA CATTAAGTAT
661 TTTGTTTTAC TGATAGTTTT TTTTTTTTTT TTTTTTTTTT TTTTGGGTG TGCATGTTTT
```

```

721 GACGTTGATG GCGCGCCTTT TATATGTGTA GTAGGCCTAT TTTCACATAAT GCATGCGATT
781 GACAATATAA GGCTCACGTA ATAAAATGCT AAAATGCATT TGTAATTGGT AACGTTAGGT
841 CCACGGGAAA TTTGGCGCCT ATTGCAGCTT TGAATAATCA TTATCATTC CTTGCTCTCAT
901 TGTGTTTGAA TTCATGCAAA ACACAAGAAA ACCAAGCGAG AAATTTTTTTT CCAAACATGT
961 TGTATTGTCA AAACGGTAAC ACTTTACAAT GAGGTTGATT AGTTCATGTA TTAACATAACA
1021 TTAAATAACC ATGAGCAATA CATTTGTTAC TGTATCTGTT AATCTTTGTT AACGTTAGTT
1081 AATAGAAATA CAGATGTTCA TTGTTTGTTT ATGTTAGTTC ACAGTGCATT AACTAATGTT
1141 AACAGATAT AAAGTATTAG TAAATGTTGA AATTAACATG TATACGTGCA GTTCATTATT
1201 AGTTCATGTT AACTAATGTA GTTAACTAAC GAACCTTATT GTAAAAGTGT TACCATCAAA
1261 ACTAATGTAA TGAAATCAAT TCACCCTGTC ATGTCAGCCT TACAGTCCTG TGTTTTTGTC
1321 AATATAATCA GAAATAAAAT TAATGTTTGA TTGTCATAA ATGCTACTGT ATTTCTAAAA
1381 TCAACAAGTA TTTAACATTA TAAAGTGTGC AATTGGCTGC AAATGTCAGT TTTATTAAAG
1441 GGTTAGTTCA CCCAAAATG AAAATAATGT CATTAATGAC TCGCCCTCAT GTCGTTCCAA
1501 GCCCGTAAGA CCTCCGTTCA TCTTCAGAAC ACAGTTTAAG ATATTTTAGA TTAGTCCGA
1561 GAGCTTTCCTG TGCCTCCATT GAGAATGTAT GTACGGTATA CTGTCCATGT CCAGAAAGGT
1621 AATAAAAACA TCAAAGTAGT CCATGTGACA TCAGTGGGTT AGTTAGAATT TTTTGAAGCA
1681 TCGAATACAT TTTGGTCCAA AAATAACAAA ACCTACGACT TTATTCGGCA TTGTATTCTC
1741 TTCCGGGTCT GTTGTCATC CGCGTTCACG ACTTCGCAGT GACGCTACAA TGCTGAATAA
1801 AGTCGTAGGT TTTGTTATTT TTGGACCAA ATGTATTTTC GATGCTTCAA ATAATTCTAC
1861 CTAACCCACT GATGTCACAT GGACTACTTT GATGTTTTTA TTACCTTTCT GGACATGGAC
1921 AGTATACCGT ACATACATTT TCAGTGGAGG GACAGAAAGC TCTCGGACTA AATCTAAAAT
1981 ATCTTAAACT GTGTTCCGAA GATGAACGGA GGTGTTACGG GCTTGGAACG ACATGAGGGT
2041 GAGTCATTAA TGACATCTTT TCATTTTGG GTGAACAAAC CCTTTAATGC TGTAAATCAGA
2101 GAGTGTATGT GTAATTCTTA CATTTATTGC ATACAATATA AATATTTATT TCTTGTTTTT
2161 ACAGAGAATG CACCCAAATT ACCTCAAAAA CTACTCTAAA TTGACAGCAC AGAAGAGAAA
2221 GATCGGGACC TCCACCCATG CTTCCAGCAG TAAGCAACTG AAAGTTGACT CAGTTTTCCC
2281 AGTCAAACAT GTGTCTCCAG TCACTGTGAA CAAAGCTATA TTAAGGTACA TCATTCAAGG
2341 ACTTCATCCT TTCAGCACTG TTGATCTGCC ATCATTTAAA GAGCTGATTA GTACACTGCA
2401 GCCTGGCATT TCTGTCATTA CAAGGCCTAC TTTACGCTCC AAGATAGCTG AAGCTGCTCT
2461 GATCATGAAA CAGAAAGTGA CTGCTGCCAT GAGTGAAGTT GAATGGATTG CAACCACAAC
2521 GGATTGTTGG ACTGCACGTA GAAAGTCATT CATTTGGTGA ACTGCTCACT GGATCAACCC
2581 TGGAAGTCTT GAAAGACATT CCGCTGCACT TGCCTGCAAA AGATTAATGG GCTCTCATAC
2641 TTTTGAGGTA CTGGCCAGTG CCATGAATGA TATCCACTCA GAGTATGAAA TACGTGACAA
2701 GGTGTTTGGT ACAACCACAG ACAGTGGTTC CAACTTTATG AAGGCTTTCA GAGTTTTTGG
2761 TGTGGAAAC AATGATATCG AGACTGAGGC AAGAAGGTGT GAAAGTGATG AACTGATTC
2821 TGAAGGCTGT GGTGAGGGAA GTGATGGTGT GGAATTCCAA GATGCCTCAC GAGTCCTGGA
2881 CCAAGACGAT GGCTTCGAAT TCCAGCTACC AAAACATCAA AAGTGTGCCT GTCACCTACT
2941 TAACCTAGTC TCAAGCGTTG ATGCCCCAAA AGCTCTCTCA AATGAACACT ACAAGAACT
3001 CTACAGATCT GTCTTTGGCA AATGCCAAGC TTTATGGAAT AAAAGCAGCC GATCGGCTCT
3061 AGCAGCTGAA GCTGTTGAAT CAGAAAGCCG GCTTCAGCTT TTAAGGCCAA ACCAAACGCG

```

3121 GTGGAATTCA ACTTTTATGG CTGTTGACAG AATTCTTCAA ATTTGCAAAG AAGCAGGAGA
 3181 AGGCGCACTT CGGAATATAT GCACCTCTCT TGAGGTTCCA ATGTAAGTGT TTTTCCCCTC
 3241 TATCGATGTA AACAAATGTG GGTGTTTTTT GTTTAATACT CTTTGATTAT GCTGATTTCT
 3301 CCTGTAGGTT TAATCCAGCA GAAATGCTGT TCTTGACAGA GTGGGCCAAC ACAATGCGTC
 3361 CAGTTGCAAA AGTACTCGAC ATCTTGCAAG CGGAAACGAA TACACAGCTG GGGTGGCTGC
 3421 TGCCTAGTGT CCATCAGTTA AGCTTGAAAC TTCAGCGACT CCACCATTCT CTCAGGTACT
 3481 GTGACCCACT TGTGGATGCC CTACAACAAG GAATCCAAAC ACGATTCAAG CATATGTTTG
 3541 AAGATCCTGA GATCATAGCA GCTGCCATCC TTCTCCCTAA ATTTCGGACC TCTTGGACAA
 3601 ATGATGAAAC CATCATAAAA CGAGGTAAAT GAATGCAAGC AACATACACT TGACGAATTC
 3661 TAATCTGGGC AACCTTTGAG CCATACCAAA ATTATTCTTT TATTTATTTA TTTTTCGACT
 3721 TTTTAGGAAT GTTATATCCC ATCTTTGGCT GTGATCTCAA TATGAATATT GATGTAAAGT
 3781 ATTCTTGCAG CAGGTTGTAG TTATCCCTCA GTGTTTCTTG AAACCAAAC CATATGTATC
 3841 ATATGTGGTT TGGAAATGCA GTTAGATTTT ATGCTAAAAT AAGGGATTTG CATGATTTTA
 3901 GATGTAGATG ACTGCACGTA AATGTAGTTA ATGACAAAAT CCATAAAATT TGTTCCCAGT
 3961 CAGAAGCCCC TCAACCAAAC TTTTCTTTGT GTCTGCTCAC TGTGCTTGTA GGCATGGACT
 4021 ACATCAGAGT GCATCTGGAG CCTTTGGACC ACAAGAAGGA ATTGGCCAAC AGTTCATCTG
 4081 ATGATGAAGA TTTTTTCGCT TCTTTGAAAC CGACAACACA TGAAGCCAGC AAAGAGTTGG
 4141 ATGGATATCT GGCCTGTGTT TCAGACACCA GGGAGTCTCT GCTCACGTTT CCTGCTATTT
 4201 GCAGCCTCTC TATCAAGACT AATACACCTC TTCCCGCATC GGCTGCCTGT GAGAGGCTTT
 4261 TCAGCACTGC AGGATTGCTT TTCAGCCCCA AAAGAGCTAG GCTTGACACT AACAATTTTG
 4321 AGAATCAGCT TCTACTGAAG TTAAATCTGA GGTTTTACAA CTTTGAGTAG CGTGTACTGG
 4381 CATTAGATTG TCTGTCTTAT AGTTTGATAA TTAAATACAA ACAGTTCTAA AGCAGGATAA
 4441 AACCTTGTAT GCATTTCAAT TAATGTTTTT TGAGATTAAA AGCTTAAACA AGAATCTCTA
 4501 CTTTTCTTTC TTGCTTTTAC TTTTACTTCC TTAATACTCA AGTACAATTT TAATGGAGTA
 4561 CTTTTTTACT TTTACTCAAG TAAGATTCTA GCCAGATACT TTTACTTTTA ATTGAGTAAA
 4621 ATTTTCCCTA AGTACTTGTA CTTTCACTTG AGTAAAATTT TTGAGTACTT TTTACACCTC
 4681 TG (SEQ ID NO: 17041).

[0124] Exemplary transposon/transposase systems of the disclosure include, but are not limited to, piggyBac and piggyBac-like transposons and transposases.

[0125] PiggyBac and piggyBac-like transposases recognizes transposon-specific inverted terminal repeat sequences (ITRs) on the ends of the transposon, and moves the contents between the ITRs into TTAA or TTAT chromosomal sites. The piggyBac or piggyBac-like transposon system has no payload limit for the genes of interest that can be included between the ITRs.

[0126] In certain embodiments, and, in particular, those embodiments wherein the transposon is a piggyBac transposon, the transposase is a piggyBac™, Super piggyBac™ (SPB) transposase. In certain embodiments, and, in particular, those embodiments wherein

the transposase is a piggyBacTM, Super piggyBacTM (SPB), the sequence encoding the transposase is an mRNA sequence.

[0127] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme.

[0128] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or a piggyBac-like transposase enzyme. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MGSSLDDEHI LSALLQSDDE LVGEDSDSEI SDHVSEDDVQ SDTEEEAFIDE VHEVQPTSSG
61 SEILDEQNVI EQPGSSLASN RILTLPQRTI RGKNKHCWST SKSTRRSRVVS ALNIVRSQRG
121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTGATFRD TNEDEIYAFF
181 GILVMTAVRK DNHMSTDDLF DRSLSMVYVS VMSRDRFDFL IRCLRMDDKS IRPTLRENDV
241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RMYIPNKPSK YGIKILMMCD
301 SGTKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCDGP LTLVSYKPKP AKMVYLLSSC
421 DEDASINEST GKPQMVMYYN QTKGGVDTLD QMCSVMTCR KTNRWPMALL YGMINIACIN
481 SFIIYSHNVS SKGEKVQSRK KFMRNLYMSL TSSEMRKRLE APTLKRYLRD NISNILPNEV
541 PGTSDDSTEE PVMKKRTYCT YCPSKIRRKA NASCKCKKV ICREHNIDMC QSCF (SEQ ID
NO: 14487).
```

[0129] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at one or more of positions 30, 165, 282, or 538 of the sequence:

```

1 MGSSLDDEHI LSALLQSDDE LVGEDSDSEI SDHVSEDDVQ SDTEEEAFIDE VHEVQPTSSG
61 SEILDEQNVI EQPGSSLASN RILTLPQRTI RGKNKHCWST SKSTRRSRVVS ALNIVRSQRG
121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTGATFRD TNEDEIYAFF
181 GILVMTAVRK DNHMSTDDLF DRSLSMVYVS VMSRDRFDFL IRCLRMDDKS IRPTLRENDV
241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RMYIPNKPSK YGIKILMMCD
301 SGTKYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCDGP LTLVSYKPKP AKMVYLLSSC
421 DEDASINEST GKPQMVMYYN QTKGGVDTLD QMCSVMTCR KTNRWPMALL YGMINIACIN
481 SFIIYSHNVS SKGEKVQSRK KFMRNLYMSL TSSEMRKRLE APTLKRYLRD NISNILPNEV
541 PGTSDDSTEE PVMKKRTYCT YCPSKIRRKA NASCKCKKV ICREHNIDMC QSCF (SEQ ID
NO: 14487).
```

[0130] In certain embodiments, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme that comprises or consists of an amino acid sequence having an amino

acid substitution at two or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at three or more of positions 30, 165, 282, or 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme that comprises or consists of an amino acid sequence having an amino acid substitution at each of the following positions 30, 165, 282, and 538 of the sequence of SEQ ID NO: 14487. In certain embodiments, the amino acid substitution at position 30 of the sequence of SEQ ID NO: 14487 is a substitution of a valine (V) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 165 of the sequence of SEQ ID NO: 14487 is a substitution of a serine (S) for a glycine (G). In certain embodiments, the amino acid substitution at position 282 of the sequence of SEQ ID NO: 14487 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 538 of the sequence of SEQ ID NO: 14487 is a substitution of a lysine (K) for an asparagine (N).

[0131] In certain embodiments of the methods of the disclosure, the transposase enzyme is a Super piggyBac™ (SPB) or piggyBac-like transposase enzyme. In certain embodiments, the Super piggyBac™ (SPB) or piggyBac-like transposase enzyme of the disclosure may comprise or consist of the amino acid sequence of the sequence of SEQ ID NO: 14487 wherein the amino acid substitution at position 30 is a substitution of a valine (V) for an isoleucine (I), the amino acid substitution at position 165 is a substitution of a serine (S) for a glycine (G), the amino acid substitution at position 282 is a substitution of a valine (V) for a methionine (M), and the amino acid substitution at position 538 is a substitution of a lysine (K) for an asparagine (N). In certain embodiments, the Super piggyBac™ (SPB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MGSSLDDEHI  LSALLQSDDE  LVGEDSDSEV  SDHVSEDDVQ  SDTEEAFIDE  VHEVQPTSSG
61  SEILDEQNVI  EQPGSSLASN  RILTLPQRTI  RGKNKHCWST  SKSTRRSRVS  ALNIVRSQRG
121 PTRMCRNIYD  PLLCFKLFFT  DEIISEIVKW  TNAEISLKRR  ESMTSATFRD  TNEDEIYAFF
181 GILVMTAVRK  DNHMSTDDLF  DRSLSMVYVS  VMSRDREDFL  IRCLRMDDKS  IRPTLRENDV
241 FTPVRKIWDL  FIHQCIQNYT  PGAHLTIDEQ  LLGFRGRCPF  RVIIPNKPSK  YGIKILMMCD
301 SGTKYMINGM  PYLGRGTQTN  GVPLGEYYVK  ELSKPVHGSC  RNITCDNWET  SIPLAKNLLQ
361 EPYKLTIVGT  VRSNKREIPE  VLKNSRSRPV  GTSMFCEFGP  LTLVSYKPKP  AKMVYLLSSC

```

421 DEDASINEST GKPQVMYYN QTKGGVDTLD QMCSVMTCR KTNRWPMALL YGMINIACIN
 481 SFIIYSHNVS SKGEKVQSRK KEMRNLYMSL TSSEMRKRLE APTLKRYLRD NISNILEKEV
 541 PGTSDSTEE PVMKKRTYCT YCPKIRRK A NASCKKCKKV ICREHNIDMC QSCF (SEQ ID NO:
 14484) .

[0132] In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™, Super piggyBac™ or piggyBac-like transposase enzyme may further comprise an amino acid substitution at one or more of positions 3, 46, 82, 103, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 258, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 486, 503, 552, 570 and 591 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™, Super piggyBac™ or piggyBac-like transposase enzyme may further comprise an amino acid substitution at one or more of positions 46, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 485, 503, 552 and 570. In certain embodiments, the amino acid substitution at position 3 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an asparagine (N) for a serine (S). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an alanine (A). In certain embodiments, the amino acid substitution at position 46 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 82 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for an isoleucine (I). In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position 119 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for an arginine (R). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) a cysteine (C). In certain embodiments, the amino acid substitution at position 125 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 177 of SEQ ID NO: 14487 or

SEQ ID NO: 14484 is a substitution of a histidine (H) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 180 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 185 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 187 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for an alanine (A). In certain embodiments, the amino acid substitution at position 200 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 207 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a valine (V). In certain embodiments, the amino acid substitution at position 209 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a valine (V). In certain embodiments, the amino acid substitution at position 226 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a methionine (M). In certain embodiments, the amino acid substitution at position 235 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a leucine (L). In certain embodiments, the amino acid substitution at position 240 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 241 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a phenylalanine (F). In certain embodiments, the amino acid substitution at position 243 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a proline (P). In certain embodiments, the amino acid substitution at position 258 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tryptophan (W) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tyrosine (Y) for a leucine (L). In certain embodiments, the amino acid substitution at position 296 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid

substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for a methionine (M). In certain embodiments, the amino acid substitution at position 298 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a proline (P). In certain embodiments, the amino acid substitution at position 311 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine for a proline (P). In certain embodiments, the amino acid substitution at position 315 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for an arginine (R). In certain embodiments, the amino acid substitution at position 319 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a threonine (T). In certain embodiments, the amino acid substitution at position 327 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 328 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a tyrosine (Y). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a cysteine (C). In certain embodiments, the amino acid substitution at position 340 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a cysteine (C). In certain embodiments, the amino acid substitution at position 421 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a histidine (H) for the aspartic acid (D). In certain embodiments, the amino acid substitution at position 436 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a valine (V). In certain embodiments, the amino acid substitution at position 456 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a tyrosine (Y) for a methionine (M). In certain embodiments, the amino acid substitution at position 470 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a phenylalanine (F) for a leucine (L). In certain embodiments, the amino acid substitution at position 485 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a serine (S). In certain embodiments, the amino acid substitution at position 503 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a leucine (L) for a methionine (M). In certain embodiments, the amino acid substitution at position 503 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an isoleucine (I) for a methionine (M). In certain

embodiments, the amino acid substitution at position 552 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a lysine (K) for a valine (V). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a threonine (T) for an alanine (A). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a glutamine (Q). In certain embodiments, the amino acid substitution at position 591 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an arginine (R) for a glutamine (Q).

[0133] In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or piggyBac-like transposase enzyme or may comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at one or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments of the methods of the disclosure, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or piggyBac-like transposase enzyme may comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at two, three, four, five, six or more of positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, including those embodiments wherein the transposase comprises the above-described mutations at positions 30, 165, 282 and/or 538, the piggyBac™ or piggyBac-like transposase enzyme may comprise or the Super piggyBac™ transposase enzyme may further comprise an amino acid substitution at positions 103, 194, 372, 375, 450, 509 and 570 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, the amino acid substitution at position 103 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a proline (P) for a serine (S). In certain embodiments, the amino acid substitution at position 194 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a valine (V) for a methionine (M). In certain embodiments, the amino acid substitution at position 372 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for an arginine (R). In certain embodiments, the amino acid substitution at position 375 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an alanine (A) for a lysine (K). In certain embodiments, the amino acid substitution at position 450 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of an asparagine (N) for an aspartic

acid (D). In certain embodiments, the amino acid substitution at position 509 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a glycine (G) for a serine (S). In certain embodiments, the amino acid substitution at position 570 of SEQ ID NO: 14487 or SEQ ID NO: 14484 is a substitution of a serine (S) for an asparagine (N). In certain embodiments, the piggyBac™ or piggyBac-like transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487. In certain embodiments, including those embodiments wherein the piggyBac™ or piggyBac-like transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, the piggyBac™ or piggyBac-like transposase enzyme may further comprise an amino acid substitution at positions 372, 375 and 450 of the sequence of SEQ ID NO: 14487 or SEQ ID NO: 14484. In certain embodiments, the piggyBac™ or piggyBac-like transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 14487, and a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 14487. In certain embodiments, the piggyBac™ or piggyBac-like transposase enzyme may comprise a substitution of a valine (V) for a methionine (M) at position 194 of SEQ ID NO: 14487, a substitution of an alanine (A) for an arginine (R) at position 372 of SEQ ID NO: 14487, a substitution of an alanine (A) for a lysine (K) at position 375 of SEQ ID NO: 14487 and a substitution of an asparagine (N) for an aspartic acid (D) at position 450 of SEQ ID NO: 14487.

[0134] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from an insect. In certain embodiments, the insect is *Trichoplusia ni* (GenBank Accession No. AAA87375; SEQ ID NO: 17083), *Argyrogramma agnata* (GenBank Accession No. GU477713; SEQ ID NO: 17084, SEQ ID NO: 17085), *Anopheles gambiae* (GenBank Accession No. XP_312615 (SEQ ID NO: 17086); GenBank Accession No. XP_320414 (SEQ ID NO: 17087); GenBank Accession No. XP_310729 (SEQ ID NO: 17088)), *Aphis gossypii* (GenBank Accession No. GU329918; SEQ ID NO: 17089, SEQ ID NO: 17090), *Acyrtosiphon pisum* (GenBank Accession No. XP_001948139; SEQ ID NO: 17091), *Agrotis ipsilon* (GenBank Accession No. GU477714; SEQ ID NO: 17092, SEQ ID NO: 17093), *Bombyx mori* (GenBank Accession No. BAD11135; SEQ ID NO: 17094), *Chilo suppressalis* (GenBank Accession No. JX294476; SEQ ID NO: 17095, SEQ ID NO: 17096), *Drosophila melanogaster* (GenBank Accession No. AAL39784; SEQ ID NO: 17097), *Helicoverpa armigera* (GenBank Accession No. ABS18391; SEQ ID NO: 17098), *Heliothis*

virescens (GenBank Accession No. ABD76335; SEQ ID NO: 17099), *Macdunnoughia crassisigna* (GenBank Accession No. EU287451; SEQ ID NO: 17100, SEQ ID NO: 17101), *Pectinophora gossypiella* (GenBank Accession No. GU270322; SEQ ID NO: 17102, SEQ ID NO: 17103), *Tribolium castaneum* (GenBank Accession No. XP_001814566; SEQ ID NO: 17104), *Ctenoplusia agnata* (also called *Argyrogramma agnata*), *Messour bouvieri*, *Megachile rotundata*, *Bombus impatiens*, *Mamestra brassicae*, *Mayetiola destructor* or *Apis mellifera*.

[0135] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from an insect. In certain embodiments, the insect is *Trichoplusia ni* (AAA87375).

[0136] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from an insect. In certain embodiments, the insect is *Bombyx mori* (BAD11135).

[0137] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from a crustacean. In certain embodiments, the crustacean is *Daphnia pulicaria* (AAM76342, SEQ ID NO: 17105).

[0138] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from a vertebrate. In certain embodiments, the vertebrate is *Xenopus tropicalis* (GenBank Accession No. BAF82026; SEQ ID NO: 17106), *Homo sapiens* (GenBank Accession No. NP_689808; SEQ ID NO: 17107), *Mus musculus* (GenBank Accession No. NP_741958; SEQ ID NO: 17108), *Macaca fascicularis* (GenBank Accession No. AB179012; SEQ ID NO: 17108, SEQ ID NO: 17109), *Rattus norvegicus* (GenBank Accession No. XP_220453; SEQ ID NO: 17110) or *Myotis lucifugus*.

[0139] In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from a urochordate. In certain embodiments, the urochordate is *Ciona intestinalis* (GenBank Accession No. XP_002123602; SEQ ID NO: 17111).

[0140] In certain embodiments, the piggyBac or piggyBac-like transposase inserts a transposon at the sequence 5'-TTAT-3' within a chromosomal site (a TTAT target sequence).

[0141] In certain embodiments, the piggyBac or piggyBac-like transposase inserts a transposon at the sequence 5'-TTAA-3' within a chromosomal site (a TTAA target sequence).

[0142] In certain embodiments, the target sequence of the piggyBac or piggyBac-like transposon comprises or consists of 5'-CTAA-3', 5'-TTAG-3', 5'-ATAA-3', 5'-TCAA-3',

5'-AGTT-3', 5'-ATTA-3', 5'-GTTA-3', 5'-TTGA-3', 5'-TTTA-3', 5'-TTAC-3', 5'-ACTA-3', 5'-AGGG-3', 5'-CTAG-3', 5'-TGAA-3', 5'-AGGT-3', 5'-ATCA-3', 5'-CTCC-3', 5'-TAAA-3', 5'-TCTC-3', 5'-TGAA-3', 5'-AAAT-3', 5'-AATC-3', 5'-ACAA-3', 5'-ACAT-3', 5'-ACTC-3', 5'-AGTG-3', 5'-ATAG-3', 5'-CAAA-3', 5'-CACA-3', 5'-CATA-3', 5'-CCAG-3', 5'-CCCA-3', 5'-CGTA-3', 5'-GTCC-3', 5'-TAAG-3', 5'-TCTA-3', 5'-TGAG-3', 5'-TGTT-3', 5'-TTCA-3', 5'-TTCT-3' and 5'-TTTT-3'.

[0143] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Bombyx mori*. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNIFVTSD MLQEILHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIALLYL AGLIKSNRQS LKDLWRDGT GVDIFRTTMS LQRFQFLQNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ CCQAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FDVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELSANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRANKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KHSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14504).

```

[0144] The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNIFVTSD MLQEILHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIALLYL AGLIKSNRQS LKDLWRDGT GVDIFRTTMS LQRFQFLQNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ CCQAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14505).

```

[0145] In certain embodiments, the piggyBac or piggyBac-like transposase is fused to a nuclear localization signal. In certain embodiments, the amino acid sequence of the piggyBac or piggyBac-like transposase fused to a nuclear localization signal is encoded by a polynucleotide sequence comprising:

```

1 atggcaccaca aaaagaaacg taaagtgatg gacattgaaa gacaggaaga aagaatcagg
61 gcgatgctcg aagaagaact gagcgactac tccgacgaat cgtcatcaga ggatgaaacc
121 gaccactgta gcgagcatga ggttaactac gacaccgagg aggagagaat cgactctgtg
181 gatgtgccct ccaactcacg ccaagaagag gccaatgcaa ttatcgcaaa cgaatcggac
241 agcgatccag acgatgatct gccactgtcc ctctgtgcgc agcgggccag cgcttcgaga
301 caagtgtcag gtccattcta cacttcgaag gacggcacta agtggtagaa gaattgccag
361 cgacctaacg tcagactccg ctccgagaat atcgtgaccg aacaggctca ggtcaagaat
421 atcgcccgcg acgcctcgac tgagtacgag tgttggaata tcttcgtgac ttccgacatg
481 ctgcaagaaa ttctgacgca caccaacagc tcgattaggc atcgccagac caagactgca
541 gcggagaact catcggccga aacctccttc tatatgcaag agactactct gtgcgaactg
601 aaggcgctga ttgcactgct gtacttggcc ggccatcatc aatcaaatag gcagagcctc
661 aaagatctct ggagaacgga tggaaactgga gtggatatct ttccgacgac tatgagcttg
721 cagcgggttc agtttctgca aaacaatatc agattcgacg acaagtccac ccgggacgaa
781 aggaacacaga ctgacaacat ggctgcgttc cggtaaatat tcgatcagtt tgtgcagtgc
841 tgccaaaacg cttatagccc atcggaattc ctgaccatcg acgaaatgct tctctccttc
901 cgggggcgct gcctgttcgg agtgtacatc ccgaacaagc cggctaaata cggaatcaaa
961 atcctggccc tgggtggacg caagaatttc tacgtcgtga atctcgaagt gtacgcagga
1021 aagcaaccgt cgggaccgta cgctgttttc aaccgcccgt ttgaagtcgt cgagcggctt
1081 attcagccgg tggccagatc ccaccgcaat gttaccttcg acaattgggt caccggctac
1141 gagctgatgc ttcaccttc gaacgagtag cggctcacta gcgtggggac tgtcaggaag
1201 aacaagcggc agatcccaga atccttcctc cgcaccgacc gccagcctaa ctctccgtg
1261 ttccgatttc aaaaggatat cagccttgct tcgtacgccc ccaagaaaaa caaggctcgtg
1321 gtcgtgatga gcaccatgca tcacgacaac agcatcgacg agtcaaccgg agaaaagcaa
1381 aagcccgaga tgatcacctt ctacaattca actaaggccg gcgtcgacgt cgtggatgaa
1441 ctgtgcgcga actataacgt gtcccggaac tctaagcggg ggccatagac tctcttctac
1501 ggagtgtgta atatggccgc aatcaacgcg tgcattatct accgcaccaa caagaacgtg
1561 accatcaagc gcaccgagtt catcagatcg ctgggtttga gcatgatcta cgagcacctc
1621 cattcacgga acaagaagaa gaatatccct acttacctga ggcagcgtat cgagaagcag
1681 ttgggagAAC caagcccgcg ccacgtgaac gtgcgggggc gctacgtgcg gtgccaagat
1741 tgcccgtaca aaaaggaccg caaaaccaa agatcgtgta acgctgccc caaacctatc
1801 tgcattggagc atgcaaatc tctgtgtgaa aattgtgctg aactcgattc ctccctg
(SEQ ID NO: 14629).

```

[0146] In certain embodiments, the piggyBac or piggyBac-like transposase is hyperactive. A hyperactive piggyBac or piggyBac-like transposase is a transposase that is more active than the naturally occurring variant from which it is derived. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Bombyx mori*. In certain embodiments, the piggyBac or piggyBac-like transposase is a hyperactive variant of SEQ ID NO: 14505. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQMSGPHYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNI EFTSD MLQEILTHTN SSIRWRQTKT AAENSSASTS
181 FYMQETTLCE LKALIGLLYI AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLQNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ SCQNAYSPSE FLTIDEMLLS FRGRCLERVY
301 IPNKPAKYGI KILALVDAKN FYVKNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCFYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSHL (SEQ ID NO: 14576).

```

[0147] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14576. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNI FVTSD MLQEILTHTN SSIRWRQTKT AAENSSAETS
181 FYMQETTLCE LKALIGLLYI AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ SCQNAISPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVHNLEVYA GKQPSGPIYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YEVMHLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAHLDS (SEQ ID NO: 14630).

```

[0148] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNI FVTSD MLQEILTHTN SSIRWRQTKT AAENSSASTS
181 FYMQETTLCE LKALIGLLYI AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ SCQNAISPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVKNLEVYA GKQPSGPIYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIAM QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14631).

```

[0149] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNI FVTSD MLQEILTHTN SSIRWRQTKT AAENSSAETS
181 FYMQETTLCE LKALIGLLYI AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ SCQNAISPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVKNLEVYA GKQPSGPIYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKTQIPENF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELQANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14632).

```

[0150] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNI FVTSD MLQEILTHTN SSIRWRQTKT AAENSSAETS
181 FYMQETTLCE LKALIGLLYI AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLQNN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ SCQNAISPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVKNLEVYA GKQPSGPIYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI

```

541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14633).

[0151] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

1 MDIERQEERI RAMLEEEELSD YSDESSSEDE TDHCSEHEVN YDTEEEERIDS VDVPSNSRQE
61 E7NAILIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QPENVRLRSE
121 NIVTEQAQVK NIARDAST EY ECWNIFVTSD MLQEILHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIALLYL AGLIKSNRQS LKDLWRTDGT GVDIFRTTMS LQRFQFLQNN
241 IRFDDKSTRD ERKQTDNMAA FRSFIDQFVQ CCQNAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN DYVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVKGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14634).

[0152] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase is more active than the transposase of SEQ ID NO: 14505. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase is at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or any percentage in between identical to SEQ ID NO: 14505.

[0153] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises an amino acid substitution at a position selected from 92, 93, 96, 97, 165, 178, 189, 196, 200, 201, 211, 215, 235, 238, 246, 253, 258, 261, 263, 271, 303, 321, 324, 330, 373, 389, 399, 402, 403, 404, 448, 473, 484, 507, 523, 527, 528, 543, 549, 550, 557, 601, 605, 607, 609, 610 or a combination thereof (relative to SEQ ID NO: 14505). In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises an amino acid substitution of Q92A, V93L, V93M, P96G, F97H, F97C, H165E, H165W, E178S, E178H, C189P, A196G, L200I, A201Q, L211A, W215Y, G219S, Q235Y, Q235G, Q238L, K246I, K253V, M258V, F261L, S263K, C271S, N303R, F321W, F321D, V324K, V324H, A330V, L373C, L373V, V389L, S399N, R402K, T403L, D404Q, D404S, D404M, N441R, G448W, E449A, V469T, C473Q, R484K, T507C, G523A, I527M, Y528K, Y543I, E549A, K550M, P557S, E601V, E605H, E605W, D607H, S609H, L610I or any combination thereof. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises an amino acid substitution of Q92A, V93L, V93M, P96G, F97H, F97C, H165E, H165W, E178S, E178H, C189P, A196G, L200I, A201Q, L211A, W215Y, G219S, Q235Y, Q235G, Q238L, K246I, K253V, M258V, F261L, S263K, C271S, N303R, F321W, F321D, V324K, V324H, A330V, L373C, L373V, V389L, S399N, R402K, T403L, D404Q, D404S, D404M,

N441R, G448W, E449A, V469T, C473Q, R484K T507C, G523A, I527M, Y528K Y543I, E549A, K550M, P557S, E601V, E605H, E605W, D607H, S609H and L610I.

[0154] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises one or more substitutions of an amino acid that is not wild type, wherein the one or more substitutions a for wild type amino acid comprises a substitution of E4X, A12X, M13X, L14X, E15X, D20X, E24X, S25X, S26X, S27X, D32X, H33X, E36X, E44X, E45X, E46X, I48X, D49X, R58X, A62X, N63X, A64X, I65X, I66X, N68X, E69X, D71X, S72X, D76X, P79X, R84X, Q85X, A87X, S88X, Q92X, V93X, S94X, G95X, P96X, F97X, Y98X, T99X, I145X, S149X, D150X, L152X, E154X, T157X, N160X, S161X, S162X, H165X, R166X, T168X, K169X, T170X, A171X, E173X, S175X, S176X, E178X, T179X, M183X, Q184X, T186X, T187X, L188X, C189X, L194X, I195X, A196X, L198X, L200X, A201X, L203X, I204X, K205X, A206X, N207X, Q209X, S210X, L211X, K212X, D213X, L214X, W215X, R216X, T217X, G219X, V222X, D223X, I224X, T227X, M229X, Q235X, L237X, Q238X, N239X, N240X, P302X, N303X, P305X, A306X, K307X, Y308X, I310X, K311X, I312X, L313X, A314X, L315X, V316X, D317X, A318X, K319X, N320X, F321X, Y322X, V323X, V324X, L326X, E327X, V328X, A330X, Q333X, P334X, S335X, G336X, P337X, A339X, V340X, S341X, N342X, R343X, P344X, F345X, E346X, V347X, E349X, I352X, Q353X, V355X, A356X, R357X, N361X, D365X, W367X, T369X, G370X, L373X, M374X, L375X, H376X, N379X, E380X, R382X, V386X, V389X, N392X, R394X, Q395X, S399X, F400X, I401X, R402X, T403X, D404X, R405X, Q406X, P407X, N408X, S409X, S410X, V411X, F412X, F414X, Q415X, I418X, T419X, L420X, N428X, V432X, M434X, D440X, N441X, S442X, I443X, D444X, E445X, G448X, E449X, Q451X, K452X, M455X, I456X, T457X, F458X, S461X, A464X, V466X, Q468X, V469X, E471X, L472X, C473X, A474X, K483X, W485X, T488X, L489X, Y491X, G492X, V493X, M496X, I499X, C502X, I503X, T507X, K509X, N510X, V511X, T512X, I513X, R515X, E517X, S521X, G523X, L524X, S525X, I527X, Y528X, E529X, H532X, S533X, N535X, K536X, K537X, N539X, I540X, T542X, Y543X, Q546X, E549X, K550X, Q551X, G553X, E554X, P555X, S556X, P557X, R558X, H559X, V560X, N561X, V562X, P563X, G564X, R565X, Y566X, V567X, Q570X, D571X, P573X, Y574X, K576X, K581X, S583X, A586X, A588X, E594X, F598X, L599X, E601X, N602X, C603X, A604X, E605X, L606X, D607X, S608X, S609X or L610X (relative to SEQ ID NO: 14505). A list of hyperactive amino acid substitutions can be found in US patent No. 10,041,077, the contents of which are incorporated herein by reference in their entirety.

[0155] In certain embodiments, the piggyBac or piggyBac-like transposase is integration deficient. In certain embodiments, an integration deficient piggyBac or piggyBac-like transposase is a transposase that can excise its corresponding transposon, but that integrates the excised transposon at a lower frequency than a corresponding wild type transposase. In certain embodiments, the piggyBac or piggyBac-like transposase is an integration deficient variant of SEQ ID NO: 14505.

[0156] In certain embodiments, the excision competent, integration deficient piggyBac or piggyBac-like transposase comprises one or more substitutions of an amino acid that is not wild type, wherein the one or more substitutions a for wild type amino acid comprises a substitution of R9X, A12X, M13X, D20X, Y21K, D23X, E24X, S25X, S26X, S27X, E28X, E30X, D32X, H33X, E36X, H37X, A39X, Y41X, D42X, T43X, E44X, E45X, E46X, R47X, D49X, S50X, S55X, A62X, N63X, A64X, I66X, A67X, N68X, E69X, D70X, D71X, S72X, D73X, P74X, D75X, D76X, D77X, I78X, S81X, V83X, R84X, Q85X, A87X, S88X, A89X, S90X, R91X, Q92X, V93X, S94X, G95X, P96X, F97X, Y98X, T99X, W012X, G103X, Y107X, K108X, L117X, I122X, Q128X, I312X, D135X, S137X, E139X, Y140X, I145X, S149X, D150X, Q153X, E154X, T157X, S161X, S162X, R164X, H165X, R166X, Q167X, T168X, K169X, T170X, A171X, A172X, E173X, R174X, S175X, S176X, A177X, E178X, T179X, S180X, Y182X, Q184X, E185X, T187X, L188X, C189X, L194X, I195X, A196X, L198X, L200X, A201X, L203X, I204X, K205X, N207X, Q209X, L211X, D213X, L214X, W215X, R216X, T217X, G219X, T220X, V222X, D223X, I224X, T227X, T228X, F234X, Q235X, L237X, Q238X, N239X, N240X, N303X, K304X, I310X, I312X, L313X, A314X, L315X, V316X, D317X, A318X, K319X, N320X, F321X, Y322X, V323X, V324X, N325X, L326X, E327X, V328X, A330X, G331X, K332X, Q333X, S335X, P337X, P344X, F345X, E349X, H359X, N361X, V362X, D365X, F368X, Y371X, E372X, L373X, H376X, E380X, R382X, R382X, V386X, G387X, T388X, V389X, K391X, N392X, R394X, Q395X, E398X, S399X, F400X, I401X, R402X, T403X, D404X, R405X, Q406X, P407X, N408X, S409X, S410X, Q415X, K416X, A424X, K426X, N428X, V430X, V432X, V433X, M434X, D436X, D440X, N441X, S442X, I443X, D444X, E445X, S446X, T447X, G448X, E449X, K450X, Q451X, E454X, M455X, I456X, T457X, F458X, S461X, A464X, V466X, Q468X, V469X, C473X, A474X, N475X, N477X, K483X, R484X, P486X, T488X, L489X, G492X, V493X, M496X, I499X, I503X, Y505X, T507X, N510X, V511X, T512X, I513X, K514X, T516X, E517X, S521X, G523X, L524X, S525X, I527X, Y528X, L531X, H532X, S533X, N535X, I540X, T542X, Y543X, R545X, Q546X, E549X, L552X, G553X, E554X, P555X,

S556X, P557X, R558X, H559X, V560X, N561X, V562X, P563X, G564X, V567X, Q570X, D571X, P573X, Y574X, K575X, K576X, N585X, A586X, M593X, K596X, E601X, N602X, A604X, E605X, L606X, D607X, S608X, S609X or L610X (relative to SEQ ID NO: 14505).

A list of integration deficient amino acid substitutions can be found in US patent No.

10,041,077, the contents of which are incorporated by reference in their entirety.

[0157] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNIFVTSD MLQEILTHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIALLYL AGLIKSNRQS LKDLWRKDGT GVDIFRTTMS LQRFQFLN
241 IRFDDISTRD ERKQTDNMAA FRSIFDQFVQ CCQNAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKKWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMMYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEVPVRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14606) .

```

In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNIFVTSD MLQEILTHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIGLLYL AGLIKSNRQS LKDLWRDGT GVDIFRTTMS LQRFYFLQNN
241 IRFDDKSTLD ERKQTDNMAA FRSIFDQFVQ SCQNAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN FYVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YELMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKRWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIYEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NYPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 VNCAELDSSL (SEQ ID NO: 14607) .

```

In certain embodiments, the piggyBac or piggyBac-like transposase that is integration deficient comprises a sequence of:

```

1 MDIERQEERI RAMLEEELSD YSDESSSEDE TDHCSEHEVN YDTEEEERIDS VDVPSNSRQE
61 EANAIIANES DSDPDDDLPL SLVRQRASAS RQVSGPFYTS KDGTKWYKNC QRPNVRLRSE
121 NIVTEQAQVK NIARDASTEY ECWNIFVTSD MLQEILTHTN SSIRHRQTKT AAENSSAETS
181 FYMQETTLCE LKALIALLYL AGLIKSNRQS LKDLWRKDGT GVDIFRTTMS LQRFQFLN
241 IRFDDKSTRD ERKQTDNMAA FRSIFDQFVQ CCQNAYSPSE FLTIDEMLLS FRGRCLFRVY
301 IPNKPAKYGI KILALVDAKN DYVVNLEVYA GKQPSGPYAV SNRPFEVVER LIQPVARSHR
361 NVTFDNWFTG YECMLHLLNE YRLTSVGTVR KNKRQIPESF IRTDRQPNSS VFGFQKDITL
421 VSYAPKKNKV VVMSTMHHD NSIDESTGEK QKPEMITFYN STKAGVDVVD ELCANYNVSR
481 NSKKWPMTLF YGVLNMAAIN ACIIYRTNKN VTIKRTEFIR SLGLSMIKEH LHSRNKKKNI
541 PTYLRQRIEK QLGEPSPRHV NVPGRYVRCQ DCPYKKDRKT KRSCNACAKP ICMEHAKFLC
601 ENCAELDSSL (SEQ ID NO: 14608) .

```

In certain embodiments, the integration deficient transposase comprises a sequence that is at least 90% identical to SEQ ID NO: 14608.

[0158] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Bombyx mori*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  ttatcccggc gagcatgagg cagggatatct cataccctgg taaaatttta aagttgtgta
61  ttttataaaa ttttcgtctg acaacactag cgcgctcagt agctggaggc aggagcgtgc
121 gggaggggat agtggcgtga tcgcagtgtg gcacgggaca ccggcgagat attcgtgtgc
181 aaacctgttt cgggtatgtt ataccctgcc tcattgttga cgtatttttt ttatgtaatt
241 tttccgatta ttaatttcaa ctgttttatt ggtattttta tgttatccat tgttcttttt
301 ttatgattta ctgtatcggc tgtctttcgt tcctttagtt gagttttttt ttattatttt
361 cagtttttga tcaaaa (SEQ ID NO: 14506).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  tcatattttt agtttaaaaa aataattata tgttttataa tgaaaagaat ctcattatct
61  ttcagtatta gtttgattta tattccaaag aataatattt ttgttaaatt gttgattttt
121 gtaaaccctct aaatgtttgt tgctaaaatt actgtgttta agaaaaagat taataaataa
181 taataatttc ataattaaaa acttctttca ttgaatgcca ttaaataaac cattatttta
241 caaaataaga tcaacataat tgagtaaata ataataagaa caatattata gtacaacaaa
301 atatgggtat gtcataccct gccacattct tgatgtaact ttttttcacc tcatgctcgc
361 cgggttat (SEQ ID NO: 14507).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  ttatcccggc gagcatgagg cagggatatct cataccctgg taaaatttta aagttgtgta
61  ttttataaaa ttttcgtctg acaacactag cgcgctcagt agctggaggc aggagcgtgc
121 gggaggggat agtggcgtga tcgcagtgtg gcacgggaca ccggcgagat attcgtgtgc
181 aaacctgttt cgggtatgtt ataccctgcc tcat (SEQ ID NO: 14508).

```

In certain embodiments, the piggyBac™ (PB) or piggyBac-like transposon comprises a sequence of:

```

1  taaataataa taatttcata attaaaaact tcttttcattg aatgccatta aataaaccat
61  tattttacaa aataagatca acataattga gtaaataata ataagaacaa tattatagta
121 caacaaaata tgggtatgtc ataccctgcc acattcttga tgtaactttt tttcacctca
181 tgctcgccgg gttat (SEQ ID NO: 14509).

```

[0159] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a left sequence corresponding to SEQ ID NO: 14506 and a right sequence corresponding to SEQ ID NO: 14507. In certain embodiments, one piggyBac or piggyBac-like transposon end is at least 85%, at least 90%, at least 95%, at least 98%, at least 99% identical or any percentage in between identical to SEQ ID NO: 14506 and the other piggyBac or piggyBac-like transposon end is at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or any percentage in between identical to SEQ ID NO: 14507. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14506 and SEQ ID NO: 14507 or SEQ ID NO: 14509. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14508 and SEQ ID NO: 14507 or SEQ ID NO: 14509. In certain embodiments, the left and right transposon ends share a 16 bp repeat sequence at their ends of

CCCGGCGAGCATGAGG (SEQ ID NO: 14510) immediately adjacent to the 5'-TTAT-3' target insertion site, which is inverted in the orientation in the two ends. In certain embodiments, left transposon end begins with a sequence comprising 5'-TTATCCCGGCGAGCATGAGG-3' (SEQ ID NO: 14511), and the right transposon ends with a sequence comprising the reverse complement of this sequence: 5'-CCTCATGCTCGCCGGGTTAT-3' (SEQ ID NO: 14512).

[0160] In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end comprising at least 14, 16, 18, 20, 30 or 40 contiguous nucleotides of SEQ ID NO: 14506 or SEQ ID NO: 14508. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end comprising at least 14, 16, 18, 20, 30 or 40 contiguous nucleotides of SEQ ID NO: 14507 or SEQ ID NO: 14509. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end with at least 90% identity to SEQ ID NO: 14506 or SEQ ID NO: 14508. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end with at least 90% identity to SEQ ID NO: 14507 or SEQ ID NO: 14509.

[0161] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 ttaacccggc gagcatgagg cagggatatct cataccctgg taaaatttta aagttgtgta
61 ttttataaaa ttttcgtctg acaacaactag cgcgctcagt agctggaggc aggagcgtgc
121 gggaggggat agtggcgtga tcgcagtgtg gcaaggagaca cggcgagat attcgtgtgc
181 aaacctgttt cgggtatgtt ataccctgcc tcattgttga cgtatttttt ttatgtaatt
241 tttccgatta ttaatttcaa ctgttttatt ggtattttta tgttatccat tgttcctttt
301 ttatgattta ctgtatcggc tgtctttcgt tccttttagtt gagttttttt ttattatttt
361 cagtttttga tcaaa (SEQ ID NO: 14515).

```

[0162] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 tcataattttt agttttaaaaa aataattata tgttttataa tgaaaagaat ctcattatct
61 ttcagtatta ggttgattta tattccaaag aataatatct ttgttaaatt gttgattttt
121 gtaaacctct aaatgtttgt tgctaaaatt actgtgttta agaaaaagat taataaataa
181 taataatttc ataattaaaa acttcttttca ttgaatgcca ttaaataatt cattattttt
241 caaaataaga tcaacataat tgagtaaata ataataagaa caatattata gtacaacaaa
301 atatgggtat gtcataccct tttttttttt tttttttttt ttttttcggg tagagggccg
361 aacctcctac gaggtccccg cgcaaaaggg gcgcgcgggg tatgtgagac tcaacgatct
421 gcatgggtgt gtgagcagac cgcgggcccc aggatttttag agcccaccca ctaaacgact
481 cctctgcact cttacacccg acgtccgac ccctccgagg tcagaacccg gatgaggtag
541 gggggctacc gcggtcaaca ctacaaccag acggcgcggc tcaccccaag gacgcccagc
601 cgacggagcc ttcgaggcga atcgaaggct ctgaaacgtc ggccgtctcg gtacggcagc
661 ccgtcgggcc gccagacgg tgccgctggt gtcccggaat accccgctgg accagaacca
721 gcctgccggg tcgggacgcg atacaccgtc gaccggtcgc tcaatcact ccacggcagc
781 gcgctagagt gctggta (SEQ ID NO: 14516).

```

[0163] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of CCCGGCGAGCATGAGG (SEQ ID NO: 14510). In certain embodiments, the

piggyBac or piggyBac-like transposon comprises an ITR sequence of SEQ ID NO: 14510. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTATCCCGGCGAGCATGAGG (SEQ ID NO: 14511). In certain embodiments, the piggyBac or piggyBac-like transposon comprises at least 16 contiguous nucleotides from SEQ ID NO: 14511. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of CCTCATGCTCGCCGGGTTAT (SEQ ID NO: 14512). In certain embodiments, the piggyBac or piggyBac-like transposon comprises at least 16 contiguous nucleotides from SEQ ID NO: 14512. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end comprising at least 16 contiguous nucleotides from SEQ ID NO: 14511 and one end comprising at least 16 contiguous nucleotides from SEQ ID NO: 14512. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14511 and SEQ ID NO: 14512. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTAACCCGGCGAGCATGAGG (SEQ ID NO: 14513). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of CCTCATGCTCGCCGGGTTAA (SEQ ID NO: 14514).

[0164] In certain embodiments, the piggyBac or piggyBac-like transposon may have ends comprising SEQ ID NO: 14506 and SEQ ID NO: 14507, or a variant of either or both of these having at least 90% sequence identity to SEQ ID NO: 14506 or SEQ ID NO: 14507, and the piggyBac or piggyBac-like transposase has the sequence of SEQ ID NO: 14504 or SEQ ID NO: 14505, or a sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identity to SEQ ID NO: 14504 or SEQ ID NO: 14505. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a heterologous polynucleotide inserted between a pair of inverted repeats, where the transposon is capable of transposition by a piggyBac or piggyBac-like transposase having at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identity to SEQ ID NO: 14504 or SEQ ID NO: 14505. In certain embodiments, the transposon comprises two transposon ends, each of which comprises SEQ ID NO: 14510 in inverted orientations in the two transposon ends. In certain embodiments, each inverted terminal repeat (ITR) is at least 90% identical to SEQ ID NO: 14510.

[0165] In certain embodiments, the piggyBac or piggyBac-like transposon is capable of insertion by a piggyBac or piggyBac-like transposase at the sequence 5'-TTAT-3' within a target nucleic acid. In certain embodiments, one end of the piggyBac or piggyBac-like

transposon comprises at least 16 contiguous nucleotides from SEQ ID NO: 14506 and the other transposon end comprises at least 16 contiguous nucleotides from SEQ ID NO: 14507. In certain embodiments, one end of the piggyBac or piggyBac-like transposon comprises at least 17, at least 18, at least 19, at least 20, at least 22, at least 25, at least 30 contiguous nucleotides from SEQ ID NO: 14506 and the other transposon end comprises at least 17, at least 18, at least 19, at least 20, at least 22, at least 25, at least 30 contiguous nucleotides from SEQ ID NO: 14507.

[0166] In certain embodiments, the piggyBac or piggyBac-like transposon comprises transposon ends (each end comprising an ITR) corresponding to SEQ ID NO: 14506 and SEQ ID NO: 14507, and has a target sequence corresponding to 5'-TTAT3'. In certain embodiments, the piggyBac or piggyBac-like transposon also comprises a sequence encoding a transposase (e.g. SEQ ID NO: 14505). In certain embodiments, the piggyBac or piggyBac-like transposon comprises one transposon end corresponding to SEQ ID NO: 14506 and a second transposon end corresponding to SEQ ID NO: 14516. SEQ ID NO: 14516 is very similar to SEQ ID NO: 14507, but has a large insertion shortly before the ITR. Although the ITR sequences for the two transposon ends are identical (they are both identical to SEQ ID NO: 14510), they have different target sequences: the second transposon has a target sequence corresponding to 5'-TTAA-3', providing evidence that no change in ITR sequence is necessary to modify the target sequence specificity. The piggyBac or piggyBac-like transposase (SEQ ID NO: 14504), which is associated with the 5'-TTAA-3' target site differs from the 5'-TTAT-3'-associated transposase (SEQ ID NO: 14505) by only 4 amino acid changes (D322Y, S473C, A507T, H582R). In certain embodiments, the piggyBac or piggyBac-like transposase (SEQ ID NO: 14504), which is associated with the 5'-TTAA-3' target site is less active than the 5'-TTAT-3'-associated piggyBac or piggyBac-like transposase (SEQ ID NO: 14505) on the transposon with 5'-TTAT-3' ends. In certain embodiments, piggyBac or piggyBac-like transposons with 5'-TTAA-3' target sites can be converted to piggyBac or piggyBac-like transposases with 5'-TTAT-3 target sites by replacing 5'-TTAA-3' target sites with 5'-TTAT-3'. Such transposons can be used either with a piggyBac or piggyBac-like transposase such as SEQ ID NO: 14504 which recognizes the 5'-TTAT-3' target sequence, or with a variant of a transposase originally associated with the 5'-TTAA-3' transposon. In certain embodiments, the high similarity between the 5'-TTAA-3' and 5'-TTAT-3' piggyBac or piggyBac-like transposases demonstrates that very few changes to the amino acid sequence of a piggyBac or piggyBac-like transposase alter target sequence

specificity. In certain embodiments, modification of any piggyBac or piggyBac-like transposon-transposase gene transfer system, in which 5'-TTAA-3' target sequences are replaced with 5'-TTAT-3'-target sequences, the ITRs remain the same, and the transposase is the original piggyBac or piggyBac-like transposase or a variant thereof resulting from using a low-level mutagenesis to introduce mutations into the transposase. In certain embodiments, piggyBac or piggyBac-like transposon transposase transfer systems can be formed by the modification of a 5'-TTAT-3'-active piggyBac or piggyBac-like transposon-transposase gene transfer systems in which 5'-TTAT-3' target sequences are replaced with 5'-TTAA-3'-target sequences, the ITRs remain the same, and the piggyBac or piggyBac-like transposase is the original transposase or a variant thereof.

[0167] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Bombyx mori*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 cccggcgagc atgaggcagg gtatctcata ccctggtaaa attttaaagt tgtgtatttt
61 ataaaatttt cgtctgacaa cactagcgcg ctcagtagct ggaggcagga gcgtgcggga
121 ggggatagtg gcgtgatcgc agtgtggcac gggacaccgg cgagatattc gtgtgcaaac
181 ctgttttcggg tatgttatac cctgcctcat tgttgacgta t (SEQ ID NO: 14577).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 tttaagaaaa agattaataa ataataataa tttcataatt aaaaacttct ttcattgaat
61 gccattaaat aaaccattat tttacaaaat aagatcaaca taattgagta aataataata
121 agaacaatat tatagtacaa caaaatatgg gtatgtcata ccctgccaca ttcttgatgt
181 aacttttttt cacctcatgc tcgccggg (SEQ ID NO: 14578).

```

In certain embodiments, the transposon comprises at least 16 contiguous bases from SEQ ID NO: 14577 and at least 16 contiguous bases from SEQ ID NO: 14578, and inverted terminal repeats that are at least 87% identical to CCCGGCGAGCATGAGG (SEQ ID NO: 14510).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 cccggcgagc atgaggcagg gtatctcata ccctggtaaa attttaaagt tgtgtatttt
61 ataaaatttt cgtctgacaa cactagcgcg ctcagtagct ggaggcagga gcgtgcggga
121 ggggatagtg gcgtgatcgc agtgtggcac gggacaccgg cgagatattc gtgtgcaaac
181 ctgttttcggg tatgttatac cctgcctcat tgttgacgta ttttttttat gtaatttttc
241 cgattattaa tttcaactgt tttattggtt tttttatggt atccattggt ctttttttat
301 gatttactgt atcggttggt tttcgttcc ttagttgagt ttttttttat tattttcagt
361 ttttgatcaa a (SEQ ID NO: 14595).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 tcatattttt agtttaaaaa aataattata tgttttataa tgaaaagaat ctcattatct
61 ttcagtatta gggtgattta tattccaaag aataatattt ttgttaaatt gttgattttt
121 gtaaaccctc aaatgtttgt tgctaaaatt actgtgttta agaaaaagat taataaataa
181 taataatttc ataattaaaa acttctttca ttgaatgcca ttaaataaac cattatttta
241 caaaataaga tcaacataat tgagtaaata ataataagaa caatattata gtacaacaaa
301 atatgggtat gtcataccct gccacattct tgatgtaact ttttttcacc tcatgctcgc
361 cggg (SEQ ID NO: 14596).

```

[0168] In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14595 and SEQ ID NO: 14596, and is transposed by the piggyBac or piggyBac-like transposase of SEQ ID NO: 14505. In certain embodiments, the ITRs of SEQ ID NO: 14595 and SEQ ID: 14596 are not flanked by a 5'-TTAA-3' sequence. In certain embodiments, the ITRs of SEQ ID NO: 14595 and SEQ ID: 14596 are flanked by a 5'-TTAT-3' sequence.

[0169] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 cccggcgagc atgaggcagg gtatctcata cccctggtaaa atttttaaagt tgtgtatttt
61 ataaaatttt cgtctgacaa cactagcgcg ctcagtagct ggaggcagga gcgtgcggga
121 ggggatagtg gcgtgatcgc agtgtggcac gggacacggg cgagatattc gtgtgcaaac
181 ctgttttcggg tatgttatac cctgcctcat tgttgacgta ttttttttat gtaatttttc
241 cgattattaa tttcaactgc tttattggta tttttatggt atccattggt ctttttttat
301 g (SEQ ID NO: 14597).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 cagggtatct cataccctgg taaaatttta agtttggtga ttttataaaa ttttcgtctg
61 acaacactag cgcgctcagt agctggaggc aggagcgtgc gggaggggat agtggcgtga
121 tcgcagtgtg gcacgggaca ccggcgagat attcgtgtgc aaacctgttt cgggtatggt
181 ataccctgcc tcattgtttga cgtatttttt ttatgtaatt ttccgatta ttaatttcaa
241 ctgtttttatt ggtattttta tggtatccat tggtcttttt ttatg (SEQ ID NO:
14598).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 cagggtatct cataccctgg taaaatttta agtttggtga ttttataaaa ttttcgtctg
61 acaacactag cgcgctcagt agctggaggc aggagcgtgc gggaggggat agtggcgtga
121 tcgcagtgtg gcacgggaca ccggcgagat attcgtgtgc aaacctgttt cgggtatggt
181 ataccctgcc tcattgtttga cgtat (SEQ ID NO: 14599).

```

In certain embodiments, the left end of the piggyBac or piggyBac-like transposon comprises a sequence of SEQ ID NO: 14577, SEQ ID NO: 14595, or SEQ ID NOs: 14597-14599. In certain embodiments, the left end of the piggyBac or piggyBac-like transposon is preceded by a left target sequence.

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 tcatattttt agtttaaaaa aataattata tgttttataa tgaaaagaat ctcattatct
61 ttcagtatta ggttgattta tattccaaag aataatattt ttgttaaatt gttgattttt
121 gtaaacctct aaatgtttgt tgctaaaatt actgtgttta agaaaaagat taataaataa
181 taataatttc ataattaaaa acttcttttc ttgaatgcc ttaaataaac cattatttta
241 caaaataaga tcaacataat tgagtaaata ataataagaa caatattata gtacaacaaa
301 atatgggtat gtcataccct gccacattct tgatgtaact ttttttcacc tcatgctcgc
361 cggg (SEQ ID NO: 14600).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 tttaagaaaa agattaataa ataataataa tttcataatt aaaaacttct ttcattgaat
61 gccattaaat aaaccattat tttaaaaaat aagatcaaca taattgagta aataataata
121 agaacaatat tatagtacaa caaaatatgg gtatgtcata cctgccaca ttcttgatgt
181 aacttttttt ca (SEQ ID NO: 14601).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  cccggcgagc atgaggcagg gtatctcata ccttggtaaa attttaaagt tgtgtatttt
61  ataaaatttt cgtctgacaa cactagcgcg cttagtagct ggaggcagga gcgtgcggga
121 ggggatagtg gcgtgatcgc agtgtggcac gggacacggg cgagatattc gtgtgcaaac
181 ctgtttcggg tatgttatac cctgcctcat tggtgacgta ttttttttat gtaatttttc
241 cgattattaa tttcaactgt tttattggtt tttttatggt atccattggt ctttttttat
301 gatttactgt atcggttggt tttcgttctt ttagttgagt ttttttttat tattttcagt
361 ttttgatcaa a (SEQ ID NO: 14602).

```

[0170] In certain embodiments, the right end of the piggyBac or piggyBac-like transposon comprises a sequence of SEQ ID NO: 14578, SEQ ID NO: 14596, or SEQ ID NOs: 14600-14601. In certain embodiments, the right end of the piggyBac or piggyBac-like transposon is followed by a right target sequence. In certain embodiments, the transposon is transposed by the transposase of SEQ ID NO: 14505. In certain embodiments, the left and right ends of the piggyBac or piggyBac-like transposon share a 16 bp repeat sequence of SEQ ID NO: 14510 in inverted orientation and immediately adjacent to the target sequence. In certain embodiments, the left transposon end begins with SEQ ID NO: 14510, and the right transposon end ends with the reverse complement of SEQ ID NO: 14510, 5'-CCTCATGCTCGCCGGG-3' (SEQ ID NO: 14603). In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR with at least 93%, at least 87%, or at least 81% or any percentage in between identity to SEQ ID NO: 14510 or SEQ ID NO: 14603. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a target sequence followed by a left transposon end comprising a sequence selected from SEQ ID NOs: 88, 105 or 107 and a right transposon end comprising SEQ ID NO: 14578 or 106 followed by a target sequence. In certain embodiments, the piggyBac or piggyBac like transposon comprises one end that comprises a sequence that is at least 90%, at least 95% or at least 99% or any percentage in between identical to SEQ ID NO: 14577 and one end that comprises a sequence that is at least 90%, at least 95% or at least 99% or any percentage in between identical to SEQ ID NO: 14578. In certain embodiments, one transposon end comprises at least 14, at least 16, at least 18 or at least 20 contiguous bases from SEQ ID NO: 14577 and one transposon end comprises at least 14, at least 16, at least 18 or at least 20 contiguous bases from SEQ ID NO: 14578.

[0171] In certain embodiments, the piggyBac or piggyBac-like transposon comprises two transposon ends wherein each transposon ends comprises a sequence that is at least 81% identical, at least 87% identical or at least 93% identical or any percentage in between identical to SEQ ID NO: 14510 in inverted orientation in the two transposon ends. One end may further comprise at least 14, at least 16, at least 18 or at least 20 contiguous bases from

SEQ ID NO: 14599, and the other end may further comprise at least 14, at least 16, at least 18 or at least 20 contiguous bases from SEQ ID NO: 14601. The piggyBac or piggyBac-like transposon may be transposed by the transposase of SEQ ID NO: 14505, and the transposase may optionally be fused to a nuclear localization signal.

[0172] In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14595 and SEQ ID NO: 14596 and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14504 or SEQ ID NO: 14505. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14597 and SEQ ID NO: 14596 and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14504 or SEQ ID NO: 14505. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14595 and SEQ ID NO: 14578 and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14504 or SEQ ID NO: 14505. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14602 and SEQ ID NO: 14600 and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14504 or SEQ ID NO: 14505.

[0173] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a left end comprising 1, 2, 3, 4, 5, 6, or 7 sequences selected from ATGAGGCAGGGTAT (SEQ ID NO: 14614), ATACCCTGCCTCAT (SEQ ID NO: 14615), GGCAGGGTAT (SEQ ID NO: 14616), ATACCCTGCC (SEQ ID NO: 14617), TAAAATTTTA (SEQ ID NO: 14618), ATTTTATAAAAT (SEQ ID NO: 14619), TCATACCCTG (SEQ ID NO: 14620) and TAAATAATAATAA (SEQ ID NO: 14621). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a right end comprising 1, 2 or 3 sequences selected from SEQ ID NO: 14617, SEQ ID NO: 14620 and SEQ ID NO: 14621.

[0174] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Xenopus tropicalis*. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MAKRFYSAEE  AAHCMASSS  EEFSGSDSEY  VPPASESDSS  TEESWCSSST  VSALEEPMEV
61  DEDVDDLEDQ  EAGDRADAAA  GGEPAWGPCC  NFPPEIPPFT  TVPGVKVDTS  NFEPINFFQL
121 FMTEAILQDM  VLYTNVYAEQ  YLTQNPLPRY  ARAHAWHPTD  IAEMKRFEVGL  TLAMGLIKAN
181 SLESYWDTTT  VLSIPVFSAT  MSRNRYQLLL  RFLHFNNNAT  AVPPDQPGHD  RLHKLRLPID
241 SLSERFAAVY  TPCQNICIDE  SLLLFKGRLO  FRQYIPSKRA  RYGIKFYKLC  ESSSGYTSYF
301 LIYEGKDSKL  DPPGCPPDLT  VSGKIVWELI  SPLLGQGFHL  YVDNFYSSIP  LFTALYCLDT

```

361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPPTP
 541 GKQRPQKGCK VCRKRGIRRD TRYPCPKCPR NPGLCFKPCF EIYHTQLHY (SEQ ID NO:
 14517).

[0175] In some embodiments, the piggyBac or piggyBac-like transposase is a hyperactive variant of SEQ ID NO: 14517. In certain embodiments, the piggyBac or piggyBac-like transposase is an integration defective variant of SEQ ID NO: 14517. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

1 MAKRFYSAEE AAAHCMAPSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
 61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
 121 FMTEAILQDM VLYTNVYAEQ YLTQNPLPRY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
 181 SLESYWNTTT VLSIPVFSAT MSRNRYQLL RFLHFNNNAT AVPPDQPDHD RLHKLRLPID
 241 SLSERFAAVY TPCQNICIDE SLLLFKGRLR FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
 301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLDT
 361 PACGTINRTR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT SAWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMLP SDNVARLIGK HFIDTLPPTP
 541 GKQRPQKGCK VCRKRGIRRD TRYPCPKCPR NPGLCFKPCF EIYHTQLHY (SEQ ID NO:
 14518).

[0176] In certain embodiments, the piggyBac or piggyBac-like transposase is isolated or derived from *Xenopus tropicalis*. In certain embodiments, the piggyBac or piggyBac-like transposase is a hyperactive piggyBac or piggyBac-like transposase. In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence at least 90% identical to:

1 MAKRFYSAEE AAAHCSASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
 61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
 121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRG ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
 181 SIESYWDTTT VLSIPVFGAT MSRNRYQLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPID
 241 SLSERFANVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSTGYTSYF
 301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLNT
 361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPD SDNVARLIGK HFIDTLPPTP
 541 GKQRPQKGCK VCRKRGIRRD TRYPCPKCPR NPGLCRKPCF EIYHTQLHY (SEQ ID NO:
 14572).

[0177] In certain embodiments, piggyBac or piggyBac-like transposase is a hyperactive piggyBac or piggyBac-like transposase. A hyperactive piggyBac or piggyBac-like transposase is a transposase that is more active than the naturally occurring variant from which it is derived. In certain embodiments, a hyperactive piggyBac or piggyBac-like transposase is more active than the transposase of SEQ ID NO: 14517. In certain

embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MAKRFYSAEE AAAHCSASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRG ARAHAWHPTD IAEMKRFBVGL TLAMGLIKAN
181 SIESYWDTTT VLSIPVFGAT MSRNRYQLLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPLD
241 SLSERFANVY TPCQNICIDE SLMLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLNT
361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLEGV EEQTVPEMPD SDNVARLIGK HFIDTLPTPT
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCRKPCF EIYHTQLHY (SEQ ID NO:
14572).

```

[0178] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MAKRFYSAEE AAAHCMASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRY ARAHAWHPTD IAEMKRFBVGL TLAMGLIKAN
181 SLESYWDTTT VLSIPVFSAT MSRNRYQLLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPLD
241 SLSERFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLNT
361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLEGV EEQTVPEMPD SDNVARLIGK HFIDTLPTPT
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCRKPCF EIYHTQLHY (SEQ ID NO:
14624) .

```

[0179] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MAKRFYSAEE AAAHCMASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRY ARAHAWHPTD IAEMKRFBVGL TLAMGLIKAN
181 SLESYWDTTT VLKIPVFSAT MSRNRYQLLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPLD
241 SLSERFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLNT
361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLEGV EEQTVPEMPD SDNVARLIGK HFIDTLPTPT
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCFKPCF EIYHTQLHY (SEQ ID NO:
14625) .

```

[0180] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MAKRFYSAEE AAAHCMASSS EQTSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRY ARAHAWHPTD IAEMKRFBVGL TLAMGLIKAN
181 SIESYWDTTT VLSIPVFGAT MSRNRYQLLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPLD
241 SLSERFANVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFYSSIP LFTALYCLNT
361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLEGV EEQTVPEMPD SDNVARLIGK HFIDTLPTPT
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCRKPCF EIYHTQLHY (SEQ ID NO:
14627) .

```

[0181] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

```

1 MAKRFYSAEE AAHCSASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAWGPPC NEFPEIPPFT TVPGVKVDTN NEFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPLTRG ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
181 SLESYWDTTT VLSIPVFGAT MSRNRYQLLL RFLHFNNNAT AVPPDQPGHD RLHKLRLPID
241 SLSERFANVY TPCQNICIDE SLMLFKGRLQ FRQYIPSKRA RYGIKFYKLC ESSTGYTSYF
301 LIYEGKDSKL DPPGCPDPLT VSGKIVWELI SPLLGQGFHL YVDNFIYSSIP LETALYCLNT
361 PACGTINRNR KGLPRALLDK KLNRGETYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RHWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTPT
541 GKQRPQKGCK VCRKRGIRRD TRYCPKPCR NGLCRKPCF EIYHTQLHY (SEQ ID NO:
14628) .

```

[0182] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises a sequence of:

(SEQ ID NO: 17042) .

[0183] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises an amino acid substitution at a position selected from amino acid 6, 7, 16, 19, 20, 21, 22, 23, 24, 26, 28, 31, 34, 67, 73, 76, 77, 88, 91, 141, 145, 146, 148, 150, 157, 162, 179, 182, 189, 192, 193, 196, 198, 200, 210, 212, 218, 248, 263, 270, 294, 297, 308, 310, 333, 336, 354, 357, 358, 359, 377, 423, 426, 428, 438, 447, 450, 462, 469, 472, 498, 502, 517, 520, 523, 533, 534, 576, 577, 582, 583 or 587 (relative to SEQ ID NO: 14517). In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises an amino acid substitution of Y6C, S7G, M16S, S19G, S20Q, S20G, S20D, E21D, E22Q, F23T, F23P, S24Y, S26V, S28Q, V31K, A34E, L67A, G73H, A76V, D77N, P88A, N91D, Y141Q, Y141A, N145E, N145V, P146T, P146V, P146K, P148T, P148H, Y150G, Y150S, Y150C, H157Y, A162C, A179K, L182I, L182V, T189G, L192H, S193N, S193K, V196I, S198G, T200W, L210H, F212N, N218E, A248N, L263M, Q270L, S294T, T297M, S308R, L310R, L333M, Q336M, A354H, C357V, L358F, D359N, L377I, V 423H, P426K, K428R, S438A, T447G, T447A, L450V, A462H, A462Q, I469V, I472L, Q498M, L502V, E517I, P520D, P520G, N523S, I533E, D534A, F576R, F576E, K577I, I582R, Y583F, L587Y or L587W, or any combination thereof including at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or all of these mutations (relative to SEQ ID NO: 14517).

[0184] In certain embodiments, the hyperactive piggyBac or piggyBac-like transposase comprises one or more substitutions of an amino acid that is not wild type, wherein the one or more substitutions a for wild type amino acid comprises a substitution of A2X, K3X, R4X, F5X, Y6X, S7X, A11X, A13X, C15X, M16X, A17X, S18X, S19X, S20X, E21X, E22X, F23X, S24X, G25X, 26X, D27X, S28X, E29X, E42X, E43X, S44X, C46X, S47X, S48X,

S49X, T50X, V51X, S52X, A53X, L54X, E55X, E56X, P57X, M58X, E59X, E62X, D63X, V64X, D65X, D66X, L67X, E68X, D69X, Q70X, E71X, A72X, G73X, D74X, R75X, A76X, D77X, A78X, A79X, A80X, G81X, G82X, E83X, P84X, A85X, W86X, G87X, P88X, P89X, C90X, N91X, F92X, P93X, E95X, I96X, P97X, P98X, F99X, T100X, T101X, P103X, G104X, V105X, K106X, V107X, D108X, T109X, N111X, P114X, I115X, N116X, F117X, F118X, Q119X, M122X, T123X, E124X, A125X, I126X, L127X, Q128X, D129X, M130X, L132X, Y133X, V126X, Y127X, A138X, E139X, Q140X, Y141X, L142X, Q144X, N145X, P146X, L147X, P148X, Y150X, A151X, A155X, H157X, P158X, I161X, A162X, V168X, T171X, L172X, A173X, M174X, I177X, A179X, L182X, D187X, T188X, T189X, T190X, L192X, S193X, I194X, P195X, V196X, S198X, A199X, T200X, S202X, L208X, L209X, L210X, R211X, F212X, F215X, N217X, N218X, A219X, T220X, A221X, V222X, P224X, D225X, Q226X, P227X, H229X, R231X, H233X, L235X, P237X, I239X, D240X, L242X, S243X, E244X, R244X, F246X, A247X, A248X, V249X, Y250X, T251X, P252X, C253X, Q254X, I256X, C257X, I258X, D259X, E260X, S261X, L262X, L263X, L264X, F265X, K266X, G267X, R268X, L269X, Q270X, F271X, R272X, Q273X, Y274X, I275X, P276X, S277X, K278X, R279X, A280X, R281X, Y282X, G283X, I284X, K285X, F286X, Y287X, K288X, L289X, C290X, E291X, S292X, S293X, S294X, G295X, Y296X, T297X, S298X, Y299X, F300X, E304X, L310X, P313X, G314X, P316X, P317X, D318X, L319X, T320X, V321X, K324X, E328X, I330X, S331X, P332X, L333X, L334X, G335X, Q336X, F338X, L340X, D343X, N344X, F345X, Y346X, S347X, L351X, F352X, A354X, L355X, Y356X, C357X, L358X, D359X, T360X, R422X, Y423X, G424X, P426X, K428X, N429X, K430X, P431X, L432X, S434X, K435X, E436X, S438X, K439X, Y440X, G443X, R446X, T447X, L450X, Q451X, N455X, T460X, R461X, A462X, K465X, V467X, G468X, I469X, Y470X, L471X, I472X, M474X, A475X, L476X, R477X, S479X, Y480X, V482X, Y483X, K484X, A485X, A486X, V487X, P488X, P490X, K491X, S493X, Y494X, Y495X, K496X, Y497T, Q498X, L499X, Q500X, I501X, L502X, P503X, A504X, L505X, L506X, F507X, G508X, G509X, V510X, E511X, E512X, Q513X, T514X, V515X, E517X, M518X, P519X, P520X, S521X, D522X, N523X, V524X, A525X, L527X, I528X, K530X, H531X, F532X, I533X, D534X, T535X, L536X, T539X, P540X, Q546X, K550X, R553X, K554X, R555X, G556X, I557X, R558X, R559X, D560X, T561X, Y564X, P566X, K567X, P569X, R570X, N571X, L574X, C575X, F576X, K577X, P578X, F580X, E581X, I582X, Y583X, T585X, Q586X, L587X, H588X or Y589X (relative to SEQ ID NO: 14517). A list of hyperactive

amino acid substitutions can be found in US patent No. 10,041,077, the contents of which are incorporated by reference in their entirety.

[0185] In certain embodiments, the piggyBac or piggyBac-like transposase is integration deficient. In certain embodiments, an integration deficient piggyBac or piggyBac-like transposase is a transposase that can excise its corresponding transposon, but that integrates the excised transposon at a lower frequency than a corresponding naturally occurring transposase. In certain embodiments, the piggyBac or piggyBac-like transposase is an integration deficient variant of SEQ ID NO: 14517. In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase is deficient relative to SEQ ID NO: 14517.

[0186] In certain embodiments, the piggyBac or piggyBac-like transposase is active for excision but deficient in integration. In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to a sequence of:

```

1 MAKRFYSAEE AAHCMASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRVDAVA GGEPAGWPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNPPLRY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
181 SLESYWDTTT VLSIPVFSAT MSRNRYQLLL KFLHFNNEAT AVPPDQPGHD RLHKLRLPID
241 SLSEFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIWELI SPLLGQGFHL YVDNFISSIP LFTALYCLDT
361 PACGTINRNR KGLPRALLDK KLNRGETYAL RKNELLAIKF FDKKNVFMLT SIHDESIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTTP
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCFKPCF EIYHTQLHYG RR (SEQ ID
NO: 14605).
```

[0187] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to a sequence of:

```

1 MAKRFYSAEE AAHCMASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAGWPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQVPLPRY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
181 SLESYWDTTT VLNIPVFSAT MSRNRYQLLL RFLHFNNEAT AVPPDQPGHD RLHKLRLPID
241 SLSEFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
301 LIYEGKDSKL DPPGCPPDLT VSGKIWELI SPLLGQGFHL YVDNFISSIP LFTALYCLDT
361 PACGTINRNR KGLPRALLDK KLNRGETYAL RKNELLAIKF FDKKNVFMLT SIHDESIRE
421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTTP
541 GKQRPQKGCK VCRKRGIRRD TRYCPKCPR NPGLCFKPCF EIYHTQLHY (SEQ ID NO:
14604).
```

[0188] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to a sequence of:

```

1 MAKRFYSAEE AAHCMASSS EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
61 DEDVDDLEDQ EAGDRADAAA GGEPAGWPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
121 FMTEAILQDM VLYTNVYAEQ YLTQNVLPY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
181 SLESYWDTTT VLSIPVFSAT MSRNRYQLLL RFLHFNNDAT AVPPDQPGHD RLHKLRLPID
241 SLTERFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
```

301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFIYSSIP LFTALYCLDT
 361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTTP
 541 GKQRPQKGCK VCRKRGIRRD TRYICPKCPR NPGLCFKPCF EIYHTQLHYG RR (SEQ ID
 NO: 14611).

[0189] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14611. In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to a sequence of:

1 MAKRFYSAEE AAHCMASST EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
 61 DEDVDDLEDQ EAGDRADAA? GGEPAGWPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
 121 FMTEAILQDM VLYTNVYAEQ YLTQVPLPRY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
 181 SLESYWDTTT VLSIPVFSAT MSRNRYQLLL RFLHENNEAT AVPPDQPGHD RLHKLRLPLD
 241 SLSEFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
 301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFIYSSIP LFTALYCLDT
 361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTTP
 541 GKQRPQKGCK VCRKRGIRRD TRYICPKCPR NPGLCFKPCF EIYHTQLHYG RR (SEQ ID
 NO: 14612).

[0190] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14612. In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises a sequence that is at least 90% identical to a sequence of:

1 MAKRFYSAEE AAHCMASST EEFSGSDSEY VPPASESDSS TEESWCSSST VSALEEPMEV
 61 DEDVDDLEDQ EAGDRADAAA GGEPAGWPPC NFPPEIPPFT TVPGVKVDTS NFEPINFFQL
 121 FMTEAILQDM VLYTNVYAEQ YLTQVPLPRY ARAHAWHPTD IAEMKRFVGL TLAMGLIKAN
 181 SLESYWDTTT VLNIPVFSAT MSRNRYQLLL RFLFNNNAT AVPPDQPGHD RLHKLRLPLD
 241 SLSEFAAVY TPCQNICIDE SLLLFKGRLO FRQYIPSKRA RYGIKFYKLC ESSSGYTSYF
 301 LIYEGKDSKL DPPGCPPDLT VSGKIVWELI SPLLGQGFHL YVDNFIYSSIP LFTALYCLDT
 361 PACGTINRNR KGLPRALLDK KLNREGTYAL RKNELLAIKF FDKKNVFMLT SIHDESVIRE
 421 QRVGRPPKNK PLCSKEYSKY MGGVDRTDQL QHYYNATRKT RAWYKKVGIY LIQMALRNSY
 481 IVYKAAVPGP KLSYYKYQLQ ILPALLFGGV EEQTVPEMPP SDNVARLIGK HFIDTLPTTP
 541 GKQRPQKGCK VCRKRGIRRD TRYICPKCPR NPGLCFKPCF EIYHTQLHYG RR (SEQ ID
 NO: 14613).

[0191] In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14613. In certain embodiments, the integration deficient piggyBac or piggyBac-like transposase comprises an amino acid substitution wherein the Asn at position 218 is replaced by a Glu or an Asp (N218D or N218E) (relative to SEQ ID NO: 14517).

[0192] In certain embodiments, the excision competent, integration deficient piggyBac or piggyBac-like transposase comprises one or more substitutions of an amino acid that is not wild type, wherein the one or more substitutions a for wild type amino acid comprises a

substitution of A2X, K3X, R4X, F5X, Y6X, S7X, A8X, E9X, E10X, A11X, A12X, A13X, H14X, C15X, M16X, A17X, S18X, S19X, S20X, E21X, E22X, F23X, S24X, G25X, 26X, D27X, S28X, E29X, V31X, P32X, P33X, A34X, S35X, E36X, S37X, D38X, S39X, S40X, T41X, E42X, E43X, S44X, W45X, C46X, S47X, S48X, S49X, T50X, V51X, S52X, A53X, L54X, E55X, E56X, P57X, M58X, E59X, V60X, M122X, T123X, E124X, A125X, L127X, Q128X, D129X, L132X, Y133X, V126X, Y127X, E139X, Q140X, Y141X, L142X, T143X, Q144X, N145X, P146X, L147X, P148X, R149X, Y150X, A151X, H154X, H157X, P158X, T159X, D160X, I161X, A162X, E163X, M164X, K165X, R166X, F167X, V168X, G169X, L170X, T171X, L172X, A173X, M174X, G175X, L176X, I177X, K178X, A179X, N180X, S181X, L182X, S184X, Y185X, D187X, T188X, T189X, T190X, V191X, L192X, S193X, I194X, P195X, V196X, F197X, S198X, A199X, T200X, M201X, S202X, R203X, N204X, R205X, Y206X, Q207X, L208X, L209X, L210X, R211X, F212X, L213X, H241X, F215X, N216X, N217X, N218X, A219X, T220X, A221X, V222X, P223X, P224X, D225X, Q226X, P227X, G228X, H229X, D230X, R231X, H233X, K234X, L235X, R236X, L238X, I239X, D240X, L242X, S243X, E244X, R244X, F246X, A247X, A248X, V249X, Y250X, T251X, P252X, C253X, Q254X, N255X, I256X, C257X, I258X, D259X, E260X, S261X, L262X, L263X, L264X, F265X, K266X, G267X, R268X, L269X, Q270X, F271X, R272X, Q273X, Y274X, I275X, P276X, S277X, K278X, R279X, A280X, R281X, Y282X, G283X, I284X, K285X, F286X, Y287X, K288X, L289X, C290X, E291X, S292X, S293X, S294X, G295X, Y296X, T297X, S298X, Y299X, F300X, I302X, E304X, G305X, K306X, D307X, S308X, K309X, L310X, D311X, P312X, P313X, G314X, C315X, P316X, P317X, D318X, L319X, T320X, V321X, S322X, G323X, K324X, I325X, V326X, W327X, E328X, L329X, I330X, S331X, P332X, L333X, L334X, G335X, Q336X, F338X, H339X, L340X, V342X, N344X, F345X, Y346X, S347X, S348X, I349X, L351X, T353X, A354X, Y356X, C357X, L358X, D359X, T360X, P361X, A362X, C363X, G364X, I366X, N367X, R368X, D369X, K371X, G372X, L373X, R375X, A376X, L377X, L378X, D379X, K380X, K381X, L382X, N383X, R384X, G385X, T387X, Y388X, A389X, L390X, K392X, N393X, E394X, A397X, K399X, F400X, F401X, D402X, N405X, L406X, L409X, R422X, Y423X, G424X, E425X, P426X, K428X, N429X, K430X, P431X, L432X, S434X, K435X, E436X, S438X, K439X, Y440X, G442X, G443X, V444X, R446X, T447X, L450X, Q451X, H452X, N455X, T457X, R458X, T460X, R461X, A462X, Y464X, K465X, V467X, G468X, I469X, L471X, I472X, Q473X, M474X, L476X, R477X, N478X, S479X, Y480X, V482X, Y483X, K484X, A485X, A486X, V487X, P488X, G489X, P490X, K491X, L492X, S493X, Y494X, Y495X, K496X, Q498X,

L499X, Q500X, I501X, L502X, P503X, A504X, L505X, L506X, F507X, G508X, G509X, V510X, E511X, E512X, Q513X, T514X, V515X, E517X, M518X, P519X, P520X, S521X, D522X, N523X, V524X, A525X, L527X, I528X, G529X, K530X, F532X, I533X, D534X, T535X, L536X, P537X, P538X, T539X, P540X, G541X, F542X, Q543X, R544X, P545X, Q546X, K547X, G548X, C549X, K550X, V551X, C552X, R553X, K554X, R555X, G556X, I557X, R558X, R559X, D560X, T561X, R562X, Y563X, Y564X, C565X, P566X, K567X, C568X, P569X, R570X, N571X, P572X, G573X, L574X, C575X, F576X, K577X, P578X, C579X, F580X, E581X, I582X, Y583X, H584X, T585X, Q586X, L587X, H588X or Y589X (relative to SEQ ID NO: 14517). A list of excision competent, integration deficient amino acid substitutions can be found in US patent No. 10,041,077, the contents of which are incorporated by reference in their entirety.

[0193] In certain embodiments, the piggyBac or piggyBac-like transposase is fused to a nuclear localization signal. In certain embodiments, SEQ ID NO: 14517 or SEQ ID NO: 14518 is fused to a nuclear localization signal. In certain embodiments, the amino acid sequence of the piggyBac or piggyBac like transposase fused to a nuclear localization signal is encoded by a polynucleotide sequence comprising:

```

1  atggcaccda aaaagaaacg taaagtgatg gccaaaagat ttacagcgc cgaagaagca
61  gcagcacatt gcatggcatc gtcatecgaa gaattctcgg ggagcgattc cgaatatgtc
121 ccaccggcct cggaaagoga ttcgagcact gaggagtctg ggtgttcctc ctcaactgtc
181 tcggctcttg aggagccgat ggaagtggat gaggatgtgg acgacttgga ggaccaggaa
241 gccggagaca gggccgacgc tgccgcggga ggggagccgg cgtggggacc tccatgcaat
301 tttcctcccg aaatcccacc gttcactact gtgccgggag tgaaggtcga cactccaac
361 ttcgaaccga tcaatttctt tcaactcttc atgactgaag cgatcctgca agatatggtg
421 ctctacacta atgtgtacgc cgagcagtac ctgactcaaa acccgctgcc tcgctacgcg
481 agagcgcgat cgtggcaccg gaccgatata gccgagatga agcgggttcg gggactgacc
541 ctcgcaatgg gctgatcaa ggccaacage ctcgagtcac actgggatac cactactgtg
601 cttagcattc cgggtgttct cgctaccatg tcccgtaacc gctaccaact cctgctgcgg
661 ttctctccact tcaacaacaa tgcgaccgct gtgccacctg accagccagg acacgacaga
721 ctccacaage tgccggccatt gatcgactcg ctgagcgagc gattcgccgc ggtgtacacc
781 ccttgccaaa acatttgcac cgacgagtcg cttctgctgt ttaaaggccg gcttcagttc
841 cgccagtaca tcccatcgaa gcgcgctcgc tatggtatca aattctacaa actctgcgag
901 tcgtccagcg gctacacgtc atacttcttg atctacgagg ggaaggactc taagctggac
961 ccaccggggg gtccaccgga tcttactgtc tccggaaaaa tcgtgtggga actcatctca
1021 cctctcctcg gacaaggctt tcatctctac gtcgacaatt tctactcacc gatccctctg
1081 ttcaccgccc tctactgccc ggatactcca gcctgtggga ccattaacag aaaccggaag
1141 ggtctgccga gagcactgct ggataagaag ttgaacaggg gagagactta cgcgctgaga
1201 aagaacgaac tcctcgccat caaattcttc gacaagaaaa atgtgtttat gctcacctcc
1261 atccacgacg aatccgtcat ccgggagcag cgcgtgggca ggccgcccga aaacaagccg
1321 ctgtgctcta aggaatactc caagtacatg ggggggtgtc accggaccga tcagctgcag
1381 cattactaca acgccactag aaagaccogg gcctggtaca agaaagtcgg catctacctg
1441 atccaaatgg cactgaggaa ttctgtatatt gctacaagg ctgccgttcc gggcccgaag
1501 ctgtcatact acaagtacca gcttcaaata ctgccggcgc tgctgttcgg tggagtggaa
1561 gaacagactg tgcccgagat gccgccatcc gacaacgtgg cccggttgat cggaaagcac
1621 ttcattgata cctgcctcc gacgcctgga aagcagcggc cacagaaggg atgcaaagtt
1681 tgccgcaagc gcggaatacg gcgcgatacc cgctactatt gcccgagtg ccccgcaat

```

1741 cccggactgt gtttcaagcc ctgttttgaa atctaccaca ccagttgca ttac (SEQ ID NO: 14626).

[0194] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Xenopus tropicalis*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```
1 ttaacctttt tactgccaat gacgcatggg atacgtcgtg gcagtaaaag ggcttaaagt
61 ccaacgacgc gtcccatagc ttgttggcat ttttaagtctt ctctctgcag cggcagcatg
121 tgccgcccgt gcagagagtt tctagcgatg acagcccctc tgggcaacga gccggggggg
181 ctgtc (SEQ ID NO: 14519).
```

[0195] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```
1 tttgcatttt tagacattta gaagcctata tcttgttaca gaattggaat tacacaaaaa
61 ttctaccata ttttgaaagc ttaggttggt ctgaaaaaaa caatatattg ttttctggg
121 taaactaaaa gtcccctcga ggaaaggccc cttaagtga acagtgcacaa acgttcaaaa
181 actgtctggc aatacaagtt ccactttgac caaacggct ggcagtaaaa gggttaa
(SEQ ID NO: 14520).
```

[0196] In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14519 and SEQ ID NO: 14520. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```
1 ttaaccccttt gctgccaat cagcgtggg atacgtcgtg gcagtaaaag ggcttaaagt
61 ccaacgacgc gtcccatagc ttgttggcat ttttaagtctt ctctctgcag cggcagcatg
121 tgccgcccgt gcagagagtt tctagcgatg acagcccctc tgggcaacga gccggggggg
181 ctgtc (SEQ ID NO: 14521).
```

[0197] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```
1 tttgcatttt tagacattta gaagcctata tcttgttaca gaattggaat tacacaaaaa
61 ttctaccata ttttgaaagc ttaggttggt ctgaaaaaaa caatatattg ttttctggg
121 taaactaaaa gtcccctcga ggaaaggccc cttaagtga acagtgcacaa acgttcaaaa
181 actgtctggc aatacaagtt ccactttggg acaaatcggc tggcagtga agggttaa
(SEQ ID NO: 14522).
```

[0198] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```
1 ttaacctttt tactgccaat gacgcatggg atacgtcgtg gcagtaaaag ggcttaaagt
61 ccaacgacgc gtcccatagc ttgttggcat ttttaattctt ctctctgcag cggcagcatg
121 tgccgcccgt gcagagagtt tctagcgatg acagcccctc tgggcaacga gccggggggg
181 ctgtc (SEQ ID NO: 14523).
```

[0199] In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14520 and SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14522 and SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523. In certain

embodiments, the piggyBac or piggyBac-like transposon comprises one end comprising at least 14, 16, 18, 20, 30 or 40 contiguous nucleotides from SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end comprising at least 14, 16, 18, 20, 30 or 40 contiguous nucleotides from SEQ ID NO: 14520 or SEQ ID NO: 14522. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end with at least 90% identity to SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523. In certain embodiments, the piggyBac or piggyBac-like transposon comprises one end with at least 90% identity to SEQ ID NO: 14520 or SEQ ID NO: 14522. In one embodiment, one transposon end is at least 90% identical to SEQ ID NO: 14519 and the other transposon end is at least 90% identical to SEQ ID NO: 14520.

[0200] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTAACCTTTTACTGCCA (SEQ ID NO: 14524). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTAACCCTTTGCCTGCCA (SEQ ID NO: 14526). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTAACCYTTTTACTGCCA (SEQ ID NO: 14527). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TGGCAGTAAAAGGGTTAA (SEQ ID NO: 14529). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TGGCAGTGAAAGGGTTAA (SEQ ID NO: 14531). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of TTAACCYTTTKMCTGCCA (SEQ ID NO: 14533). In certain embodiments, one end of the piggyBac or piggyBac-like transposon comprises a sequence selected from SEQ ID NO: 14524, SEQ ID NO: 14526 and SEQ ID NO: 14527. In certain embodiments, one end of the piggyBac™ (PB) or piggyBac-like transposon comprises a sequence selected from SEQ ID NO: 14529 and SEQ ID NO: 14531. In certain embodiments, each inverted terminal repeat of the piggyBac or piggyBac-like transposon comprises a sequence of ITR sequence of CCYTTTKMCTGCCA (SEQ ID NO: 14563). In certain embodiments, each end of the piggyBac™ (PB) or piggyBac-like transposon comprises SEQ ID NO: 14563 in inverted orientations. In certain embodiments, one ITR of the piggyBac or piggyBac-like transposon comprises a sequence selected from SEQ ID NO: 14524, SEQ ID NO: 14526 and SEQ ID NO: 14527. In certain embodiments, one ITR of the piggyBac or piggyBac-like transposon comprises a sequence selected from SEQ ID NO: 14529 and SEQ ID NO: 14531. In certain embodiments, the piggyBac or

piggyBac like transposon comprises SEQ ID NO: 14533 in inverted orientation in the two transposon ends.

[0201] In certain embodiments, The piggyBac or piggyBac-like transposon may have ends comprising SEQ ID NO: 14519 and SEQ ID NO: 14520 or a variant of either or both of these having at least 90% sequence identity to SEQ ID NO: 14519 or SEQ ID NO: 14520, and the piggyBac or piggyBac-like transposase has the sequence of SEQ ID NO: 14517 or a variant showing at least %, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between sequence identity to SEQ ID NO: 14517 or SEQ ID NO: 14518. In certain embodiments, one piggyBac or piggyBac-like transposon end comprises at least 14 contiguous nucleotides from SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523, and the other transposon end comprises at least 14 contiguous nucleotides from SEQ ID NO: 14520 or SEQ ID NO: 14522. In certain embodiments, one transposon end comprises at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 22, at least 25, at least 30 contiguous nucleotides from SEQ ID NO: 14519, SEQ ID NO: 14521 or SEQ ID NO: 14523, and the other transposon end comprises at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 22, at least 25 or at least 30 contiguous nucleotides from SEQ ID NO: 14520 or SEQ ID NO: 14522.

[0202] In certain embodiments, the piggyBac or piggyBac-like transposase recognizes a transposon end with a left sequence corresponding to SEQ ID NO: 14519, and a right sequence corresponding to SEQ ID NO: 14520. It will excise the transposon from one DNA molecule by cutting the DNA at the 5'-TTAA-3' sequence at the left end of one transposon end to the 5'-TTAA-3' at the right end of the second transposon end, including any heterologous DNA that is placed between them, and insert the excised sequence into a second DNA molecule. In certain embodiments, truncated and modified versions of the left and right transposon ends will also function as part of a transposon that can be transposed by the piggyBac or piggyBac-like transposase. For example, the left transposon end can be replaced by a sequence corresponding to SEQ ID NO: 14521 or SEQ ID NO: 14523, the right transposon end can be replaced by a shorter sequence corresponding to SEQ ID NO: 14522. In certain embodiments, the left and right transposon ends share an 18 bp almost perfectly repeated sequence at their ends (5'-TTAACCYTTTKMCTGCCA: SEQ ID NO: 14533) that includes the 5'-TTAA-3' insertion site, which sequence is inverted in the orientation in the two ends. That is in SEQ ID NO: 14519 and SEQ ID NO: 14523 the left transposon end

begins with the sequence 5'-TTAACCTTTTACTGCCA-3' (SEQ ID NO: 14524), or in SEQ ID NO: 14521 the left transposon end begins with the sequence 5'-TTAACCTTTTGCCTGCCA-3' (SEQ ID NO: 14526); the right transposon ends with approximately the reverse complement of this sequence: in SEQ ID NO: 14520 it ends 5'-TGGCAGTAAAAGGGTTAA-3' (SEQ ID NO: 14529), in SEQ ID NO: 14522 it ends 5'-TGGCAGTGAAAGGGTTAA-3' (SEQ ID NO: 14531.) One embodiment of the invention is a transposon that comprises a heterologous polynucleotide inserted between two transposon ends each comprising SEQ ID NO: 14533 in inverted orientations in the two transposon ends. In certain embodiments, one transposon end comprises a sequence selected from SEQ ID NOS: 14524, SEQ ID NO: 14526 and SEQ ID NO: 14527. In some embodiments, one transposon end comprises a sequence selected from SEQ ID NO: 14529 and SEQ ID NO: 14531.

[0203] In certain embodiments, the piggyBac™ (PB) or piggyBac-like transposon is isolated or derived from *Xenopus tropicalis*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```
1 ccctttgcct gccaatcacg catgggatac gtcgtggcag taaaagggt taaatgccaa
61 cgacgcgtcc catacgtt (SEQ ID NO: 14573).
```

[0204] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```
1 cctgggtaaa ctaaaagtcc cctcgaggaa aggccctaa agtgaaacag tgcaaaacgt
61 tcaaaaactg tctggcaata caagttccac tttgggacaa atcggctggc agtgaaagg
```

(SEQ ID NO: 14574).

[0205] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at least 16 contiguous bases from SEQ ID NO: 14573 or SEQ ID NO: 14574, and inverted terminal repeat of CCYTTTBMCTGCCA (SEQ ID NO: 14575).

[0206] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```
1 ccctttgcct gccaatcacg catgggatac gtcgtggcag taaaagggt taaatgccaa
61 cgacgcgtcc catacgttgc tggcatttta agtcttctct ctgcagcggc agcatgtgcc
121 gccgctgcag agagtttcta gcatgacag cccctctggg caacgagccg ggggggctgt
181 c (SEQ ID NO: 14579).
```

[0207] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```
1 cctttttact gccaatgacg catgggatac gtcgtggcag taaaagggt taaatgccaa
61 cgacgcgtcc catacgttgc tggcatttta attcttctct ctgcagcggc agcatgtgcc
121 gccgctgcag agagtttcta gcatgacag cccctctggg caacgagccg ggggggctgt
181 c (SEQ ID NO: 14580).
```

[0208] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 ccttttttact gccaatgacg catgggatac gtcgtggcag taaaagggct taaatgccaa
61 cgacgcgtcc catacgttgt tggcatttta agtcttctct ctgcagcggc agcatgtgcc
121 gccgctgcag agagtttcta gcgatgacag cccctctggg caacgagccg ggggggctgt
181 c (SEQ ID NO: 14581).

```

[0209] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 ccttttttact gccaatgacg catgggatac gtcgtggcag taaaagggct taaatgccaa
61 cgacgcgtcc catacgttgt tggcatttta agtcttctct ctgcagcggc agcatgtgcc
121 gccgctgcag agag (SEQ ID NO: 14582).

```

[0210] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 ccttttttact gccaatgacg catgggatac gtcgtggcag taaaagggct taaatgccaa
61 cgacgcgtcc catacgttgt tggcatttta agtctt (SEQ ID NO: 14583).

```

[0211] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 ccctttgcct gccaatcacg catgggatac gtcgtggcag taaaagggct taaatgccaa
61 cgacgcgtcc catacgttgt tggcatttta agtctt (SEQ ID NO: 14584).

```

[0212] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 ttatcctttt tactgccaat gacgcatggg atacgtcgtg gcagtataag ggcttaaatg
61 ccaacgacgc gtccatacag ttgttggcat ttttaagtct ctctctgcag cggcagcatg
121 tgccgccgct gcagagagtt tctagcgatg acagcccctc tgggcaacga gccggggggg
181 ctgtc (SEQ ID NO: 14585).

```

[0213] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 tttgcatttt tagacattta gaagcctata tcttgttaca gaattggaat tacacaaaaa
61 ttctaccata ttttgaaagc ttaggttgtt ctgaaaaaaa caatatattg ttttcctggg
121 taaactaaaa gtcccctcga ggaaaggccc ctaaagtga acagtgcata acgttcaaaa
181 actgtctggc aatacaagtt ccactttggg acaaatcggc tggcagtga agg (SEQ
ID NO: 14586).

```

[0214] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a left transposon end sequence selected from SEQ ID NO: 14573 and SEQ ID NOs: 14579-14585. In certain embodiments, the left transposon end sequence is preceded by a left target sequence. In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

```

1 tttgcatttt tagacattta gaagcctata tcttgttaca gaattggaat tacacaaaaa
61 ttctaccata ttttgaaagc ttaggttgtt ctgaaaaaaa caatatattg ttttcctggg

```

121 taaactaaaa gtcccctcga ggaaaggccc ctaaagtga acagtgc aaa acgttcaaaa
 181 actgtctggc aatacaagtc ccactttgac caaacggct ggcagtaaaa ggg (SEQ ID
 NO: 14587).

[0215] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

1 ttgttctgaa aaaaacaata tattgttttc ctgggttaaac taaaagtccc ctgaggaaa
 61 ggcccctaaa gtgaaacagt gcaaacggt caaaaactgt ctggcaatac aagttccact
 121 ttgacaaaa cggctggcag taaaagg (SEQ ID NO: 14588).

[0216] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

1 ttgtcatttt tagacattta gaagcctata tcttgttaca gaattggaat tacacaaaaa
 61 ttctaccata ttttgaaagc ttagggttgt ctgaaaaaaa caatatattg ttttcctggg
 121 taaactaaaa gtcccctcga ggaaaggccc ctaaagtga acagtgc aaa acgttcaaaa
 181 actgtctggc aatacaagtc ccactttgac caaacggct ggcagtaaaa gggttat
 (SEQ ID NO: 14589).

[0217] In certain embodiments, the piggyBac or piggyBac-like transposon comprises at a sequence of:

1 ttgttctgaa aaaaacaata tattgttttc ctgggttaaac taaaagtccc ctgaggaaa
 61 ggcccctaaa gtgaaacagt gcaaacggt caaaaactgt ctggcaatac aagttccact
 121 ttgggacaaa tcggctggca gtgaaagg (SEQ ID NO: 14590).

[0218] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a right transposon end sequence selected from SEQ ID NO: 14574 and SEQ ID NOs: 14587-14590.

In certain embodiments, the right transposon end sequence is followed by a right target sequence. In certain embodiments, the left and right transposon ends share a 14 repeated sequence inverted in orientation in the two ends (SEQ ID NO: 14575) adjacent to the target sequence. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a left transposon end comprising a target sequence and a sequence that is selected from SEQ ID NOs: 14582-14584 and 14573, and a right transposon end comprising a sequence selected from SEQ ID NOs: 14588-14590 and 14574 followed by a right target sequence.

[0219] In certain embodiments, the left transposon end of the piggyBac or piggyBac-like transposon comprises

1 atcacgcatg ggatacgtcg tggcagtaaa agggcttaaa tgccaacgac gcgtcccata
 61 cggtt

(SEQ ID NO: 14591), and an ITR. In certain embodiments, the left transposon end comprises

1 atgacgcatg ggatacgtcg tggcagtaaa agggcttaaa tgccaacgac gcgtcccata
 61 cggtgttggc attttaagtc tt

(SEQ ID NO: 14592) and an ITR. In certain embodiments, the right transposon end of the piggyBac or piggyBac-like transposon comprises

```
1 cctgggtaaa ctaaaagtcc cctcgaggaa aggcccttaa agtgaaacag tgcaaaacgt
61 tcaaaaactg tctggcaata caagttccac tttgggacaa atcggc
```

(SEQ ID NO: 14593) and an ITR. In certain embodiments, the right transposon end comprises

```
1 ttgtttctgaa aaaaacaata tattgttttc ctgggtaaac taaaagtccc ctcgaggaaa
61 ggccctaaa gtgaaacagt gcaaaacgtt caaaaactgt ctggcaatac aagttccact
121 ttgaccacaaa cggc
```

(SEQ ID NO: 14594) and an ITR.

[0220] In certain embodiments, one transposon end comprises a sequence that is at least 90%, at least 95%, at least 99% or any percentage in between identical to SEQ ID NO: 14573 and the other transposon end comprises a sequence that is at least 90%, at least 95%, at least 99% or any percentage in between identical to SEQ ID NO: 14574. In certain embodiments, one transposon end comprises at least 14, at least 16, at least 18, at least 20 or at least 25 contiguous nucleotides from SEQ ID NO: 14573 and one transposon end comprises at least 14, at least 16, at least 18, at least 20 or at least 25 contiguous nucleotides from SEQ ID NO: 14574. In certain embodiments, one transposon end comprises at least 14, at least 16, at least 18, at least 20 from SEQ ID NO: 14591, and the other end comprises at least 14, at least 16, at least 18, at least 20 from SEQ ID NO: 14593. In certain embodiments, each transposon end comprises SEQ ID NO: 14575 in inverted orientations.

[0221] In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence selected from of SEQ ID NO: 14573, SEQ ID NO: 14579, SEQ ID NO: 14581, SEQ ID NO: 14582, SEQ ID NO: 14583, and SEQ ID NO: 14588, and a sequence selected from SEQ ID NO: 14587, SEQ ID NO: 14588, SEQ ID NO: 14589 and SEQ ID NO: 14586 and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14517 or SEQ ID NO: 14518.

[0222] In certain embodiments, the piggyBac or piggyBac-like transposon comprises ITRs of CCCTTTGCCTGCCA (SEQ ID NO: 14622) (left ITR) and TGGCAGTGAAAGGG (SEQ ID NO: 14623) (right ITR) adjacent to the target sequences.

[0223] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Helicoverpa armigera*. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid

sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MASRQRLNHD EIATILENDD DYSPLDSESE KEDCVVEDDV WSDNEDAIVD FVEDTSAQED
61 PDNNIASRES PNLEVTSITS HRIITLPQRS IRGKNNHVWS TTKGRTTGRT SAINIIRTNR
121 GPTRMCRNIV DPLLCFQLFI TDEIIHEIVK WTNVEIIVKR QNLKDISASY RDTNTMEIWA
181 LVGILTLTAV MKDNHLSTDE LFDATFSGTR YVSVMSRERF EFLIRCIRMD DKTLRPTLRS
241 DDAFLPVRKI WEIFINQCRQ NHVPGSNLTV DEQLLGFRGR CPFRMYIPNK PDKYGIKFPM
301 MCAAATKYMI DAIPYL GKST KTNGLPLGEF YVKDLTKTVH GTNRNITCDN WETSIPLAKN
361 MLQAPYNLTI VGTIRSNKRE MPEEIKNSRS RPYGSSMFCF DGPLTLVSYK PKPSKMVFLI
421 SSCDENAVIN ESNGKPDML FYNQTKGGVD SEDQMCKSMS ANRKTNRWPM AVFYGMLNMA
481 FVNSYIIYCH NKINKQEKPI SRKEFMKKLS IQLTTPWMQE RLQAPTLKRT LRDNITNVLK
541 NVVPASSENI SNEPEPKRRR YCGVCSYKKR RMTRAQCKKC KKAICGEHNI DVCQDCI
(SEQ ID NO: 14525).

```

[0224] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Helicoverpa armigera*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 ttaaccctag aagcccaatc tacgtaaatt tgacgtatac cgcggcgaaa tatctctgtc
61 tctttcatgt ttaccgtcgg atcgccgcta acttctgaac caactcagta gccattggga
121 cctcgcagga cacagttgcg tcatctcggg aagtgcgcgc atttgttgtt actctctatt
181 acaacacacg tcacgtcacg tcgttgacgc tcatttttgac gtataattgg gctttgtgta
241 acttttgaat ttgtttcaaa ttttttatgt ttgtgattta ttgagtttaa tcgtattgtt
301 tcgttacatt tttcatataa taataatatt ttcagggttg gtacaaa (SEQ ID NO:

```

14570). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 agactgtttt tttctaagag acttctaata tattattacg agttgattta attttatgaa
61 aacattttaa actagttgat tttttttata attacataat ttaagaaaa agtggttagag
121 gcttgatttt tttgttgatt ttttctaaga ttgattataa gtgccataat agtattaata
181 aagagtattt ttaacttaa aatgtatttt atttattaat taaaacttca attatgataa
241 ctcatgcaaa aatatagttc attaacagaa aaaaatagga aaactttgaa gttttgtttt
301 tacacgtcat ttttacgtat gattgggctt tatagctagt taaatatgat tgggcttcta
361 ggggttaa (SEQ ID NO: 14528).

```

[0225] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Pectinophora gossypiella*. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MDLRKQDEKI RQWLEQDIEE DSKGESDNSS SETEDIVEME VHKNTSSESE VSSES DYEPV
61 CPSKRQRTQI IESEESDNSE SIRPSRRQTS RVIDSDETDE DVMSSTPQNI PRNPNIQPS
121 SRFLYGKNKH KWSSAAKPSS VRTSRNIIH FIPGPKERAR EVSEPIDIFS LEISEDMLQQ
181 VVFTTNAEML IRKNKYKTET FTVSPTNLEE IRALLGLLEN AAAMKSNHLP TRMLENTHRS
241 GTIFKACMSA ERLNFLIKCL RFDDKLTRNV RQRDDRFAPL RDLWQALISN FQKWYTPGSY
301 ITVDEQLVGF RGRCSFRMYI PNKPNKYGIK LVMAADVNSK YIVNAIPYLG KGTDPQNQPL
361 ATFFIKEITS TLHGTRNIT MDNWFTSVPL ANELLMAPYN LTLVGTLRSN KREIPEKLKN
421 SKSRAIGTSM FCYDGDKTLV SYKAKSNKVV FILSTIHDQP DINQETGKPE MIHFYNSTKG
481 AVDTVDQMCS SISTNRKTQR WPLCVFYNML NLSIINAYVV YVYNNVRNNK KPMSRRDFVI

```

541 KLGDQLMEPW LRQRLQTVTL REDIKVMIQD ILGESSDLEA PVPSVSNVRK IYYLCPSKAR
601 RMTKHRCIKC KQAICGPHNI DICSRCIE (SEQ ID NO: 14530).

[0226] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Pectinophora gossypiella*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ttaaccctag ataactaaac attcgtccgc tcgacgacgc gctatgccgc gaaattgaag
61 tttacctatt attccgcgctc ccccgccccc gccgcttttt ctagcttcct gatttgcaaa
121 atagtgcacg gcgtgacacg ctcgagggtca cagacaatt aggtcgaaag ttacaggaat
181 ttcgtcgtcc gctcgacgaa agtttagtaa ttacgtaagt ttggcaaagg taagtgaatg
241 aagtattttt ttataattat tttttaattc tttatagtga taacgtaagg tttattttaa
301 tttattactt ttatagttat ttagccaatt gttataaatt ccttgttatt gctgaaaaat
361 ttgcctgttt tagtcaaaat ttattaactt ttcgatcggt ttttag (SEQ ID NO:

14532). In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 tttcactaag taattttgtt cctatttagt agataagtaa cacataatta ttgtgatatt
61 caaaacttaa gaggtttaat aaataataat aaaaaaaaaa tgggttttat ttcttagtct
121 gctcgacgaa tgttttagtta ttacgtaacc gtgaatatag tttagtagtc taggggttaa
(SEQ ID NO: 14571).

[0227] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Ctenopplusia agnata*. The piggyBac or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

1 MASRQHLYQD EIAAILENED DYSPHDTSE MEDCVTQDDV RSDVEDEMVD NINGNTSPAS
61 RHEDPETPDP SSEASNLEVT LSSHRIILP QRSIREKNNH IWSTTKGOSS GRTAAINIVR
121 TNRGPTRMCR NIVDPLLCFQ LFIKEEIVVE IVKWTNVEMV QKRVNLKDIS ASYRDTNEME
181 IWAIIISMLTL SAVMKDNHLS TDELFNVSYG TRYVSVMSE RFEFLRLLR MGDKLLRPNL
241 RQEDAFTPVR KIWEIFINQC RLNYVPGTNL TVDEQLLGFR GRCPFRMYIP NKPDKYGIKE
301 PMVCDAATKY MVDAIPYLGK STKTQGLPLG EFYVKELTQT VHGTNRNVTC DNWFTSVPLA
361 KSLNNSPYNL TLVGTIRSNK REIPEEVKNS RSRQVGSSMF CFDGPLTLVS YKPKPSKMF
421 LLSSCNEDAV VNQSNKPKDM ILFYNQTKGG VDSFDQMCSS MSTNRKTNRW PMAVFYGMLN
481 MAFVNSYIIY CHNMLAKKEK PLSRKDFMKK LSTDLTTPSM QKRLEAPTLK RSLRDNITNV
541 LKIVPQAAID TSFDEPEPKK RRYCGFCSYK KKRMTKTQCF KCKKPVCGEH NIDVCQDCI
(SEQ ID NO: 14534).

[0228] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Ctenopplusia agnata*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ttaaccctag aagcccaatc tacgtcattc tgacgtgtat gtcgccgaaa atactctgtc
61 tctttctcct gcacgatcgg attgccgcga acgctcgatt caaccaggt ggccgccgaga
121 tctattggag gactgcggcg ttgattcggg aagtcccgcg attttgtcat agtaacagta
181 ttgcaogtca gcttgacgta tatttgggct ttgtgttatt ttgttaaatt ttcaacgtta
241 gtttattatt gcatcttttt gttacattac tggtttattt gcatgtatta ctcaaatatt
301 atttttattt tagcgtagaa aataca (SEQ ID NO: 14535).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 agactgtttt ttttgtattt gcattatata ttatatctta aagttgattt aattcctaaga
61 aaaacattaa aataagtttc tttttgtaaa atttaattaa ttataagaaa aagtttaagt
121 tgatctcatt ttttataaaa atttgcaatg tttccaaagt tattattgta aaagaataaa
181 taaaagtaaa ctgagtttta attgatgttt tattatatca ttatactata tattacttaa
241 ataaaacaat aactgaatgt atttcctaaa ggaatcacta gaaaatatag tgatcaaaaa
301 ttacacgtc atttttgcgt atgattgggc tttatagggt ctaaaaatat gattgggcct
361 ctaggggttaa (SEQ ID NO: 14536).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCCTAGAAGCCCAATC (SEQ ID NO: 14564).

[0229] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Agrotis ipsilon*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MESRQRLNQD EIATILENDD DYSPLSDSE AEDRVVEDDV WSDNEDAMID YVEDTSRQED
61 PDNNIASQES ANLEVTSLS HRIISLPQRS ICGKNNHVWS TTKGRRTGRT SAINIIRTNR
121 GPTRMCRNIV DPLLCFQLEI TDEIIHEIVK WTNVEMIVKR QNLIDISASY RDTNTMEMWA
181 LVGILTLTAV MKDNHLSTDE LFDATFSGTR YVSVMSRERF EFLIRC MRMD DKT LRPT LRS
241 DDAFIPVRKL WEIFINQCR L NYVPGGNLTV DEQLLGFRGR CPFRMYIPNK PDKYGI RFP M
301 MCDAATKYMI DAIPYL GKST KTNGLPLGEF YVKELTKTVH GTNRNV TCDN WFTSIPLAKN
361 MLQAPYNLTI VGTIRSNKRE IPEEIKNSRS RPYGSSMFCF DGPLTLVSYK PKPSRMV FLL
421 SSCDENAVIN ESNGKPD MIL FYNQTKGGVD SFDQMCKSMS ANRKTNRWPM AVFYGMLNMA
481 FVNSYIIYCH NKINKQKKPI NRKEFMKNLS TDLTTPWMQE RLKAPT LKRT LRDNITNV LK
541 NVVPPSPANN SEEPGPKKRS YCGFCSYKKR RMTKTQFYKC KKAICGEHNI DVCQDCV
(SEQ ID NO: 14537).

```

[0230] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Agrotis ipsilon*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 ttaaccctag aagcccaatc tacgtaaatt tgacgtatac cgcggcgaaa tatatctgtc
61 tctttcacgt ttaccgtcgg attcccgtca acttcggaac caactcagta gccattgaga
121 actcccagga cacagttgcg tcatctcggg aagtgccgcc attttgttgt aatagacagg
181 ttgcacgtca ttttgacgta taattgggct ttgtgtaact ttgaaatta tttataattt
241 ttattgatgt gatttatatt agttaatcgt attgtttcgt tacatttttc atatgatatt
301 aatattttca gattgaatat aaa (SEQ ID NO: 14538). In certain embodiments,

```

the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 agactgtttt ttttaaaagg cttataaagt attactattg cgtgatttaa ttttataaaa
61 atatttataaa ccagttgatt tttttaataa ttacctaat ttaagaaaaa atgtagaag
121 cttgatattt ttgttgattt ttttctaaga tttgattaaa aggccataat tgtattaata
181 aagagtattt ttaacttcaa atttatttta tttattaatt aaaacttcaa ttatgataat
241 acatgcaaaa atatagttca tcaacagaaa aatataggaa aactctaata gttttatttt
301 tacacgtcat ttttacgtat gattgggctt tatagctagt caaatatgat tgggcttcta
361 ggggttaa (SEQ ID NO: 14539).

```

[0231] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Megachile rotundata*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MNGKDSLGEF YLDDLSDCLD CRSASSTDDE SDSSNIAIRK RCPIPLIYSD SEDEDMNNNV
61 EDNNHFVKES NRYHYQIVEK YKITSKTKKW KDVTVTEMKK FLGLIILMGQ VKKDVLYDYW
121 STDPSIETPF FSKVMSRNR F LQIMQSWHFY NNNDISPNSH RLVKIQPVID YFKEKFNNVY
181 KSDQQLSLDE CLIPWRGRSL IKTYNPAKIT KYGILVRVLS EARTGYVSNF CVYAADGKKI
241 EETVLSVIGP YKNMWHHVYQ DNYNSVNIA KIFLKNKLRV CGTIRKNRSL PQILQTVKLS
301 RGQHQLFLNG HTLLEVWNNG KRNVMISTI HSAQMAESRN RSRTSDCPIQ KPISIIDYNK
361 YMRGVDRADQ YLSYYSIFRK TKKWTKRVM FFINCALENS FKVYTTLNGQ KITYKNFLHK
421 AALSLIEDCG TEEQGTDLN SEPTTTRTTS RVDHPGRLEN FGKHKLVNIV TSGQCKKPLR
481 QCRVCASKKK LSRTGFAKY CNVPLHKGDC FERYHSLKKY (SEQ ID NO: 14540).

```

[0232] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Megachile rotundata*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 ttaaataatg ccactctag atgaacttaa cactttaccg accggccgtc gattattcga
61 cgtttgctcc ccagcgctta ccgaccggcc atcgattatt cgacgtttgc ttcccagcgc
121 ttaccgaccg gtcatcgact tttgatcttt ccgtagatt tggtaggtc agattgacaa
181 gtagcaagca ttccgcattc tttattcaaa taatcggtgc ttttttctaa gcttagccc
241 ttagaa (SEQ ID NO: 14541).

```

In certain embodiments, the the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1 acaacttctt ttttcaacaa atattgttat atggattatt tactttattta tttatttatg
61 gtatatattta tgtttattta tttatggta ttatgggata ttttatgtaa ataataaact
121 gaaaacgatt gtaatagatg aaataaatat tgttttaaca ctaatataat taaagtaaaa
181 gatttttaata aatttcgtta ccctacaata acacgaagcg tacaatttta ccagagtta
241 ttaa (SEQ ID NO: 14542).

```

[0233] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Bombus impatiens*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1 MNEKNGIGEF YLDDLSDCPD SYSRSNSGDE SDGSDTIIRK RGSVLPPRYS DSEDDEINNV
61 EDNANNVENN DDIWSTNDEA IILEPFEGSP GLKIMPSSAE SVTDNVNLEF GDDFFEHLVR
121 ESNRYHYQVM EKYKIPSKAK KWTDTVP EM KKFLGLIVLM GQIKKDVLYD YWSTDPSIET
181 PFFSQVMSRN RFVQIMQSWH FCNNDNIPHD SHRLAKIQPV IDYFRRKFND VYKPCQQLSL
241 DESIIPWRGR LSIKTYNEAK ITKYGILVRV LSEAVTGYVC NFDVYAADGK KLEDTAVIEP
301 YKNIWHQIYQ DNYNSVKMA RILLKNKVRV CGTIRKNRGL PRSLKTIQLS RGQYEFRRNH
361 QILLEWNNNG RRNVNMISTI HSAQLMESRS KSKRSDVPIQ KPNSIIDYNK YMRGVDRADQ

```

421 YLAYYSIFRK TKKWTKRVVM FFINCALFNS FRVYTILNGK NITYKNFLHK VAVSWIEDGE
 481 TNCTEQDDNL PNSEPTRRAP RLDHFGRLSN YGKHKLINIV TSGRSLKPQR QCRVCAVQKK
 541 RSRTCFVCKF CNVPLHKGDC FERYHTLKKY (SEQ ID NO: 14543).

[0234] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Bombus impatiens*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ttaattttttt aacatttttac cgaccgatag ccgattaatc gggttttttgc cgctgacgct
 61 taccgaccga taacctatta atcggttttt tgctgctgaa gcttaccacac ctatagccta
 121 cctatagtta atcggttgcc atggcgataa acaatctttc tcattatatg agcagtaatt
 181 tgttatttag tactaaggta ccttgctcag ttgcgtcagt tgcgttgctt tgtaagctcc
 241 cacagtttta taccaattcg aaaaacttac cgttcgcg (SEQ ID NO: 14544).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 actattttcac atttgaacta aaaaccgttg taatagataa aataaatata atttagtatt
 61 aatattatgg aaacaaaaga ttttattcaa ttttaattatc ctatagtaac aaaaagcggc
 121 caattttatc tgagcatatg aaaagcacag atactcccg cgcacagtct aaaccgaaac
 181 agagccggcg ccagggagaa tctgcgcctg agcagccggg cggacgtgcg tttgctgttg
 241 aaccgctagt ggtcagtaaa ccagaaccag tcagtaagcc agtaactgat cagttaacta
 301 gattgtatag ttcaaattga acttaattca gtttttaagc gtttgaatgt tgtctaactt
 361 cgttatatat tatattcttt ttaa (SEQ ID NO: 14545).

[0235] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Mamestra brassicae*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

1 MFSFVPNKEQ TRTVLIFCFH LKTTAAESHR PLVEAFGEQV PTVKTCERWF QRFKSGDFDV
 61 DDKEHGKPPK RYEDAELOAL LDEDDAQTQK QLAEQLEVSQ QAVSNRLREG GKIQKVGRWV
 121 PHELNERQRE RRKNTCEILL SRYKRKSEFLH RIVTGEKWI FFVNPKRKKS YVDPGQPATS
 181 TARPNNRFGKK TRLCVWWDQS GVIYYELLKP GETVNTARYQ QQLINLNRAL QRKRPEYQKR
 241 QHRVIFLHDN APSHTARAVR DTLETNLNEV LPHAAYSDDL APSDYHLEAS MGHALAEQRF
 301 DSYESVEEWL DEWFAAKDDE FYWRGIHKLP ERWDNCVASD GKYFE (SEQ ID NO:
 14546).

[0236] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Mamestra brassicae*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ttatttgggtt gcccaaaaag taattgcgga tttttcatat acctgtcttt taaacgtaca
 61 tagggatcga actcagtaaa actttgacct tgtgaaataa caaacttgac tgtccaacca
 121 ccatagtttg gcgcgaattg agcgtcataa ttgttttgac tttttgcagt caac (SEQ

ID NO: 14547).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 atgattttttt ctttttaaac caatttttaat tagttaattg atataaaaaat ccgcaattac
 61 tttttgggca acccaataa (SEQ ID NO: 14548).

[0237] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Mayetiola destructor*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MENFENWRKR RHLREVLLGH FFAKKTAAES HRLLEVEVYGE HALAKTQCFE WFQRFKSGDF
61  DTEDKERPGQ PKKFEDEELE ALLDEDCCQT QEELAKSLGV TQQAISKRLK AAGYIQKQGN
121 WVPHELKPRD VERRFCMSEM LLQRHKKKSF LSRIITGDEK WIHYDNSKRK KSYVKRGGRA
181 KSTPKSNLHG AKVMLCIWWD QRGVLYYELL EPGQTITGDL YRTQLIRLKQ ALAEKRPEYA
241 KRHGAVIFHH DNARPHVALP VKNYLENSGW EVLPHPPYSP DLAPSDYHLF RSMQNDLAGK
301 RFTSEQGIRK WLDSEFLAAK AKFFEKGIHE LSERWEKVIA SDGQYFE (SEQ ID NO:
14549).

```

[0238] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Mayetiola destructor*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  taagacttcc aaaatttcca cccgaacttt accttccccg cgcattatgt ctctcttttc
61  accctctgat ccttggtatt gttgtcgagc acgatttata ttgggtgtac aacttaaaaa
121 ccggaattgg acgctagatg tccacactaa cgaatagtgt aaaagcacia atttcatata
181 tacgtcattt tgaaggtaca ttgacagct atcaaaatca gtcaataaaa ctattctatc
241 tgtgtgcata atattttttt attaact (SEQ ID NO: 14550).

```

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  tgcattcatt cattttgtta tcgaaataaa gcattaattt tcaactaaaaa attccgggtt
61  ttaagttgta caccacaata catccttagt gacaattttc aaatggcttt ccattgagc
121 tgaaaccgtg gctctagtaa gaaaaacgcc caaccgctca tcatatgcct tttttttctc
181 aacatccg (SEQ ID NO: 14551).

```

[0239] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Apis mellifera*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

```

1  MENQKEHYRH ILLFYFRKGK NASQAHKKLC AVYGDEALKE RQCQNWFDKF RSGDFSLEKDE
61  KRSGRPVEVD DDLIKAIIDS DRHSTTREIA EKLHVSHTCI ENHLKQLGYV QKLDTWVPHE
121 LKEKHLTQRI NSCDLLKKPN ENDPFLKRLI TGDEKWVYIN NIKRKRSWSR PREPAQTTSK
181 AGIHRKKVLL SVWWDYKGIV YFELLPPNRT INSVVYIEQL TKLNNAVEEK RPELTNRKGV
241 VFHHDNARPH TSLVTRQKLL ELGWDVLPHP PYSPDLAPSD YFLFRSLQNS LNGKNFNDD
301 DIKSYLIQFF ANKNQKEYER GIMMLPERWQ KVIDQNGQHI TE (SEQ ID NO: 14552).

```

[0240] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Apis mellifera*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

```

1  ttgggttggc aactaagtaa ttgcggattt cactcataga tggcttcagt tgaattttta

```

61 ggtttgctgg cgtagtccaa atgtaaaaca cattttgtta ttgatagtt ggcaattcag
 121 ctgtcaatca gtaaaaaaag ttttttgatc ggttgcgtag ttttcgtttg gcgttcgttg
 181 aaaa (SEQ ID NO: 14553).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 agttatttag ttccatgaaa aaattgtctt tgattttcta aaaaaaatcc gcaattactt
 61 agttgccaat ccaa (SEQ ID NO: 14554).

[0241] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Messor bouvieri*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

1 MSSFVPENVH LRHALLFLEH QKKRAAESHRLVETYGEHA PTIRTCETWF RQFKCGDFNV
 61 QDKERPGRPK TFEDAELQEL LDEDSTQTQK QLAEKLNVSR VAICERLQAM GKIQKMGRWV
 121 PHELNDRQME NRKIVSEMLL QRYERKSFLH RIVTGDEKWI YFENPKRKKS WLSPEAGPS
 181 TARPNRFRGK TMLCVWWDQI GVVYYELLKP GETVNTDRYP QQMINLNCAL IEKRPQYAQR
 241 HDKVILQHDN APSHTAKPVK EMLKSLGWEV LSHPPYSPDL APSDYHLFAS MGHALAEQHF
 301 ADFEEVKKWL DEWFSSKEKL FFWNGIHKLS ERWTKCIESN GQYFE (SEQ ID NO: 14555).

[0242] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Messor bouvieri*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 agtcagaaat gacacctcga tcgacgacta atcgacgtct aatcgacgtc gattttatgt
 61 caacatgtta ccaggtgtgt cggtaattcc tttccggttt ttccggcaga tgtcactagc
 121 cataagtatg aatgttatg atttgataca tatgtcattt tattctactg acattaacct
 181 taaaactaca caagttacgt tccgccaaaa taacagcggt atagatttat aattttttga
 241 aa (SEQ ID NO: 14556).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ataaatttga actatccatt ctaagtaacg tgttttcttt aacgaaaaaa ccggaaaaga
 61 attaccgaca ctctgggtat gtcaacatgt tattttcgac attgaatcgc gtcgattcga
 121 agtcgatcga ggtgtcattt ctgact (SEQ ID NO: 14557).

[0243] In certain embodiments of the methods of the disclosure, the transposase enzyme is a piggyBac or piggyBac-like transposase enzyme. In certain embodiments, the piggyBac or piggyBac-like transposase enzyme is isolated or derived from *Trichoplusia ni*. The piggyBac (PB) or piggyBac-like transposase enzyme may comprise or consist of an amino acid sequence at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between identical to:

1 MGSSLDDEHI LSALLQSDDE LVGEDSDSEV SDHVSDDVQ SDTEEAFIDE VHEVQPTSSG
 61 SEILDEQNVI EQPGSSLASN RILTLPQRTI RGKNKHCWST SKSTRRSRVS ALNIVRSQRG
 121 PTRMCRNIYD PLLCFKLFFT DEIISEIVKW TNAEISLKRR ESMTSATFRD TNEDEIYAFF
 181 GILVMTAVRK DNHMSTDDLF DRSLSMVYVS VMSRDRDFEL IRCLRMDDKS IRPTLRENDV
 241 FTPVRKIWDL FIHQCIQNYT PGAHLTIDEQ LLGFRGRCPF RVYIPNKPSK YGIKILMMCD

301 SGT KYMINGM PYLGRGTQTN GVPLGEYYVK ELSKPVHGSC RNITCDNWFT SIPLAKNLLQ
 361 EPYKLTIVGT VRSNKREIPE VLKNSRSRPV GTSMFCDGP LTLVSYKPKP AKMVYLLSSC
 421 DEDASINEST GKPMVMYYN QTKGGVDTLD QMCSVMTCR KTNRWPMALL YGMINIACIN
 481 SFIIYSHNVS SKGEKVQSRK KFMRNLYMSL TSSEMRKRLE APTLKRYLRD NISNILEKEV
 541 PGTSDDSTEE PVMKKRTYCT YCPSKIRKKA NASCKKCKKV ICREHNIDMC QSCF (SEQ
 ID NO: 14558).

[0244] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Trichoplusia ni*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ttaaccctag aaagatagtc tgcgtaaaat tgacgcattgc attcttgaaa tattgctctc
 61 tctttctaaa tagcgcgaaat ccgtcgtgtg gcatttagga catctcagtc gccgcttgga
 121 gctcccggtg ggcggtgcttg tcaatgcggt aagtgtcact gattttgaac tataacgacc
 181 gcgtgagtc aaatgacgca tgattatctt ttacgtgact tttaagattt aactcatagc
 241 ataattatat tggtatttca tggtctactt acgtgataac ttattatata tatattttct
 301 tggtatagat atc (SEQ ID NO: 14559).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 tttgttactt tatagaagaa attttgagtt tttgtttttt tttaataaat aaataaacat
 61 aaataaattg tttgttgaat ttattattag tatgtaagt taaatataat aaaacttaat
 121 atctattcaa attaataaat aaacctcgat atacagaccg ataaaacaca tgcgtcaatt
 181 ttacgcattg ttatctttta cgtacgtcac aatatgatta tctttctagg gttaa (SEQ
 ID NO: 14560).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 ccctagaaag atagtctgcg taaaattgac gcatgcattc ttgaaatatt gctctctctt
 61 tctaaatagc gcgaatccgt ccgtgtgcat ttaggacatc tcagtcgccg cttggagctc
 121 ccgtgaggcg tgcttgctaa tgcggtaagt gtcactgatt ttgaactata acgaccgcgt
 181 gagtcaaaat gacgcattg tatctttttac gtgactttta agatttaact catacgataa
 241 ttatattggt atttcatggt ctacttacgt gataacttat tatatatata ttttcttggt
 301 atagatattc (SEQ ID NO: 14561).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 tttgttactt tatagaagaa attttgagtt tttgtttttt tttaataaat aaataaacat
 61 aaataaattg tttgttgaat ttattattag tatgtaagt taaatataat aaaacttaat
 121 atctattcaa attaataaat aaacctcgat atacagaccg ataaaacaca tgcgtcaatt
 181 ttacgcattg ttatctttta cgtacgtcac aatatgatta tctttctagg g (SEQ ID
 NO: 14562).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 tctaaatagc gcgaatccgt ccgtgtgcat ttaggacatc tcagtcgccg cttggagctc
 61 ccgtgaggcg tgcttgctaa tgcggtaagt gtcactgatt ttgaactata acgaccgcgt
 121 gagtcaaaat gacgcattg tatctttttac gtgactttta agatttaact catacgataa
 181 ttatattggt atttcatggt ctacttacgt gataacttat tatatatata ttttcttggt
 241 atagatattc (SEQ ID NO: 14609).

In certain embodiments, the piggyBac or piggyBac-like transposon comprises a sequence of:

1 tttgttactt tatagaagaa attttgagtt tttgtttttt tttaataaat aaataaacat
 61 aaataaattg tttgttgaat ttattattag tatgtaagt taaatataat aaaacttaat
 121 atctattcaa attaataaat aaacctcgat atacagaccg ataaaacaca tgcgtcaatt
 181 ttacgcattg ttatctttta cgtacgtcac aatatgatta tctttctagg g (SEQ ID
 NO: 14610).

[0245] In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14561 and SEQ ID NO: 14562, and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14558. In certain embodiments, the piggyBac or piggyBac-like transposon comprises SEQ ID NO: 14609 and SEQ ID NO: 14610, and the piggyBac or piggyBac-like transposase comprises SEQ ID NO: 14558.

[0246] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Aphis gossypii*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCTTCCAGCGGGCGCGC (SEQ ID NO: 14565).

[0247] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Chilo suppressalis*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCCAGATTAGCCT (SEQ ID NO: 14566).

[0248] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Heliothis virescens*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCCTTAATTACTCGCG (SEQ ID NO: 14567).

[0249] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Pectinophora gossypiella*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCCTAGATAACTAAAC (SEQ ID NO: 14568).

[0250] In certain embodiments, the piggyBac or piggyBac-like transposon is isolated or derived from *Anopheles stephensi*. In certain embodiments, the piggyBac or piggyBac-like transposon comprises an ITR sequence of CCCTAGAAAGATA (SEQ ID NO: 14569).

Immune and Immune Precursor Cells

[0251] In certain embodiments, immune cells of the disclosure comprise lymphoid progenitor cells, natural killer (NK) cells, T lymphocytes (T-cell), stem memory T cells (T_{SCM} cells), central memory T cells (T_{CM}), stem cell-like T cells, B lymphocytes (B-cells), myeloid progenitor cells, neutrophils, basophils, eosinophils, monocytes, macrophages, platelets, erythrocytes, red blood cells (RBCs), megakaryocytes or osteoclasts.

[0252] In certain embodiments, immune precursor cells comprise any cells which can differentiate into one or more types of immune cells. In certain embodiments, immune precursor cells comprise multipotent stem cells that can self renew and develop into immune cells. In certain embodiments, immune precursor cells comprise hematopoietic stem cells (HSCs) or descendants thereof. In certain embodiments, immune precursor cells comprise

precursor cells that can develop into immune cells. In certain embodiments, the immune precursor cells comprise hematopoietic progenitor cells (HPCs).

Hematopoietic Stem Cells (HSCs)

[0253] Hematopoietic stem cells (HSCs) are multipotent, self-renewing cells. All differentiated blood cells from the lymphoid and myeloid lineages arise from HSCs. HSCs can be found in adult bone marrow, peripheral blood, mobilized peripheral blood, peritoneal dialysis effluent and umbilical cord blood.

[0254] HSCs of the disclosure may be isolated or derived from a primary or cultured stem cell. HSCs of the disclosure may be isolated or derived from an embryonic stem cell, a multipotent stem cell, a pluripotent stem cell, an adult stem cell, or an induced pluripotent stem cell (iPSC).

[0255] Immune precursor cells of the disclosure may comprise an HSC or an HSC descendent cell. Exemplary HSC descendent cells of the disclosure include, but are not limited to, multipotent stem cells, lymphoid progenitor cells, natural killer (NK) cells, T lymphocyte cells (T-cells), B lymphocyte cells (B-cells), myeloid progenitor cells, neutrophils, basophils, eosinophils, monocytes, and macrophages.

[0256] HSCs produced by the methods of the disclosure may retain features of “primitive” stem cells that, while isolated or derived from an adult stem cell and while committed to a single lineage, share characteristics of embryonic stem cells. For example, the “primitive” HSCs produced by the methods of the disclosure retain their “stemness” following division and do not differentiate. Consequently, as an adoptive cell therapy, the “primitive” HSCs produced by the methods of the disclosure not only replenish their numbers, but expand in vivo. “Primitive” HSCs produced by the methods of the disclosure may be therapeutically-effective when administered as a single dose. In some embodiments, primitive HSCs of the disclosure are CD34+. In some embodiments, primitive HSCs of the disclosure are CD34+ and CD38-. In some embodiments, primitive HSCs of the disclosure are CD34+, CD38- and CD90+. In some embodiments, primitive HSCs of the disclosure are CD34+, CD38-, CD90+ and CD45RA-. In some embodiments, primitive HSCs of the disclosure are CD34+, CD38-, CD90+, CD45RA-, and CD49f+. In some embodiments, the most primitive HSCs of the disclosure are CD34+, CD38-, CD90+, CD45RA-, and CD49f+.

[0257] In some embodiments of the disclosure, primitive HSCs, HSCs, and/or HSC descendent cells may be modified according to the methods of the disclosure to express an exogenous sequence (e.g. a chimeric antigen receptor or therapeutic protein). In some

embodiments of the disclosure, modified primitive HSCs, modified HSCs, and/or modified HSC descendent cells may be forward differentiated to produce a modified immune cell including, but not limited to, a modified T cell, a modified natural killer cell and/or a modified B-cell of the disclosure.

T Cells

[0258] Modified T cells of the disclosure may be derived from modified hematopoietic stem and progenitor cells (HSPCs) or modified HSCs.

[0259] Unlike traditional biologics and chemotherapeutics, modified-T cells of the disclosure possess the capacity to rapidly reproduce upon antigen recognition, thereby potentially obviating the need for repeat treatments. To achieve this, in some embodiments, modified-T cells of the disclosure not only drive an initial response, but also persist in the patient as a stable population of viable memory T cells to prevent potential relapses. Alternatively, in some embodiments, when it is not desired, modified-T cells of the disclosure do not persist in the patient.

[0260] Intensive efforts have been focused on the development of antigen receptor molecules that do not cause T cell exhaustion through antigen-independent (tonic) signaling, as well as of a modified-T cell product containing early memory T cells, especially stem cell memory (T_{SCM}) or stem cell-like T cells. Stem cell-like modified-T cells of the disclosure exhibit the greatest capacity for self-renewal and multipotent capacity to derive central memory (T_{CM}) T cells or T_{CM} like cells, effector memory (T_{EM}) and effector T cells (T_E), thereby producing better tumor eradication and long-term modified-T cell engraftment. A linear pathway of differentiation may be responsible for generating these cells: Naïve T cells (T_N) $>$ T_{SCM} $>$ T_{CM} $>$ T_{EM} $>$ T_E $>$ T_{TE} , whereby T_N is the parent precursor cell that directly gives rise to T_{SCM} , which then, in turn, directly gives rise to T_{CM} , etc. Compositions of T cells of the disclosure may comprise one or more of each parental T cell subset with T_{SCM} cells being the most abundant (e.g. T_{SCM} $>$ T_{CM} $>$ T_{EM} $>$ T_E $>$ T_{TE}).

[0261] In some embodiments of the methods of the disclosure, the immune cell precursor is differentiated into or is capable of differentiating into an early memory T cell, a stem cell like T-cell, a Naïve T cells (T_N), a T_{SCM} , a T_{CM} , a T_{EM} , a T_E , or a T_{TE} . In some embodiments, the immune cell precursor is a primitive HSC, an HSC, or a HSC descendent cell of the disclosure.

[0262] In some embodiments of the methods of the disclosure, the immune cell is an early memory T cell, a stem cell like T-cell, a Naïve T cells (T_N), a T_{SCM}, a T_{CM}, a T_{EM}, a T_E, or a T_{TE}.

[0263] In some embodiments of the methods of the disclosure, the immune cell is an early memory T cell.

[0264] In some embodiments of the methods of the disclosure, the immune cell is a stem cell like T-cell.

[0265] In some embodiments of the methods of the disclosure, the immune cell is a T_{SCM}.

[0266] In some embodiments of the methods of the disclosure, the immune cell is a T_{CM}.

[0267] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of an early memory T cell. In certain embodiments, the plurality of modified early memory T cells comprises at least one modified stem cell-like T cell. In certain embodiments, the plurality of modified early memory T cells comprises at least one modified T_{SCM}. In certain embodiments, the plurality of modified early memory T cells comprises at least one modified T_{CM}.

[0268] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of a stem cell-like T cell. In certain embodiments, the plurality of modified stem cell-like T cells comprises at least one modified T_{SCM}. In certain embodiments, the plurality of modified stem cell-like T cells comprises at least one modified T_{CM}.

[0269] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of a stem memory T cell (T_{SCM}). In certain embodiments, the cell-surface markers comprise CD62L and CD45RA. In certain embodiments, the cell-surface markers comprise one or more of CD62L, CD45RA, CD28, CCR7, CD127, CD45RO, CD95, CD95

and IL-2R β . In certain embodiments, the cell-surface markers comprise one or more of CD45RA, CD95, IL-2R β , CCR7, and CD62L.

[0270] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of a central memory T cell (T_{CM}). In certain embodiments, the cell-surface markers comprise one or more of CD45RO, CD95, IL-2R β , CCR7, and CD62L.

[0271] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of a naïve T cell (T_N). In certain embodiments, the cell-surface markers comprise one or more of CD45RA, CCR7 and CD62L.

[0272] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of an effector T-cell (modified T_{EFF}). In certain embodiments, the cell-surface markers comprise one or more of CD45RA, CD95, and IL-2R β .

[0273] In some embodiments of the methods of the disclosure, the methods modify and/or the methods produce a plurality of modified T cells, wherein at least 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or any percentage in between of the plurality of modified T cells expresses one or more cell-surface marker(s) of a stem cell-like T cell, a stem memory T cell (T_{SCM}) or a central memory T cell (T_{CM}).

[0274] In some embodiments of the methods of the disclosure, a buffer comprises the immune cell or precursor thereof. The buffer maintains or enhances a level of cell viability and/or a stem-like phenotype of the immune cell or precursor thereof, including T-cells. In certain embodiments, the buffer maintains or enhances a level of cell viability and/or a stem-like phenotype of the primary human T cells prior to the nucleofection. In certain embodiments, the buffer maintains or enhances a level of cell viability and/or a stem-like phenotype of the primary human T cells during the nucleofection. In certain embodiments,

the buffer maintains or enhances a level of cell viability and/or a stem-like phenotype of the primary human T cells following the nucleofection. In certain embodiments, the buffer comprises one or more of KCl, MgCl₂, ClNa, Glucose and Ca(NO₃)₂ in any absolute or relative abundance or concentration, and, optionally, the buffer further comprises a supplement selected from the group consisting of HEPES, Tris/HCl, and a phosphate buffer. In certain embodiments, the buffer comprises 5 mM KCl, 15 mM MgCl₂, 90 mM ClNa, 10 mM Glucose and 0.4 mM Ca(NO₃)₂. In certain embodiments, the buffer comprises 5 mM KCl, 15 mM MgCl₂, 90 mM ClNa, 10 mM Glucose and 0.4 mM Ca(NO₃)₂ and a supplement comprising 20 mM HEPES and 75 mM Tris/HCl. In certain embodiments, the buffer comprises 5 mM KCl, 15 mM MgCl₂, 90 mM ClNa, 10 mM Glucose and 0.4 mM Ca(NO₃)₂ and a supplement comprising 40 mM Na₂HPO₄/NaH₂PO₄ at pH 7.2. In certain embodiments, the composition comprising primary human T cells comprises 100 μ l of the buffer and between 5x10⁶ and 25x10⁶ cells. In certain embodiments, the composition comprises a scalable ratio of 250x10⁶ primary human T cells per milliliter of buffer or other media during the introduction step.

[0275] In some embodiments of the methods of the disclosure, the methods comprise contacting an immune cell of the disclosure, including a T cell of the disclosure, and a T-cell expansion composition. In some embodiments of the methods of the disclosure, the step of introducing a transposon and/or transposase of the disclosure into an immune cell of the disclosure may further comprise contacting the immune cell and a T-cell expansion composition. In some embodiments, including those in which the introducing step of the methods comprises an electroporation or a nucleofection step, the electroporation or a nucleofection step may be performed with the immune cell contacting T-cell expansion composition of the disclosure.

[0276] In some embodiments of the methods of the disclosure, the T-cell expansion composition comprises, consists essentially of or consists of phosphorus; one or more of an octanoic acid, a palmitic acid, a linoleic acid, and an oleic acid; a sterol; and an alkane.

[0277] In certain embodiments of the methods of producing a modified T cell of the disclosure, the expansion supplement comprises one or more cytokine(s). The one or more cytokine(s) may comprise any cytokine, including but not limited to, lymphokines. Exemplary lymphokines include, but are not limited to, interleukin-2 (IL-2), interleukin-3 (IL-3), interleukin-4 (IL-4), interleukin-5 (IL-5), interleukin-6 (IL-6), interleukin-7 (IL-7), interleukin-15 (IL-15), interleukin-21 (IL-21), granulocyte-macrophage colony-stimulating

factor (GM-CSF) and interferon-gamma (INF γ). The one or more cytokine(s) may comprise IL-2.

[0278] In some embodiments of the methods of the disclosure, the T-cell expansion composition comprises human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid, nicotinamide, 2,4,7,9-tetramethyl-5-decyn-4,7-diol (TMDD), diisopropyl adipate (DIPA), n-butyl-benzenesulfonamide, 1,2-benzenedicarboxylic acid, bis(2-methylpropyl) ester, palmitic acid, linoleic acid, oleic acid, stearic acid hydrazide, oleamide, a sterol and an alkane. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid, palmitic acid, linoleic acid, oleic acid and a sterol. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid at a concentration of between 0.9 mg/kg to 90 mg/kg, inclusive of the endpoints; palmitic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; linoleic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; oleic acid at a concentration of 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; and a sterol at a concentration of about 0.1 mg/kg to 10 mg/kg, inclusive of the endpoints. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid at a concentration of about 9 mg/kg, palmitic acid at a concentration of about 2 mg/kg, linoleic acid at a concentration of about 2 mg/kg, oleic acid at a concentration of about 2 mg/kg and a sterol at a concentration of about 1 mg/kg. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid at a concentration of between 6.4 μ mol/kg and 640 μ mol/kg, inclusive of the endpoints; palmitic acid at a concentration of between 0.7 μ mol/kg and 70 μ mol/kg, inclusive of the endpoints; linoleic acid at a concentration of between 0.75 μ mol/kg and 75 μ mol/kg, inclusive of the endpoints; oleic acid at a concentration of between 0.75 μ mol/kg and 75 μ mol/kg, inclusive of the endpoints; and a sterol at a concentration of between 0.25 μ mol/kg and 25 μ mol/kg, inclusive of the endpoints. In certain embodiments of this method, the T-cell expansion composition further comprises one or more of octanoic acid at a concentration of about 64 μ mol/kg, palmitic acid at a concentration of about 7 μ mol/kg, linoleic acid at a concentration of about 7.5 μ mol/kg, oleic acid at a concentration of about 7.5 μ mol/kg and a sterol at a concentration of about 2.5 μ mol/kg.

[0279] In certain embodiments, the T-cell expansion composition comprises one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement to produce a plurality of expanded modified T-cells, wherein at least 2% of the plurality of modified T-cells expresses one or more cell-surface marker(s) of an early memory T cell, a stem cell-like T cell, a stem memory T cell (T_{SCM}) and/or a central memory T cell (T_{CM}). In certain embodiments, the T-cell expansion composition comprises or further comprises one or more of octanoic acid, nicotinamide, 2,4,7,9-tetramethyl-5-decyn-4,7-diol (TMDD), diisopropyl adipate (DIPA), n-butyl-benzenesulfonamide, 1,2-benzenedicarboxylic acid, bis(2-methylpropyl) ester, palmitic acid, linoleic acid, oleic acid, stearic acid hydrazide, oleamide, a sterol and an alkane. In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid, palmitic acid, linoleic acid, oleic acid and a sterol (e.g. cholesterol). In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of between 0.9 mg/kg to 90 mg/kg, inclusive of the endpoints; palmitic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; linoleic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; oleic acid at a concentration of 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; and a sterol at a concentration of about 0.1 mg/kg to 10 mg/kg, inclusive of the endpoints (wherein mg/kg = parts per million). In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of about 9 mg/kg, palmitic acid at a concentration of about 2 mg/kg, linoleic acid at a concentration of about 2 mg/kg, oleic acid at a concentration of about 2 mg/kg, and a sterol at a concentration of about 1 mg/kg (wherein mg/kg = parts per million). In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of 9.19 mg/kg, palmitic acid at a concentration of 1.86 mg/kg, linoleic acid at a concentration of about 2.12 mg/kg, oleic acid at a concentration of about 2.13 mg/kg, and a sterol at a concentration of about 1.01 mg/kg (wherein mg/kg = parts per million). In certain embodiments, the T-cell expansion composition comprises octanoic acid at a concentration of 9.19 mg/kg, palmitic acid at a concentration of 1.86 mg/kg, linoleic acid at a concentration of 2.12 mg/kg, oleic acid at a concentration of about 2.13 mg/kg, and a sterol at a concentration of 1.01 mg/kg (wherein mg/kg = parts per million). In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of between 6.4 μ mol/kg and 640 μ mol/kg, inclusive of the endpoints; palmitic acid at a concentration of between 0.7 μ mol/kg and 70 μ mol/kg, inclusive of the endpoints; linoleic acid at a

concentration of between 0.75 $\mu\text{mol/kg}$ and 75 $\mu\text{mol/kg}$, inclusive of the endpoints; oleic acid at a concentration of between 0.75 $\mu\text{mol/kg}$ and 75 $\mu\text{mol/kg}$, inclusive of the endpoints; and a sterol at a concentration of between 0.25 $\mu\text{mol/kg}$ and 25 $\mu\text{mol/kg}$, inclusive of the endpoints. In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of about 64 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.5 $\mu\text{mol/kg}$, oleic acid at a concentration of about 7.5 $\mu\text{mol/kg}$ and a sterol at a concentration of about 2.5 $\mu\text{mol/kg}$. In certain embodiments, the T-cell expansion composition comprises one or more of octanoic acid at a concentration of about 63.75 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7.27 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.57 $\mu\text{mol/kg}$, oleic acid at a concentration of about 7.56 $\mu\text{mol/kg}$ and a sterol at a concentration of about 2.61 $\mu\text{mol/kg}$. In certain embodiments, the T-cell expansion composition comprises octanoic acid at a concentration of about 63.75 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7.27 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.57 $\mu\text{mol/kg}$, oleic acid at a concentration of 7.56 $\mu\text{mol/kg}$ and a sterol at a concentration of 2.61 $\mu\text{mol/kg}$.

[0280] As used herein, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of phosphorus, an octanoic fatty acid, a palmitic fatty acid, a linoleic fatty acid and an oleic acid. In certain embodiments, the media comprises an amount of phosphorus that is 10-fold higher than may be found in, for example, Iscove's Modified Dulbecco's Medium (IMDM); available at ThermoFisher Scientific as Catalog number 12440053).

[0281] As used herein, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, Iscove's MDM, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following elements: boron, sodium, magnesium, phosphorus, potassium, and calcium. In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be

used interchangeably with a media comprising one or more of the following elements present in the corresponding average concentrations: boron at 3.7 mg/L, sodium at 3000 mg/L, magnesium at 18 mg/L, phosphorus at 29 mg/L, potassium at 15 mg/L and calcium at 4 mg/L.

[0282] As used herein, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following components: octanoic acid (CAS No. 124-07-2), nicotinamide (CAS No. 98-92-0), 2,4,7,9-tetramethyl-5-decyn-4,7-diol (TMDD) (CAS No. 126-86-3), diisopropyl adipate (DIPA) (CAS No. 6938-94-9), n-butyl-benzenesulfonamide (CAS No. 3622-84-2), 1,2-benzenedicarboxylic acid, bis(2-methylpropyl) ester (CAS No. 84-69-5), palmitic acid (CAS No. 57-10-3), linoleic acid (CAS No. 60-33-3), oleic acid (CAS No. 112-80-1), stearic acid hydrazide (CAS No. 4130-54-5), oleamide (CAS No. 3322-62-1), sterol (e.g., cholesterol) (CAS No. 57-88-5), and alkanes (e.g., nonadecane) (CAS No. 629-92-5). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following components: octanoic acid (CAS No. 124-07-2), nicotinamide (CAS No. 98-92-0), 2,4,7,9-tetramethyl-5-decyn-4,7-diol (TMDD) (CAS No. 126-86-3), diisopropyl adipate (DIPA) (CAS No. 6938-94-9), n-butyl-benzenesulfonamide (CAS No. 3622-84-2), 1,2-benzenedicarboxylic acid, bis(2-methylpropyl) ester (CAS No. 84-69-5), palmitic acid (CAS No. 57-10-3), linoleic acid (CAS No. 60-33-3), oleic acid (CAS No. 112-80-1), stearic acid hydrazide (CAS No. 4130-54-5), oleamide (CAS No. 3322-62-1), sterol (e.g., cholesterol) (CAS No. 57-88-5), alkanes (e.g., nonadecane) (CAS No. 629-92-5), and phenol red (CAS No. 143-74-8). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following components: octanoic acid (CAS No. 124-07-2), nicotinamide (CAS No. 98-92-0), 2,4,7,9-tetramethyl-5-decyn-4,7-diol (TMDD) (CAS No. 126-86-3), diisopropyl adipate (DIPA) (CAS No. 6938-94-9), n-butyl-benzenesulfonamide (CAS No. 3622-84-2), 1,2-benzenedicarboxylic acid, bis(2-methylpropyl) ester (CAS No. 84-69-5), palmitic acid (CAS No. 57-10-3), linoleic acid (CAS No. 60-33-3), oleic acid (CAS No. 112-80-1), stearic acid

hydrazide (CAS No. 4130-54-5), oleamide (CAS No. 3322-62-1), phenol red (CAS No. 143-74-8) and lanolin alcohol.

[0283] In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following ions: sodium, ammonium, potassium, magnesium, calcium, chloride, sulfate and phosphate.

[0284] As used herein, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following free amino acids: histidine, asparagine, serine, glutamate, arginine, glycine, aspartic acid, glutamic acid, threonine, alanine, proline, cysteine, lysine, tyrosine, methionine, valine, isoleucine, leucine, phenylalanine and tryptophan. In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following free amino acids in the corresponding average mole percentages: histidine (about 1%), asparagine (about 0.5%), serine (about 1.5%), glutamine (about 67%), arginine (about 1.5%), glycine (about 1.5%), aspartic acid (about 1%), glutamic acid (about 2%), threonine (about 2%), alanine (about 1%), proline (about 1.5%), cysteine (about 1.5%), lysine (about 3%), tyrosine (about 1.5%), methionine (about 1%), valine (about 3.5%), isoleucine (about 3%), leucine (about 3.5%), phenylalanine (about 1.5%) and tryptophan (about 0.5%). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of the following free amino acids in the corresponding average mole percentages: histidine (about .78%), asparagine (about 0.4%), serine (about 1.6%), glutamine (about 67.01%), arginine (about 1.67%), glycine (about 1.72%), aspartic acid (about 1.00%), glutamic acid (about 1.93%), threonine (about 2.38%), alanine (about 1.11%), proline (about 1.49%), cysteine (about 1.65%), lysine (about 2.84%), tyrosine (about 1.62%), methionine (about 0.85%), valine (about 3.45%), isoleucine

(about 3.14%), leucine (about 3.3%), phenylalanine (about 1.64%) and tryptophan (about 0.37%).

[0285] As used herein, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of human serum albumin, recombinant human insulin, human transferrin, 2-Mercaptoethanol, Iscove’s MDM, and an expansion supplement at 37°C. Alternatively, or in addition, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of phosphorus, an octanoic fatty acid, a palmitic fatty acid, a linoleic fatty acid and an oleic acid. In certain embodiments, the media comprises an amount of phosphorus that is 10-fold higher than may be found in, for example, Iscove's Modified Dulbecco's Medium ((IMDM); available at ThermoFisher Scientific as Catalog number 12440053).

[0286] In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid, palmitic acid, linoleic acid, oleic acid and a sterol (e.g. cholesterol). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of between 0.9 mg/kg to 90 mg/kg, inclusive of the endpoints; palmitic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; linoleic acid at a concentration of between 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; oleic acid at a concentration of 0.2 mg/kg to 20 mg/kg, inclusive of the endpoints; and a sterol at a concentration of about 0.1 mg/kg to 10 mg/kg, inclusive of the endpoints (wherein mg/kg = parts per million). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of about 9 mg/kg, palmitic acid at a concentration of about 2 mg/kg, linoleic acid at a concentration of about 2 mg/kg, oleic acid at a concentration of about 2 mg/kg, and a sterol at a concentration of about 1 mg/kg (wherein mg/kg = parts per million). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of 9.19 mg/kg, palmitic acid at a concentration of 1.86 mg/kg, linoleic acid at a concentration of about 2.12 mg/kg, oleic acid at a concentration of about 2.13 mg/kg, and a sterol at a concentration of about 1.01 mg/kg (wherein mg/kg = parts per

million). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of 9.19 mg/kg, palmitic acid at a concentration of 1.86 mg/kg, linoleic acid at a concentration of 2.12 mg/kg, oleic acid at a concentration of about 2.13 mg/kg, and a sterol at a concentration of 1.01 mg/kg (wherein mg/kg = parts per million). In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of between 6.4 $\mu\text{mol/kg}$ and 640 $\mu\text{mol/kg}$, inclusive of the endpoints; palmitic acid at a concentration of between 0.7 $\mu\text{mol/kg}$ and 70 $\mu\text{mol/kg}$, inclusive of the endpoints; linoleic acid at a concentration of between 0.75 $\mu\text{mol/kg}$ and 75 $\mu\text{mol/kg}$, inclusive of the endpoints; oleic acid at a concentration of between 0.75 $\mu\text{mol/kg}$ and 75 $\mu\text{mol/kg}$, inclusive of the endpoints; and a sterol at a concentration of between 0.25 $\mu\text{mol/kg}$ and 25 $\mu\text{mol/kg}$, inclusive of the endpoints. In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of about 64 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.5 $\mu\text{mol/kg}$, oleic acid at a concentration of about 7.5 $\mu\text{mol/kg}$ and a sterol at a concentration of about 2.5 $\mu\text{mol/kg}$.

[0287] In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of about 63.75 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7.27 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.57 $\mu\text{mol/kg}$, oleic acid at a concentration of about 7.56 $\mu\text{mol/kg}$ and a sterol at a concentration of about 2.61 $\mu\text{mol/kg}$. In certain embodiments, the terms “supplemented T-cell expansion composition” or “T-cell expansion composition” may be used interchangeably with a media comprising one or more of octanoic acid at a concentration of about 63.75 $\mu\text{mol/kg}$, palmitic acid at a concentration of about 7.27 $\mu\text{mol/kg}$, linoleic acid at a concentration of about 7.57 $\mu\text{mol/kg}$, oleic acid at a concentration of 7.56 $\mu\text{mol/kg}$ and a sterol at a concentration of 2.61 $\mu\text{mol/kg}$.

[0288] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the method comprises contacting a modified T cell and an inhibitor of the P13K-Akt-mTOR pathway. Modified T-cells of the disclosure, including modified stem cell-like T cells, T_{SCM} and/or T_{CM} of the disclosure, may

be incubated, cultured, grown, stored, or otherwise, combined at any step in the methods of the procedure with a growth medium comprising one or more inhibitors a component of a PI3K pathway. Exemplary inhibitors a component of a PI3K pathway include, but are not limited to, an inhibitor of GSK3 β such as TWS119 (also known as GSK 3B inhibitor XII; CAS Number 601514-19-6 having a chemical formula C₁₈H₁₄N₄O₂). Exemplary inhibitors of a component of a PI3K pathway include, but are not limited to, bb007 (BLUEBIRDBIO™). Additional Exemplary inhibitors of a component of a PI3K pathway include, but are not limited to, an allosteric Akt inhibitor VIII (also referred to as Akti-1/2 having Compound number 10196499), ATP competitive inhibitors (Orthosteric inhibitors targeting the ATP-binding pocket of the protein kinase B (Akt)), Isoquinoline-5-sulfonamides (H-8, H-89, and NL-71-101), Azepane derivatives (A series of structures derived from (-)-balanol), Aminofurazans (GSK690693), Heterocyclic rings (7-azaindole, 6-phenylpurine derivatives, pyrrolo[2,3-d]pyrimidine derivatives, CCT128930, 3-aminopyrrolidine, anilino triazole derivatives, spiroindoline derivatives, AZD5363, ipatasertib (GDC-0068, RG7440), A-674563, and A-443654), Phenylpyrazole derivatives (AT7867 and AT13148), Thiophenecarboxamide derivatives (Afuresertib (GSK2110183), 2-pyrimidyl-5-amidothiophene derivative (DC120), uprosertib (GSK2141795)), Allosteric inhibitors (Superior to orthosteric inhibitors providing greater specificity, reduced side-effects and less toxicity), 2,3-diphenylquinoxaline analogues (2,3-diphenylquinoxaline derivatives, triazolo[3,4-f][1,6]naphthyridin-3(2H)-one derivative (MK-2206)), Alkylphospholipids (Edelfosine (1-O-octadecyl-2-O-methyl-rac-glycero-3-phosphocholine, ET-18-OCH₃) ilmofosine (BM 41.440), miltefosine (hexadecylphosphocholine, HePC), perifosine (D-21266), erucylphosphocholine (ErPC), erufosine (ErPC3, erucylphosphohomocholine), Indole-3-carbinol analogues (Indole-3-carbinol, 3-chloroacetylindole, diindolylmethane, diethyl 6-methoxy-5,7-dihydroindolo [2,3-b]carbazole-2,10-dicarboxylate (SR13668), OSU-A9), Sulfonamide derivatives (PH-316 and PHT-427), Thiourea derivatives (PIT-1, PIT-2, DM-PIT-1, N-[(1-methyl-1H-pyrazol-4-yl)carbonyl]-N'-(3-bromophenyl)-thiourea), Purine derivatives (Triciribine (TCN, NSC 154020), triciribine mono-phosphate active analogue (TCN-P), 4-amino-pyrido[2,3-d]pyrimidine derivative API-1, 3-phenyl-3H-imidazo[4,5-b]pyridine derivatives, ARQ 092), BAY 1125976, 3-methyl-xanthine, quinoline-4-carboxamide and 2-[4-(cyclohexa-1,3-dien-1-yl)-1H-pyrazol-3-yl]phenol, 3-oxo-tirucallic acid, 3 α - and 3 β -acetoxy-tirucallic acids, acetoxy-tirucallic acid, and irreversible inhibitors (antibiotics, Lactoquinomycin, Frenolicin B, kalafungin, medermycin, Boc-Phe-vinyl ketone,

4-hydroxynonenal (4-HNE), 1,6-naphthyridinone derivatives, and imidazo-1,2-pyridine derivatives).

[0289] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the method comprises contacting a modified T cell and an inhibitor of T cell effector differentiation. Exemplary inhibitors of T cell effector differentiation include, but are not limited to, a BET inhibitor (e.g. JQ1, a hienotriazolodiazepine) and/or an inhibitor of the BET family of proteins (e.g. BRD2, BRD3, BRD4, and BRDT).

[0290] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the method comprises contacting a modified T cell and an agent that reduces nucleocytoplasmic Acetyl-CoA. Exemplary agents that reduce nucleocytoplasmic Acetyl-CoA include, but are not limited to, 2-hydroxy-citrate (2-HC) as well as agents that increase expression of *Acsc1*.

[0291] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the method comprises contacting a modified T cell and a composition comprising a histone deacetylase (HDAC) inhibitor. In some embodiments, the composition comprising an HDAC inhibitor comprises or consists of valproic acid, Sodium Phenylbutyrate (NaPB) or a combination thereof. In some embodiments, the composition comprising an HDAC inhibitor comprises or consists of valproic acid. In some embodiments, the composition comprising an HDAC inhibitor comprises or consists of Sodium Phenylbutyrate (NaPB).

[0292] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the activation supplement may comprise one or more cytokine(s). The one or more cytokine(s) may comprise any cytokine, including but not limited to, lymphokines. Exemplary lymphokines include, but are not limited to, interleukin-2 (IL-2), interleukin-3 (IL-3), interleukin-4 (IL-4), interleukin-5 (IL-5), interleukin-6 (IL-6), interleukin-7 (IL-7), interleukin-15 (IL-15), interleukin-21 (IL-21), granulocyte-macrophage colony-stimulating factor (GM-CSF) and interferon-gamma (INF γ). The one or more cytokine(s) may comprise IL-2.

[0293] In certain embodiments of the methods of producing a modified T cell (e.g. a stem cell-like T cell, a T_{SCM} and/or a T_{CM}) of the disclosure, the activation supplement may comprise one or more activator complexes. Exemplary and nonlimiting activator complexes may comprise a monomeric, dimeric, trimeric or tetrameric antibody complex that binds one

or more of CD3, CD28, and CD2. In some embodiments, the activation supplement comprises or consists of an activator complex that comprises a human, a humanized or a recombinant or a chimeric antibody. In some embodiments, the activation supplement comprises or consists of an activator complex that binds CD3 and CD28. In some embodiments, the activation supplement comprises or consists of an activator complex that binds CD3, CD28 and CD2.

Natural Killer (NK) cells

[0294] In certain embodiments, the modified immune or immune precursor cells of the disclosure are natural killer (NK) cells. In certain embodiments, NK cells are cytotoxic lymphocytes that differentiate from lymphoid progenitor cells.

[0295] Modified NK cells of the disclosure may be derived from modified hematopoietic stem and progenitor cells (HSPCs) or modified HSCs.

[0296] In certain embodiments, non-activated NK cells are derived from CD3-depleted leukopheresis (containing CD14/CD19/CD56+ cells).

[0297] In certain embodiments, NK cells are electroporated using a Lonza 4D nucleofector or BTX ECM 830 (500V, 700 usec pulse length, 0.2 mm electrode gap, one pulse). All Lonza 4D nucleofector programs are contemplated as within the scope of the methods of the disclosure.

[0298] In certain embodiments, 5×10^6 cells were electroporated per electroporation in 100 μ L P3 buffer in cuvettes. However, this ratio of cells per volume is scalable for commercial manufacturing methods.

[0299] In certain embodiments, NK cells were stimulated by co-culture with an additional cell line. In certain embodiments, the additional cell line comprises artificial antigen presenting cells (aAPCs). In certain embodiments, stimulation occurs at day 1, 2, 3, 4, 5, 6, or 7 following electroporation. In certain embodiments, stimulation occurs at day 2 following electroporation.

[0300] In certain embodiments, NK cells express CD56.

B cells

[0301] In certain embodiments, the modified immune or immune precursor cells of the disclosure are B cells. B cells are a type of lymphocyte that express B cell receptors on the cell surface. B cell receptors bind to specific antigens.

[0302] Modified B cells of the disclosure may be derived from modified hematopoietic stem and progenitor cells (HSPCs) or modified HSCs.

[0303] In certain embodiments, HSPCs are modified using the methods of the disclosure, and then primed for B cell differentiation in presence of human IL-3, Flt3L, TPO, SCF, and G-CSF for at least 3 days, at least 4 days, at least 5 days, at least 6 days or at least 7 days. In certain embodiments, HSPCs are modified using the methods of the disclosure, and then primed for B cell differentiation in presence of human IL-3, Flt3L, TPO, SCF, and G-CSF for 5 days.

[0304] In certain embodiments, following priming, modified HSPC cells are transferred to a layer of feeder cells and fed bi-weekly, along with transfer to a fresh layer of feeders once per week. In certain embodiments, the feeder cells are MS-5 feeder cells.

[0305] In certain embodiments, modified HSPC cells are cultured with MS-5 feeder cells for at least 7, 14, 21, 28, 30, 33, 35, 42 or 48 days. In certain embodiments, modified HSPC cells are cultured with MS-5 feeder cells for 33 days.

Methods of Cell Modification

[0306] In some embodiments of the methods of the disclosure, a composition comprises a scalable ratio of 250×10^6 primary human T cells per milliliter of buffer or other media during a delivery or an introduction step.

[0307] In some embodiments of the methods of the disclosure, a composition is delivered or introduced to a cell by electroporation or nucleofection. In some embodiments, a delivery or introduction step comprises electroporation or nucleofection.

[0308] In some embodiments of the methods of the disclosure, a composition is delivered or introduced to a cell by a method other than electroporation or nucleofection.

[0309] In some embodiments of the methods of the disclosure, a composition is delivered or introduced by one or more of topical delivery, adsorption, absorption, electroporation, spinfection, co-culture, transfection, mechanical delivery, sonic delivery, vibrational delivery, magnetofection or by nanoparticle-mediated delivery. In some embodiments, a delivery or introduction step comprises one or more of topical delivery, adsorption, absorption, electroporation, spinfection, co-culture, transfection, mechanical delivery, sonic delivery, vibrational delivery, magnetofection or by nanoparticle-mediated delivery.

[0310] In some embodiments of the methods of the disclosure, a composition is delivered or introduced by liposomal transfection, calcium phosphate transfection, fugene transfection, and dendrimer-mediated transfection. In some embodiments, a delivery or introduction step comprises one or more of liposomal transfection, calcium phosphate transfection, fugene transfection, and dendrimer-mediated transfection.

[0311] In some embodiments of the methods of the disclosure, a composition is delivered or introduced by mechanical transfection comprises cell squeezing, cell bombardment, or gene gun techniques. In some embodiments, a delivery or introduction step comprises one or more of mechanical transfection comprises cell squeezing, cell bombardment, or gene gun techniques.

[0312] In some embodiments of the methods of the disclosure, a composition is delivered or introduced by nanoparticle-mediated transfection comprises liposomal delivery, delivery by micelles, and delivery by polymerosomes. In some embodiments, a delivery or introduction step comprises one or more of liposomal delivery, delivery by micelles, and delivery by polymerosomes.

Non-transposition Methods of Delivery

[0313] In some embodiments of the compositions and methods of the disclosure, a modified cell of the disclosure may be produced by introducing a sequence into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure. The introducing step may comprise delivery of a sequence and/or a gene editing composition via a non-transposition delivery system. The introduction step may be performed *ex vivo*, *in vivo*, *in vitro* or *in situ*.

[0314] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprises one or more of topical delivery, adsorption, absorption, electroporation, spin-fecton, co-culture, transfection, mechanical delivery, sonic delivery, vibrational delivery, magnetofection and nanoparticle-mediated delivery.

[0315] In some embodiments of the compositions and methods of the disclosure, introducing a nucleic acid sequence and/or a gene editing construct into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprises liposomal transfection, calcium phosphate transfection, fugene transfection, and dendrimer-mediated transfection.

[0316] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure by mechanical transfection comprises cell squeezing, cell bombardment, or gene gun techniques.

[0317] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure by nanoparticle-mediated transfection comprises one or more of a liposome, a micelle, a polymer and a polymerosome.

[0318] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprises a non-viral vector. In some embodiments, the non-viral vector comprises the sequence and/or the gene editing composition. In some embodiments, the non-viral vector comprises plasmid DNA, linear double-stranded DNA (dsDNA), linear single-stranded DNA (ssDNA), DoggyBone™ DNA, nanoplasids, minicircle DNA, single-stranded oligodeoxynucleotides (ssODN), DDNA oligonucleotides, single-stranded mRNA (ssRNA), and double-stranded mRNA (dsRNA).

[0319] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprises a viral vector. In some embodiments, the viral vector is a non-integrating and/or non-chromosomal vector. Exemplary non-integrating non-chromosomal vectors include, but are not limited to, adeno-associated virus (AAV), adenovirus, and herpes viruses. In some embodiments, the viral vector is an integrating chromosomal vector. Integrating chromosomal vectors include, but are not limited to, adeno-associated vectors (AAV), Lentiviruses, and gamma-retroviruses. In some embodiments, the viral vector comprises the sequence and/or the gene editing composition.

[0320] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprises a combination of vectors of the disclosure. Exemplary, non-limiting vector combinations include: viral and non-viral vectors, a plurality of non-viral vectors, or a plurality of viral vectors. Exemplary but non-limiting vectors combinations include: a combination of a DNA-derived and an RNA-derived vector, a combination of non-viral expression vector and a viral delivery vector, a combination of a non-viral expression vector and a nanoparticle delivery vector, a combination of two distinct non-viral expression vectors, a combination of a non-viral expression vector and a mechanical or chemical method of transfection.

[0321] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure stably integrates a sequence, transiently integrates a sequence, produces site-specific integration of a sequence, or produces a biased integration of a sequence. In some embodiments, the sequence is a nucleic acid sequence. In some embodiments, the nucleic acid sequence comprises a transgene.

[0322] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure stably integrates a sequence. In some embodiments, the sequence is a nucleic acid sequence. In some embodiments, the stable chromosomal integration can be a random integration, a site-specific integration, or a biased integration. In some embodiments, the site-specific integration can be non-assisted or assisted. In some embodiments, the assisted site-specific integration is co-delivered with a site-directed nuclease. In some embodiments, the site-directed nuclease comprises a transgene with 5' and 3' nucleotide sequence extensions that contain a percentage homology to upstream and downstream regions of the site of genomic integration. In some embodiments, the transgene with homologous nucleotide extensions enable genomic integration by homologous recombination, microhomology-mediated end joining, or nonhomologous end-joining. In some embodiments the site-specific integration occurs at a safe harbor site. Genomic safe harbor sites are able to accommodate the integration of new genetic material in a manner that ensures that the newly inserted genetic elements function reliably (for example, are expressed at a therapeutically effective level of expression) and do not cause deleterious alterations to the host genome that cause a risk to the host organism. Potential genomic safe harbors include, but are not limited to, intronic sequences of the human albumin gene, the adeno-associated virus site 1 (AAVS1), a naturally occurring site of integration of AAV virus on chromosome 19, the site of the chemokine (C-C motif) receptor 5 (CCR5) gene and the site of the human ortholog of the mouse Rosa26 locus.

[0323] In some embodiments, the site-specific transgene integration occurs at a site that disrupts expression of a target gene. In some embodiments, disruption of target gene expression occurs by site-specific integration at introns, exons, promoters, genetic elements, enhancers, suppressors, start codons, stop codons, and response elements. In some embodiments, exemplary target genes targeted by site-specific integration include but are not

limited to TRAC, TRAB, PDI, any immunosuppressive gene, and genes involved in allo-rejection.

[0324] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure site-specific transgene integration occurs at a site that results in enhanced expression of a target gene. In some embodiments, enhancement of target gene expression occurs by site-specific integration at introns, exons, promoters, genetic elements, enhancers, suppressors, start codons, stop codons, and response elements.

[0325] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure enzymes may be used to create strand breaks in the host genome to facilitate delivery or integration of the transgene. In some embodiments, enzymes create single-strand breaks. In some embodiments, enzymes create double-strand breaks. In some embodiments, examples of break-inducing enzymes include but are not limited to: transposases, integrases, endonucleases, CRISPR-Cas9, transcription activator-like effector nucleases (TALEN), zinc finger nucleases (ZFN), Cas-CLOVER™, and CPF1. In some embodiments, break-inducing enzymes can be delivered to the cell encoded in DNA, encoded in mRNA, as a protein, as a nucleoprotein complex with a guide RNA (gRNA).

[0326] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure site-specific transgene integration is controlled by a vector-mediated integration site bias. In some embodiments vector-mediated integration site bias is controlled by the chosen lentiviral vector. In some embodiments vector-mediated integration site bias is controlled by the chosen gamma-retroviral vector.

[0327] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure site-specific transgene integration site is a non-stable chromosomal insertion. In some embodiments, the integrated transgene may become silenced, removed, excised, or further modified.

[0328] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure genome modification comprises a non-stable integration of a transgene. In some embodiments, the non-stable integration can be a transient non-chromosomal integration, a semi-stable non chromosomal integration, a semi-persistent non-chromosomal insertion, or a non-stable chromosomal insertion. In some embodiments, the transient non-chromosomal insertion can be epi-chromosomal or cytoplasmic.

[0329] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure the transient non-chromosomal insertion of a transgene does not integrate into a chromosome and the modified genetic material is not replicated during cell division.

[0330] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure genome modification comprises a semi-stable or persistent non-chromosomal integration of a transgene. In some embodiments, a DNA vector encodes a Scaffold/matrix attachment region (S-MAR) module that binds to nuclear matrix proteins for episomal retention of a non-viral vector allowing for autonomous replication in the nucleus of dividing cells.

[0331] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure genome modification is a non-stable chromosomal integration of a transgene. In some embodiments, the integrated transgene may become silenced, removed, excised, or further modified.

[0332] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure modification to the genome by transgene insertion can occur via host cell-directed double-strand breakage repair (homology-directed repair) by homologous recombination (HR), microhomology-mediated end joining (MMEJ), nonhomologous end joining (NHEJ), transposase enzyme-mediated modification, integrase enzyme-mediated modification, endonuclease enzyme-mediated modification, or recombinant enzyme-mediated modification. In some embodiments, the modification to the

genome by transgene insertion can occur via CRISPR-Cas9, TALEN, ZFNs, Cas-CLOVER, and cpfl.

Nanoparticle Delivery

[0333] In some embodiments of the compositions and methods of the disclosure, introducing a sequence and/or a gene editing composition into an HSC, an HSC descendent cell, an immune cell or an immune precursor cell of the disclosure comprise a nanoparticle vector. Nanoparticle vectors may encapsulate a composition of the disclosure. Alternatively, or in addition, a surface of a nanoparticle vector may comprise a composition of the disclosure. In some embodiments, the surface is an interior surface. In some embodiments, the surface is an exterior surface. In some embodiments, the surface comprises a composition of the disclosure integrated therein or thereon.

[0334] Nonlimiting examples of nanoparticle vectors of the disclosure may comprise one or more of a hydrophilic block, a hydrophobic block, and a charged block. In some embodiments, the hydrophilic block may be poly(ethylene oxide) (PEO), and the charged block may be poly(L-histidine).

[0335] The disclosure provides nanoparticle vectors comprising di-block and tri-block co-polymers. Exemplary di-block co-polymers may comprise one or more of a hydrophilic block, a hydrophobic block, and a charged block. In some embodiments, the hydrophilic block may be poly(ethylene oxide) (PEO), and the charged block may be poly(L-histidine). Exemplary tri-block co-polymers may comprise one or more of a hydrophilic block, a hydrophobic block, and a charged block. In some embodiments, the hydrophilic block may be poly(ethylene oxide) (PEO), and the charged block may be poly(L-histidine).

[0336] An exemplary tri-block copolymer that may be used in various embodiments is a PEO-b-PLA-b-PHIS, with variable numbers of repeating units in each block varying by design.

[0337] Poly(histidine) (i.e., poly(L-histidine)), is a pH-sensitive polymer due to the imidazole ring providing an electron lone pair on the unsaturated nitrogen. That is, poly(histidine) has amphoteric properties through protonation-deprotonation. The various embodiments enable intracellular delivery of compositions of the disclosure, including gene editing compositions, by, for example, complexing with poly(histidine)-based micelles.

[0338] Diblock copolymers that may be used as intermediates for making triblock copolymers of the embodiment micelles may have hydrophilic biocompatible poly(ethylene oxide) (PEO), which is chemically synonymous with PEG, coupled to various hydrophobic

aliphatic poly(anhydrides), poly(nucleic acids), poly(esters), poly(ortho esters), poly(peptides), poly(phosphazenes) and poly(saccharides), including but not limited by poly(lactide) (PLA), poly(glycolide) (PLGA), poly(lactic-co-glycolic acid) (PLGA), poly(ϵ -caprolactone) (PCL), and poly(trimethylene carbonate) (PTMC).

[0339] Polymeric micelles comprised of 100% PEGylated surfaces possess improved in vitro chemical stability, augmented in vivo bioavailability, and prolonged blood circulatory half-lives. For example, aliphatic polyesters, constituting the polymeric micelle's membrane portions, are degraded by hydrolysis of their ester linkages in physiological conditions such as in the human body. Because of their biodegradable nature, aliphatic polyesters have received a great deal of attention for use as implantable biomaterials in drug delivery devices, bioresorbable sutures, adhesion barriers, and as scaffolds for injury repair via tissue engineering.

[0340] Without wishing to be bound by a particular theory, it is believed that in the micelles that are formed by the various embodiment triblock copolymers, the hydrophobic blocks aggregate to form a core, leaving the hydrophilic blocks and poly(histidine) blocks on the ends to form one or more surrounding layer.

Scaffold Proteins

[0341] Protein scaffolds of the disclosure may be derived from a fibronectin type III (FN3) repeat protein, encoding or complementary nucleic acids, vectors, host cells, compositions, combinations, formulations, devices, and methods of making and using them. In a preferred embodiment, the protein scaffold is comprised of a consensus sequence of multiple FN3 domains from human Tenascin-C (hereinafter "Tenascin"). In a further preferred embodiment, the protein scaffold of the present disclosure is a consensus sequence of 15 FN3 domains. The protein scaffolds of the disclosure can be designed to bind various molecules, for example, a cellular target protein. In a preferred embodiment, the protein scaffolds of the disclosure can be designed to bind an epitope of a wild type and/or variant form of a ligand or an antigen.

[0342] Protein scaffolds of the disclosure may include additional molecules or moieties, for example, the Fc region of an antibody, albumin binding domain, or other moiety influencing half-life. In further embodiments, the protein scaffolds of the disclosure may be bound to a nucleic acid molecule that may encode the protein scaffold.

[0343] The disclosure provides at least one method for expressing at least one protein scaffold based on a consensus sequence of multiple FN3 domains, in a host cell, comprising

culturing a host cell as described herein under conditions wherein at least one protein scaffold is expressed in detectable and/or recoverable amounts.

[0344] The disclosure provides at least one composition comprising (a) a protein scaffold based on a consensus sequence of multiple FN3 domains and/or encoding nucleic acid as described herein; and (b) a suitable and/or pharmaceutically acceptable carrier or diluent.

[0345] The disclosure provides a method of generating libraries of a protein scaffold based on a fibronectin type III (FN3) repeat protein, preferably, a consensus sequence of multiple FN3 domains and, more preferably, a consensus sequence of multiple FN3 domains from human Tenascin. The library is formed by making successive generations of scaffolds by altering (by mutation) the amino acids or the number of amino acids in the molecules in particular positions in portions of the scaffold, e.g., loop regions. Libraries can be generated by altering the amino acid composition of a single loop or the simultaneous alteration of multiple loops or additional positions of the scaffold molecule. The loops that are altered can be lengthened or shortened accordingly. Such libraries can be generated to include all possible amino acids at each position, or a designed subset of amino acids. The library members can be used for screening by display, such as in vitro or CIS display (DNA, RNA, ribosome display, etc.), yeast, bacterial, and phage display.

[0346] Protein scaffolds of the disclosure provide enhanced biophysical properties, such as stability under reducing conditions and solubility at high concentrations; they may be expressed and folded in prokaryotic systems, such as *E. coli*, in eukaryotic systems, such as yeast, and in in vitro transcription/translation systems, such as the rabbit reticulocyte lysate system.

[0347] The disclosure provides a method of generating a scaffold molecule that binds to a particular target by panning the scaffold library of the invention with the target and detecting binders. In other related aspects, the disclosure comprises screening methods that may be used to generate or affinity mature protein scaffolds with the desired activity, e.g., capable of binding to target proteins with a certain affinity. Affinity maturation can be accomplished by iterative rounds of mutagenesis and selection using systems, such as phage display or in vitro display. Mutagenesis during this process may be the result of site directed mutagenesis to specific scaffold residues, random mutagenesis due to error-prone PCR, DNA shuffling, and/or a combination of these techniques.

[0348] The disclosure provides an isolated, recombinant and/or synthetic protein scaffold based on a consensus sequence of fibronectin type III (FN3) repeat protein, including,

without limitation, mammalian-derived scaffold, as well as compositions and encoding nucleic acid molecules comprising at least one polynucleotide encoding protein scaffold based on the consensus FN3 sequence. The disclosure further includes, but is not limited to, methods of making and using such nucleic acids and protein scaffolds, including diagnostic and therapeutic compositions, methods and devices.

[0349] The protein scaffolds of the disclosure offer advantages over conventional therapeutics, such as ability to administer locally, orally, or cross the blood-brain barrier, ability to express in E. Coli allowing for increased expression of protein as a function of resources versus mammalian cell expression ability to be engineered into bispecific or tandem molecules that bind to multiple targets or multiple epitopes of the same target, ability to be conjugated to drugs, polymers, and probes, ability to be formulated to high concentrations, and the ability of such molecules to effectively penetrate diseased tissues and tumors.

[0350] Moreover, the protein scaffolds possess many of the properties of antibodies in relation to their fold that mimics the variable region of an antibody. This orientation enables the FN3 loops to be exposed similar to antibody complementarity determining regions (CDRs). They should be able to bind to cellular targets and the loops can be altered, e.g., affinity matured, to improve certain binding or related properties.

[0351] Three of the six loops of the protein scaffold of the disclosure correspond topologically to the complementarity determining regions (CDRs 1-3), i.e., antigen-binding regions, of an antibody, while the remaining three loops are surface exposed in a manner similar to antibody CDRs. These loops span at or about residues 13-16, 22-28, 38-43, 51-54, 60-64, and 75-81 of the consensus sequence. Preferably, the loop regions at or about residues 22-28, 51-54, and 75-81 are altered for binding specificity and affinity. One or more of these loop regions are randomized with other loop regions and/or other strands maintaining their sequence as backbone portions to populate a library and potent binders can be selected from the library having high affinity for a particular protein target. One or more of the loop regions can interact with a target protein similar to an antibody CDR interaction with the protein.

Discovery of Antigen/Ligand Recognition Region Sequences

[0352] The disclosure provides a method of generating libraries of antigen/ligand recognition region (ARR/LRR) sequences for binding antigens and/or ligands of the disclosure. The library is formed by making successive generations of ARR/LRR sequences by altering (by mutation) the amino acids or the number of amino acids in the sequences at

particular positions of the ARR/LRR. In some embodiments, the ARR/LRR comprises one or more of a protein scaffolded, an antibody mimetic, a Centyrin, a single chain antibody (scFv), a single domain antibody, a VHH and a VH of the disclosure. In some embodiments, the library is formed by making successive generations of ARR/LRR sequences by altering (by mutation) the amino acids or the number of amino acids in the sequences at particular positions of an antibody, an ScFv, VHH or VH, e.g., one or more complementarity determining regions (CDR) and/or framework regions of a variable domain.

[0353] Libraries can be generated by altering the amino acid composition of a single CDR or the simultaneous alteration of multiple CDRs or additional positions of an antibody, an scFv, VHH or VH (e.g. a framework sequence of the variable region). The CDR and/or framework sequence of the variable domain that are altered can be lengthened or shortened accordingly.

[0354] Libraries can be generated by altering the amino acid composition of a loop of a scaffold protein or a Centyrin. The loop sequences that are altered can be lengthened or shortened accordingly.

[0355] Libraries can be generated by altering the amino acid composition of an antigen or ligand-binding or specificity-determining region of an antibody mimetic.

[0356] Such libraries can be generated to include all possible amino acids at each position, or a designed subset of amino acids. The library members can be used for screening by display, such as in vitro or CIS display (DNA, RNA, ribosome display, etc.), yeast, bacterial, and phage display.

[0357] ARRs/LRRs of the disclosure provide enhanced biophysical properties, such as stability under reducing conditions and solubility at high concentrations; they may be expressed and folded in prokaryotic systems, such as *E. coli*, in eukaryotic systems, such as yeast, and in in vitro transcription/translation systems, such as the rabbit reticulocyte lysate system.

[0358] The disclosure provides a method of generating an ARR/LRR or a portion thereof that binds to a particular target by panning a library of the invention with the target and detecting binders. In other related aspects, the disclosure comprises screening methods that may be used to generate or affinity mature ARRs/LRRs with the desired activity, e.g., capable of binding to target proteins with a certain affinity. Affinity maturation can be accomplished by iterative rounds of mutagenesis and selection using systems, such as phage display or in vitro display. Mutagenesis during this process may be the result of site directed

mutagenesis to specific protein residues, random mutagenesis due to error-prone PCR, DNA shuffling, and/or a combination of these techniques.

[0359] The disclosure provides an isolated, recombinant and/or synthetic protein scaffold comprising at least one VHH. The disclosure further includes, but is not limited to, methods of making and using such nucleic acids and protein scaffolds, including diagnostic and therapeutic compositions, methods and devices.

[0360] The compositions of the disclosure offer advantages over conventional therapeutics, such as ability to administer locally, orally, or cross the blood-brain barrier, ability to express in *E. Coli* allowing for increased expression of protein as a function of resources versus mammalian cell expression ability to be engineered into bispecific or tandem molecules that bind to multiple targets or multiple epitopes of the same target, ability to be conjugated to drugs, polymers, and probes, ability to be formulated to high concentrations, and the ability of such molecules to effectively penetrate diseased tissues and tumors.

Production and Generation of Proteins

[0361] Proteins of the disclosure can be optionally produced by a cell line, a mixed cell line, an immortalized cell or clonal population of immortalized cells, as well known in the art. See, e.g., Ausubel, et al., ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., NY, N.Y. (1987-2001); Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor, N.Y. (1989); Harlow and Lane, *Antibodies, a Laboratory Manual*, Cold Spring Harbor, N.Y. (1989); Colligan, et al., eds., *Current Protocols in Immunology*, John Wiley & Sons, Inc., NY (1994-2001); Colligan et al., *Current Protocols in Protein Science*, John Wiley & Sons, NY, N.Y., (1997-2001).

[0362] Amino acids encoding a protein can be altered, added and/or deleted to reduce immunogenicity or reduce, enhance or modify binding, affinity, on-rate, off-rate, avidity, specificity, half-life, stability, solubility or any other suitable characteristic, as known in the art.

[0363] Proteins can be engineered with retention of high affinity for an antigen or a ligand as well as other favorable biological properties. To achieve this goal, the proteins can be optionally prepared by a process of analysis of the parental sequences and various conceptual engineered products using three-dimensional models of the parental and engineered sequences. Three-dimensional models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate sequences and can

measure possible immunogenicity (e.g., Immunofilter program of Xencor, Inc. of Monrovia, Calif.). Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate sequence, i.e., the analysis of residues that influence the ability of the protein to bind its antigen. In this way, residues can be selected and combined from the parent and reference sequences so that the desired characteristic, such as affinity for the target antigen(s)/ligand(s), is achieved. Alternatively, or in addition to, the above procedures, other suitable methods of engineering can be used.

Screening of ARR/LRRs

[0364] Screening protein ARR/LRRs or any portion thereof for specific binding to similar proteins or fragments can be conveniently achieved using nucleotide (DNA or RNA display) or peptide display libraries, for example, in vitro display. This method involves the screening of large collections of peptides for individual members having the desired function or structure. The displayed nucleotide or peptide sequences can be from 3 to 5000 or more nucleotides or amino acids in length, frequently from 5-100 amino acids long, and often from about 8 to 25 amino acids long. In addition to direct chemical synthetic methods for generating peptide libraries, several recombinant DNA methods have been described. One type involves the display of a peptide sequence on the surface of a bacteriophage or cell. Each bacteriophage or cell contains the nucleotide sequence encoding the particular displayed peptide sequence. Such methods are described in PCT Patent Publication Nos. 91/17271, 91/18980, 91/19818, and 93/08278.

[0365] Other systems for generating libraries of peptides have aspects of both in vitro chemical synthesis and recombinant methods. See, PCT Patent Publication Nos. 92/05258, 92/14843, and 96/19256. See also, U.S. Pat. Nos. 5,658,754; and 5,643,768. Peptide display libraries, vector, and screening kits are commercially available from such suppliers as Invitrogen (Carlsbad, Calif.), and Cambridge Antibody Technologies (Cambridgeshire, UK). See, e.g., U.S. Pat. Nos. 4,704,692, 4,939,666, 4,946,778, 5,260,203, 5,455,030, 5,518,889, 5,534,621, 5,656,730, 5,763,733, 5,767,260, 5,856,456, assigned to Enzon; 5,223,409, 5,403,484, 5,571,698, 5,837,500, assigned to Dyax; 5,427,908, 5,580,717, assigned to Affymax; 5,885,793, assigned to Cambridge Antibody Technologies; 5,750,373, assigned to Genentech; 5,618,920, 5,595,898, 5,576,195, 5,698,435, 5,693,493, 5,698,417, assigned to Xoma, Colligan, supra; Ausubel, supra; or Sambrook, supra.

[0366] The ARR/LRRs of the disclosure comprising one or more of a protein scaffold, an antibody, an ScFv, a Centyrin, a single domain antibody, a VHH or a VH of the disclosure

can bind human or other mammalian proteins with a wide range of affinities (KD). In a preferred embodiment, at least one ARR/LRR can optionally bind to a target protein with high affinity, for example, with a KD equal to or less than about 10^{-7} M, such as but not limited to, 0.1-9.9 (or any range or value therein) $\times 10^{-8}$, 10^{-9} , 10^{-10} , 10^{-11} , 10^{-12} , 10^{-13} , 10^{-14} , 10^{-15} or any range or value therein, as determined by surface plasmon resonance or the Kinexa method, as practiced by those of skill in the art. In a preferred embodiment, at least one protein scaffold, antibody, ScFv, Centyrin, single domain antibody, VHH or VH of the disclosure can optionally bind to a target protein with high affinity, for example, with a KD equal to or less than about 10^{-7} M, such as but not limited to, 0.1-9.9 (or any range or value therein) $\times 10^{-8}$, 10^{-9} , 10^{-10} , 10^{-11} , 10^{-12} , 10^{-13} , 10^{-14} , 10^{-15} or any range or value therein, as determined by surface plasmon resonance or the Kinexa method, as practiced by those of skill in the art.

[0367] The affinity or avidity of a protein scaffold, an antibody, an ScFv, a Centyrin, a single domain antibody, a VHH or a VH of the disclosure for an antigen/ligand can be determined experimentally using any suitable method. (See, for example, Berzofsky, et al., "Antibody-Antigen Interactions," In Fundamental Immunology, Paul, W. E., Ed., Raven Press: New York, N.Y. (1984); Kuby, Janis Immunology, W.H. Freeman and Company: New York, N.Y. (1992); and methods described herein). The measured affinity of a particular protein-antigen/ligand interaction can vary if measured under different conditions (e.g., salt concentration, pH). Thus, measurements of affinity and other antigen-binding parameters (e.g., KD, K_{on} , K_{off}) are preferably made with standardized solutions of protein scaffold (e.g. VHH) and antigen, and a standardized buffer, such as the buffer described herein.

[0368] Competitive assays can be performed with the protein scaffold, antibody, ScFv, Centyrin, single domain antibody, VHH or VH of the disclosure in order to determine what proteins, antibodies, and other antagonists compete for binding to a target protein and/or share the epitope region. These assays as readily known to those of ordinary skill in the art evaluate competition between antagonists or ligands for a limited number of binding sites on a protein. The protein and/or antibody is immobilized or insolubilized before or after the competition and the sample bound to the target protein is separated from the unbound sample, for example, by decanting (where the protein/antibody was preinsolubilized) or by centrifuging (where the protein/antibody was precipitated after the competitive reaction). Also, the competitive binding may be determined by whether function is altered by the binding or lack of binding of the protein scaffold, antibody, ScFv, Centyrin, single domain

antibody, VHH or VH to the target protein, e.g., whether protein scaffold, antibody, ScFv, Centyrin, single domain antibody, VHH or VH inhibits or potentiates the enzymatic activity of, for example, a label. ELISA and other functional assays may be used, as well known in the art.

Therapeutic Proteins

[0369] In certain embodiments of the disclosure, T cells are modified to express therapeutic proteins, including secreted human proteins. These secreted proteins may be used as a monotherapy or in combination with another therapy in the treatment or prevention of any disease or disorder. These secreted proteins may be used as a monotherapy or in combination with another therapy for enzyme replacement and/or administration of biologic therapeutics.

A database of human secreted proteins can be found at

proteinatlas.org/search/protein__class:Predicted%20secreted%20proteins, the contents of which are incorporated herein by reference. Exemplary human therapeutic proteins can be found, but are not limited to the human proteins in Table 1.

[0370] **Table 1.** Exemplary therapeutic proteins (and proteins to enhance CAR-T efficacy). Compositions of the disclosure may comprise a promoter of one or more of the proteins of Table 1 driving expression of any sequence of the disclosure.

Gene Name	Gene Description	Protein SEQ ID NO
A1BG	Alpha-1-B glycoprotein	SEQ ID NOS: 1 - 2
A2M	Alpha-2-macroglobulin	SEQ ID NOS: 3 - 6
A2ML1	Alpha-2-macroglobulin-like 1	SEQ ID NOS: 7 - 12
A4GNT	Alpha-1,4-N-acetylglucosaminyltransferase	SEQ ID NO: 13
AADACL2	Arylacetamide deacetylase-like 2	SEQ ID NOS: 14 - 15
AANAT	Aralkylamine N-acetyltransferase	SEQ ID NOS: 16 - 19
ABCG1	ATP-binding cassette, sub-family G (WHITE), member 1	SEQ ID NOS: 20 - 26
ABHD1	Abhydrolase domain containing 1	SEQ ID NOS: 27 - 31
ABHD10	Abhydrolase domain containing 10	SEQ ID NOS: 32 - 35
ABHD14A	Abhydrolase domain containing 14A	SEQ ID NOS: 36 - 40
ABHD15	Abhydrolase domain containing 15	SEQ ID NO: 41
ABI3BP	ABI family, member 3 (NESH) binding protein	SEQ ID NOS: 42 - 63
FAM175A	Family with sequence similarity 175, member A	SEQ ID NOS: 64 - 71
LA16c-380H5.3		SEQ ID NO: 72
AC008641.1		SEQ ID NO: 73
CTB-60B18.6		SEQ ID NOS: 74 - 75
AC009133.22		SEQ ID NO: 76
AC009491.2		SEQ ID NO: 77

RP11-977G19.10		SEQ ID NOS: 78 - 80
CTD-2370N5.3		SEQ ID NOS: 81 - 84
RP11-196G11.1		SEQ ID NOS: 85 - 87
AC136352.5		SEQ ID NO: 88
RP11-812E19.9		SEQ ID NO: 89
AC145212.4	MaFF-interacting protein	SEQ ID NO: 90
AC233755.1		SEQ ID NO: 91
AC011513.3		SEQ ID NOS: 92 - 93
ACACB	Acetyl-CoA carboxylase beta	SEQ ID NOS: 94 - 100
ACAN	Aggrecan	SEQ ID NOS: 101 - 108
ACE	Angiotensin I converting enzyme	SEQ ID NOS: 109 - 121
ACHE	Acetylcholinesterase (Yt blood group)	SEQ ID NOS: 122 - 134
ACP2	Acid phosphatase 2, lysosomal	SEQ ID NOS: 135 - 142
ACP5	Acid phosphatase 5, tartrate resistant	SEQ ID NOS: 143 - 151
ACP6	Acid phosphatase 6, lysophosphatidic	SEQ ID NOS: 152 - 158
PAPL	Iron/zinc purple acid phosphatase-like protein	SEQ ID NOS: 159 - 162
ACPP	Acid phosphatase, prostate	SEQ ID NOS: 163 - 167
ACR	Acrosin	SEQ ID NOS: 168 - 169
ACRBP	Acrosin binding protein	SEQ ID NOS: 170 - 174
ACRV1	Acrosomal vesicle protein 1	SEQ ID NOS: 175 - 178
ACSF2	Acyl-CoA synthetase family member 2	SEQ ID NOS: 179 - 187
ACTL10	Actin-like 10	SEQ ID NO: 188
ACVR1	Activin A receptor, type I	SEQ ID NOS: 189 - 197
ACVR1C	Activin A receptor, type IC	SEQ ID NOS: 198 - 201
ACVRL1	Activin A receptor type II-like 1	SEQ ID NOS: 202 - 207
ACYP1	Acylphosphatase 1, erythrocyte (common) type	SEQ ID NOS: 208 - 213
ACYP2	Acylphosphatase 2, muscle type	SEQ ID NOS: 214 - 221
CECR1	Cat eye syndrome chromosome region, candidate 1	SEQ ID NOS: 222 - 229
ADAM10	ADAM metallopeptidase domain 10	SEQ ID NOS: 230 - 237
ADAM12	ADAM metallopeptidase domain 12	SEQ ID NOS: 238 - 240
ADAM15	ADAM metallopeptidase domain 15	SEQ ID NOS: 241 - 252
ADAM17	ADAM metallopeptidase domain 17	SEQ ID NOS: 253 - 255
ADAM18	ADAM metallopeptidase domain 18	SEQ ID NOS: 256 - 260
ADAM22	ADAM metallopeptidase domain 22	SEQ ID NOS: 261 - 269
ADAM28	ADAM metallopeptidase domain 28	SEQ ID NOS: 270 - 275
ADAM29	ADAM metallopeptidase domain 29	SEQ ID NOS: 276 - 284
ADAM32	ADAM metallopeptidase domain 32	SEQ ID NOS: 285 - 291
ADAM33	ADAM metallopeptidase domain 33	SEQ ID NOS: 292 - 296
ADAM7	ADAM metallopeptidase domain 7	SEQ ID NOS: 297 - 300
ADAM8	ADAM metallopeptidase domain 8	SEQ ID NOS: 301 - 305
ADAM9	ADAM metallopeptidase domain 9	SEQ ID NOS: 306 - 311
ADAMDEC1	ADAM-like, decysin 1	SEQ ID NOS: 312 - 314

ADAMTS1	ADAM metalloproteinase with thrombospondin type 1 motif, 1	SEQ ID NOS: 315 - 318
ADAMTS10	ADAM metalloproteinase with thrombospondin type 1 motif, 10	SEQ ID NOS: 319 - 324
ADAMTS12	ADAM metalloproteinase with thrombospondin type 1 motif, 12	SEQ ID NOS: 325 - 327
ADAMTS13	ADAM metalloproteinase with thrombospondin type 1 motif, 13	SEQ ID NOS: 328 - 335
ADAMTS14	ADAM metalloproteinase with thrombospondin type 1 motif, 14	SEQ ID NOS: 336 - 337
ADAMTS15	ADAM metalloproteinase with thrombospondin type 1 motif, 15	SEQ ID NO: 338
ADAMTS16	ADAM metalloproteinase with thrombospondin type 1 motif, 16	SEQ ID NOS: 339 - 340
ADAMTS17	ADAM metalloproteinase with thrombospondin type 1 motif, 17	SEQ ID NOS: 341 - 344
ADAMTS18	ADAM metalloproteinase with thrombospondin type 1 motif, 18	SEQ ID NOS: 345 - 348
ADAMTS19	ADAM metalloproteinase with thrombospondin type 1 motif, 19	SEQ ID NOS: 349 - 352
ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2	SEQ ID NOS: 353 - 355
ADAMTS20	ADAM metalloproteinase with thrombospondin type 1 motif, 20	SEQ ID NOS: 356 - 359
ADAMTS3	ADAM metalloproteinase with thrombospondin type 1 motif, 3	SEQ ID NOS: 360 - 361
ADAMTS5	ADAM metalloproteinase with thrombospondin type 1 motif, 5	SEQ ID NO: 362
ADAMTS6	ADAM metalloproteinase with thrombospondin type 1 motif, 6	SEQ ID NOS: 363 - 364
ADAMTS7	ADAM metalloproteinase with thrombospondin type 1 motif, 7	SEQ ID NO: 365
ADAMTS8	ADAM metalloproteinase with thrombospondin type 1 motif, 8	SEQ ID NO: 366
ADAMTS9	ADAM metalloproteinase with thrombospondin type 1 motif, 9	SEQ ID NOS: 367 - 371
ADAMTSL1	ADAMTS-like 1	SEQ ID NOS: 372 - 382
ADAMTSL2	ADAMTS-like 2	SEQ ID NOS: 383 - 385
ADAMTSL3	ADAMTS-like 3	SEQ ID NOS: 386 - 387
ADAMTSL4	ADAMTS-like 4	SEQ ID NOS: 388 - 391
ADAMTSL5	ADAMTS-like 5	SEQ ID NOS: 392 - 397
ADCK1	AarF domain containing kinase 1	SEQ ID NOS: 398 - 402
ADCYAP1	Adenylate cyclase activating polypeptide 1 (pituitary)	SEQ ID NOS: 403 - 404
ADCYAP1R1	Adenylate cyclase activating polypeptide 1 (pituitary) receptor type I	SEQ ID NOS: 405 - 411
ADGRA3	Adhesion G protein-coupled receptor A3	SEQ ID NOS: 412 - 416
ADGRB2	Adhesion G protein-coupled receptor B2	SEQ ID NOS: 417 - 425
ADGRD1	Adhesion G protein-coupled receptor D1	SEQ ID NOS: 426 - 431
ADGRE3	Adhesion G protein-coupled receptor E3	SEQ ID NOS: 432 - 436

ADGRE5	Adhesion G protein-coupled receptor E5	SEQ ID NOS: 437 - 442
ADGRF1	Adhesion G protein-coupled receptor F1	SEQ ID NOS: 443 - 447
ADGRG1	Adhesion G protein-coupled receptor G1	SEQ ID NOS: 448 - 512
ADGRG5	Adhesion G protein-coupled receptor G5	SEQ ID NOS: 513 - 515
ADGRG6	Adhesion G protein-coupled receptor G6	SEQ ID NOS: 516 - 523
ADGRV1	Adhesion G protein-coupled receptor V1	SEQ ID NOS: 524 - 540
ADI1	Acireductone dioxygenase 1	SEQ ID NOS: 541 - 543
ADIG	Adipogenin	SEQ ID NOS: 544 - 547
ADIPOQ	Adiponectin, C1Q and collagen domain containing	SEQ ID NOS: 548 - 549
ADM	Adrenomedullin	SEQ ID NOS: 550 - 557
ADM2	Adrenomedullin 2	SEQ ID NOS: 558 - 559
ADM5	Adrenomedullin 5 (putative)	SEQ ID NO: 560
ADPGK	ADP-dependent glucokinase	SEQ ID NOS: 561 - 570
ADPRHL2	ADP-ribosylhydrolase like 2	SEQ ID NO: 571
AEBP1	AE binding protein 1	SEQ ID NOS: 572 - 579
LACE1	Lactation elevated 1	SEQ ID NOS: 580 - 583
AFM	Afamin	SEQ ID NO: 584
AFP	Alpha-fetoprotein	SEQ ID NOS: 585 - 586
AGA	Aspartylglucosaminidase	SEQ ID NOS: 587 - 589
AGER	Advanced glycosylation end product-specific receptor	SEQ ID NOS: 590 - 600
AGK	Acylglycerol kinase	SEQ ID NOS: 601 - 606
AGPS	Alkylglycerone phosphate synthase	SEQ ID NOS: 607 - 610
AGR2	Anterior gradient 2, protein disulphide isomerase family member	SEQ ID NOS: 611 - 614
AGR3	Anterior gradient 3, protein disulphide isomerase family member	SEQ ID NOS: 615 - 617
AGRN	Agrin	SEQ ID NOS: 618 - 621
AGRP	Agouti related neuropeptide	SEQ ID NO: 622
AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	SEQ ID NO: 623
AGTPBP1	ATP/GTP binding protein 1	SEQ ID NOS: 624 - 627
AGTRAP	Angiotensin II receptor-associated protein	SEQ ID NOS: 628 - 635
AHCYL2	Adenosylhomocysteinase-like 2	SEQ ID NOS: 636 - 642
AHSG	Alpha-2-HS-glycoprotein	SEQ ID NOS: 643 - 644
AIG1	Androgen-induced 1	SEQ ID NOS: 645 - 653
AK4	Adenylate kinase 4	SEQ ID NOS: 654 - 657
AKAP10	A kinase (PRKA) anchor protein 10	SEQ ID NOS: 658 - 666
AKR1C1	Aldo-keto reductase family 1, member C1	SEQ ID NOS: 667 - 669
RP4-576H24.4		SEQ ID NOS: 670 - 672
SERPINA3	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	SEQ ID NO: 673
RP11-14J7.7		SEQ ID NOS: 674 - 675
RP11-903H12.5		SEQ ID NO: 676
AL356289.1		SEQ ID NO: 677

AL589743.1		SEQ ID NO: 678
XXbac-BPG116M5.17		SEQ ID NOS: 679 - 680
XXbac-BPG181M17.5		SEQ ID NO: 681
XXbac-BPG32J3.20		SEQ ID NO: 682
RP11-350O14.18		SEQ ID NO: 683
ALAS2	5'-aminolevulinate synthase 2	SEQ ID NOS: 684 - 691
ALB	Albumin	SEQ ID NOS: 692 - 701
ALDH9A1	Aldehyde dehydrogenase 9 family, member A1	SEQ ID NO: 702
ALDOA	Aldolase A, fructose-bisphosphate	SEQ ID NOS: 703 - 717
ALG1	ALG1, chitobiosyldiphosphodolichol beta-mannosyltransferase	SEQ ID NOS: 718 - 723
ALG5	ALG5, dolichyl-phosphate beta-glucosyltransferase	SEQ ID NOS: 724 - 725
ALG9	ALG9, alpha-1,2-mannosyltransferase	SEQ ID NOS: 726 - 736
FAM150A	Family with sequence similarity 150, member A	SEQ ID NOS: 737 - 738
FAM150B	Family with sequence similarity 150, member B	SEQ ID NOS: 739 - 745
ALKBH1	AlkB homolog 1, histone H2A dioxygenase	SEQ ID NOS: 746 - 748
ALKBH5	AlkB homolog 5, RNA demethylase	SEQ ID NOS: 749 - 750
ALPI	Alkaline phosphatase, intestinal	SEQ ID NOS: 751 - 752
ALPL	Alkaline phosphatase, liver/bone/kidney	SEQ ID NOS: 753 - 757
ALPP	Alkaline phosphatase, placental	SEQ ID NO: 758
ALPPL2	Alkaline phosphatase, placental-like 2	SEQ ID NO: 759
AMBN	Ameloblastin (enamel matrix protein)	SEQ ID NOS: 760 - 762
AMBP	Alpha-1-microglobulin/bikunin precursor	SEQ ID NOS: 763 - 765
AMELX	Amelogenin, X-linked	SEQ ID NOS: 766 - 768
AMELY	Amelogenin, Y-linked	SEQ ID NOS: 769 - 770
AMH	Anti-Mullerian hormone	SEQ ID NO: 771
AMPD1	Adenosine monophosphate deaminase 1	SEQ ID NOS: 772 - 774
AMTN	Amelotin	SEQ ID NOS: 775 - 776
AMY1A	Amylase, alpha 1A (salivary)	SEQ ID NOS: 777 - 779
AMY1B	Amylase, alpha 1B (salivary)	SEQ ID NOS: 780 - 783
AMY1C	Amylase, alpha 1C (salivary)	SEQ ID NO: 784
AMY2A	Amylase, alpha 2A (pancreatic)	SEQ ID NOS: 785 - 787
AMY2B	Amylase, alpha 2B (pancreatic)	SEQ ID NOS: 788 - 792
ANG	Angiogenin, ribonuclease, RNase A family, 5	SEQ ID NOS: 793 - 794
ANGEL1	Angel homolog 1 (Drosophila)	SEQ ID NOS: 795 - 798
ANGPT1	Angiopoietin 1	SEQ ID NOS: 799 - 803
ANGPT2	Angiopoietin 2	SEQ ID NOS: 804 - 807
ANGPT4	Angiopoietin 4	SEQ ID NO: 808
ANGPTL1	Angiopoietin-like 1	SEQ ID NOS: 809 - 811
ANGPTL2	Angiopoietin-like 2	SEQ ID NOS: 812 - 813

ANGPTL3	Angiopoietin-like 3	SEQ ID NO: 814
ANGPTL4	Angiopoietin-like 4	SEQ ID NOS: 815 - 822
ANGPTL5	Angiopoietin-like 5	SEQ ID NOS: 823 - 824
ANGPTL6	Angiopoietin-like 6	SEQ ID NOS: 825 - 827
ANGPTL7	Angiopoietin-like 7	SEQ ID NO: 828
C19orf80	Chromosome 19 open reading frame 80	SEQ ID NOS: 829 - 832
ANK1	Ankyrin 1, erythrocytic	SEQ ID NOS: 833 - 843
ANKDD1A	Ankyrin repeat and death domain containing 1A	SEQ ID NOS: 844 - 850
ANKRD54	Ankyrin repeat domain 54	SEQ ID NOS: 851 - 859
ANKRD60	Ankyrin repeat domain 60	SEQ ID NO: 860
ANO7	Anoctamin 7	SEQ ID NOS: 861 - 864
ANOS1	Anosmin 1	SEQ ID NO: 865
ANTXR1	Anthrax toxin receptor 1	SEQ ID NOS: 866 - 869
AOAH	Acyloxyacyl hydrolase (neutrophil)	SEQ ID NOS: 870 - 874
AOC1	Amine oxidase, copper containing 1	SEQ ID NOS: 875 - 880
AOC2	Amine oxidase, copper containing 2 (retina-specific)	SEQ ID NOS: 881 - 882
AOC3	Amine oxidase, copper containing 3	SEQ ID NOS: 883 - 889
AP000721.4		SEQ ID NO: 890
APBB1	Amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	SEQ ID NOS: 891 - 907
APCDD1	Adenomatosis polyposis coli down-regulated 1	SEQ ID NOS: 908 - 913
APCS	Amyloid P component, serum	SEQ ID NO: 914
APELA	Apelin receptor early endogenous ligand	SEQ ID NOS: 915 - 917
APLN	Apelin	SEQ ID NO: 918
APLP2	Amyloid beta (A4) precursor-like protein 2	SEQ ID NOS: 919 - 928
APOA1	Apolipoprotein A-I	SEQ ID NOS: 929 - 933
APOA2	Apolipoprotein A-II	SEQ ID NOS: 934 - 942
APOA4	Apolipoprotein A-IV	SEQ ID NO: 943
APOA5	Apolipoprotein A-V	SEQ ID NOS: 944 - 946
APOB	Apolipoprotein B	SEQ ID NOS: 947 - 948
APOC1	Apolipoprotein C-I	SEQ ID NOS: 949 - 957
APOC2	Apolipoprotein C-II	SEQ ID NOS: 958 - 962
APOC3	Apolipoprotein C-III	SEQ ID NOS: 963 - 966
APOC4	Apolipoprotein C-IV	SEQ ID NOS: 967 - 968
APOC4-APOC2	APOC4-APOC2 readthrough (NMD candidate)	SEQ ID NOS: 969 - 970
APOD	Apolipoprotein D	SEQ ID NOS: 971 - 974
APOE	Apolipoprotein E	SEQ ID NOS: 975 - 978
APOF	Apolipoprotein F	SEQ ID NO: 979
APOH	Apolipoprotein H (beta-2-glycoprotein I)	SEQ ID NOS: 980 - 983
APOL1	Apolipoprotein L, 1	SEQ ID NOS: 984 - 994
APOL3	Apolipoprotein L, 3	SEQ ID NOS: 995 - 1009
APOM	Apolipoprotein M	SEQ ID NOS: 1010 - 1012
APOOL	Apolipoprotein O-like	SEQ ID NOS: 1013 - 1015
ARCNI	Archain 1	SEQ ID NOS: 1016 - 1020

ARFIP2	ADP-ribosylation factor interacting protein 2	SEQ ID NOS: 1021 - 1027
ARHGAP36	Rho GTPase activating protein 36	SEQ ID NOS: 1028 - 1033
HMHA1	Histocompatibility (minor) HA-1	SEQ ID NOS: 1034 - 1042
ARHGAP6	Rho GTPase activating protein 6	SEQ ID NOS: 1043 - 1048
ARHGEF4	Rho guanine nucleotide exchange factor (GEF) 4	SEQ ID NOS: 1049 - 1059
ARL16	ADP-ribosylation factor-like 16	SEQ ID NOS: 1060 - 1068
ARMC5	Armadillo repeat containing 5	SEQ ID NOS: 1069 - 1075
ARNTL	Aryl hydrocarbon receptor nuclear translocator-like	SEQ ID NOS: 1076 - 1090
ARSA	Arylsulfatase A	SEQ ID NOS: 1091 - 1096
ARSB	Arylsulfatase B	SEQ ID NOS: 1097 - 1100
ARSE	Arylsulfatase E (chondrodysplasia punctata 1)	SEQ ID NOS: 1101 - 1104
ARSG	Arylsulfatase G	SEQ ID NOS: 1105 - 1108
ARSI	Arylsulfatase family, member I	SEQ ID NOS: 1109 - 1111
ARSK	Arylsulfatase family, member K	SEQ ID NOS: 1112 - 1116
ART3	ADP-ribosyltransferase 3	SEQ ID NOS: 1117 - 1124
ART4	ADP-ribosyltransferase 4 (Dombrock blood group)	SEQ ID NOS: 1125 - 1128
ART5	ADP-ribosyltransferase 5	SEQ ID NOS: 1129 - 1133
ARTN	Artemin	SEQ ID NOS: 1134 - 1144
ASAH1	N-acylsphingosine amidohydrolase (acid ceramidase) 1	SEQ ID NOS: 1145 - 1195
ASAH2	N-acylsphingosine amidohydrolase (non-lysosomal ceramidase) 2	SEQ ID NOS: 1196 - 1201
ASCL1	Achaete-scute family bHLH transcription factor 1	SEQ ID NO: 1202
ASIP	Agouti signaling protein	SEQ ID NOS: 1203 - 1204
ASPN	Asporin	SEQ ID NOS: 1205 - 1206
ASTL	Astacin-like metallo-endopeptidase (M12 family)	SEQ ID NO: 1207
ATAD5	ATPase family, AAA domain containing 5	SEQ ID NOS: 1208 - 1209
ATAT1	Alpha tubulin acetyltransferase 1	SEQ ID NOS: 1210 - 1215
ATG2A	Autophagy related 2A	SEQ ID NOS: 1216 - 1218
ATG5	Autophagy related 5	SEQ ID NOS: 1219 - 1227
ATMIN	ATM interactor	SEQ ID NOS: 1228 - 1231
ATP13A1	ATPase type 13A1	SEQ ID NOS: 1232 - 1234
ATP5F1	ATP synthase, H ⁺ transporting, mitochondrial Fo complex, subunit B1	SEQ ID NOS: 1235 - 1236
ATP6AP1	ATPase, H ⁺ transporting, lysosomal accessory protein 1	SEQ ID NOS: 1237 - 1244
ATP6AP2	ATPase, H ⁺ transporting, lysosomal accessory protein 2	SEQ ID NOS: 1245 - 1267
ATPAF1	ATP synthase mitochondrial F1 complex assembly factor 1	SEQ ID NOS: 1268 - 1278
AUH	AU RNA binding protein/enoyl-CoA hydratase	SEQ ID NOS: 1279 - 1280
AVP	Arginine vasopressin	SEQ ID NO: 1281
AXIN2	Axin 2	SEQ ID NOS: 1282 - 1289
AZGP1	Alpha-2-glycoprotein 1, zinc-binding	SEQ ID NOS: 1290 - 1292
AZU1	Azurocidin 1	SEQ ID NOS: 1293 - 1294
B2M	Beta-2-microglobulin	SEQ ID NOS: 1295 - 1301

B3GALNT1	Beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	SEQ ID NOS: 1302 - 1314
B3GALNT2	Beta-1,3-N-acetylgalactosaminyltransferase 2	SEQ ID NOS: 1315 - 1317
B3GALT1	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 1	SEQ ID NO: 1318
B3GALT4	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 4	SEQ ID NO: 1319
B3GALT5	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 5	SEQ ID NOS: 1320 - 1324
B3GALT6	UDP-Gal:betaGal beta 1,3-galactosyltransferase polypeptide 6	SEQ ID NO: 1325
B3GAT3	Beta-1,3-glucuronyltransferase 3	SEQ ID NOS: 1326 - 1330
B3GLCT	Beta 3-glucosyltransferase	SEQ ID NO: 1331
B3GNT3	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3	SEQ ID NOS: 1332 - 1335
B3GNT4	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 4	SEQ ID NOS: 1336 - 1339
B3GNT6	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 6	SEQ ID NOS: 1340 - 1341
B3GNT7	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7	SEQ ID NO: 1342
B3GNT8	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 8	SEQ ID NO: 1343
B3GNT9	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9	SEQ ID NO: 1344
B4GALNT1	Beta-1,4-N-acetyl-galactosaminyl transferase 1	SEQ ID NOS: 1345 - 1356
B4GALNT3	Beta-1,4-N-acetyl-galactosaminyl transferase 3	SEQ ID NOS: 1357 - 1358
B4GALNT4	Beta-1,4-N-acetyl-galactosaminyl transferase 4	SEQ ID NOS: 1359 - 1361
B4GALT4	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 4	SEQ ID NOS: 1362 - 1375
B4GALT5	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5	SEQ ID NO: 1376
B4GALT6	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 6	SEQ ID NOS: 1377 - 1380
B4GAT1	Beta-1,4-glucuronyltransferase 1	SEQ ID NO: 1381
B9D1	B9 protein domain 1	SEQ ID NOS: 1382 - 1398
BACE2	Beta-site APP-cleaving enzyme 2	SEQ ID NOS: 1399 - 1401
BAGE5	B melanoma antigen family, member 5	SEQ ID NO: 1402
BCAM	Basal cell adhesion molecule (Lutheran blood group)	SEQ ID NOS: 1403 - 1406
BCAN	Brevican	SEQ ID NOS: 1407 - 1413
BCAP29	B-cell receptor-associated protein 29	SEQ ID NOS: 1414 - 1426
BCAR1	Breast cancer anti-estrogen resistance 1	SEQ ID NOS: 1427 - 1444
BCHE	Butyrylcholinesterase	SEQ ID NOS: 1445 - 1449
BCKDHB	Branched chain keto acid dehydrogenase E1, beta polypeptide	SEQ ID NOS: 1450 - 1452
BDNF	Brain-derived neurotrophic factor	SEQ ID NOS: 1453 - 1470
BGLAP	Bone gamma-carboxy glutamate (gla) protein	SEQ ID NO: 1471
BGN	Biglycan	SEQ ID NOS: 1472 - 1473

BLVRB	Biliverdin reductase B	SEQ ID NOS: 1474 - 1478
BMP1	Bone morphogenetic protein 1	SEQ ID NOS: 1479 - 1490
BMP10	Bone morphogenetic protein 10	SEQ ID NO: 1491
BMP15	Bone morphogenetic protein 15	SEQ ID NO: 1492
BMP2	Bone morphogenetic protein 2	SEQ ID NO: 1493
BMP3	Bone morphogenetic protein 3	SEQ ID NO: 1494
BMP4	Bone morphogenetic protein 4	SEQ ID NOS: 1495 - 1502
BMP6	Bone morphogenetic protein 6	SEQ ID NO: 1503
BMP7	Bone morphogenetic protein 7	SEQ ID NOS: 1504 - 1507
BMP8A	Bone morphogenetic protein 8a	SEQ ID NO: 1508
BMP8B	Bone morphogenetic protein 8b	SEQ ID NO: 1509
BMPER	BMP binding endothelial regulator	SEQ ID NOS: 1510 - 1513
BNC1	Basonuclin 1	SEQ ID NOS: 1514 - 1515
BOC	BOC cell adhesion associated, oncogene regulated	SEQ ID NOS: 1516 - 1526
BOD1	Biorientation of chromosomes in cell division 1	SEQ ID NOS: 1527 - 1531
BOLA1	BolA family member 1	SEQ ID NOS: 1532 - 1534
BPI	Bactericidal/permeability-increasing protein	SEQ ID NOS: 1535 - 1538
BPIFA1	BPI fold containing family A, member 1	SEQ ID NOS: 1539 - 1542
BPIFA2	BPI fold containing family A, member 2	SEQ ID NOS: 1543 - 1544
BPIFA3	BPI fold containing family A, member 3	SEQ ID NOS: 1545 - 1546
BPIFB1	BPI fold containing family B, member 1	SEQ ID NOS: 1547 - 1548
BPIFB2	BPI fold containing family B, member 2	SEQ ID NO: 1549
BPIFB3	BPI fold containing family B, member 3	SEQ ID NO: 1550
BPIFB4	BPI fold containing family B, member 4	SEQ ID NOS: 1551 - 1552
BPIFB6	BPI fold containing family B, member 6	SEQ ID NOS: 1553 - 1554
BPIFC	BPI fold containing family C	SEQ ID NOS: 1555 - 1558
BRF1	BRF1, RNA polymerase III transcription initiation factor 90 kDa subunit	SEQ ID NOS: 1559 - 1574
BRINP1	Bone morphogenetic protein/retinoic acid inducible neural-specific 1	SEQ ID NOS: 1575 - 1576
BRINP2	Bone morphogenetic protein/retinoic acid inducible neural-specific 2	SEQ ID NO: 1577
BRINP3	Bone morphogenetic protein/retinoic acid inducible neural-specific 3	SEQ ID NOS: 1578 - 1580
BSG	Basigin (Ok blood group)	SEQ ID NOS: 1581 - 1591
BSPH1	Binder of sperm protein homolog 1	SEQ ID NO: 1592
BST1	Bone marrow stromal cell antigen 1	SEQ ID NOS: 1593 - 1597
BTBD17	BTB (POZ) domain containing 17	SEQ ID NO: 1598
BTD	Biotinidase	SEQ ID NOS: 1599 - 1608
BTN2A2	Butyrophilin, subfamily 2, member A2	SEQ ID NOS: 1609 - 1622
BTN3A1	Butyrophilin, subfamily 3, member A1	SEQ ID NOS: 1623 - 1629
BTN3A2	Butyrophilin, subfamily 3, member A2	SEQ ID NOS: 1630 - 1640
BTN3A3	Butyrophilin, subfamily 3, member A3	SEQ ID NOS: 1641 - 1649
RP4-608O15.3	Complement factor H-related protein 2	SEQ ID NO: 1650
C10orf99	Chromosome 10 open reading frame 99	SEQ ID NO: 1651

C11orf1	Chromosome 11 open reading frame 1	SEQ ID NOS: 1652 - 1656
C11orf24	Chromosome 11 open reading frame 24	SEQ ID NOS: 1657 - 1659
C11orf45	Chromosome 11 open reading frame 45	SEQ ID NOS: 1660 - 1661
C11orf94	Chromosome 11 open reading frame 94	SEQ ID NO: 1662
C12orf10	Chromosome 12 open reading frame 10	SEQ ID NOS: 1663 - 1666
C12orf49	Chromosome 12 open reading frame 49	SEQ ID NOS: 1667 - 1670
C12orf73	Chromosome 12 open reading frame 73	SEQ ID NOS: 1671 - 1680
C12orf76	Chromosome 12 open reading frame 76	SEQ ID NOS: 1681 - 1688
C14orf93	Chromosome 14 open reading frame 93	SEQ ID NOS: 1689 - 1704
C16orf89	Chromosome 16 open reading frame 89	SEQ ID NOS: 1705 - 1707
C16orf90	Chromosome 16 open reading frame 90	SEQ ID NOS: 1708 - 1709
C17orf67	Chromosome 17 open reading frame 67	SEQ ID NO: 1710
C17orf75	Chromosome 17 open reading frame 75	SEQ ID NOS: 1711 - 1719
C17orf99	Chromosome 17 open reading frame 99	SEQ ID NOS: 1720 - 1722
C18orf54	Chromosome 18 open reading frame 54	SEQ ID NOS: 1723 - 1727
C19orf47	Chromosome 19 open reading frame 47	SEQ ID NOS: 1728 - 1735
C19orf70	Chromosome 19 open reading frame 70	SEQ ID NOS: 1736 - 1739
C1GALT1	Core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1	SEQ ID NOS: 1740 - 1744
C1orf127	Chromosome 1 open reading frame 127	SEQ ID NOS: 1745 - 1748
C1orf159	Chromosome 1 open reading frame 159	SEQ ID NOS: 1749 - 1761
C1orf198	Chromosome 1 open reading frame 198	SEQ ID NOS: 1762 - 1766
C1orf54	Chromosome 1 open reading frame 54	SEQ ID NOS: 1767 - 1769
C1orf56	Chromosome 1 open reading frame 56	SEQ ID NO: 1770
C1QA	Complement component 1, q subcomponent, A chain	SEQ ID NOS: 1771 - 1773
C1QB	Complement component 1, q subcomponent, B chain	SEQ ID NOS: 1774 - 1777
C1QC	Complement component 1, q subcomponent, C chain	SEQ ID NOS: 1778 - 1780
C1QL1	Complement component 1, q subcomponent-like 1	SEQ ID NO: 1781
C1QL2	Complement component 1, q subcomponent-like 2	SEQ ID NO: 1782
C1QL3	Complement component 1, q subcomponent-like 3	SEQ ID NOS: 1783 - 1784
C1QL4	Complement component 1, q subcomponent-like 4	SEQ ID NO: 1785
C1QTNF1	C1q and tumor necrosis factor related protein 1	SEQ ID NOS: 1786 - 1795
FAM132A	Family with sequence similarity 132, member A	SEQ ID NO: 1796
C1QTNF2	C1q and tumor necrosis factor related protein 2	SEQ ID NO: 1797
C1QTNF3	C1q and tumor necrosis factor related protein 3	SEQ ID NOS: 1798 - 1799
C1QTNF4	C1q and tumor necrosis factor related protein 4	SEQ ID NOS: 1800 - 1801
C1QTNF5	C1q and tumor necrosis factor related protein 5	SEQ ID NOS: 1802 - 1804
C1QTNF7	C1q and tumor necrosis factor related protein 7	SEQ ID NOS: 1805 - 1809
C1QTNF8	C1q and tumor necrosis factor related protein 8	SEQ ID NOS: 1810 - 1811
C1QTNF9	C1q and tumor necrosis factor related protein 9	SEQ ID NOS: 1812 - 1813
C1QTNF9B	C1q and tumor necrosis factor related protein 9B	SEQ ID NOS: 1814 - 1816
C1R	Complement component 1, r subcomponent	SEQ ID NOS: 1817 - 1825
C1RL	Complement component 1, r subcomponent-like	SEQ ID NOS: 1826 - 1834

C1S	Complement component 1, s subcomponent	SEQ ID NOS: 1835 - 1844
C2	Complement component 2	SEQ ID NOS: 1845 - 1859
C21orf33	Chromosome 21 open reading frame 33	SEQ ID NOS: 1860 - 1868
C21orf62	Chromosome 21 open reading frame 62	SEQ ID NOS: 1869 - 1872
C22orf15	Chromosome 22 open reading frame 15	SEQ ID NOS: 1873 - 1875
C22orf46	Chromosome 22 open reading frame 46	SEQ ID NO: 1876
C2CD2	C2 calcium-dependent domain containing 2	SEQ ID NOS: 1877 - 1879
C2orf40	Chromosome 2 open reading frame 40	SEQ ID NOS: 1880 - 1882
C2orf66	Chromosome 2 open reading frame 66	SEQ ID NO: 1883
C2orf69	Chromosome 2 open reading frame 69	SEQ ID NO: 1884
C2orf78	Chromosome 2 open reading frame 78	SEQ ID NO: 1885
C3	Complement component 3	SEQ ID NOS: 1886 - 1890
C3orf33	Chromosome 3 open reading frame 33	SEQ ID NOS: 1891 - 1895
C3orf58	Chromosome 3 open reading frame 58	SEQ ID NOS: 1896 - 1899
C4A	Complement component 4A (Rodgers blood group)	SEQ ID NOS: 1900 - 1901
C4B	Complement component 4B (Chido blood group)	SEQ ID NOS: 1902 - 1903
C4BPA	Complement component 4 binding protein, alpha	SEQ ID NOS: 1904 - 1906
C4BPB	Complement component 4 binding protein, beta	SEQ ID NOS: 1907 - 1911
C4orf48	Chromosome 4 open reading frame 48	SEQ ID NOS: 1912 - 1913
C5	Complement component 5	SEQ ID NO: 1914
C5orf46	Chromosome 5 open reading frame 46	SEQ ID NOS: 1915 - 1916
C6	Complement component 6	SEQ ID NOS: 1917 - 1920
C6orf120	Chromosome 6 open reading frame 120	SEQ ID NO: 1921
C6orf15	Chromosome 6 open reading frame 15	SEQ ID NO: 1922
C6orf58	Chromosome 6 open reading frame 58	SEQ ID NO: 1923
C7	Complement component 7	SEQ ID NO: 1924
C7orf57	Chromosome 7 open reading frame 57	SEQ ID NOS: 1925 - 1929
C8A	Complement component 8, alpha polypeptide	SEQ ID NO: 1930
C8B	Complement component 8, beta polypeptide	SEQ ID NOS: 1931 - 1933
C8G	Complement component 8, gamma polypeptide	SEQ ID NOS: 1934 - 1935
C9	Complement component 9	SEQ ID NO: 1936
C9orf47	Chromosome 9 open reading frame 47	SEQ ID NOS: 1937 - 1939
CA10	Carbonic anhydrase X	SEQ ID NOS: 1940 - 1946
CA11	Carbonic anhydrase XI	SEQ ID NOS: 1947 - 1948
CA6	Carbonic anhydrase VI	SEQ ID NOS: 1949 - 1953
CA9	Carbonic anhydrase IX	SEQ ID NOS: 1954 - 1955
CABLES1	Cdk5 and Abl enzyme substrate 1	SEQ ID NOS: 1956 - 1961
CABP1	Calcium binding protein 1	SEQ ID NOS: 1962 - 1965
CACNA2D1	Calcium channel, voltage-dependent, alpha 2/delta subunit 1	SEQ ID NOS: 1966 - 1969
CACNA2D4	Calcium channel, voltage-dependent, alpha 2/delta subunit 4	SEQ ID NOS: 1970 - 1983
CADM3	Cell adhesion molecule 3	SEQ ID NOS: 1984 - 1986
CALCA	Calcitonin-related polypeptide alpha	SEQ ID NOS: 1987 - 1991
CALCB	Calcitonin-related polypeptide beta	SEQ ID NOS: 1992 - 1994

CALCR	Calcitonin receptor	SEQ ID NOS: 1995 - 2001
CALCRL	Calcitonin receptor-like	SEQ ID NOS: 2002 - 2006
FAM26D	Family with sequence similarity 26, member D	SEQ ID NOS: 2007 - 2011
CALR	Calreticulin	SEQ ID NOS: 2012 - 2015
CALR3	Calreticulin 3	SEQ ID NOS: 2016 - 2017
CALU	Calumenin	SEQ ID NOS: 2018 - 2023
CAMK2D	Calcium/calmodulin-dependent protein kinase II delta	SEQ ID NOS: 2024 - 2035
CAMP	Cathelicidin antimicrobial peptide	SEQ ID NO: 2036
CANX	Calnexin	SEQ ID NOS: 2037 - 2051
CARM1	Coactivator-associated arginine methyltransferase 1	SEQ ID NOS: 2052 - 2059
CARNS1	Carnosine synthase 1	SEQ ID NOS: 2060 - 2062
CARTPT	CART prepropeptide	SEQ ID NO: 2063
CASQ1	Calsequestrin 1 (fast-twitch, skeletal muscle)	SEQ ID NOS: 2064 - 2065
CASQ2	Calsequestrin 2 (cardiac muscle)	SEQ ID NO: 2066
CATSPERG	Catsper channel auxiliary subunit gamma	SEQ ID NOS: 2067 - 2074
CBLN1	Cerebellin 1 precursor	SEQ ID NOS: 2075 - 2077
CBLN2	Cerebellin 2 precursor	SEQ ID NOS: 2078 - 2081
CBLN3	Cerebellin 3 precursor	SEQ ID NOS: 2082 - 2083
CBLN4	Cerebellin 4 precursor	SEQ ID NO: 2084
CCBE1	Collagen and calcium binding EGF domains 1	SEQ ID NOS: 2085 - 2087
CCDC112	Coiled-coil domain containing 112	SEQ ID NOS: 2088 - 2091
CCDC129	Coiled-coil domain containing 129	SEQ ID NOS: 2092 - 2099
CCDC134	Coiled-coil domain containing 134	SEQ ID NOS: 2100 - 2101
CCDC149	Coiled-coil domain containing 149	SEQ ID NOS: 2102 - 2105
CCDC3	Coiled-coil domain containing 3	SEQ ID NOS: 2106 - 2107
CCDC80	Coiled-coil domain containing 80	SEQ ID NOS: 2108 - 2111
CCDC85A	Coiled-coil domain containing 85A	SEQ ID NO: 2112
CCDC88B	Coiled-coil domain containing 88B	SEQ ID NOS: 2113 - 2115
CCER2	Coiled-coil glutamate-rich protein 2	SEQ ID NOS: 2116 - 2117
CCK	Cholecystokinin	SEQ ID NOS: 2118 - 2120
CCL1	Chemokine (C-C motif) ligand 1	SEQ ID NO: 2121
CCL11	Chemokine (C-C motif) ligand 11	SEQ ID NO: 2122
CCL13	Chemokine (C-C motif) ligand 13	SEQ ID NOS: 2123 - 2124
CCL14	Chemokine (C-C motif) ligand 14	SEQ ID NOS: 2125 - 2128
CCL15	Chemokine (C-C motif) ligand 15	SEQ ID NOS: 2129 - 2130
CCL16	Chemokine (C-C motif) ligand 16	SEQ ID NOS: 2131 - 2133
CCL17	Chemokine (C-C motif) ligand 17	SEQ ID NOS: 2134 - 2135
CCL18	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)	SEQ ID NO: 2136
CCL19	Chemokine (C-C motif) ligand 19	SEQ ID NOS: 2137 - 2138
CCL2	Chemokine (C-C motif) ligand 2	SEQ ID NOS: 2139 - 2140
CCL20	Chemokine (C-C motif) ligand 20	SEQ ID NOS: 2141 - 2143
CCL21	Chemokine (C-C motif) ligand 21	SEQ ID NOS: 2144 - 2145
CCL22	Chemokine (C-C motif) ligand 22	SEQ ID NO: 2146

CCL23	Chemokine (C-C motif) ligand 23	SEQ ID NOS: 2147 - 2149
CCL24	Chemokine (C-C motif) ligand 24	SEQ ID NOS: 2150 - 2151
CCL25	Chemokine (C-C motif) ligand 25	SEQ ID NOS: 2152 - 2155
CCL26	Chemokine (C-C motif) ligand 26	SEQ ID NOS: 2156 - 2157
CCL27	Chemokine (C-C motif) ligand 27	SEQ ID NO: 2158
CCL28	Chemokine (C-C motif) ligand 28	SEQ ID NOS: 2159 - 2161
CCL3	Chemokine (C-C motif) ligand 3	SEQ ID NO: 2162
CCL3L3	Chemokine (C-C motif) ligand 3-like 3	SEQ ID NO: 2163
CCL4	Chemokine (C-C motif) ligand 4	SEQ ID NOS: 2164 - 2165
CCL4L2	Chemokine (C-C motif) ligand 4-like 2	SEQ ID NOS: 2166 - 2175
CCL5	Chemokine (C-C motif) ligand 5	SEQ ID NOS: 2176 - 2178
CCL7	Chemokine (C-C motif) ligand 7	SEQ ID NOS: 2179 - 2181
CCL8	Chemokine (C-C motif) ligand 8	SEQ ID NO: 2182
CCNB1IP1	Cyclin B1 interacting protein 1, E3 ubiquitin protein ligase	SEQ ID NOS: 2183 - 2194
CCNL1	Cyclin L1	SEQ ID NOS: 2195 - 2203
CCNL2	Cyclin L2	SEQ ID NOS: 2204 - 2211
CD14	CD14 molecule	SEQ ID NOS: 2212 - 2216
CD160	CD160 molecule	SEQ ID NOS: 2217 - 2221
CD164	CD164 molecule, sialomucin	SEQ ID NOS: 2222 - 2227
CD177	CD177 molecule	SEQ ID NOS: 2228 - 2230
CD1E	CD1e molecule	SEQ ID NOS: 2231 - 2244
CD2	CD2 molecule	SEQ ID NOS: 2245 - 2246
CD200	CD200 molecule	SEQ ID NOS: 2247 - 2253
CD200R1	CD200 receptor 1	SEQ ID NOS: 2254 - 2258
CD22	CD22 molecule	SEQ ID NOS: 2259 - 2276
CD226	CD226 molecule	SEQ ID NOS: 2277 - 2284
CD24	CD24 molecule	SEQ ID NOS: 2285 - 2291
CD276	CD276 molecule	SEQ ID NOS: 2292 - 2307
CD300A	CD300a molecule	SEQ ID NOS: 2308 - 2312
CD300LB	CD300 molecule-like family member b	SEQ ID NOS: 2313 - 2314
CD300LF	CD300 molecule-like family member f	SEQ ID NOS: 2315 - 2323
CD300LG	CD300 molecule-like family member g	SEQ ID NOS: 2324 - 2329
CD3D	CD3d molecule, delta (CD3-TCR complex)	SEQ ID NOS: 2330 - 2333
CD4	CD4 molecule	SEQ ID NOS: 2334 - 2336
CD40	CD40 molecule, TNF receptor superfamily member 5	SEQ ID NOS: 2337 - 2340
CD44	CD44 molecule (Indian blood group)	SEQ ID NOS: 2341 - 2367
CD48	CD48 molecule	SEQ ID NOS: 2368 - 2370
CD5	CD5 molecule	SEQ ID NOS: 2371 - 2372
CD55	CD55 molecule, decay accelerating factor for complement (Cromer blood group)	SEQ ID NOS: 2373 - 2383
CD59	CD59 molecule, complement regulatory protein	SEQ ID NOS: 2384 - 2394
CD5L	CD5 molecule-like	SEQ ID NO: 2395
CD6	CD6 molecule	SEQ ID NOS: 2396 - 2403
CD68	CD68 molecule	SEQ ID NOS: 2404 - 2407

CD7	CD7 molecule	SEQ ID NOS: 2408 - 2413
CD79A	CD79a molecule, immunoglobulin-associated alpha	SEQ ID NOS: 2414 - 2416
CD80	CD80 molecule	SEQ ID NOS: 2417 - 2419
CD86	CD86 molecule	SEQ ID NOS: 2420 - 2426
CD8A	CD8a molecule	SEQ ID NOS: 2427 - 2430
CD8B	CD8b molecule	SEQ ID NOS: 2431 - 2436
CD99	CD99 molecule	SEQ ID NOS: 2437 - 2445
CDC23	Cell division cycle 23	SEQ ID NOS: 2446 - 2450
CDC40	Cell division cycle 40	SEQ ID NOS: 2451 - 2453
CDC45	Cell division cycle 45	SEQ ID NOS: 2454 - 2460
CDCP1	CUB domain containing protein 1	SEQ ID NOS: 2461 - 2462
CDCP2	CUB domain containing protein 2	SEQ ID NOS: 2463 - 2464
CDH1	Cadherin 1, type 1	SEQ ID NOS: 2465 - 2472
CDH11	Cadherin 11, type 2, OB-cadherin (osteoblast)	SEQ ID NOS: 2473 - 2482
CDH13	Cadherin 13	SEQ ID NOS: 2483 - 2492
CDH17	Cadherin 17, LI cadherin (liver-intestine)	SEQ ID NOS: 2493 - 2497
CDH18	Cadherin 18, type 2	SEQ ID NOS: 2498 - 2504
CDH19	Cadherin 19, type 2	SEQ ID NOS: 2505 - 2509
CDH23	Cadherin-related 23	SEQ ID NOS: 2510 - 2525
CDH5	Cadherin 5, type 2 (vascular endothelium)	SEQ ID NOS: 2526 - 2533
CDHR1	Cadherin-related family member 1	SEQ ID NOS: 2534 - 2539
CDHR4	Cadherin-related family member 4	SEQ ID NOS: 2540 - 2544
CDHR5	Cadherin-related family member 5	SEQ ID NOS: 2545 - 2551
CDKN2A	Cyclin-dependent kinase inhibitor 2A	SEQ ID NOS: 2552 - 2562
CDFN	Cerebral dopamine neurotrophic factor	SEQ ID NOS: 2563 - 2564
CDON	Cell adhesion associated, oncogene regulated	SEQ ID NOS: 2565 - 2572
CDSN	Corneodesmosin	SEQ ID NO: 2573
CEACAM16	Carcinoembryonic antigen-related cell adhesion molecule 16	SEQ ID NOS: 2574 - 2575
CEACAM18	Carcinoembryonic antigen-related cell adhesion molecule 18	SEQ ID NO: 2576
CEACAM19	Carcinoembryonic antigen-related cell adhesion molecule 19	SEQ ID NOS: 2577 - 2583
CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5	SEQ ID NOS: 2584 - 2591
CEACAM7	Carcinoembryonic antigen-related cell adhesion molecule 7	SEQ ID NOS: 2592 - 2594
CEACAM8	Carcinoembryonic antigen-related cell adhesion molecule 8	SEQ ID NOS: 2595 - 2596
CEL	Carboxyl ester lipase	SEQ ID NO: 2597
CELA2A	Chymotrypsin-like elastase family, member 2A	SEQ ID NO: 2598
CELA2B	Chymotrypsin-like elastase family, member 2B	SEQ ID NOS: 2599 - 2600
CELA3A	Chymotrypsin-like elastase family, member 3A	SEQ ID NOS: 2601 - 2603
CELA3B	Chymotrypsin-like elastase family, member 3B	SEQ ID NOS: 2604 - 2606
CEMIP	Cell migration inducing protein, hyaluronan binding	SEQ ID NOS: 2607 - 2611

CEP89	Centrosomal protein 89kDa	SEQ ID NOS: 2612 - 2617
CER1	Cerberus 1, DAN family BMP antagonist	SEQ ID NO: 2618
CERCAM	Cerebral endothelial cell adhesion molecule	SEQ ID NOS: 2619 - 2626
CERS1	Ceramide synthase 1	SEQ ID NOS: 2627 - 2631
CES1	Carboxylesterase 1	SEQ ID NOS: 2632 - 2637
CES3	Carboxylesterase 3	SEQ ID NOS: 2638 - 2642
CES4A	Carboxylesterase 4A	SEQ ID NOS: 2643 - 2648
CES5A	Carboxylesterase 5A	SEQ ID NOS: 2649 - 2656
CETP	Cholesteryl ester transfer protein, plasma	SEQ ID NOS: 2657 - 2659
CCDC108	Coiled-coil domain containing 108	SEQ ID NOS: 2660 - 2669
CFB	Complement factor B	SEQ ID NOS: 2670 - 2674
CFC1	Cripto, FRL-1, cryptic family 1	SEQ ID NOS: 2675 - 2677
CFC1B	Cripto, FRL-1, cryptic family 1B	SEQ ID NOS: 2678 - 2680
CFD	Complement factor D (adipsin)	SEQ ID NOS: 2681 - 2682
CFDP1	Craniofacial development protein 1	SEQ ID NOS: 2683 - 2686
CFH	Complement factor H	SEQ ID NOS: 2687 - 2689
CFHR1	Complement factor H-related 1	SEQ ID NOS: 2690 - 2691
CFHR2	Complement factor H-related 2	SEQ ID NOS: 2692 - 2693
CFHR3	Complement factor H-related 3	SEQ ID NOS: 2694 - 2698
CFHR4	Complement factor H-related 4	SEQ ID NOS: 2699 - 2702
CFHR5	Complement factor H-related 5	SEQ ID NO: 2703
CFI	Complement factor I	SEQ ID NOS: 2704 - 2708
CFP	Complement factor properdin	SEQ ID NOS: 2709 - 2712
CGA	Glycoprotein hormones, alpha polypeptide	SEQ ID NOS: 2713 - 2717
CGB1	Chorionic gonadotropin, beta polypeptide 1	SEQ ID NOS: 2718 - 2719
CGB2	Chorionic gonadotropin, beta polypeptide 2	SEQ ID NOS: 2720 - 2721
CGB	Chorionic gonadotropin, beta polypeptide	SEQ ID NO: 2722
CGB5	Chorionic gonadotropin, beta polypeptide 5	SEQ ID NO: 2723
CGB7	Chorionic gonadotropin, beta polypeptide 7	SEQ ID NOS: 2724 - 2726
CGB8	Chorionic gonadotropin, beta polypeptide 8	SEQ ID NO: 2727
CGREF1	Cell growth regulator with EF-hand domain 1	SEQ ID NOS: 2728 - 2735
CHAD	Chondroadherin	SEQ ID NOS: 2736 - 2738
CHADL	Chondroadherin-like	SEQ ID NOS: 2739 - 2741
CHEK2	Checkpoint kinase 2	SEQ ID NOS: 2742 - 2763
CHGA	Chromogranin A	SEQ ID NOS: 2764 - 2766
CHGB	Chromogranin B	SEQ ID NOS: 2767 - 2768
CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)	SEQ ID NOS: 2769 - 2770
CHI3L2	Chitinase 3-like 2	SEQ ID NOS: 2771 - 2784
CHIA	Chitinase, acidic	SEQ ID NOS: 2785 - 2793
CHID1	Chitinase domain containing 1	SEQ ID NOS: 2794 - 2812
CHIT1	Chitinase 1 (chitotriosidase)	SEQ ID NOS: 2813 - 2816
CHL1	Cell adhesion molecule L1-like	SEQ ID NOS: 2817 - 2825
CHN1	Chimerin 1	SEQ ID NOS: 2826 - 2836
CHPF	Chondroitin polymerizing factor	SEQ ID NOS: 2837 - 2839
CHPF2	Chondroitin polymerizing factor 2	SEQ ID NOS: 2840 - 2843

CHRD	Chordin	SEQ ID NOS: 2844 - 2849
CHRD1	Chordin-like 1	SEQ ID NOS: 2850 - 2854
CHRD2	Chordin-like 2	SEQ ID NOS: 2855 - 2863
CHRNA2	Cholinergic receptor, nicotinic, alpha 2 (neuronal)	SEQ ID NOS: 2864 - 2872
CHRNA5	Cholinergic receptor, nicotinic, alpha 5 (neuronal)	SEQ ID NOS: 2873 - 2876
CHRNA1	Cholinergic receptor, nicotinic, beta 1 (muscle)	SEQ ID NOS: 2877 - 2882
CHRNA3	Cholinergic receptor, nicotinic, delta (muscle)	SEQ ID NOS: 2883 - 2888
CHST1	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	SEQ ID NO: 2889
CHST10	Carbohydrate sulfotransferase 10	SEQ ID NOS: 2890 - 2897
CHST11	Carbohydrate (chondroitin 4) sulfotransferase 11	SEQ ID NOS: 2898 - 2902
CHST13	Carbohydrate (chondroitin 4) sulfotransferase 13	SEQ ID NOS: 2903 - 2904
CHST4	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4	SEQ ID NOS: 2905 - 2906
CHST5	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 5	SEQ ID NOS: 2907 - 2908
CHST6	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6	SEQ ID NOS: 2909 - 2910
CHST7	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 7	SEQ ID NO: 2911
CHST8	Carbohydrate (N-acetylgalactosamine 4-O) sulfotransferase 8	SEQ ID NOS: 2912 - 2915
CHSY1	Chondroitin sulfate synthase 1	SEQ ID NOS: 2916 - 2917
CHSY3	Chondroitin sulfate synthase 3	SEQ ID NO: 2918
CHTF8	Chromosome transmission fidelity factor 8	SEQ ID NOS: 2919 - 2929
CILP	Cartilage intermediate layer protein, nucleotide pyrophosphohydrolase	SEQ ID NO: 2930
CILP2	Cartilage intermediate layer protein 2	SEQ ID NOS: 2931 - 2932
CKLF	Chemokine-like factor	SEQ ID NOS: 2933 - 2938
CKMT1A	Creatine kinase, mitochondrial 1A	SEQ ID NOS: 2939 - 2944
CKMT1B	Creatine kinase, mitochondrial 1B	SEQ ID NOS: 2945 - 2954
CLCA1	Chloride channel accessory 1	SEQ ID NOS: 2955 - 2956
CLCF1	Cardiotrophin-like cytokine factor 1	SEQ ID NOS: 2957 - 2958
CLDN15	Claudin 15	SEQ ID NOS: 2959 - 2964
CLDN7	Claudin 7	SEQ ID NOS: 2965 - 2971
CLDND1	Claudin domain containing 1	SEQ ID NOS: 2972 - 2997
CLEC11A	C-type lectin domain family 11, member A	SEQ ID NOS: 2998 - 3000
CLEC16A	C-type lectin domain family 16, member A	SEQ ID NOS: 3001 - 3006
CLEC18A	C-type lectin domain family 18, member A	SEQ ID NOS: 3007 - 3012
CLEC18B	C-type lectin domain family 18, member B	SEQ ID NOS: 3013 - 3016
CLEC18C	C-type lectin domain family 18, member C	SEQ ID NOS: 3017 - 3023
CLEC19A	C-type lectin domain family 19, member A	SEQ ID NOS: 3024 - 3027
CLEC2B	C-type lectin domain family 2, member B	SEQ ID NOS: 3028 - 3029
CLEC3A	C-type lectin domain family 3, member A	SEQ ID NOS: 3030 - 3031
CLEC3B	C-type lectin domain family 3, member B	SEQ ID NOS: 3032 - 3033
CLGN	Calmegin	SEQ ID NOS: 3034 - 3036
CLN5	Ceroid-lipofuscinosis, neuronal 5	SEQ ID NOS: 3037 - 3048

CLPS	Colipase, pancreatic	SEQ ID NOS: 3049 - 3051
CLPSL1	Colipase-like 1	SEQ ID NOS: 3052 - 3053
CLPSL2	Colipase-like 2	SEQ ID NOS: 3054 - 3055
CLPX	Casainolytic mitochondrial matrix peptidase chaperone subunit	SEQ ID NOS: 3056 - 3058
CLSTN3	Calsyntenin 3	SEQ ID NOS: 3059 - 3065
CLU	Clusterin	SEQ ID NOS: 3066 - 3079
CLUL1	Clusterin-like 1 (retinal)	SEQ ID NOS: 3080 - 3087
CMA1	Chymase 1, mast cell	SEQ ID NOS: 3088 - 3089
CMPK1	Cytidine monophosphate (UMP-CMP) kinase 1, cytosolic	SEQ ID NOS: 3090 - 3093
CNBD1	Cyclic nucleotide binding domain containing 1	SEQ ID NOS: 3094 - 3097
CNDP1	Carnosine dipeptidase 1 (metallopeptidase M20 family)	SEQ ID NOS: 3098 - 3100
RQCD1	RCD1 required for cell differentiationI homolog (S. pombe)	SEQ ID NOS: 3101 - 3107
CNPY2	Canopy FGF signaling regulator 2	SEQ ID NOS: 3108 - 3112
CNPY3	Canopy FGF signaling regulator 3	SEQ ID NOS: 3113 - 3114
CNPY4	Canopy FGF signaling regulator 4	SEQ ID NOS: 3115 - 3117
CNTR	Ciliary neurotrophic factor receptor	SEQ ID NOS: 3118 - 3121
CNTN1	Contactin 1	SEQ ID NOS: 3122 - 3131
CNTN2	Contactin 2 (axonal)	SEQ ID NOS: 3132 - 3143
CNTN3	Contactin 3 (plasmacytoma associated)	SEQ ID NO: 3144
CNTN4	Contactin 4	SEQ ID NOS: 3145 - 3153
CNTN5	Contactin 5	SEQ ID NOS: 3154 - 3159
CNTNAP2	Contactin associated protein-like 2	SEQ ID NOS: 3160 - 3163
CNTNAP3	Contactin associated protein-like 3	SEQ ID NOS: 3164 - 3168
CNTNAP3B	Contactin associated protein-like 3B	SEQ ID NOS: 3169 - 3177
COASY	CoA synthase	SEQ ID NOS: 3178 - 3187
COCH	Cochlin	SEQ ID NOS: 3188 - 3199
COG3	Component of oligomeric golgi complex 3	SEQ ID NOS: 3200 - 3203
COL10A1	Collagen, type X, alpha 1	SEQ ID NOS: 3204 - 3207
COL11A1	Collagen, type XI, alpha 1	SEQ ID NOS: 3208 - 3218
COL11A2	Collagen, type XI, alpha 2	SEQ ID NOS: 3219 - 3223
COL12A1	Collagen, type XII, alpha 1	SEQ ID NOS: 3224 - 3231
COL14A1	Collagen, type XIV, alpha 1	SEQ ID NOS: 3232 - 3239
COL15A1	Collagen, type XV, alpha 1	SEQ ID NOS: 3240 - 3241
COL16A1	Collagen, type XVI, alpha 1	SEQ ID NOS: 3242 - 3246
COL18A1	Collagen, type XVIII, alpha 1	SEQ ID NOS: 3247 - 3251
COL19A1	Collagen, type XIX, alpha 1	SEQ ID NOS: 3252 - 3254
COL1A1	Collagen, type I, alpha 1	SEQ ID NOS: 3255 - 3256
COL1A2	Collagen, type I, alpha 2	SEQ ID NOS: 3257 - 3258
COL20A1	Collagen, type XX, alpha 1	SEQ ID NOS: 3259 - 3262
COL21A1	Collagen, type XXI, alpha 1	SEQ ID NOS: 3263 - 3268
COL22A1	Collagen, type XXII, alpha 1	SEQ ID NOS: 3269 - 3271
COL24A1	Collagen, type XXIV, alpha 1	SEQ ID NOS: 3272 - 3275

COL26A1	Collagen, type XXVI, alpha 1	SEQ ID NOS: 3276 - 3277
COL27A1	Collagen, type XXVII, alpha 1	SEQ ID NOS: 3278 - 3280
COL28A1	Collagen, type XXVIII, alpha 1	SEQ ID NOS: 3281 - 3285
COL2A1	Collagen, type II, alpha 1	SEQ ID NOS: 3286 - 3287
COL3A1	Collagen, type III, alpha 1	SEQ ID NOS: 3288 - 3290
COL4A1	Collagen, type IV, alpha 1	SEQ ID NOS: 3291 - 3293
COL4A2	Collagen, type IV, alpha 2	SEQ ID NOS: 3294 - 3296
COL4A3	Collagen, type IV, alpha 3 (Goodpasture antigen)	SEQ ID NOS: 3297 - 3300
COL4A4	Collagen, type IV, alpha 4	SEQ ID NOS: 3301 - 3302
COL4A5	Collagen, type IV, alpha 5	SEQ ID NOS: 3303 - 3309
COL4A6	Collagen, type IV, alpha 6	SEQ ID NOS: 3310 - 3315
COL5A1	Collagen, type V, alpha 1	SEQ ID NOS: 3316 - 3318
COL5A2	Collagen, type V, alpha 2	SEQ ID NOS: 3319 - 3320
COL5A3	Collagen, type V, alpha 3	SEQ ID NO: 3321
COL6A1	Collagen, type VI, alpha 1	SEQ ID NOS: 3322 - 3323
COL6A2	Collagen, type VI, alpha 2	SEQ ID NOS: 3324 - 3329
COL6A3	Collagen, type VI, alpha 3	SEQ ID NOS: 3330 - 3338
COL6A5	Collagen, type VI, alpha 5	SEQ ID NOS: 3339 - 3343
COL6A6	Collagen, type VI, alpha 6	SEQ ID NOS: 3344 - 3346
COL7A1	Collagen, type VII, alpha 1	SEQ ID NOS: 3347 - 3348
COL8A1	Collagen, type VIII, alpha 1	SEQ ID NOS: 3349 - 3352
COL8A2	Collagen, type VIII, alpha 2	SEQ ID NOS: 3353 - 3355
COL9A1	Collagen, type IX, alpha 1	SEQ ID NOS: 3356 - 3359
COL9A2	Collagen, type IX, alpha 2	SEQ ID NOS: 3360 - 3363
COL9A3	Collagen, type IX, alpha 3	SEQ ID NOS: 3364 - 3365
COLEC10	Collectin sub-family member 10 (C-type lectin)	SEQ ID NO: 3366
COLEC11	Collectin sub-family member 11	SEQ ID NOS: 3367 - 3376
COLGALT1	Collagen beta(1-O)galactosyltransferase 1	SEQ ID NOS: 3377 - 3379
COLGALT2	Collagen beta(1-O)galactosyltransferase 2	SEQ ID NOS: 3380 - 3382
COLQ	Collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase	SEQ ID NOS: 3383 - 3387
COMP	Cartilage oligomeric matrix protein	SEQ ID NOS: 3388 - 3390
COPS6	COP9 signalosome subunit 6	SEQ ID NOS: 3391 - 3394
COQ6	Coenzyme Q6 monooxygenase	SEQ ID NOS: 3395 - 3402
CORT	Cortistatin	SEQ ID NO: 3403
CP	Ceruloplasmin (ferroxidase)	SEQ ID NOS: 3404 - 3408
CPA1	Carboxypeptidase A1 (pancreatic)	SEQ ID NOS: 3409 - 3413
CPA2	Carboxypeptidase A2 (pancreatic)	SEQ ID NOS: 3414 - 3415
CPA3	Carboxypeptidase A3 (mast cell)	SEQ ID NO: 3416
CPA4	Carboxypeptidase A4	SEQ ID NOS: 3417 - 3422
CPA6	Carboxypeptidase A6	SEQ ID NOS: 3423 - 3425
CPAMD8	C3 and PZP-like, alpha-2-macroglobulin domain containing 8	SEQ ID NOS: 3426 - 3431
CPB1	Carboxypeptidase B1 (tissue)	SEQ ID NOS: 3432 - 3436
CPB2	Carboxypeptidase B2 (plasma)	SEQ ID NOS: 3437 - 3439
CPE	Carboxypeptidase E	SEQ ID NOS: 3440 - 3444

CPM	Carboxypeptidase M	SEQ ID NOS: 3445 - 3454
CPN1	Carboxypeptidase N, polypeptide 1	SEQ ID NOS: 3455 - 3456
CPN2	Carboxypeptidase N, polypeptide 2	SEQ ID NOS: 3457 - 3458
CPO	Carboxypeptidase O	SEQ ID NO: 3459
CPQ	Carboxypeptidase Q	SEQ ID NOS: 3460 - 3465
CPVL	Carboxypeptidase, vitellogenic-like	SEQ ID NOS: 3466 - 3476
CPXM1	Carboxypeptidase X (M14 family), member 1	SEQ ID NO: 3477
CPXM2	Carboxypeptidase X (M14 family), member 2	SEQ ID NOS: 3478 - 3479
CPZ	Carboxypeptidase Z	SEQ ID NOS: 3480 - 3483
CR1L	Complement component (3b/4b) receptor 1-like	SEQ ID NOS: 3484 - 3485
CRB2	Crumbs family member 2	SEQ ID NOS: 3486 - 3488
CREG1	Cellular repressor of E1A-stimulated genes 1	SEQ ID NO: 3489
CREG2	Cellular repressor of E1A-stimulated genes 2	SEQ ID NO: 3490
CRELD1	Cysteine-rich with EGF-like domains 1	SEQ ID NOS: 3491 - 3496
CRELD2	Cysteine-rich with EGF-like domains 2	SEQ ID NOS: 3497 - 3501
CRH	Corticotropin releasing hormone	SEQ ID NO: 3502
CRHBP	Corticotropin releasing hormone binding protein	SEQ ID NOS: 3503 - 3504
CRHR1	Corticotropin releasing hormone receptor 1	SEQ ID NOS: 3505 - 3516
CRHR2	Corticotropin releasing hormone receptor 2	SEQ ID NOS: 3517 - 3523
CRISP1	Cysteine-rich secretory protein 1	SEQ ID NOS: 3524 - 3527
CRISP2	Cysteine-rich secretory protein 2	SEQ ID NOS: 3528 - 3530
CRISP3	Cysteine-rich secretory protein 3	SEQ ID NOS: 3531 - 3534
CRISPLD2	Cysteine-rich secretory protein LCCL domain containing 2	SEQ ID NOS: 3535 - 3542
CRLF1	Cytokine receptor-like factor 1	SEQ ID NOS: 3543 - 3544
CRP	C-reactive protein, pentraxin-related	SEQ ID NOS: 3545 - 3549
CRTAC1	Cartilage acidic protein 1	SEQ ID NOS: 3550 - 3554
CRTAP	Cartilage associated protein	SEQ ID NOS: 3555 - 3556
CRY2	Cryptochrome circadian clock 2	SEQ ID NOS: 3557 - 3560
CSAD	Cysteine sulfinic acid decarboxylase	SEQ ID NOS: 3561 - 3573
CSF1	Colony stimulating factor 1 (macrophage)	SEQ ID NOS: 3574 - 3581
CSF1R	Colony stimulating factor 1 receptor	SEQ ID NOS: 3582 - 3586
CSF2	Colony stimulating factor 2 (granulocyte-macrophage)	SEQ ID NO: 3587
CSF2RA	Colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	SEQ ID NOS: 3588 - 3599
CSF3	Colony stimulating factor 3 (granulocyte)	SEQ ID NOS: 3600 - 3606
CSGALNACT1	Chondroitin sulfate N-acetylgalactosaminyltransferase 1	SEQ ID NOS: 3607 - 3615
CSH1	Chorionic somatomammotropin hormone 1 (placental lactogen)	SEQ ID NOS: 3616 - 3619
CSH2	Chorionic somatomammotropin hormone 2	SEQ ID NOS: 3620 - 3624
CSHL1	Chorionic somatomammotropin hormone-like 1	SEQ ID NOS: 3625 - 3631
CSN1S1	Casein alpha s1	SEQ ID NOS: 3632 - 3637
CSN2	Casein beta	SEQ ID NO: 3638
CSN3	Casein kappa	SEQ ID NO: 3639

CST1	Cystatin SN	SEQ ID NOS: 3640 - 3641
CST11	Cystatin 11	SEQ ID NOS: 3642 - 3643
CST2	Cystatin SA	SEQ ID NO: 3644
CST3	Cystatin C	SEQ ID NOS: 3645 - 3647
CST4	Cystatin S	SEQ ID NO: 3648
CST5	Cystatin D	SEQ ID NO: 3649
CST6	Cystatin E/M	SEQ ID NO: 3650
CST7	Cystatin F (leukocystatin)	SEQ ID NO: 3651
CST8	Cystatin 8 (cystatin-related epididymal specific)	SEQ ID NOS: 3652 - 3653
CST9	Cystatin 9 (testatin)	SEQ ID NO: 3654
CST9L	Cystatin 9-like	SEQ ID NO: 3655
CSTL1	Cystatin-like 1	SEQ ID NOS: 3656 - 3658
CT55	Cancer/testis antigen 55	SEQ ID NOS: 3659 - 3660
CTBS	Chitobiase, di-N-acetyl-	SEQ ID NOS: 3661 - 3663
CTGF	Connective tissue growth factor	SEQ ID NO: 3664
CTHRC1	Collagen triple helix repeat containing 1	SEQ ID NOS: 3665 - 3668
CTLA4	Cytotoxic T-lymphocyte-associated protein 4	SEQ ID NOS: 3669 - 3672
CTNS	Cystinosin, lysosomal cystine transporter	SEQ ID NOS: 3673 - 3680
CTRB1	Chymotrypsinogen B1	SEQ ID NOS: 3681 - 3683
CTRB2	Chymotrypsinogen B2	SEQ ID NOS: 3684 - 3687
CTRC	Chymotrypsin C (caldecrin)	SEQ ID NOS: 3688 - 3689
CTRL	Chymotrypsin-like	SEQ ID NOS: 3690 - 3692
CTSA	Cathepsin A	SEQ ID NOS: 3693 - 3701
CTSB	Cathepsin B	SEQ ID NOS: 3702 - 3726
CTSC	Cathepsin C	SEQ ID NOS: 3727 - 3731
CTSD	Cathepsin D	SEQ ID NOS: 3732 - 3742
CTSE	Cathepsin E	SEQ ID NOS: 3743 - 3744
CTSF	Cathepsin F	SEQ ID NOS: 3745 - 3748
CTSG	Cathepsin G	SEQ ID NO: 3749
CTSH	Cathepsin H	SEQ ID NOS: 3750 - 3755
CTSK	Cathepsin K	SEQ ID NOS: 3756 - 3757
CTSL	Cathepsin L	SEQ ID NOS: 3758 - 3760
CTSO	Cathepsin O	SEQ ID NO: 3761
CTSS	Cathepsin S	SEQ ID NOS: 3762 - 3766
CTSV	Cathepsin V	SEQ ID NOS: 3767 - 3768
CTSW	Cathepsin W	SEQ ID NOS: 3769 - 3771
CTSZ	Cathepsin Z	SEQ ID NO: 3772
CUBN	Cubilin (intrinsic factor-cobalamin receptor)	SEQ ID NOS: 3773 - 3776
CUTA	CutA divalent cation tolerance homolog (E. coli)	SEQ ID NOS: 3777 - 3786
CX3CL1	Chemokine (C-X3-C motif) ligand 1	SEQ ID NOS: 3787 - 3790
CXADR	Coxsackie virus and adenovirus receptor	SEQ ID NOS: 3791 - 3795
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	SEQ ID NO: 3796
CXCL10	Chemokine (C-X-C motif) ligand 10	SEQ ID NO: 3797
CXCL11	Chemokine (C-X-C motif) ligand 11	SEQ ID NOS: 3798 - 3799

CXCL12	Chemokine (C-X-C motif) ligand 12	SEQ ID NOS: 3800 - 3805
CXCL13	Chemokine (C-X-C motif) ligand 13	SEQ ID NO: 3806
CXCL14	Chemokine (C-X-C motif) ligand 14	SEQ ID NOS: 3807 - 3808
CXCL17	Chemokine (C-X-C motif) ligand 17	SEQ ID NOS: 3809 - 3810
CXCL2	Chemokine (C-X-C motif) ligand 2	SEQ ID NO: 3811
CXCL3	Chemokine (C-X-C motif) ligand 3	SEQ ID NO: 3812
CXCL5	Chemokine (C-X-C motif) ligand 5	SEQ ID NO: 3813
CXCL6	Chemokine (C-X-C motif) ligand 6	SEQ ID NOS: 3814 - 3815
CXCL8	Chemokine (C-X-C motif) ligand 8	SEQ ID NOS: 3816 - 3817
CXCL9	Chemokine (C-X-C motif) ligand 9	SEQ ID NO: 3818
CXorf36	Chromosome X open reading frame 36	SEQ ID NOS: 3819 - 3820
CYB5D2	Cytochrome b5 domain containing 2	SEQ ID NOS: 3821 - 3824
CYHR1	Cysteine/histidine-rich 1	SEQ ID NOS: 3825 - 3832
CYP17A1	Cytochrome P450, family 17, subfamily A, polypeptide 1	SEQ ID NOS: 3833 - 3837
CYP20A1	Cytochrome P450, family 20, subfamily A, polypeptide 1	SEQ ID NOS: 3838 - 3844
CYP21A2	Cytochrome P450, family 21, subfamily A, polypeptide 2	SEQ ID NOS: 3845 - 3852
CYP26B1	Cytochrome P450, family 26, subfamily B, polypeptide 1	SEQ ID NOS: 3853 - 3857
CYP2A6	Cytochrome P450, family 2, subfamily A, polypeptide 6	SEQ ID NOS: 3858 - 3859
CYP2A7	Cytochrome P450, family 2, subfamily A, polypeptide 7	SEQ ID NOS: 3860 - 3862
CYP2B6	Cytochrome P450, family 2, subfamily B, polypeptide 6	SEQ ID NOS: 3863 - 3866
CYP2C18	Cytochrome P450, family 2, subfamily C, polypeptide 18	SEQ ID NOS: 3867 - 3868
CYP2C19	Cytochrome P450, family 2, subfamily C, polypeptide 19	SEQ ID NOS: 3869 - 3870
CYP2C8	Cytochrome P450, family 2, subfamily C, polypeptide 8	SEQ ID NOS: 3871 - 3878
CYP2C9	Cytochrome P450, family 2, subfamily C, polypeptide 9	SEQ ID NOS: 3879 - 3881
CYP2E1	Cytochrome P450, family 2, subfamily E, polypeptide 1	SEQ ID NOS: 3882 - 3887
CYP2F1	Cytochrome P450, family 2, subfamily F, polypeptide 1	SEQ ID NOS: 3888 - 3891
CYP2J2	Cytochrome P450, family 2, subfamily J, polypeptide 2	SEQ ID NO: 3892
CYP2R1	Cytochrome P450, family 2, subfamily R, polypeptide 1	SEQ ID NOS: 3893 - 3898
CYP2S1	Cytochrome P450, family 2, subfamily S, polypeptide 1	SEQ ID NOS: 3899 - 3904
CYP2W1	Cytochrome P450, family 2, subfamily W, polypeptide 1	SEQ ID NOS: 3905 - 3907
CYP46A1	Cytochrome P450, family 46, subfamily A, polypeptide 1	SEQ ID NOS: 3908 - 3912

CYP4F11	Cytochrome P450, family 4, subfamily F, polypeptide 11	SEQ ID NOS: 3913 - 3917
CYP4F2	Cytochrome P450, family 4, subfamily F, polypeptide 2	SEQ ID NOS: 3918 - 3922
CYR61	Cysteine-rich, angiogenic inducer, 61	SEQ ID NO: 3923
CYTL1	Cytokine-like 1	SEQ ID NOS: 3924 - 3926
D2HGDH	D-2-hydroxyglutarate dehydrogenase	SEQ ID NOS: 3927 - 3935
DAG1	Dystroglycan 1 (dystrophin-associated glycoprotein 1)	SEQ ID NOS: 3936 - 3950
DAND5	DAN domain family member 5, BMP antagonist	SEQ ID NOS: 3951 - 3952
DAO	D-amino-acid oxidase	SEQ ID NOS: 3953 - 3958
DAZAP2	DAZ associated protein 2	SEQ ID NOS: 3959 - 3967
DBH	Dopamine beta-hydroxylase (dopamine beta-monooxygenase)	SEQ ID NOS: 3968 - 3969
DBNL	Drebrin-like	SEQ ID NOS: 3970 - 3987
DCD	Dermcidin	SEQ ID NOS: 3988 - 3990
DCN	Decorin	SEQ ID NOS: 3991 - 4009
DDIAS	DNA damage-induced apoptosis suppressor	SEQ ID NOS: 4010 - 4019
DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit (non-catalytic)	SEQ ID NOS: 4020 - 4023
DDR1	Discoidin domain receptor tyrosine kinase 1	SEQ ID NOS: 4024 - 4069
DDR2	Discoidin domain receptor tyrosine kinase 2	SEQ ID NOS: 4070 - 4075
DDT	D-dopachrome tautomerase	SEQ ID NOS: 4076 - 4081
DDX17	DEAD (Asp-Glu-Ala-Asp) box helicase 17	SEQ ID NOS: 4082 - 4086
DDX20	DEAD (Asp-Glu-Ala-Asp) box polypeptide 20	SEQ ID NOS: 4087 - 4089
DDX25	DEAD (Asp-Glu-Ala-Asp) box helicase 25	SEQ ID NOS: 4090 - 4096
DDX28	DEAD (Asp-Glu-Ala-Asp) box polypeptide 28	SEQ ID NO: 4097
DEAF1	DEAF1 transcription factor	SEQ ID NOS: 4098 - 4100
DEF8	Differentially expressed in FDCP 8 homolog (mouse)	SEQ ID NOS: 4101 - 4120
DEFA1	Defensin, alpha 1	SEQ ID NOS: 4121 - 4122
DEFA1B	Defensin, alpha 1B	SEQ ID NO: 4123
DEFA3	Defensin, alpha 3, neutrophil-specific	SEQ ID NO: 4124
DEFA4	Defensin, alpha 4, corticostatin	SEQ ID NO: 4125
DEFA5	Defensin, alpha 5, Paneth cell-specific	SEQ ID NO: 4126
DEFA6	Defensin, alpha 6, Paneth cell-specific	SEQ ID NO: 4127
DEFB1	Defensin, beta 1	SEQ ID NO: 4128
DEFB103A	Defensin, beta 103A	SEQ ID NO: 4129
DEFB103B	Defensin, beta 103B	SEQ ID NO: 4130
DEFB104A	Defensin, beta 104A	SEQ ID NO: 4131
DEFB104B	Defensin, beta 104B	SEQ ID NO: 4132
DEFB105A	Defensin, beta 105A	SEQ ID NO: 4133
DEFB105B	Defensin, beta 105B	SEQ ID NO: 4134
DEFB106A	Defensin, beta 106A	SEQ ID NO: 4135
DEFB106B	Defensin, beta 106B	SEQ ID NO: 4136
DEFB107A	Defensin, beta 107A	SEQ ID NO: 4137
DEFB107B	Defensin, beta 107B	SEQ ID NO: 4138

DEFB108B	Defensin, beta 108B	SEQ ID NO: 4139
DEFB110	Defensin, beta 110	SEQ ID NOS: 4140 - 4141
DEFB113	Defensin, beta 113	SEQ ID NO: 4142
DEFB114	Defensin, beta 114	SEQ ID NO: 4143
DEFB115	Defensin, beta 115	SEQ ID NO: 4144
DEFB116	Defensin, beta 116	SEQ ID NO: 4145
DEFB118	Defensin, beta 118	SEQ ID NO: 4146
DEFB119	Defensin, beta 119	SEQ ID NOS: 4147 - 4149
DEFB121	Defensin, beta 121	SEQ ID NO: 4150
DEFB123	Defensin, beta 123	SEQ ID NO: 4151
DEFB124	Defensin, beta 124	SEQ ID NO: 4152
DEFB125	Defensin, beta 125	SEQ ID NO: 4153
DEFB126	Defensin, beta 126	SEQ ID NO: 4154
DEFB127	Defensin, beta 127	SEQ ID NO: 4155
DEFB128	Defensin, beta 128	SEQ ID NO: 4156
DEFB129	Defensin, beta 129	SEQ ID NO: 4157
DEFB130	Defensin, beta 130	SEQ ID NO: 4158
RP11-1236K1.1		SEQ ID NO: 4159
DEFB131	Defensin, beta 131	SEQ ID NO: 4160
CTD-2313N18.7		SEQ ID NO: 4161
DEFB132	Defensin, beta 132	SEQ ID NO: 4162
DEFB133	Defensin, beta 133	SEQ ID NO: 4163
DEFB134	Defensin, beta 134	SEQ ID NOS: 4164 - 4165
DEFB135	Defensin, beta 135	SEQ ID NO: 4166
DEFB136	Defensin, beta 136	SEQ ID NO: 4167
DEFB4A	Defensin, beta 4A	SEQ ID NO: 4168
DEFB4B	Defensin, beta 4B	SEQ ID NO: 4169
C10orf10	Chromosome 10 open reading frame 10	SEQ ID NOS: 4170 - 4171
DGCR2	DiGeorge syndrome critical region gene 2	SEQ ID NOS: 4172 - 4175
DHH	Desert hedgehog	SEQ ID NO: 4176
DHRS4	Dehydrogenase/reductase (SDR family) member 4	SEQ ID NOS: 4177 - 4184
DHRS4L2	Dehydrogenase/reductase (SDR family) member 4 like 2	SEQ ID NOS: 4185 - 4194
DHRS7	Dehydrogenase/reductase (SDR family) member 7	SEQ ID NOS: 4195 - 4202
DHRS7C	Dehydrogenase/reductase (SDR family) member 7C	SEQ ID NOS: 4203 - 4205
DHRS9	Dehydrogenase/reductase (SDR family) member 9	SEQ ID NOS: 4206 - 4213
DHRSX	Dehydrogenase/reductase (SDR family) X-linked	SEQ ID NOS: 4214 - 4218
DHX29	DEAH (Asp-Glu-Ala-His) box polypeptide 29	SEQ ID NOS: 4219 - 4221
DHX30	DEAH (Asp-Glu-Ala-His) box helicase 30	SEQ ID NOS: 4222 - 4229
DHX8	DEAH (Asp-Glu-Ala-His) box polypeptide 8	SEQ ID NOS: 4230 - 4234
DIO2	Deiodinase, iodothyronine, type II	SEQ ID NOS: 4235 - 4244
DIXDC1	DIX domain containing 1	SEQ ID NOS: 4245 - 4248
DKK1	Dickkopf WNT signaling pathway inhibitor 1	SEQ ID NO: 4249

DKK2	Dickkopf WNT signaling pathway inhibitor 2	SEQ ID NOS: 4250 - 4252
DKK3	Dickkopf WNT signaling pathway inhibitor 3	SEQ ID NOS: 4253 - 4258
DKK4	Dickkopf WNT signaling pathway inhibitor 4	SEQ ID NO: 4259
DKKL1	Dickkopf-like 1	SEQ ID NOS: 4260 - 4265
DLG4	Discs, large homolog 4 (Drosophila)	SEQ ID NOS: 4266 - 4274
DLK1	Delta-like 1 homolog (Drosophila)	SEQ ID NOS: 4275 - 4278
DLL1	Delta-like 1 (Drosophila)	SEQ ID NOS: 4279 - 4280
DLL3	Delta-like 3 (Drosophila)	SEQ ID NOS: 4281 - 4283
DMBT1	Deleted in malignant brain tumors 1	SEQ ID NOS: 4284 - 4290
DMKN	Dermokine	SEQ ID NOS: 4291 - 4337
DMP1	Dentin matrix acidic phosphoprotein 1	SEQ ID NOS: 4338 - 4339
DMRTA2	DMRT-like family A2	SEQ ID NOS: 4340 - 4341
DNAAF5	Dynein, axonemal, assembly factor 5	SEQ ID NOS: 4342 - 4345
DNAH14	Dynein, axonemal, heavy chain 14	SEQ ID NOS: 4346 - 4360
DNAJB11	DnaJ (Hsp40) homolog, subfamily B, member 11	SEQ ID NOS: 4361 - 4362
DNAJB9	DnaJ (Hsp40) homolog, subfamily B, member 9	SEQ ID NO: 4363
DNAJC25-GNG10	DNAJC25-GNG10 readthrough	SEQ ID NO: 4364
DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	SEQ ID NOS: 4365 - 4366
DNASE1	Deoxyribonuclease I	SEQ ID NOS: 4367 - 4377
DNASE1L1	Deoxyribonuclease I-like 1	SEQ ID NOS: 4378 - 4388
DNASE1L2	Deoxyribonuclease I-like 2	SEQ ID NOS: 4389 - 4394
DNASE1L3	Deoxyribonuclease I-like 3	SEQ ID NOS: 4395 - 4400
DNASE2	Deoxyribonuclease II, lysosomal	SEQ ID NOS: 4401 - 4402
DNASE2B	Deoxyribonuclease II beta	SEQ ID NOS: 4403 - 4404
DPEP1	Dipeptidase 1 (renal)	SEQ ID NOS: 4405 - 4409
DPEP2	Dipeptidase 2	SEQ ID NOS: 4410 - 4416
DPEP3	Dipeptidase 3	SEQ ID NO: 4417
DPF3	D4, zinc and double PHD fingers, family 3	SEQ ID NOS: 4418 - 4424
DPP4	Dipeptidyl-peptidase 4	SEQ ID NOS: 4425 - 4429
DPP7	Dipeptidyl-peptidase 7	SEQ ID NOS: 4430 - 4435
DPT	Dermatopontin	SEQ ID NO: 4436
DRAXIN	Dorsal inhibitory axon guidance protein	SEQ ID NO: 4437
DSE	Dermatan sulfate epimerase	SEQ ID NOS: 4438 - 4446
DSG2	Desmoglein 2	SEQ ID NOS: 4447 - 4448
DSPP	Dentin sialophosphoprotein	SEQ ID NOS: 4449 - 4450
DST	Dystonin	SEQ ID NOS: 4451 - 4469
DUOX1	Dual oxidase 1	SEQ ID NOS: 4470 - 4474
DYNLT3	Dynein, light chain, Tctex-type 3	SEQ ID NOS: 4475 - 4477
E2F5	E2F transcription factor 5, p130-binding	SEQ ID NOS: 4478 - 4484
EBAG9	Estrogen receptor binding site associated, antigen, 9	SEQ ID NOS: 4485 - 4493
EBI3	Epstein-Barr virus induced 3	SEQ ID NO: 4494
ECHDC1	Ethylmalonyl-CoA decarboxylase 1	SEQ ID NOS: 4495 - 4513
ECM1	Extracellular matrix protein 1	SEQ ID NOS: 4514 - 4516

ECM2	Extracellular matrix protein 2, female organ and adipocyte specific	SEQ ID NOS: 4517 - 4520
ECSIT	ECSIT signalling integrator	SEQ ID NOS: 4521 - 4532
EDDM3A	Epididymal protein 3A	SEQ ID NO: 4533
EDDM3B	Epididymal protein 3B	SEQ ID NO: 4534
EDEM2	ER degradation enhancer, mannosidase alpha-like 2	SEQ ID NOS: 4535 - 4536
EDEM3	ER degradation enhancer, mannosidase alpha-like 3	SEQ ID NOS: 4537 - 4539
EDIL3	EGF-like repeats and discoidin I-like domains 3	SEQ ID NOS: 4540 - 4541
EDN1	Endothelin 1	SEQ ID NO: 4542
EDN2	Endothelin 2	SEQ ID NO: 4543
EDN3	Endothelin 3	SEQ ID NOS: 4544 - 4549
EDNRB	Endothelin receptor type B	SEQ ID NOS: 4550 - 4558
EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	SEQ ID NOS: 4559 - 4569
EFEMP2	EGF containing fibulin-like extracellular matrix protein 2	SEQ ID NOS: 4570 - 4581
EFNA1	Ephrin-A1	SEQ ID NOS: 4582 - 4583
EFNA2	Ephrin-A2	SEQ ID NO: 4584
EFNA4	Ephrin-A4	SEQ ID NOS: 4585 - 4587
EGFL6	EGF-like-domain, multiple 6	SEQ ID NOS: 4588 - 4589
EGFL7	EGF-like-domain, multiple 7	SEQ ID NOS: 4590 - 4594
EGFL8	EGF-like-domain, multiple 8	SEQ ID NOS: 4595 - 4597
EGFLAM	EGF-like, fibronectin type III and laminin G domains	SEQ ID NOS: 4598 - 4606
EGFR	Epidermal growth factor receptor	SEQ ID NOS: 4607 - 4614
EHBP1	EH domain binding protein 1	SEQ ID NOS: 4615 - 4626
EHF	Ets homologous factor	SEQ ID NOS: 4627 - 4636
EHMT1	Euchromatic histone-lysine N-methyltransferase 1	SEQ ID NOS: 4637 - 4662
EHMT2	Euchromatic histone-lysine N-methyltransferase 2	SEQ ID NOS: 4663 - 4667
EIF2AK1	Eukaryotic translation initiation factor 2-alpha kinase 1	SEQ ID NOS: 4668 - 4671
ELANE	Elastase, neutrophil expressed	SEQ ID NOS: 4672 - 4673
ELN	Elastin	SEQ ID NOS: 4674 - 4696
ELP2	Elongator acetyltransferase complex subunit 2	SEQ ID NOS: 4697 - 4709
ELSPBP1	Epididymal sperm binding protein 1	SEQ ID NOS: 4710 - 4715
EMC1	ER membrane protein complex subunit 1	SEQ ID NOS: 4716 - 4722
EMC10	ER membrane protein complex subunit 10	SEQ ID NOS: 4723 - 4729
EMC9	ER membrane protein complex subunit 9	SEQ ID NOS: 4730 - 4733
EMCN	Endomucin	SEQ ID NOS: 4734 - 4738
EMID1	EMI domain containing 1	SEQ ID NOS: 4739 - 4745
EMILIN1	Elastin microfibril interfacer 1	SEQ ID NOS: 4746 - 4747
EMILIN2	Elastin microfibril interfacer 2	SEQ ID NO: 4748
EMILIN3	Elastin microfibril interfacer 3	SEQ ID NO: 4749
ENAM	Enamelin	SEQ ID NO: 4750
ENDOG	Endonuclease G	SEQ ID NO: 4751

ENDOU	Endonuclease, polyU-specific	SEQ ID NOS: 4752 - 4754
ENHO	Energy homeostasis associated	SEQ ID NO: 4755
ENO4	Enolase family member 4	SEQ ID NOS: 4756 - 4760
ENPP6	Ectonucleotide pyrophosphatase/phosphodiesterase 6	SEQ ID NOS: 4761 - 4762
ENPP7	Ectonucleotide pyrophosphatase/phosphodiesterase 7	SEQ ID NOS: 4763 - 4764
ENTPD5	Ectonucleoside triphosphate diphosphohydrolase 5	SEQ ID NOS: 4765 - 4769
ENTPD8	Ectonucleoside triphosphate diphosphohydrolase 8	SEQ ID NOS: 4770 - 4773
EOGT	EGF domain-specific O-linked N-acetylglucosamine (GlcNAc) transferase	SEQ ID NOS: 4774 - 4781
EPCAM	Epithelial cell adhesion molecule	SEQ ID NOS: 4782 - 4785
EPDR1	Ependymin related 1	SEQ ID NOS: 4786 - 4789
EPGN	Epithelial mitogen	SEQ ID NOS: 4790 - 4798
EPHA10	EPH receptor A10	SEQ ID NOS: 4799 - 4806
EPHA3	EPH receptor A3	SEQ ID NOS: 4807 - 4809
EPHA4	EPH receptor A4	SEQ ID NOS: 4810 - 4819
EPHA7	EPH receptor A7	SEQ ID NOS: 4820 - 4821
EPHA8	EPH receptor A8	SEQ ID NOS: 4822 - 4823
EPHB2	EPH receptor B2	SEQ ID NOS: 4824 - 4828
EPHB4	EPH receptor B4	SEQ ID NOS: 4829 - 4831
EPHX3	Epoxide hydrolase 3	SEQ ID NOS: 4832 - 4835
EPO	Erythropoietin	SEQ ID NO: 4836
EPPIN	Epididymal peptidase inhibitor	SEQ ID NOS: 4837 - 4839
EPPIN-WFDC6	EPPIN-WFDC6 readthrough	SEQ ID NO: 4840
EPS15	Epidermal growth factor receptor pathway substrate 15	SEQ ID NOS: 4841 - 4843
EPS8L1	EPS8-like 1	SEQ ID NOS: 4844 - 4849
EPX	Eosinophil peroxidase	SEQ ID NO: 4850
EPYC	Epiphycan	SEQ ID NOS: 4851 - 4852
EQTN	Equatorin, sperm acrosome associated	SEQ ID NOS: 4853 - 4855
ERAP1	Endoplasmic reticulum aminopeptidase 1	SEQ ID NOS: 4856 - 4861
ERAP2	Endoplasmic reticulum aminopeptidase 2	SEQ ID NOS: 4862 - 4869
ERBB3	Erb-b2 receptor tyrosine kinase 3	SEQ ID NOS: 4870 - 4883
FAM132B	Family with sequence similarity 132, member B	SEQ ID NOS: 4884 - 4886
ERLIN1	ER lipid raft associated 1	SEQ ID NOS: 4887 - 4889
ERLIN2	ER lipid raft associated 2	SEQ ID NOS: 4890 - 4898
ERN1	Endoplasmic reticulum to nucleus signaling 1	SEQ ID NOS: 4899 - 4900
ERN2	Endoplasmic reticulum to nucleus signaling 2	SEQ ID NOS: 4901 - 4905
ERO1A	Endoplasmic reticulum oxidoreductase alpha	SEQ ID NOS: 4906 - 4912
ERO1B	Endoplasmic reticulum oxidoreductase beta	SEQ ID NOS: 4913 - 4915
ERP27	Endoplasmic reticulum protein 27	SEQ ID NOS: 4916 - 4917
ERP29	Endoplasmic reticulum protein 29	SEQ ID NOS: 4918 - 4921
ERP44	Endoplasmic reticulum protein 44	SEQ ID NO: 4922
ERV3-1	Endogenous retrovirus group 3, member 1	SEQ ID NO: 4923

ESM1	Endothelial cell-specific molecule 1	SEQ ID NOS: 4924 - 4926
ESRP1	Epithelial splicing regulatory protein 1	SEQ ID NOS: 4927 - 4935
EXOG	Endo/exonuclease (5'-3'), endonuclease G-like	SEQ ID NOS: 4936 - 4949
EXTL1	Exostosin-like glycosyltransferase 1	SEQ ID NO: 4950
EXTL2	Exostosin-like glycosyltransferase 2	SEQ ID NOS: 4951 - 4955
F10	Coagulation factor X	SEQ ID NOS: 4956 - 4959
F11	Coagulation factor XI	SEQ ID NOS: 4960 - 4964
F12	Coagulation factor XII (Hageman factor)	SEQ ID NO: 4965
F13B	Coagulation factor XIII, B polypeptide	SEQ ID NO: 4966
F2	Coagulation factor II (thrombin)	SEQ ID NOS: 4967 - 4969
F2R	Coagulation factor II (thrombin) receptor	SEQ ID NOS: 4970 - 4971
F2RL3	Coagulation factor II (thrombin) receptor-like 3	SEQ ID NOS: 4972 - 4973
F5	Coagulation factor V (proaccelerin, labile factor)	SEQ ID NOS: 4974 - 4975
F7	Coagulation factor VII (serum prothrombin conversion accelerator)	SEQ ID NOS: 4976 - 4979
F8	Coagulation factor VIII, procoagulant component	SEQ ID NOS: 4980 - 4985
F9	Coagulation factor IX	SEQ ID NOS: 4986 - 4987
FABP6	Fatty acid binding protein 6, ileal	SEQ ID NOS: 4988 - 4990
FAM107B	Family with sequence similarity 107, member B	SEQ ID NOS: 4991 - 5012
FAM131A	Family with sequence similarity 131, member A	SEQ ID NOS: 5013 - 5021
FAM171A1	Family with sequence similarity 171, member A1	SEQ ID NOS: 5022 - 5023
FAM171B	Family with sequence similarity 171, member B	SEQ ID NOS: 5024 - 5025
FAM172A	Family with sequence similarity 172, member A	SEQ ID NOS: 5026 - 5030
FAM177A1	Family with sequence similarity 177, member A1	SEQ ID NOS: 5031 - 5040
FAM180A	Family with sequence similarity 180, member A	SEQ ID NOS: 5041 - 5043
FAM189A1	Family with sequence similarity 189, member A1	SEQ ID NOS: 5044 - 5045
FAM198A	Family with sequence similarity 198, member A	SEQ ID NOS: 5046 - 5048
FAM19A1	Family with sequence similarity 19 (chemokine (C-C motif)-like), member A1	SEQ ID NOS: 5049 - 5051
FAM19A2	Family with sequence similarity 19 (chemokine (C-C motif)-like), member A2	SEQ ID NOS: 5052 - 5059
FAM19A3	Family with sequence similarity 19 (chemokine (C-C motif)-like), member A3	SEQ ID NOS: 5060 - 5061
FAM19A4	Family with sequence similarity 19 (chemokine (C-C motif)-like), member A4	SEQ ID NOS: 5062 - 5064
FAM19A5	Family with sequence similarity 19 (chemokine (C-C motif)-like), member A5	SEQ ID NOS: 5065 - 5068
FAM20A	Family with sequence similarity 20, member A	SEQ ID NOS: 5069 - 5072
FAM20C	Family with sequence similarity 20, member C	SEQ ID NO: 5073
FAM213A	Family with sequence similarity 213, member A	SEQ ID NOS: 5074 - 5079
FAM46B	Family with sequence similarity 46, member B	SEQ ID NO: 5080
FAM57A	Family with sequence similarity 57, member A	SEQ ID NOS: 5081 - 5086
FAM78A	Family with sequence similarity 78, member A	SEQ ID NOS: 5087 - 5089
FAM96A	Family with sequence similarity 96, member A	SEQ ID NOS: 5090 - 5094
FAM9B	Family with sequence similarity 9, member B	SEQ ID NOS: 5095 - 5098
FAP	Fibroblast activation protein, alpha	SEQ ID NOS: 5099 - 5105
FAS	Fas cell surface death receptor	SEQ ID NOS: 5106 - 5115

FAT1	FAT atypical cadherin 1	SEQ ID NOS: 5116 - 5122
FBLN1	Fibulin 1	SEQ ID NOS: 5123 - 5135
FBLN2	Fibulin 2	SEQ ID NOS: 5136 - 5141
FBLN5	Fibulin 5	SEQ ID NOS: 5142 - 5147
FBLN7	Fibulin 7	SEQ ID NOS: 5148 - 5153
FBN1	Fibrillin 1	SEQ ID NOS: 5154 - 5157
FBN2	Fibrillin 2	SEQ ID NOS: 5158 - 5163
FBN3	Fibrillin 3	SEQ ID NOS: 5164 - 5168
FBXW7	F-box and WD repeat domain containing 7, E3 ubiquitin protein ligase	SEQ ID NOS: 5169 - 5179
FCAR	Fc fragment of IgA receptor	SEQ ID NOS: 5180 - 5189
FCGBP	Fc fragment of IgG binding protein	SEQ ID NOS: 5190 - 5192
FCGR1B	Fc fragment of IgG, high affinity Ib, receptor (CD64)	SEQ ID NOS: 5193 - 5198
FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	SEQ ID NOS: 5199 - 5205
FCGRT	Fc fragment of IgG, receptor, transporter, alpha	SEQ ID NOS: 5206 - 5216
FCMR	Fc fragment of IgM receptor	SEQ ID NOS: 5217 - 5223
FCN1	Ficolin (collagen/fibrinogen domain containing) 1	SEQ ID NOS: 5224 - 5225
FCN2	Ficolin (collagen/fibrinogen domain containing lectin) 2	SEQ ID NOS: 5226 - 5227
FCN3	Ficolin (collagen/fibrinogen domain containing) 3	SEQ ID NOS: 5228 - 5229
FCRL1	Fc receptor-like 1	SEQ ID NOS: 5230 - 5232
FCRL3	Fc receptor-like 3	SEQ ID NOS: 5233 - 5238
FCRL5	Fc receptor-like 5	SEQ ID NOS: 5239 - 5241
FCRLA	Fc receptor-like A	SEQ ID NOS: 5242 - 5253
FCRLB	Fc receptor-like B	SEQ ID NOS: 5254 - 5258
FDCSP	Follicular dendritic cell secreted protein	SEQ ID NO: 5259
FETUB	Fetuin B	SEQ ID NOS: 5260 - 5266
FGA	Fibrinogen alpha chain	SEQ ID NOS: 5267 - 5269
FGB	Fibrinogen beta chain	SEQ ID NOS: 5270 - 5272
FGF10	Fibroblast growth factor 10	SEQ ID NOS: 5273 - 5274
FGF17	Fibroblast growth factor 17	SEQ ID NOS: 5275 - 5276
FGF18	Fibroblast growth factor 18	SEQ ID NO: 5277
FGF19	Fibroblast growth factor 19	SEQ ID NO: 5278
FGF21	Fibroblast growth factor 21	SEQ ID NOS: 5279 - 5280
FGF22	Fibroblast growth factor 22	SEQ ID NOS: 5281 - 5282
FGF23	Fibroblast growth factor 23	SEQ ID NO: 5283
FGF3	Fibroblast growth factor 3	SEQ ID NO: 5284
FGF4	Fibroblast growth factor 4	SEQ ID NO: 5285
FGF5	Fibroblast growth factor 5	SEQ ID NOS: 5286 - 5288
FGF7	Fibroblast growth factor 7	SEQ ID NOS: 5289 - 5293
FGF8	Fibroblast growth factor 8 (androgen-induced)	SEQ ID NOS: 5294 - 5299
FGFBP1	Fibroblast growth factor binding protein 1	SEQ ID NO: 5300
FGFBP2	Fibroblast growth factor binding protein 2	SEQ ID NO: 5301
FGFBP3	Fibroblast growth factor binding protein 3	SEQ ID NO: 5302

FGFR1	Fibroblast growth factor receptor 1	SEQ ID NOS: 5303 - 5325
FGFR2	Fibroblast growth factor receptor 2	SEQ ID NOS: 5326 - 5347
FGFR3	Fibroblast growth factor receptor 3	SEQ ID NOS: 5348 - 5355
FGFR4	Fibroblast growth factor receptor 4	SEQ ID NOS: 5356 - 5365
FGFRL1	Fibroblast growth factor receptor-like 1	SEQ ID NOS: 5366 - 5371
FGG	Fibrinogen gamma chain	SEQ ID NOS: 5372 - 5377
FGL1	Fibrinogen-like 1	SEQ ID NOS: 5378 - 5384
FGL2	Fibrinogen-like 2	SEQ ID NOS: 5385 - 5386
FHL1	Four and a half LIM domains 1	SEQ ID NOS: 5387 - 5414
FHOD3	Formin homology 2 domain containing 3	SEQ ID NOS: 5415 - 5421
FIBIN	Fin bud initiation factor homolog (zebrafish)	SEQ ID NO: 5422
FICD	FIC domain containing	SEQ ID NOS: 5423 - 5426
FJX1	Four jointed box 1	SEQ ID NO: 5427
FKBP10	FK506 binding protein 10, 65 kDa	SEQ ID NOS: 5428 - 5433
FKBP11	FK506 binding protein 11, 19 kDa	SEQ ID NOS: 5434 - 5440
FKBP14	FK506 binding protein 14, 22 kDa	SEQ ID NOS: 5441 - 5443
FKBP2	FK506 binding protein 2, 13kDa	SEQ ID NOS: 5444 - 5447
FKBP7	FK506 binding protein 7	SEQ ID NOS: 5448 - 5453
FKBP9	FK506 binding protein 9, 63 kDa	SEQ ID NOS: 5454 - 5457
FLT1	Fms-related tyrosine kinase 1	SEQ ID NOS: 5458 - 5466
FLT4	Fms-related tyrosine kinase 4	SEQ ID NOS: 5467 - 5471
FMO1	Flavin containing monooxygenase 1	SEQ ID NOS: 5472 - 5476
FMO2	Flavin containing monooxygenase 2 (non-functional)	SEQ ID NOS: 5477 - 5479
FMO3	Flavin containing monooxygenase 3	SEQ ID NOS: 5480 - 5482
FMO5	Flavin containing monooxygenase 5	SEQ ID NOS: 5483 - 5489
FMOD	Fibromodulin	SEQ ID NO: 5490
FN1	Fibronectin 1	SEQ ID NOS: 5491 - 5503
FNDC1	Fibronectin type III domain containing 1	SEQ ID NOS: 5504 - 5505
FNDC7	Fibronectin type III domain containing 7	SEQ ID NOS: 5506 - 5507
FOCAD	Focadhesin	SEQ ID NOS: 5508 - 5514
FOLR2	Folate receptor 2 (fetal)	SEQ ID NOS: 5515 - 5524
FOLR3	Folate receptor 3 (gamma)	SEQ ID NOS: 5525 - 5529
FOXRED2	FAD-dependent oxidoreductase domain containing 2	SEQ ID NOS: 5530 - 5533
FP325331.1	Uncharacterized protein UNQ6126/PRO20091	SEQ ID NO: 5534
CH507-9B2.3		SEQ ID NOS: 5535 - 5541
FPGS	Folypoly glutamate synthase	SEQ ID NOS: 5542 - 5548
FRAS1	Fraser extracellular matrix complex subunit 1	SEQ ID NOS: 5549 - 5554
FREM1	FRAS1 related extracellular matrix 1	SEQ ID NOS: 5555 - 5559
FREM3	FRAS1 related extracellular matrix 3	SEQ ID NO: 5560
FRMPD2	FERM and PDZ domain containing 2	SEQ ID NOS: 5561 - 5564
FRZB	Frizzled-related protein	SEQ ID NO: 5565
FSHB	Follicle stimulating hormone, beta polypeptide	SEQ ID NOS: 5566 - 5568
FSHR	Follicle stimulating hormone receptor	SEQ ID NOS: 5569 - 5572
FST	Follistatin	SEQ ID NOS: 5573 - 5576

FSTL1	Follistatin-like 1	SEQ ID NOS: 5577 - 5580
FSTL3	Follistatin-like 3 (secreted glycoprotein)	SEQ ID NOS: 5581 - 5586
FSTL4	Follistatin-like 4	SEQ ID NOS: 5587 - 5589
FSTL5	Follistatin-like 5	SEQ ID NOS: 5590 - 5592
FTCDNL1	Formiminotransferase cyclodeaminase N-terminal like	SEQ ID NOS: 5593 - 5596
FUCA1	Fucosidase, alpha-L- 1, tissue	SEQ ID NO: 5597
FUCA2	Fucosidase, alpha-L- 2, plasma	SEQ ID NOS: 5598 - 5599
FURIN	Furin (paired basic amino acid cleaving enzyme)	SEQ ID NOS: 5600 - 5606
FUT10	Fucosyltransferase 10 (alpha (1,3) fucosyltransferase)	SEQ ID NOS: 5607 - 5609
FUT11	Fucosyltransferase 11 (alpha (1,3) fucosyltransferase)	SEQ ID NOS: 5610 - 5611
FXN	Frataxin	SEQ ID NOS: 5612 - 5619
FXR1	Fragile X mental retardation, autosomal homolog 1	SEQ ID NOS: 5620 - 5632
FXYP3	FXYP domain containing ion transport regulator 3	SEQ ID NOS: 5633 - 5645
GABBR1	Gamma-aminobutyric acid (GABA) B receptor, 1	SEQ ID NOS: 5646 - 5657
GABRA1	Gamma-aminobutyric acid (GABA) A receptor, alpha 1	SEQ ID NOS: 5658 - 5673
GABRA2	Gamma-aminobutyric acid (GABA) A receptor, alpha 2	SEQ ID NOS: 5674 - 5688
GABRA5	Gamma-aminobutyric acid (GABA) A receptor, alpha 5	SEQ ID NOS: 5689 - 5697
GABRG3	Gamma-aminobutyric acid (GABA) A receptor, gamma 3	SEQ ID NOS: 5698 - 5703
GABRP	Gamma-aminobutyric acid (GABA) A receptor, pi	SEQ ID NOS: 5704 - 5712
GAL	Galanin/GMAP prepropeptide	SEQ ID NO: 5713
GAL3ST1	Galactose-3-O-sulfotransferase 1	SEQ ID NOS: 5714 - 5735
GAL3ST2	Galactose-3-O-sulfotransferase 2	SEQ ID NO: 5736
GAL3ST3	Galactose-3-O-sulfotransferase 3	SEQ ID NOS: 5737 - 5738
GALC	Galactosylceramidase	SEQ ID NOS: 5739 - 5748
GALNS	Galactosamine (N-acetyl)-6-sulfatase	SEQ ID NOS: 5749 - 5754
GALNT10	Polypeptide N-acetylgalactosaminyltransferase 10	SEQ ID NOS: 5755 - 5758
GALNT12	Polypeptide N-acetylgalactosaminyltransferase 12	SEQ ID NOS: 5759 - 5760
GALNT15	Polypeptide N-acetylgalactosaminyltransferase 15	SEQ ID NOS: 5761 - 5764
GALNT2	Polypeptide N-acetylgalactosaminyltransferase 2	SEQ ID NO: 5765
GALNT6	Polypeptide N-acetylgalactosaminyltransferase 6	SEQ ID NOS: 5766 - 5777
GALNT8	Polypeptide N-acetylgalactosaminyltransferase 8	SEQ ID NOS: 5778 - 5781
GALNTL6	Polypeptide N-acetylgalactosaminyltransferase-like 6	SEQ ID NOS: 5782 - 5785
GALP	Galanin-like peptide	SEQ ID NOS: 5786 - 5788
GANAB	Glucosidase, alpha; neutral AB	SEQ ID NOS: 5789 - 5797
GARS	Glycyl-tRNA synthetase	SEQ ID NOS: 5798 - 5801
GAS1	Growth arrest-specific 1	SEQ ID NO: 5802
GAS6	Growth arrest-specific 6	SEQ ID NO: 5803
GAST	Gastrin	SEQ ID NO: 5804
PDDC1	Parkinson disease 7 domain containing 1	SEQ ID NOS: 5805 - 5813

GBA	Glucosidase, beta, acid	SEQ ID NOS: 5814 - 5817
GBGT1	Globoside alpha-1,3-N-acetylgalactosaminyltransferase 1	SEQ ID NOS: 5818 - 5826
GC	Group-specific component (vitamin D binding protein)	SEQ ID NOS: 5827 - 5831
GCG	Glucagon	SEQ ID NOS: 5832 - 5833
GCGR	Glucagon receptor	SEQ ID NOS: 5834 - 5836
GCNT7	Glucosaminyl (N-acetyl) transferase family member 7	SEQ ID NOS: 5837 - 5838
GCSH	Glycine cleavage system protein H (aminomethyl carrier)	SEQ ID NOS: 5839 - 5847
GDF1	Growth differentiation factor 1	SEQ ID NO: 5848
GDF10	Growth differentiation factor 10	SEQ ID NO: 5849
GDF11	Growth differentiation factor 11	SEQ ID NOS: 5850 - 5851
GDF15	Growth differentiation factor 15	SEQ ID NOS: 5852 - 5854
GDF2	Growth differentiation factor 2	SEQ ID NO: 5855
GDF3	Growth differentiation factor 3	SEQ ID NO: 5856
GDF5	Growth differentiation factor 5	SEQ ID NOS: 5857 - 5858
GDF6	Growth differentiation factor 6	SEQ ID NOS: 5859 - 5861
GDF7	Growth differentiation factor 7	SEQ ID NO: 5862
GDF9	Growth differentiation factor 9	SEQ ID NOS: 5863 - 5867
GDNF	Glial cell derived neurotrophic factor	SEQ ID NOS: 5868 - 5875
GFOD2	Glucose-fructose oxidoreductase domain containing 2	SEQ ID NOS: 5876 - 5881
GFPT2	Glutamine-fructose-6-phosphate transaminase 2	SEQ ID NOS: 5882 - 5884
GFRA2	GDNF family receptor alpha 2	SEQ ID NOS: 5885 - 5891
GFRA4	GDNF family receptor alpha 4	SEQ ID NOS: 5892 - 5894
GGA2	Golgi-associated, gamma adaptin ear containing, ARF binding protein 2	SEQ ID NOS: 5895 - 5903
GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	SEQ ID NO: 5904
GGT1	Gamma-glutamyltransferase 1	SEQ ID NOS: 5905 - 5927
GGT5	Gamma-glutamyltransferase 5	SEQ ID NOS: 5928 - 5932
GH1	Growth hormone 1	SEQ ID NOS: 5933 - 5937
GH2	Growth hormone 2	SEQ ID NOS: 5938 - 5942
GHDC	GH3 domain containing	SEQ ID NOS: 5943 - 5950
GHRH	Growth hormone releasing hormone	SEQ ID NOS: 5951 - 5953
GHRHR	Growth hormone releasing hormone receptor	SEQ ID NOS: 5954 - 5959
GHRL	Ghrelin/obestatin prepropeptide	SEQ ID NOS: 5960 - 5970
GIF	Gastric intrinsic factor (vitamin B synthesis)	SEQ ID NOS: 5971 - 5972
GIP	Gastric inhibitory polypeptide	SEQ ID NO: 5973
GKN1	Gastrokin 1	SEQ ID NO: 5974
GKN2	Gastrokin 2	SEQ ID NOS: 5975 - 5976
GLA	Galactosidase, alpha	SEQ ID NOS: 5977 - 5978
GLB1	Galactosidase, beta 1	SEQ ID NOS: 5979 - 5987
GLB1L	Galactosidase, beta 1-like	SEQ ID NOS: 5988 - 5995
GLB1L2	Galactosidase, beta 1-like 2	SEQ ID NOS: 5996 - 5997

GLCE	Glucuronic acid epimerase	SEQ ID NOS: 5998 - 5999
GLG1	Golgi glycoprotein 1	SEQ ID NOS: 6000 - 6007
GLIPR1	GLI pathogenesis-related 1	SEQ ID NOS: 6008 - 6011
GLIPR1L1	GLI pathogenesis-related 1 like 1	SEQ ID NOS: 6012 - 6015
GLIS3	GLIS family zinc finger 3	SEQ ID NOS: 6016 - 6024
GLMP	Glycosylated lysosomal membrane protein	SEQ ID NOS: 6025 - 6033
GLRB	Glycine receptor, beta	SEQ ID NOS: 6034 - 6039
GLS	Glutaminase	SEQ ID NOS: 6040 - 6047
GLT6D1	Glycosyltransferase 6 domain containing 1	SEQ ID NOS: 6048 - 6049
GLTPD2	Glycolipid transfer protein domain containing 2	SEQ ID NO: 6050
GLUD1	Glutamate dehydrogenase 1	SEQ ID NO: 6051
GM2A	GM2 ganglioside activator	SEQ ID NOS: 6052 - 6054
GML	Glycosylphosphatidylinositol anchored molecule like	SEQ ID NOS: 6055 - 6056
GNAS	GNAS complex locus	SEQ ID NOS: 6057 - 6078
GNLY	Granulysin	SEQ ID NOS: 6079 - 6082
GNPTG	N-acetylglucosamine-1-phosphate transferase, gamma subunit	SEQ ID NOS: 6083 - 6087
GNRH1	Gonadotropin-releasing hormone 1 (luteinizing-releasing hormone)	SEQ ID NOS: 6088 - 6089
GNRH2	Gonadotropin-releasing hormone 2	SEQ ID NOS: 6090 - 6093
GNS	Glucosamine (N-acetyl)-6-sulfatase	SEQ ID NOS: 6094 - 6099
GOLM1	Golgi membrane protein 1	SEQ ID NOS: 6100 - 6104
GORAB	Golgin, RAB6-interacting	SEQ ID NOS: 6105 - 6107
GOT2	Glutamic-oxaloacetic transaminase 2, mitochondrial	SEQ ID NOS: 6108 - 6110
GP2	Glycoprotein 2 (zymogen granule membrane)	SEQ ID NOS: 6111 - 6119
GP6	Glycoprotein VI (platelet)	SEQ ID NOS: 6120 - 6123
GPC2	Glypican 2	SEQ ID NOS: 6124 - 6125
GPC5	Glypican 5	SEQ ID NOS: 6126 - 6128
GPC6	Glypican 6	SEQ ID NOS: 6129 - 6130
GPD2	Glycerol-3-phosphate dehydrogenase 2 (mitochondrial)	SEQ ID NOS: 6131 - 6139
GPED1	G protein-coupled estrogen receptor 1	SEQ ID NOS: 6140 - 6146
GPHA2	Glycoprotein hormone alpha 2	SEQ ID NOS: 6147 - 6149
GPHB5	Glycoprotein hormone beta 5	SEQ ID NOS: 6150 - 6151
GPIHBP1	Glycosylphosphatidylinositol anchored high density lipoprotein binding protein 1	SEQ ID NO: 6152
GPLD1	Glycosylphosphatidylinositol specific phospholipase D1	SEQ ID NO: 6153
GPNMB	Glycoprotein (transmembrane) nmb	SEQ ID NOS: 6154 - 6156
GPR162	G protein-coupled receptor 162	SEQ ID NOS: 6157 - 6160
GPX3	Glutathione peroxidase 3	SEQ ID NOS: 6161 - 6168
GPX4	Glutathione peroxidase 4	SEQ ID NOS: 6169 - 6179
GPX5	Glutathione peroxidase 5	SEQ ID NOS: 6180 - 6181
GPX6	Glutathione peroxidase 6	SEQ ID NOS: 6182 - 6184
GPX7	Glutathione peroxidase 7	SEQ ID NO: 6185

GREM1	Gremlin 1, DAN family BMP antagonist	SEQ ID NOS: 6186 - 6188
GREM2	Gremlin 2, DAN family BMP antagonist	SEQ ID NO: 6189
GRHL3	Grainyhead-like transcription factor 3	SEQ ID NOS: 6190 - 6195
GRIA2	Glutamate receptor, ionotropic, AMPA 2	SEQ ID NOS: 6196 - 6207
GRIA3	Glutamate receptor, ionotropic, AMPA 3	SEQ ID NOS: 6208 - 6213
GRIA4	Glutamate receptor, ionotropic, AMPA 4	SEQ ID NOS: 6214 - 6225
GRIK2	Glutamate receptor, ionotropic, kainate 2	SEQ ID NOS: 6226 - 6234
GRIN2B	Glutamate receptor, ionotropic, N-methyl D-aspartate 2B	SEQ ID NOS: 6235 - 6238
GRM2	Glutamate receptor, metabotropic 2	SEQ ID NOS: 6239 - 6242
GRM3	Glutamate receptor, metabotropic 3	SEQ ID NOS: 6243 - 6247
GRM5	Glutamate receptor, metabotropic 5	SEQ ID NOS: 6248 - 6252
GRN	Granulin	SEQ ID NOS: 6253 - 6268
GRP	Gastrin-releasing peptide	SEQ ID NOS: 6269 - 6273
DFNA5	Deafness, autosomal dominant 5	SEQ ID NOS: 6274 - 6282
GSG1	Germ cell associated 1	SEQ ID NOS: 6283 - 6291
GSN	Gelsolin	SEQ ID NOS: 6292 - 6300
GTDC1	Glycosyltransferase-like domain containing 1	SEQ ID NOS: 6301 - 6314
GTPBP10	GTP-binding protein 10 (putative)	SEQ ID NOS: 6315 - 6323
GUCA2A	Guanylate cyclase activator 2A (guanylin)	SEQ ID NO: 6324
GUCA2B	Guanylate cyclase activator 2B (uroguanylin)	SEQ ID NO: 6325
GUSB	Glucuronidase, beta	SEQ ID NOS: 6326 - 6330
GVQW1	GVQW motif containing 1	SEQ ID NO: 6331
GXYLT1	Glucoside xylosyltransferase 1	SEQ ID NOS: 6332 - 6333
GXYLT2	Glucoside xylosyltransferase 2	SEQ ID NOS: 6334 - 6336
GYPB	Glycophorin B (MNS blood group)	SEQ ID NOS: 6337 - 6345
GZMA	Granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)	SEQ ID NO: 6346
GZMB	Granzyme B (granzyme 2, cytotoxic T-lymphocyte-associated serine esterase 1)	SEQ ID NOS: 6347 - 6355
GZMH	Granzyme H (cathepsin G-like 2, protein h-CCPX)	SEQ ID NOS: 6356 - 6358
GZMK	Granzyme K (granzyme 3; tryptase II)	SEQ ID NO: 6359
GZMM	Granzyme M (lymphocyte met-ase 1)	SEQ ID NOS: 6360 - 6361
H6PD	Hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)	SEQ ID NOS: 6362 - 6363
HABP2	Hyaluronan binding protein 2	SEQ ID NOS: 6364 - 6365
HADHB	Hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), beta subunit	SEQ ID NOS: 6366 - 6372
HAMP	Hepcidin antimicrobial peptide	SEQ ID NOS: 6373 - 6374
HAPLN1	Hyaluronan and proteoglycan link protein 1	SEQ ID NOS: 6375 - 6381
HAPLN2	Hyaluronan and proteoglycan link protein 2	SEQ ID NOS: 6382 - 6383
HAPLN3	Hyaluronan and proteoglycan link protein 3	SEQ ID NOS: 6384 - 6387
HAPLN4	Hyaluronan and proteoglycan link protein 4	SEQ ID NO: 6388
HARS2	Histidyl-tRNA synthetase 2, mitochondrial	SEQ ID NOS: 6389 - 6404
HAVCR1	Hepatitis A virus cellular receptor 1	SEQ ID NOS: 6405 - 6409
HCCS	Holocytochrome c synthase	SEQ ID NOS: 6410 - 6412

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 162

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 162

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

CLAIMS

What is claimed is:

1. A composition comprising
 - (a) an inducible transgene construct, comprising a sequence encoding an inducible promoter and a sequence encoding a transgene, and
 - (b) a receptor construct, comprising a sequence encoding a constitutive promoter and a sequence encoding an exogenous receptor,wherein, upon integration of the construct of (a) and the construct of (b) into a genomic sequence of a cell, the exogenous receptor is expressed, and
wherein the exogenous receptor, upon binding a ligand, transduces an intracellular signal that targets the inducible promoter of (a) to modify gene expression.
2. The composition of claim 1, wherein the exogenous receptor of (b) comprises a non-naturally occurring receptor.
3. The composition of claim 1 or 2, wherein the non-naturally occurring receptor is a synthetic, modified, recombinant, mutant or chimeric receptor.
4. The composition of any one of claims 1-3, wherein the non-naturally occurring receptor comprises one or more sequences isolated or derived from a T-cell receptor (TCR).
5. The composition of any one of claims 1-4, wherein the non-naturally occurring receptor comprises one or more sequences isolated or derived from a scaffold protein.
6. The composition of any one of claims 1-5, wherein the non-naturally occurring receptor comprises a transmembrane domain.
7. The composition of any one of claims 1-6, wherein the non-naturally occurring receptor interacts with an intracellular receptor that transduces an intracellular signal.
8. The composition of any one of claims 1-6, wherein the non-naturally occurring receptor comprises an intracellular signalling domain.

9. The composition of any one of claims 1-8, wherein the non-naturally occurring receptor is a chimeric ligand receptor (CLR).
10. The composition of claim 9, wherein the CLR is a chimeric antigen receptor.
11. The composition of claim 9 or 10, wherein the chimeric ligand receptor comprises
 - (a) an ectodomain comprising a ligand recognition region, wherein the ligand recognition region comprises at least scaffold protein;
 - (b) a transmembrane domain, and
 - (c) an endodomain comprising at least one costimulatory domain.
12. The composition of claim 11, wherein the ectodomain of (a) further comprises a signal peptide.
13. The composition of claim 11 or 12, wherein the ectodomain of (a) further comprises a hinge between the ligand recognition region and the transmembrane domain.
14. The composition of claim 12 or 13, wherein the signal peptide comprises a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR signal peptide.
15. The composition of claim 12 or 13, wherein the signal peptide comprises a sequence encoding a human CD8 α signal peptide.
16. The composition of claim 15, wherein the signal peptide comprises an amino acid sequence comprising MALPVTALLPLALLHAARP (SEQ ID NO: 17000).
17. The composition of claim 15 or 16, wherein the signal peptide is encoded by a nucleic acid sequence comprising atggcactgccagtcaccgccctgctgctgcctctggctctgctgctgcacgcagctagacca (SEQ ID NO: 17001).

18. The composition of any one of claims 11-17, wherein the transmembrane domain comprises a sequence encoding a human CD2, CD3 δ , CD3 ϵ , CD3 γ , CD3 ζ , CD4, CD8 α , CD19, CD28, 4-1BB or GM-CSFR transmembrane domain.
19. The composition of any one of claims 11-17, wherein the transmembrane domain comprises a sequence encoding a human CD8 α transmembrane domain.
20. The composition of claim 19, wherein the transmembrane domain comprises an amino acid sequence comprising IYIWAPLAGTCGVLLLSLVITLYC (SEQ ID NO: 17002).
21. The composition of claim 19 or 20, wherein the transmembrane domain is encoded by a nucleic acid sequence comprising
atctacatttgggcaccactggccgggacctgtggagtgtctgtctgagcctgggtcatcacactgtactgc (SEQ ID NO: 17003).
22. The composition of any one of claims 11-21, wherein the endodomain comprises a human CD3 ζ endodomain.
23. The composition of any one of claims 11-22, wherein the at least one costimulatory domain comprises a human 4-1BB, CD28, CD3 ζ , CD40, ICOS, MyD88, OX-40 intracellular segment, or any combination thereof.
24. The composition of any one of claims 11-22, wherein the at least one costimulatory domain comprises a human CD3 ζ and/or a 4-1BB costimulatory domain.
25. The composition of claim 23 or 24, wherein the CD3 ζ costimulatory domain comprises an amino acid sequence comprising
RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPPEMGGKPRRKNPQ
EGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALP
PR (SEQ ID NO: 17004).
26. The composition of claim 25, wherein the CD3 ζ costimulatory domain is encoded by a nucleic acid sequence comprising

cgcgtagaagttagtcgatcagcagatgccccagcttacaacagggacagaaccagctgtataacgagctgaatctgggcccgcga
gaggaatatgacgtgctggataagcggagaggacgcgaccccgaaatgggaggcaagcccaggcgcaaaaaccctcaggaagg
cctgtataacgagctgcagaaggacaaaatggcagaagcctattctgagatcgccatgaagggggagcgacggagaggcaaagg
gcacgatgggctgtaccagggactgagcaccgccacaaaggacacctatgatgctctgcataatgcaggcactgcctccaagg
(SEQ ID NO: 17005).

27. The composition of claim 23 or 24, wherein the 4-1BB costimulatory domain comprises an amino acid sequence comprising
KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL (SEQ ID NO: 17006).

28. The composition of claim 27, wherein the 4-1BB costimulatory domain is encoded by a nucleic acid sequence comprising
aagagaggcaggaagaaactgctgtatatatttcaaacagccctcatgcgccccgtgcagactaccaggaggaagacgggtgctcc
tgatgattccctgaggaagaggaaggcgggtgtgagctg (SEQ ID NO: 17007).

29. The composition of claims 24-28, wherein the 4-1BB costimulatory domain is located between the transmembrane domain and the CD3 ζ costimulatory domain.

30. The composition of any one of claims 11-29, wherein the hinge comprises a sequence derived from a human CD8 α , IgG4, and/or CD4 sequence.

31. The composition of any one of claims 11-29, wherein the hinge comprises a sequence derived from a human CD8 α sequence.

32. The composition of claim 30 or 31, wherein the hinge comprises an amino acid sequence comprising TTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD (SEQ ID NO: 17008).

33. The composition of claim 32, wherein the hinge is encoded by a nucleic acid sequence comprising
actaccacaccagcacctagaccaccaactccagctccaaccatcgcgagtcagcccctgagtctgagacctgaggcctgcaggec
agctgcaggaggagctgtgcacaccaggggcctggacttcgcctgcgac (SEQ ID NO: 17028) or a nucleic acid sequence comprising

ACCACAACCCCTGCCCCCAGACCTCCCACACCCGCCCCTACCATCGCGAGTCAGCCCCTGAGTCTGA
 GACCTGAGGCCTGCAGGCCAGCTGCAGGAGGAGCTGTGCACACCAGGGGCCTGGACTTCGCCTGC
 GAC (SEQ ID NO: 17009).

34. The composition of any one of claims 11-33, wherein the at least one protein scaffold specifically binds the ligand.

35. The composition of any one of claims 11-33, wherein the at least one protein scaffold comprises an antibody, an antibody fragment, a single domain antibody, a single chain antibody, an antibody mimetic, or a Centyrin.

36. The composition of any one of claims 11-33, wherein the ligand recognition region comprises one or more of an antibody, an antibody fragment, a single domain antibody, a single chain antibody, an antibody mimetic, and a Centyrin.

37. The composition of claim 35 or 36, wherein the single domain antibody comprises or consists of a VHH.

38. The composition of claim 35 or 36, wherein the antibody mimetic comprises or consists of an affibody, an affililn, an affimer, an affitin, an alphabody, an anticalin, an avimer, a DARPin, a Fynomer, a Kunitz domain peptide or a monobody.

39. The composition of claim 35 or 36, wherein the Centyrin comprises or consists of a consensus sequence of at least one fibronectin type III (FN3) domain.

40. The composition of claim 39, wherein the at least one fibronectin type III (FN3) domain is derived from a human protein.

41. The composition of claim 40, wherein the human protein is Tenascin-C.

42. The composition of any one of claims 39-41, wherein the consensus sequence comprises

LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYDL
TGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17010).

43. The composition of any one of claims 39-41, wherein the consensus sequence comprises

MLPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVGEAINLTVPGSERSYD
LTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 17011).

44. The composition of any one of claims 39-43, wherein the consensus sequence is modified at one or more positions within

(a) a A-B loop comprising or consisting of the amino acid residues TEDS at positions 13-16 of the consensus sequence;

(b) a B-C loop comprising or consisting of the amino acid residues TAPDAAF at positions 22-28 of the consensus sequence;

(c) a C-D loop comprising or consisting of the amino acid residues SEKVGE at positions 38-43 of the consensus sequence;

(d) a D-E loop comprising or consisting of the amino acid residues GSER at positions 51-54 of the consensus sequence;

(e) a E-F loop comprising or consisting of the amino acid residues GLKPG at positions 60-64 of the consensus sequence;

(f) a F-G loop comprising or consisting of the amino acid residues KGGHRSN at positions 75-81 of the consensus sequence; or

(g) any combination of (a)-(f).

45. The composition of any one of claims 39-44, comprising a consensus sequence of at least 5 fibronectin type III (FN3) domains.

46. The composition of any one of claims 39-44, comprising a consensus sequence of at least 10 fibronectin type III (FN3) domains.

47. The composition of any one of claims 39-44, comprising a consensus sequence of at least 15 fibronectin type III (FN3) domains.

48. The composition of any one of claims 39-47, wherein the scaffold binds an antigen with at least one affinity selected from a K_D of less than or equal to $10^{-9}M$, less than or equal to $10^{-10}M$, less than or equal to $10^{-11}M$, less than or equal to $10^{-12}M$, less than or equal to $10^{-13}M$, less than or equal to $10^{-14}M$, and less than or equal to $10^{-15}M$.
49. The CAR of claim 48, wherein the K_D is determined by surface plasmon resonance.
50. The composition of any one of claims 1-49, wherein the sequence encoding the constitutive promoter of (b) comprises a sequence encoding an EF1 α promoter.
51. The composition of any one of claims 1-49, the sequence encoding the constitutive promoter of (b) comprises a sequence encoding a CMV promoter, a U6 promoter, a SV40 promoter, a PGK1 promoter, a Ubc promoter, a human beta actin promoter, a CAG promoter, or an EF1 α promoter.
52. The composition of any one of claims 1-51, wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding an NF κ B promoter.
53. The composition of any one of claims 1-51, wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding an interferon (IFN) promoter or a sequence encoding an interleukin-2 promoter.
54. The composition of claim 53, wherein the interferon (IFN) promoter is an IFN γ promoter.
55. The composition of any one of claims 1-51, wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding an NR4A1 promoter or a sequence encoding an NR4A1 promoter.
56. The composition of any one of claims 1-51, wherein the sequence encoding the inducible promoter of (a) comprises a sequence encoding a CD5 promoter or a sequence encoding a CD5 promoter.

57. The composition of any one of claims 1-56, wherein the transgene comprises a sequence that is endogenous with respect to the genomic sequence of the cell.
58. The composition of any one of claims 1-56, wherein the transgene comprises a sequence that is exogenous with respect to the genomic sequence of the cell.
59. The composition of claim 58, wherein the exogenous sequence is a sequence variant of an endogenous sequence within the genome of the cell.
60. The composition of claim 58, wherein the exogenous sequence is a wild type sequence of gene that is entirely or partially absent in the cell, and wherein the gene is entirely present in the genome of a healthy cell.
61. The composition of claim 58, wherein the exogenous sequence is a synthetic, modified, recombinant, chimeric or non-naturally occurring sequence with respect to the genome of the cell.
62. The composition of any one of claims 57-61, wherein the transgene encodes a secreted protein.
63. The composition of claim 62, wherein the secreted protein is produced and/or secreted from the cell at a level that is therapeutically effective to treat a disease or disorder in a subject in need thereof.
64. The composition of any one of claims 1-63, wherein a first transposon comprises the inducible transgene construct of (a) and a second transposon comprises the receptor construct of (b).
65. The composition of claim 64, wherein a first vector comprises the first transposon and a second vector comprises the second transposon.
66. The composition of any one of claims 1-63, wherein a vector comprises the first transposon and the second transposon.

67. The composition of any one of claims 1-61, wherein a transposon comprises the inducible transgene construct of (a) and the receptor construct of (b).
68. The composition of claim 67, wherein a vector comprises the transposon.
69. The composition of claim 67 or 68, wherein the inducible transgene construct of (a) and the receptor construct of (b) are oriented in the same direction.
70. The composition of claim 65 or 66, wherein the inducible transgene construct of (a) and the receptor construct of (b) are not oriented in the same direction.
71. The composition of 70, wherein the inducible transgene construct of (a) and the receptor construct of (b) are oriented in opposite directions.
72. The composition of any one of claims 65-71, wherein the vector is a plasmid or a nanoplasmid.
73. The composition of any one of claims 65-72, wherein the transposon is a piggyBac, Sleeping Beauty, Helraiser or Tol2 transposon.
74. The composition of any one of claims 65-72, wherein the transposon is a piggyBac transposon.
75. The composition of any one of claims 64-74, wherein the composition further comprises a plasmid or a nanoplasmid comprising a sequence encoding a transposase enzyme.
76. The composition of claim 75, wherein the sequence encoding a transposase enzyme is an mRNA sequence.
77. The composition of claim 76, wherein the transposase is a piggyBac transposase.

78. The composition of claim 77, wherein the piggyBac transposase comprises an amino acid sequence comprising SEQ ID NO: 17029.

79. The composition of claim 77 or 78, wherein the piggyBac transposase is a hyperactive variant and wherein the hyperactive variant comprises an amino acid substitution at one or more of positions 30, 165, 282 and 538 of SEQ ID NO: 17029.

80. The composition of claim 79, wherein the amino acid substitution at position 30 of SEQ ID NO: 17029 is a substitution of a valine (V) for an isoleucine (I) (I30V).

81. The composition of claim 79, wherein the amino acid substitution at position 165 of SEQ ID NO: 17029 is a substitution of a serine (S) for a glycine (G) (G165S).

82. The composition of claim 79, wherein the amino acid substitution at position 282 of SEQ ID NO: 17029 is a substitution of a valine (V) for a methionine (M) (M282V).

83. The composition of claim 79, wherein the amino acid substitution at position 538 of SEQ ID NO: 17029 is a substitution of a lysine (K) for an asparagine (N) (N538K).

84. The composition of any one of claims 78-83, wherein the transposase is a Super piggyBac (SPB) transposase.

85. The composition of claim 84, wherein the Super piggyBac (SPB) transposase comprises an amino acid sequence comprising SEQ ID NO: 17030.

86. The composition of any one of claims 64-85, wherein the first transposon and/or the second transposon further comprises a selection gene.

87. The composition of claim 86, wherein the selection gene comprises *neo*, DHFR (Dihydrofolate Reductase), TYMS (Thymidylate Synthetase), MGMT (O(6)-methylguanine-DNA methyltransferase), multidrug resistance gene (MDR1), ALDH1 (Aldehyde dehydrogenase 1 family, member A1), FRAXE, RAD51C (RAD51 Paralog C), GCS (glucosylceramide synthase), NKX2.2 (NK2 Homeobox 2) or any combination thereof.

88. The composition of claim 86, wherein the selection gene comprises DHFR.
89. The composition of any one of claims 64-88, wherein the first transposon and or the second transposon comprises an inducible caspase polypeptide comprising
- (a) a ligand binding region,
 - (b) a linker, and
 - (c) a truncated caspase 9 polypeptide,
- wherein the inducible caspase polypeptide does not comprise a non-human sequence.
90. The composition of claim 89, wherein the non-human sequence is a restriction site.
91. The composition of claim 89 or 90, wherein the ligand binding region inducible caspase polypeptide comprises a FK506 binding protein 12 (FKBP12) polypeptide.
92. The composition of claim 91, wherein the amino acid sequence of the FK506 binding protein 12 (FKBP12) polypeptide comprises a modification at position 36 of the sequence.
93. The composition of claim 91, wherein the modification is a substitution of valine (V) for phenylalanine (F) at position 36 (F36V).
94. The composition of any one of claims 91-93, wherein the FKBP12 polypeptide is encoded by an amino acid sequence comprising
- GVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKVDSSRDRNKPFFKFMLGKQEVI
RGWEEGVAQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELLKLE (SEQ ID
NO: 17012).
95. The composition of claim 94, wherein the FKBP12 polypeptide is encoded by a nucleic acid sequence comprising
- GGGGTCCAGGTCGAGACTATTTACACAGGGGATGGGCGAACATTTCCAAAAAGG
GGCCAGACTTGCGTCGTGCATTACACCGGGATGCTGGAGGACGGGAAGAAAGTG
GACAGCTCCAGGGATCGCAACAAGCCCTTCAAGTTCATGCTGGGAAAGCAGGAA
GTGATCCGAGGATGGGAGGAAGGCGTGGCACAGATGTCAGTCGGCCAGCGGGCC

AAACTGACCATTAGCCCTGACTACGCTTATGGAGCAACAGGCCACCCAGGGATC
ATTCCCCCTCATGCCACCCTGGTCTTCGAT GTGGAAGCTGCTGAAGCTGGAG (SEQ
ID NO: 17013).

96. The composition of any one of claims 89-95, wherein the linker region of the inducible proapoptotic polypeptide is encoded by an amino acid comprising GGGGS (SEQ ID NO: 17014).

97. The composition of claim 96, wherein the linker region of the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising GGAGGAGGAGGATCC (SEQ ID NO: 17015).

98. The composition of any one of claims 89-97, wherein the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid sequence that does not comprise an arginine (R) at position 87 of the sequence.

99. The composition of any one of claims 89-98, wherein the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid sequence that does not comprise an alanine (A) at position 282 the sequence.

100. The composition of any one of claims 89-99, wherein the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by an amino acid comprising

GFGDVGALESLRGNADLAYILSMEPCGHCLIINNVNFCRESGLRTRTGSNIDCEKLRR
RFSSLHFMVEVKGDLTAKKMVLALLELAQQDHGALDCCVVVILSHGCQASHLQFPG
AVYGTGDCPVSVKIVNIFNGTSCPSLGGKPKLFFIQACGGEQKDHGFEVASTSPEDE
SPGSNPEPDATPFQEGLRFTDQLDAISSLPTPSDIFVSYSTFPGFVSWRDPKSGSWYVE
TLDDIFEQWAHSEDLQSLLLRVANAVSVKGIYKQMPGCFNFLRKKLFFKTS (SEQ ID
NO: 17016).

101. The composition of claim 100, wherein the truncated caspase 9 polypeptide of the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising

TTTGGGGACGTGGGGGGCCCTGGAGTCTCTGCGAGGAAATGCCGATCTGGCTTAC
 ATCCTGAGCATGGAACCCTGCGGCCACTGTCTGATCATTAACAATGTGAACTTCT
 GCAGAGAAAGCGGACTGCGAACACGGACTGGCTCCAATATTGACTGTGAGAAGC
 TGCGGAGAAGGTTCTCTAGTCTGCACTTTATGGTTCGAAGTGAAAGGGGATCTGAC
 CGCCAAGAAAATGGTGCTGGCCCTGCTGGAGCTGGCTCAGCAGGACCATGGAGC
 TCTGGATTGCTGCGTGGTCGTGATCCTGTCCCACGGGTGCCAGGCTTCTCATCTG
 CAGTTCCCCGGAGCAGTGTACGGAACAGACGGCTGTCCTGTCAGCGTGGAGAAG
 ATCGTCAACATCTTCAACGGCACTTCTTGCCCTAGTCTGGGGGGGAAAGCCAAAAC
 TGTTCTTTATCCAGGCCTGTGGCGGGGAACAGAAAGATCACGGCTTCGAGGTGG
 CCAGCACCAAGCCCTGAGGACGAATCACCAAGGGAGCAACCCTGAACCAGATGCAA
 CTCCATTCCAGGAGGGACTGAGGACCTTTGACCAGCTGGATGCTATCTCAAGCCT
 GCCCACTCCTAGTGACATTTTCGTGTCTTACAGTACCTTCCCAGGCTTTGTCTCAT
 GGCGCGATCCCAAGTCAGGGAGCTGGTACGTGGAGACACTGGACGACATCTTTG
 AACAGTGGGCCCATTTCAGAGGACCTGCAGAGCCTGCTGCTGCGAGTGGCAAACG
 CTGTCTCTGTGAAGGGCATCTACAAACAGATGCCCGGGTGCTTCAATTTTCTGAG
 AAAGAAACTGTTCTTTAAGACTTCC (SEQ ID NO: 17017).

102. The composition of any one of claims 89-101, wherein the inducible proapoptotic polypeptide is encoded by an amino acid sequence comprising
 GVQVETISPGDGRTPKRGQTCVVHYTGMLEDGKKVDSSRDRNKPFFKMLGKQEVI
 RGWEEGVAQMSVGQRAKLTI SPDYAYGATGHPGIIPPHATLVFDVELLKLEGGGGS
 GFGDVGALESLRGNADLAYILSMEPCGHCLIINNVNFCRESGLRTRTGSNIDCEKLRR
 RFSSLHFMVEVKGDLTAKKMVLALLELAQQDHGALDCCVVVILSHGCQASHLQFPG
 AVYGTDGCPVSVEKIVNIFNGTSCPSLGGKPKLFFIQACGGEQKDHGFEVASTSPEDE
 SPGSNPEPDATPFQEGLRTFDQLDAISSLPTPSDIFVSYSTFPGFVSWRDPKSGSWYVE
 TLDDIFEQWAHSEDLQSLLLRVANAVSVKGIYKQMPGCFNFLRKKLFFKTS (SEQ ID
 NO: 17018).

103. The composition of claim 102, wherein the inducible proapoptotic polypeptide is encoded by a nucleic acid sequence comprising
 Ggggtccaggtcgagactatttcaccaggggatgggcgaacatttcaaaaaggggccagacttgcgtcgtgcattacaccgggat
 gctggaggacgggaagaaagtggacagctccagggatcgcaacaagcccttcaagttcatgctgggaaagcaggaagtgatccga
 ggatgggaggaaggcgtggcacagatgtcagtcggccagcgggccaactgaccattagccctgactacgcttatggagcaacag

gccacccagggatcatccccctcatgccaccctgggtcttcgatgtggaactgctgaagctggaggaggaggaggatccggatttg
 gggacgtggggggccctggagtctctgcgaggaaatgccgatctggcttacatcctgagcatggaacctgcggccactgtctgatcat
 taacaatgtgaacttctgcagagaaagcggactgcgaacacggactggctccaatattgactgtgagaagctgcggagaaggttctct
 agtctgcactttatggtcgaagtgaaggggatctgaccgccaagaaaatgggtgctggccctgctggagctggctcagcaggacat
 ggagctctggattgctgcgtggctgctgatcctgtcccacgggtgccaggcttctcatctgcagtccccggagcagtgtacggaacag
 acggctgtcctgtcagcgtggagaagatcgtcaacatcttcaacggcacttcttgcctagctgggggggaaagccaaaactgttcttta
 tccaggcctgtggcgggggaacagaaagatcacggcttcgaggtggccagcaccagccctgaggacgaatcaccaggaggcaacc
 ctgaaccagatgcaactccattccaggaggaggactgaggacctttgaccagctggatgctatctcaagcctgccactcctagtacatt
 ttcgtgtcttacagtaccttcccaggctttgtctcatggcgcgatcccaagtcaggagctggtagctggagacactggacgacatcttt
 gaacagtgggcccattcagaggacctgcagagcctgtctgtcgcgagtggcaaacgctgtctctgtgaagggaatctacaacagatg
 cccgggtgcttcaattttctgagaaagaaactgttctttaagacttcc (SEQ ID NO: 17019).

104. The composition of any one of claims 64-103, wherein the first transposon and/or the second transposon comprises at least one self-cleaving peptide.

105. The composition of claim 104, wherein the at least one self-cleaving peptide comprises a T2A peptide, a GSG-T2A peptide, an E2A peptide, a GSG-E2A peptide, an F2A peptide, a GSG-F2A peptide, a P2A peptide, or a GSG-P2A peptide.

106. The composition of claim 105, wherein the at least one self-cleaving peptide comprises a T2A peptide.

107. The composition of claim 105 or 106, wherein the T2A peptide comprises an amino acid sequence comprising EGRGSLLTCGDVEENPGP (SEQ ID NO: 17020).

108. The composition of claim 105 or 106, wherein the GSG-T2A peptide comprises an amino acid sequence comprising GSGEGRGSLLTCGDVEENPGP (SEQ ID NO: 17021).

109. The composition of claim 105 or 106, wherein the E2A peptide comprises an amino acid sequence comprising QCTNYALLKLAGDVESNPGP (SEQ ID NO: 17022).

110. The composition of claim 105 or 106, wherein the GSG-E2A peptide comprises an amino acid sequence comprising GSGQCTNYALLKLAGDVESNPGP (SEQ ID NO: 17023).
111. The composition of claim 105 or 106, wherein the F2A peptide comprises an amino acid sequence comprising VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 17024).
112. The composition of claim 105 or 106, wherein the GSG-F2A peptide comprises an amino acid sequence comprising GSGVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 17025).
113. The composition of claim 105 or 106, wherein the P2A peptide comprises an amino acid sequence comprising ATNFSLLKQAGDVEENPGP (SEQ ID NO: 17026).
114. The composition of claim 105 or 106, wherein the GSG-P2A peptide comprises an amino acid sequence comprising GSGATNFSLLKQAGDVEENPGP (SEQ ID NO: 17027).
115. The composition any one of claims 104-114, wherein, the at least one self-cleaving peptide is positioned between
- (a) the selection gene and the inducible transgene construct or
 - (b) the inducible transgene construct and the inducible caspase polypeptide.
116. The composition any one of claims 104-115, wherein, the at least one self-cleaving peptide is positioned between
- (a) the selection gene and the reporter construct or
 - (b) the reporter construct and the inducible caspase polypeptide.
117. A cell comprising the composition of any one of claims 1-116.
118. A method of inducing conditional gene expression in a cell comprising
- (a) contacting the cell with a composition of any one of claims 1-117, under conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter and

(b) contacting the exogenous receptor and a ligand that specifically binds thereto, to transduce an intracellular signal that targets the inducible promoter, thereby modifying gene expression.

119. The method of claim 118, wherein the cell is in vivo, ex vivo, in vitro or in situ.

120. The method of claim 118 or 119, wherein the cell is an immune cell.

121. The method of claim 120, wherein the immune cell is a T-cell, a Natural Killer (NK) cell, a Natural Killer (NK)-like cell, a hematopoietic progenitor cell, a peripheral blood (PB) derived T cell or an umbilical cord blood (UCB) derived T-cell.

122. The method of claim 121, wherein the immune cell is a T-cell.

123. The method of any one of claims 118-122, wherein the cell is autologous.

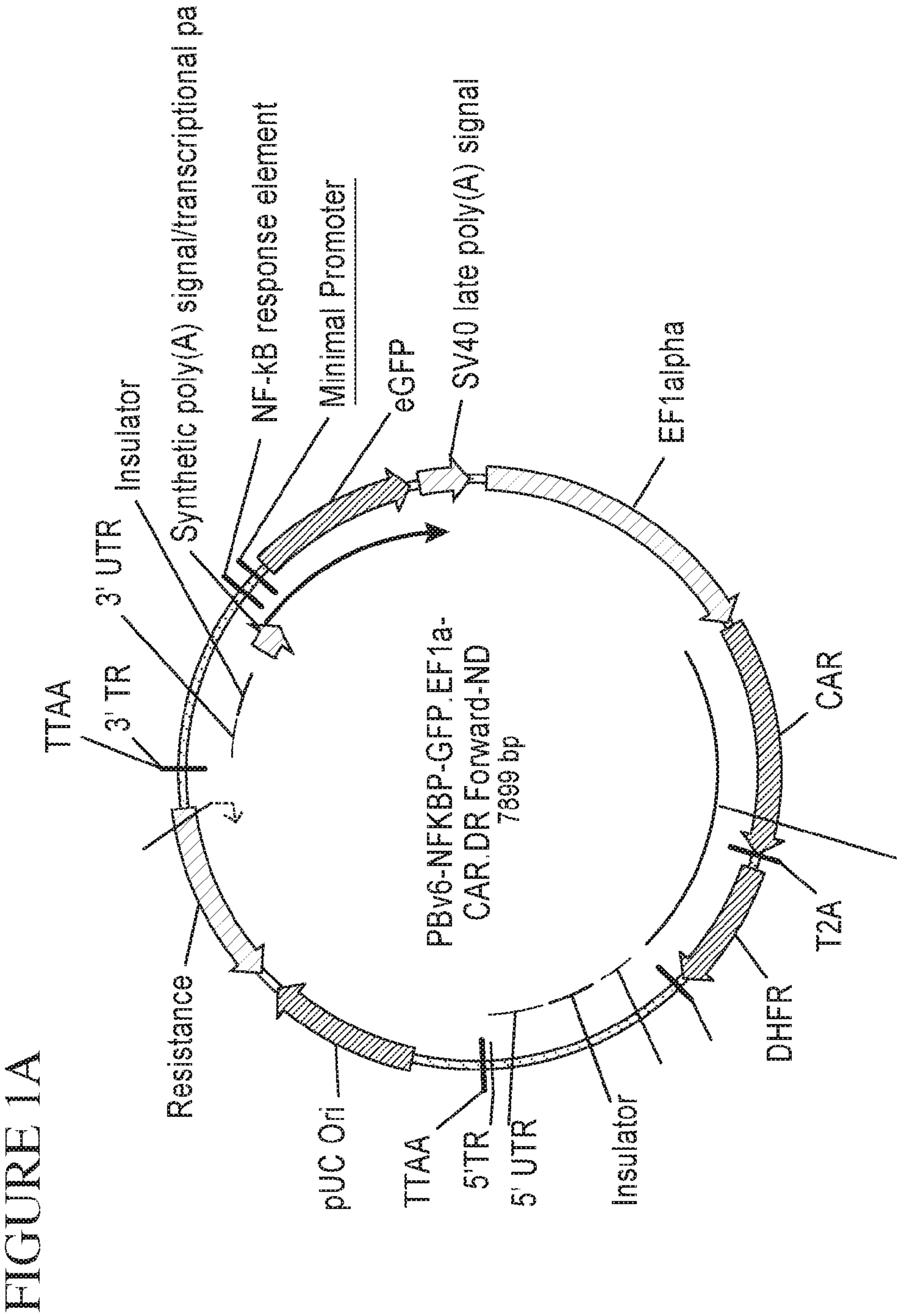
124. The method of any one of claims 118-122, wherein the cell is allogeneic.

125. The method of any one of claims 118-122, wherein conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter comprise in vivo conditions.

126. The method of any one of claims 118-122, wherein conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter comprise a temperature substantially similar to an internal temperature of a human body, a CO₂ level substantially similar to an internal CO₂ levels of a human body, an O₂ level substantially similar to an internal O₂ levels of a human body, an aqueous or saline environment with a level of electrolytes substantially similar to a level of electrolytes of an interior of a human body.

127. The method of any one of claims 118-122, wherein the ligand is non-naturally occurring.

128. A method of treating a disease or disorder in a subject in need thereof, comprising administering to the subject a composition of any one of claims 1-115, under conditions suitable to allow for integration of the inducible transgene construct into the genome of the cell and for the expression of the exogenous reporter, and administering a ligand to which the exogenous receptor selectively binds, wherein the binding of the ligand to the exogenous receptor transduces an intracellular signal to target the inducible promoter controlling the transgene, wherein the transgene is expressed, and wherein the product of the transgene is therapeutically-effective for treating the disease or disorder.
129. The method of claim 128, wherein the product of the transgene is a secreted protein.
130. The method of claim 129, wherein the secreted protein is a clotting factor.
131. The method of claim 130, wherein the clotting factor is factor IX.
132. The method of any one of claims 128-131, wherein the disease or disorder is a clotting disorder.



pNFkB-GFP & EF1a-CAR in the same direction

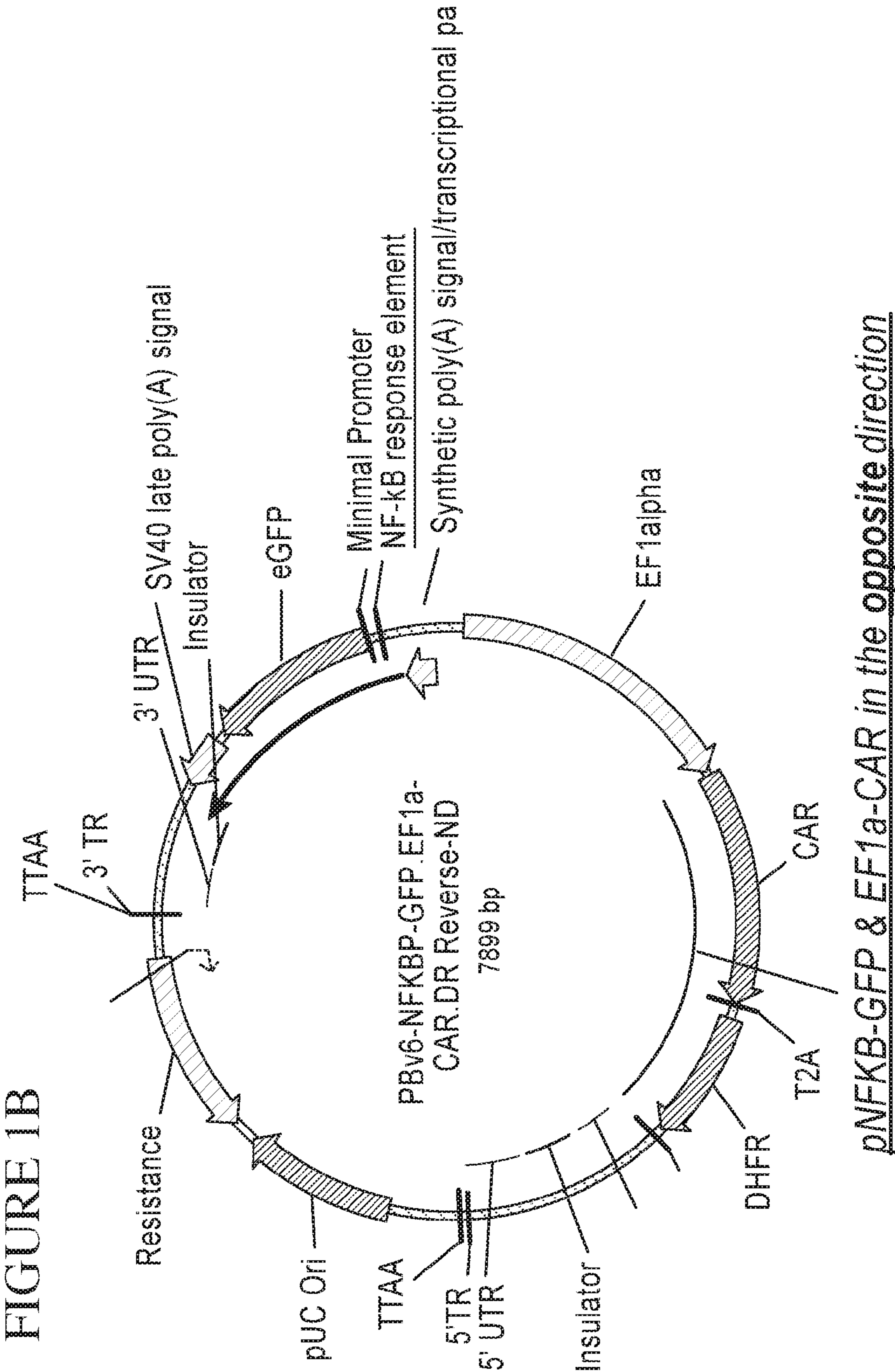
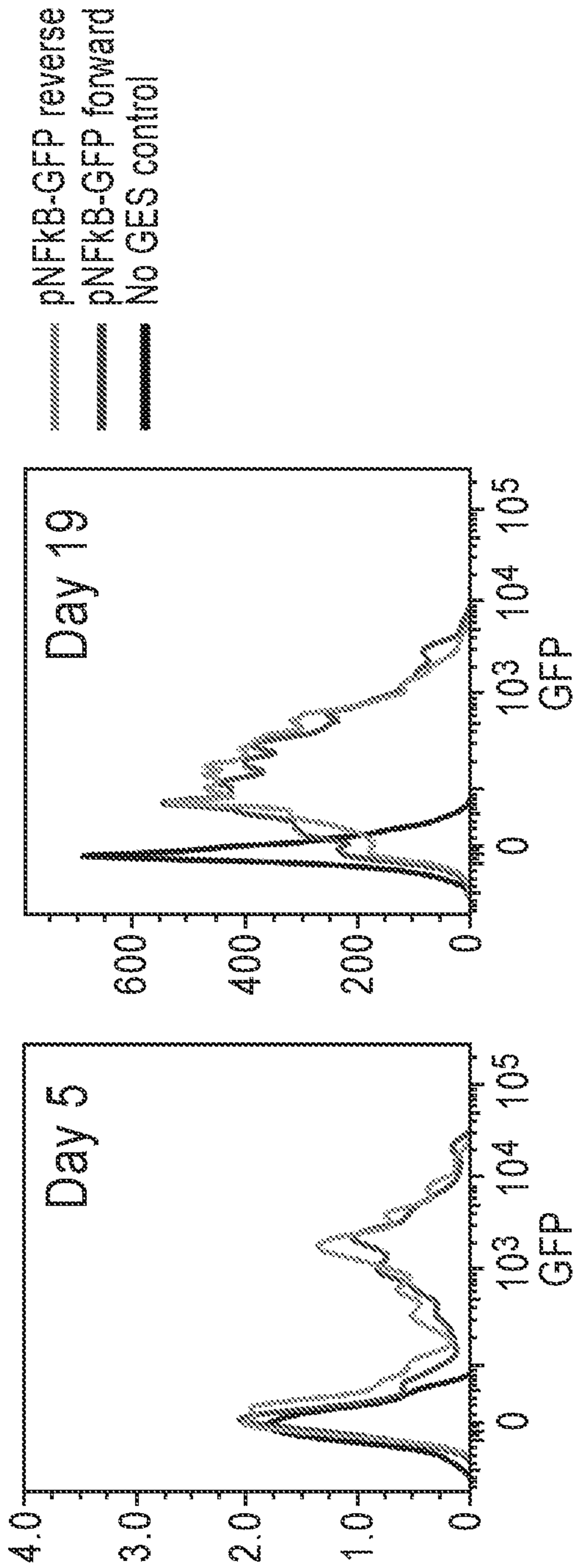


FIGURE 2



- At Day 5-post T cell activation (Day 5), T cells are proliferating and are highly stimulated; GFP expression is strong since NFkB activity is high
- By Day 19, T cells are almost fully resting; GFP expression is significantly lower than Day 5 (~1/8 MFI), since NFkB activity is low
- GFP expression is still observed at Day 19, which may be due to the long half life of GFP protein (~30hr), or, basal level of NFkB activity through TCR, CAR, cytokine, growth factor, etc...

FIGURE 3

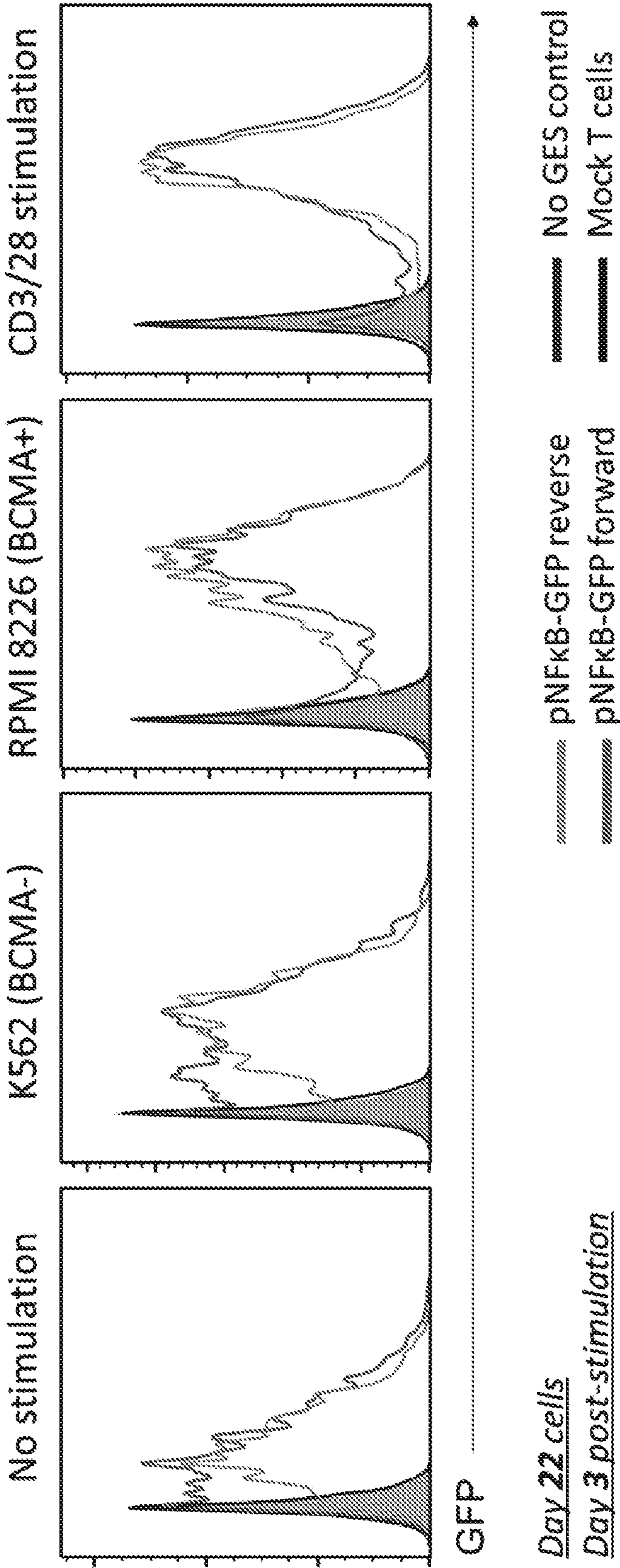


FIGURE 4

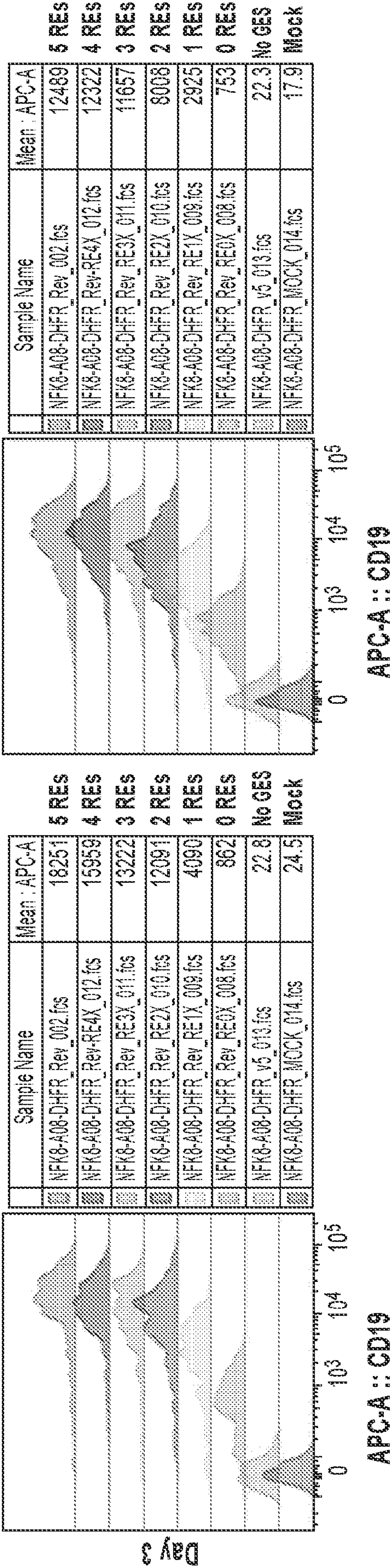
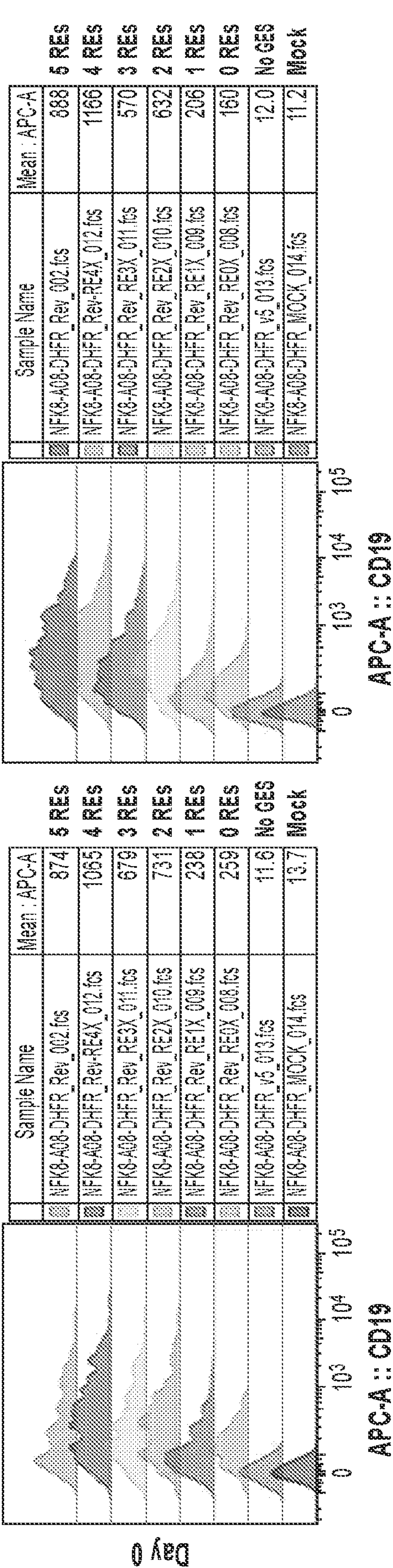


FIGURE 4 (cont.)

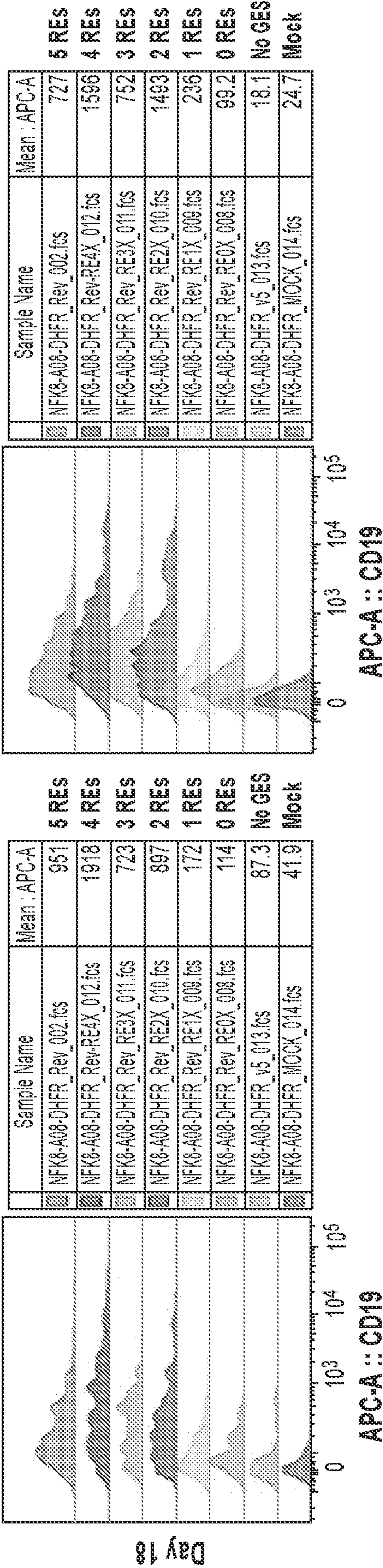


FIGURE 4 (cont.)

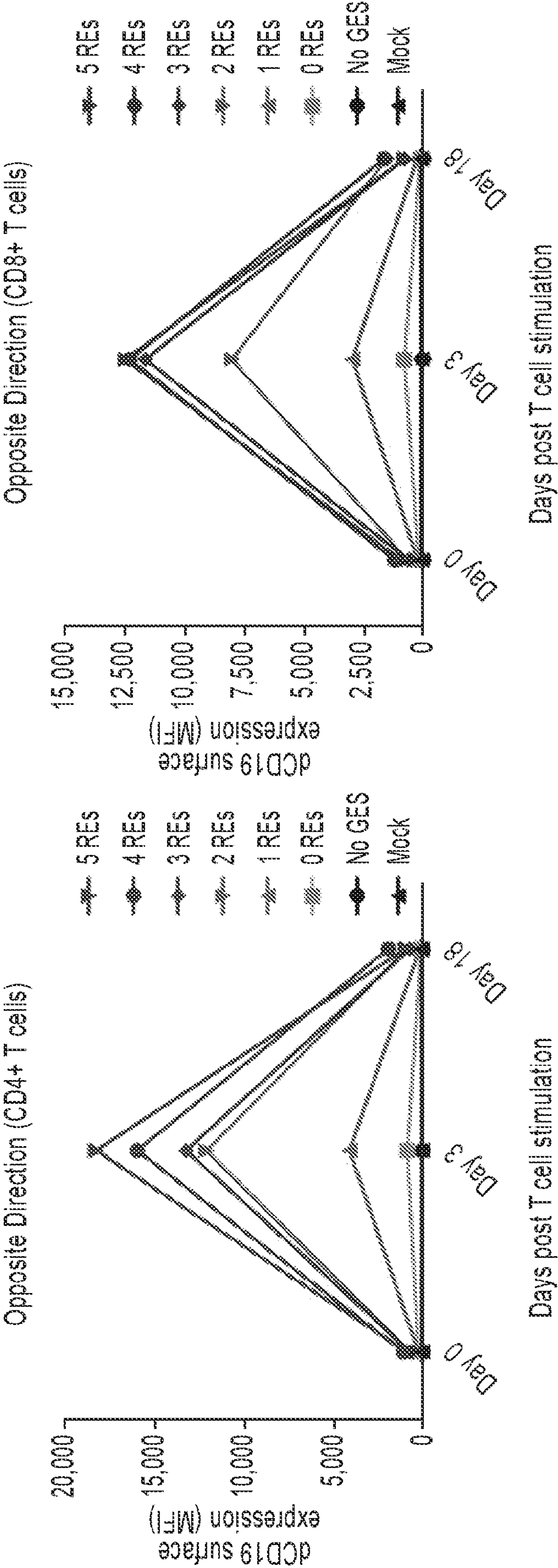


FIGURE 5

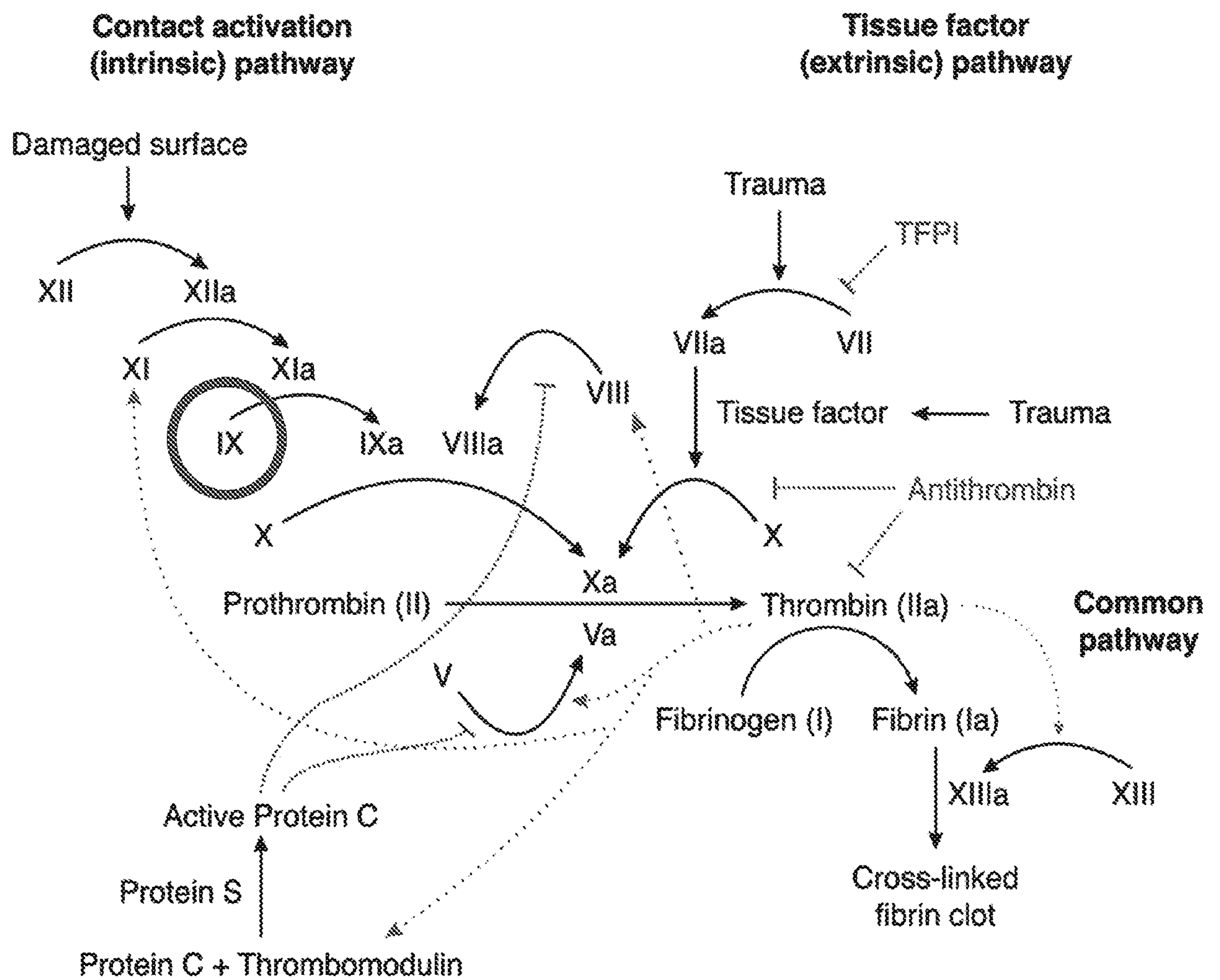
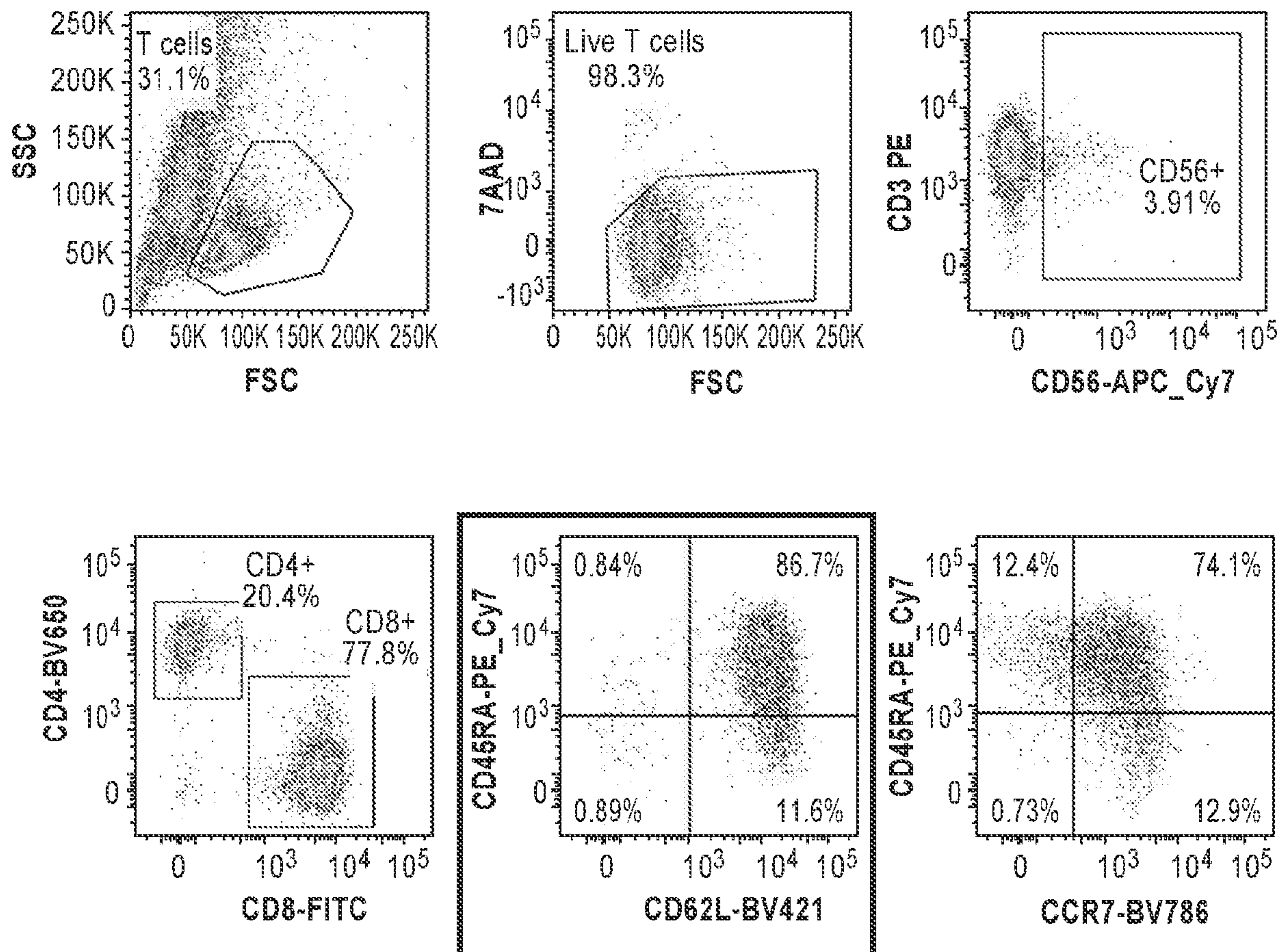


FIGURE 6



10/20

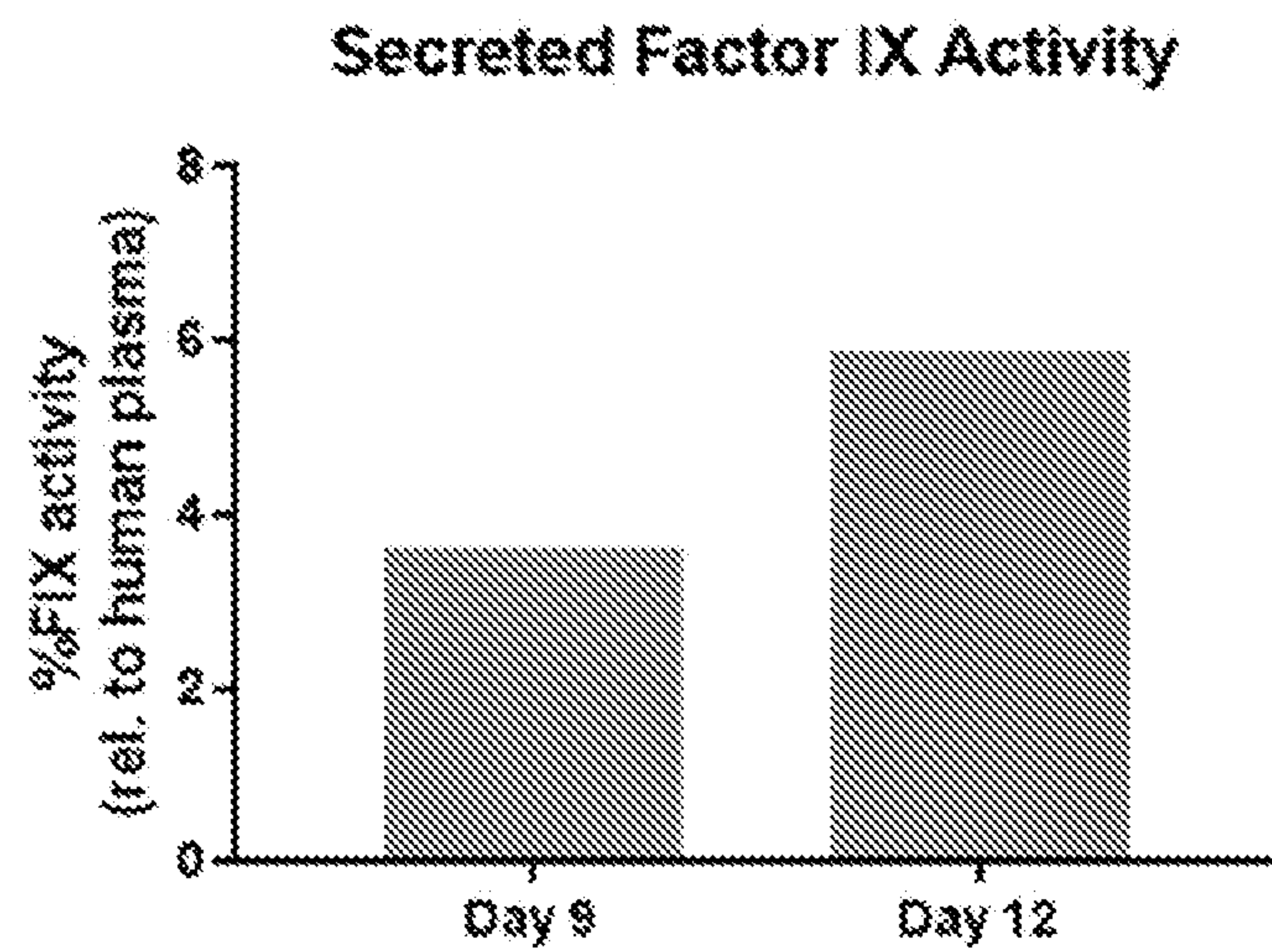
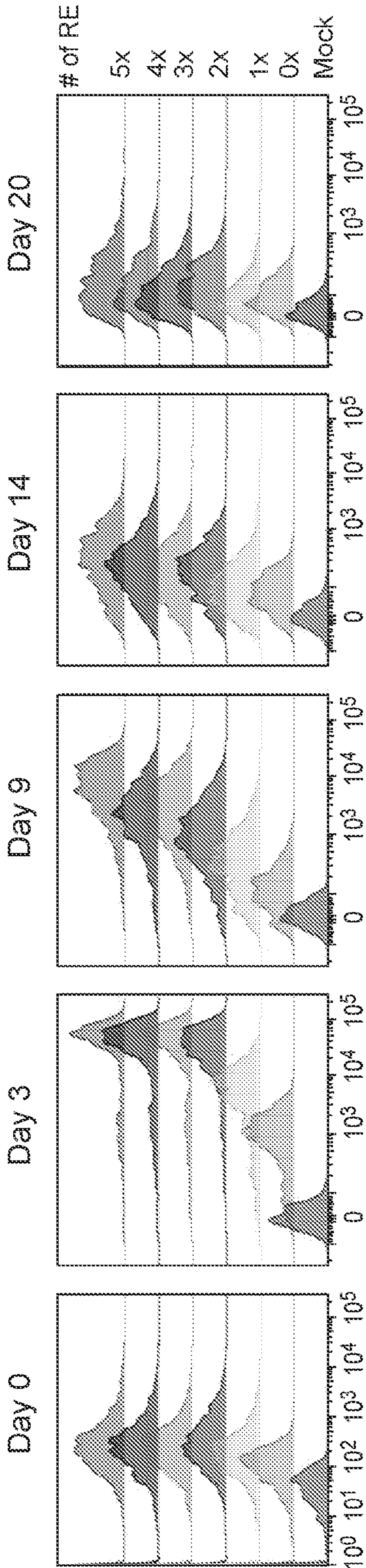
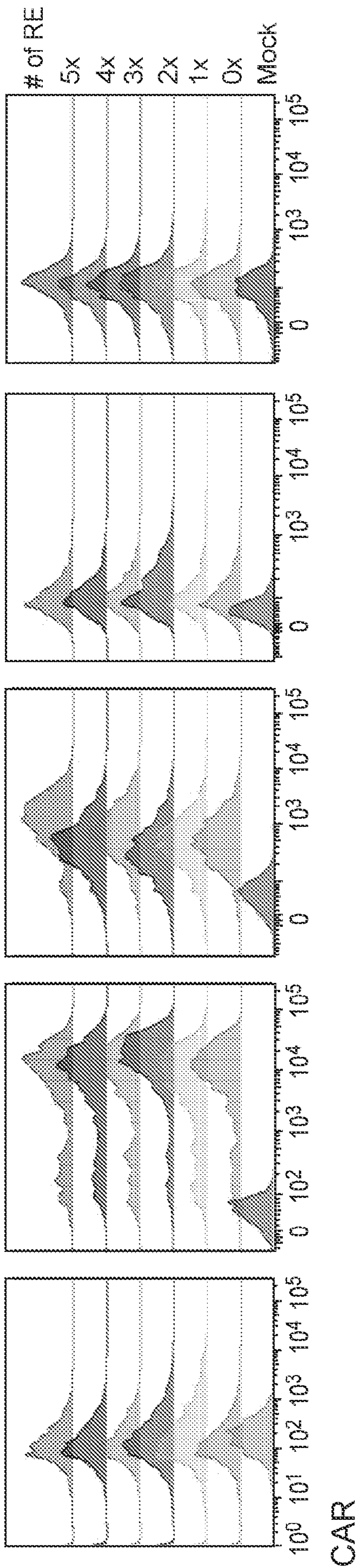
FIGURE 7A**FIGURE 7B**

FIGURE 8

CD4+ Cells



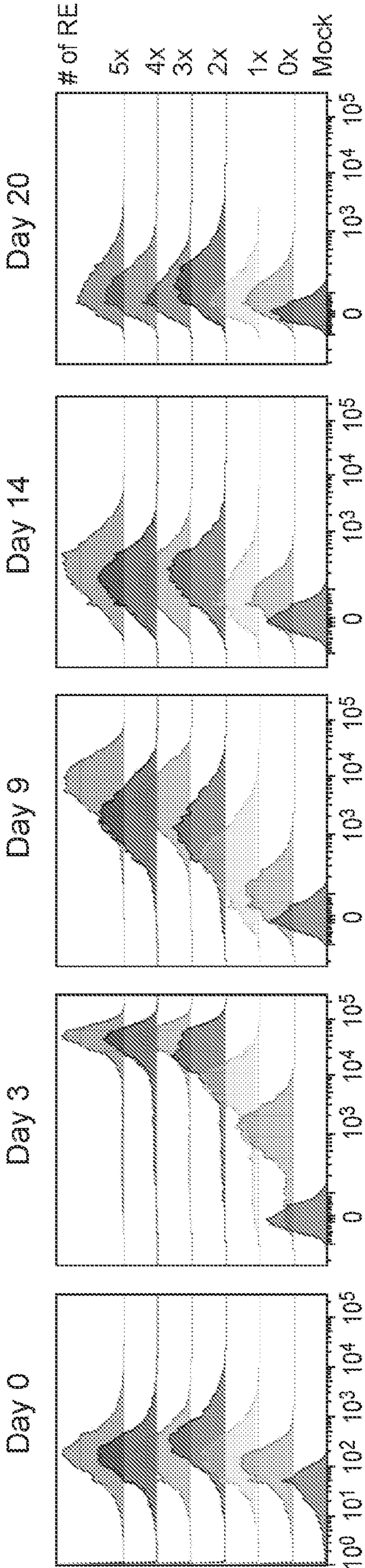
dCD19



CAR

FIGURE 9

CD8+ Cells



dCD19

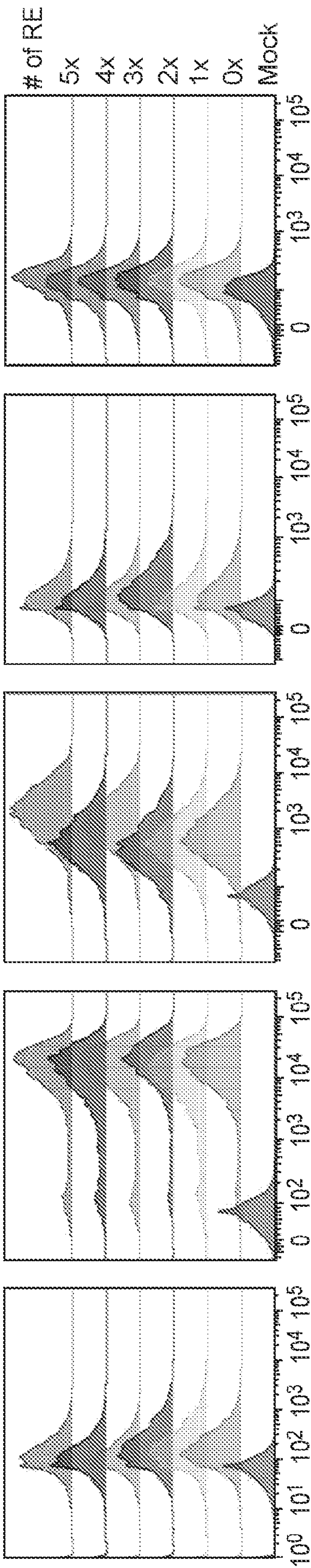


FIGURE 10

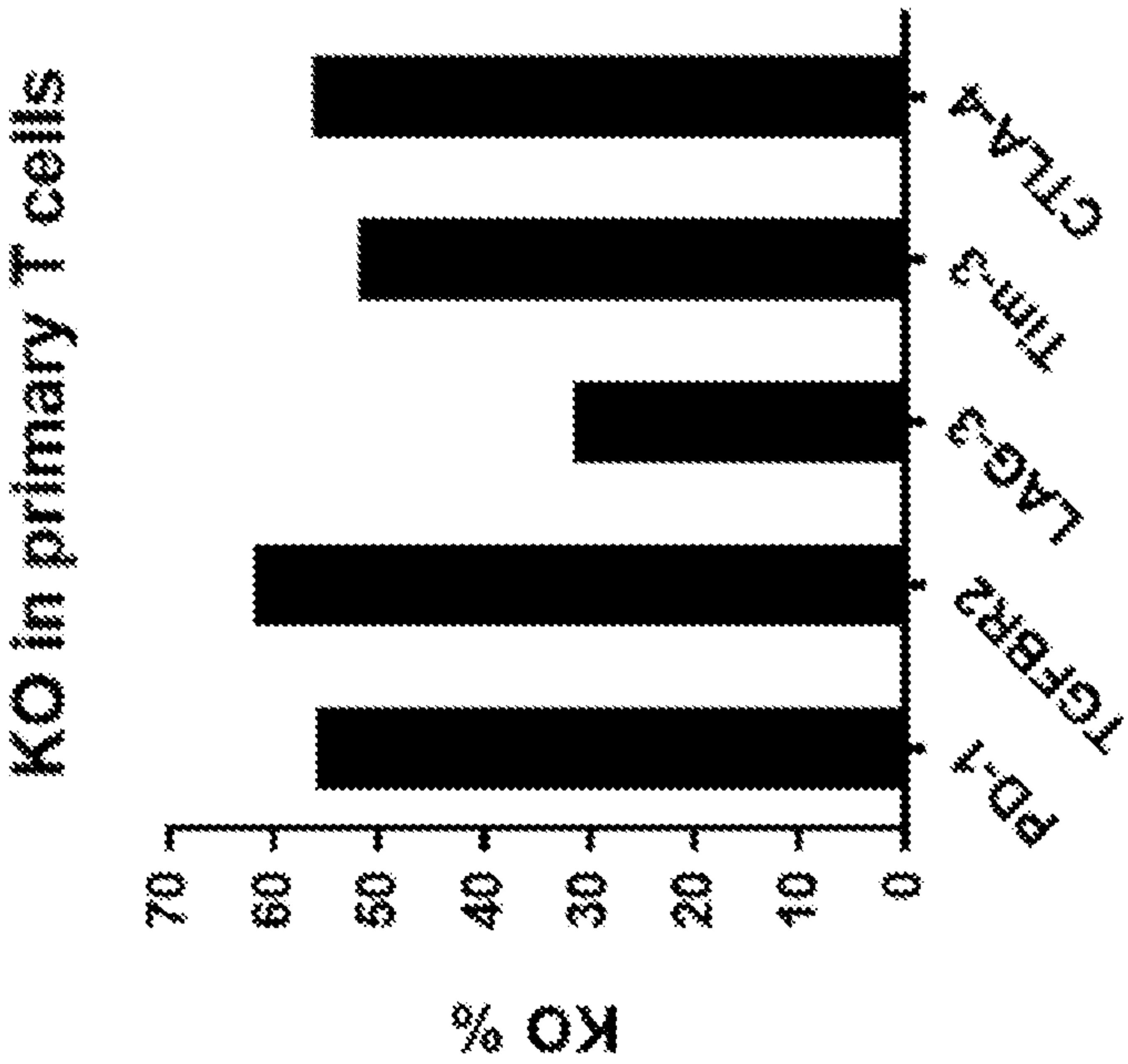


FIGURE 11

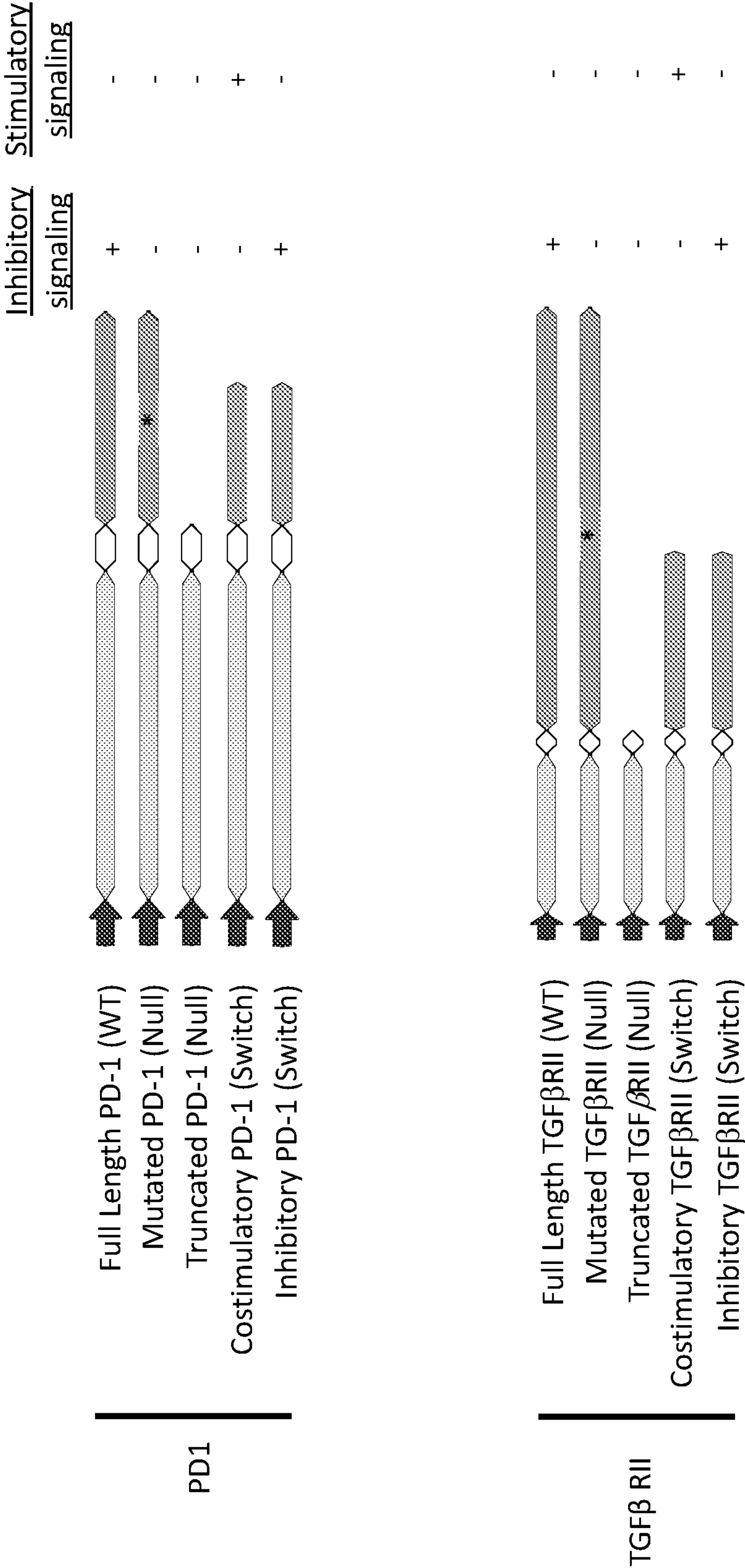


FIGURE 12

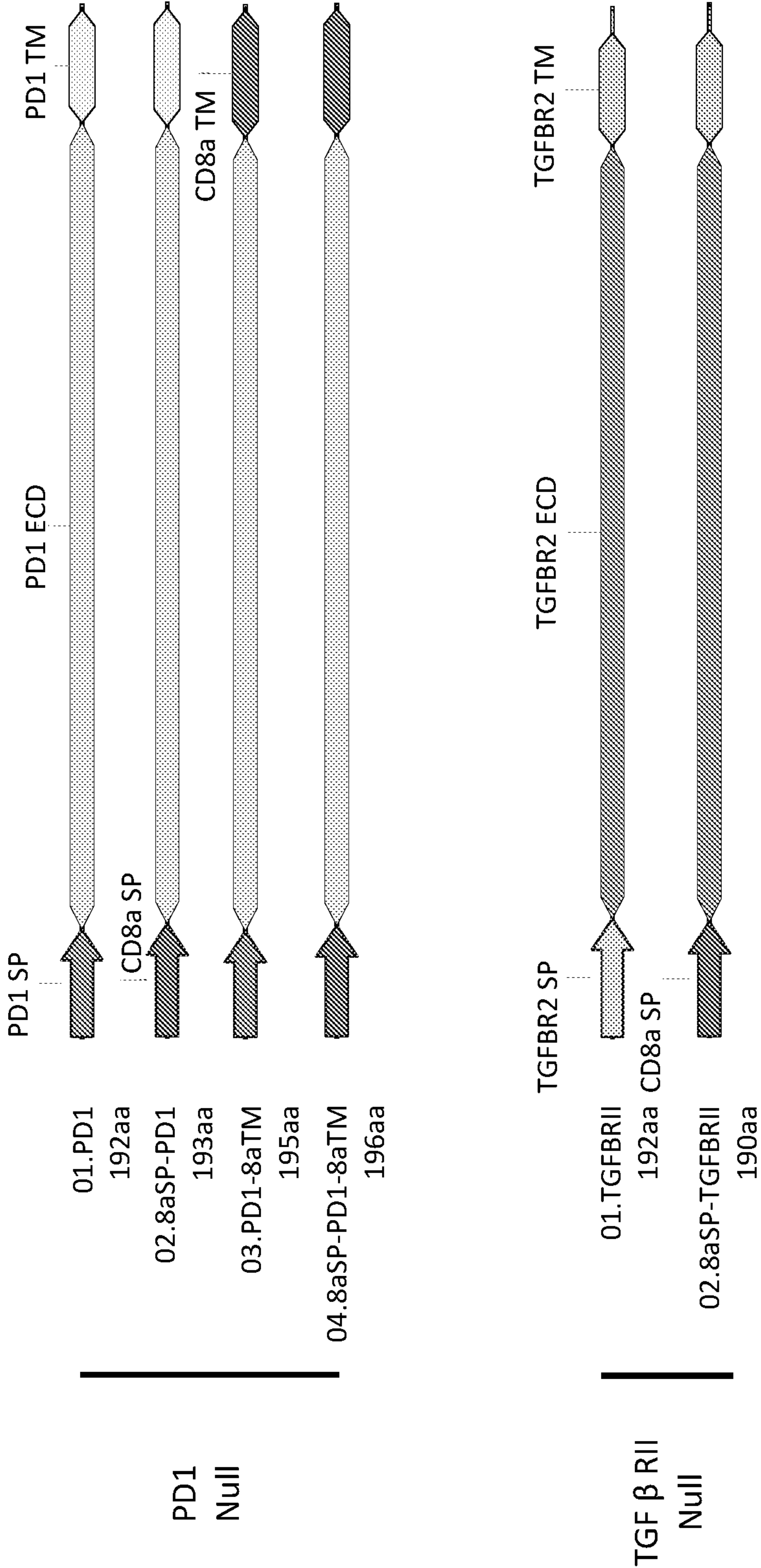


FIGURE 13

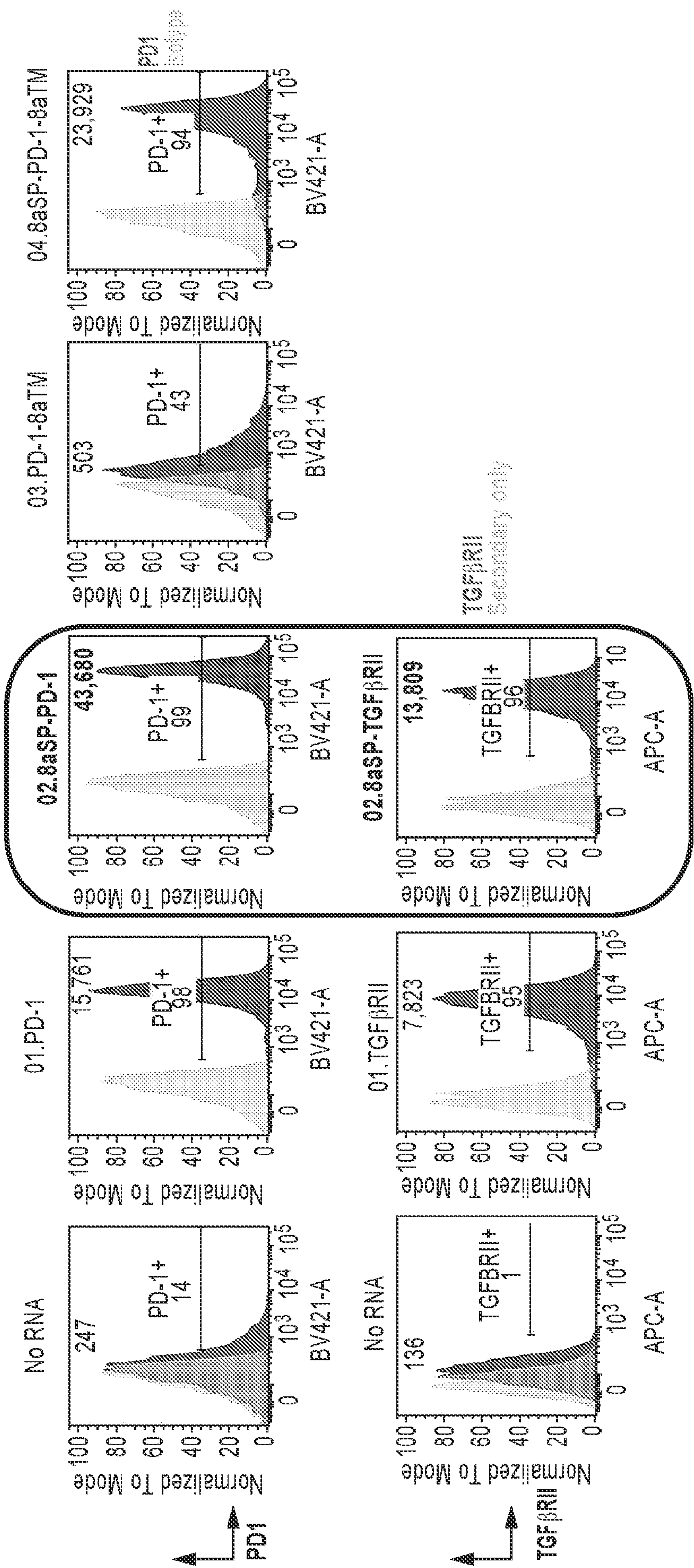


FIGURE 14

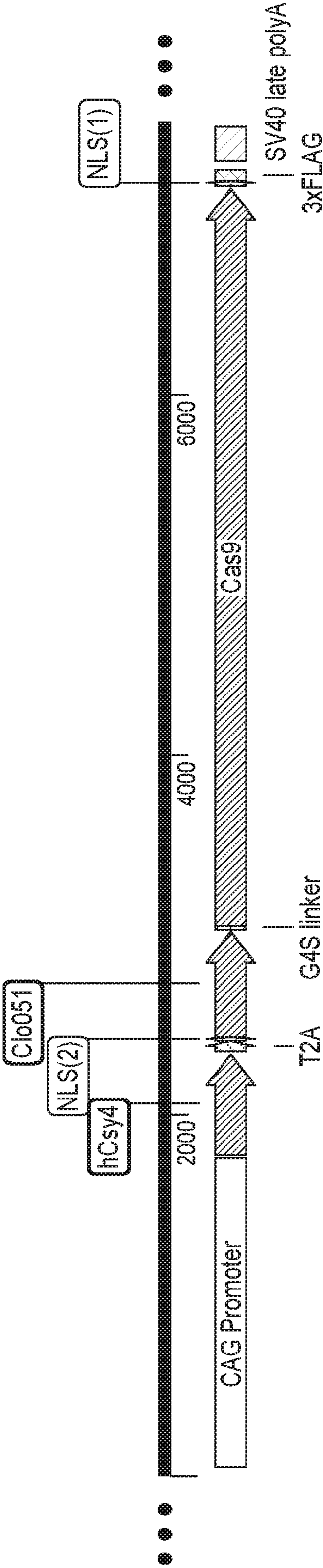
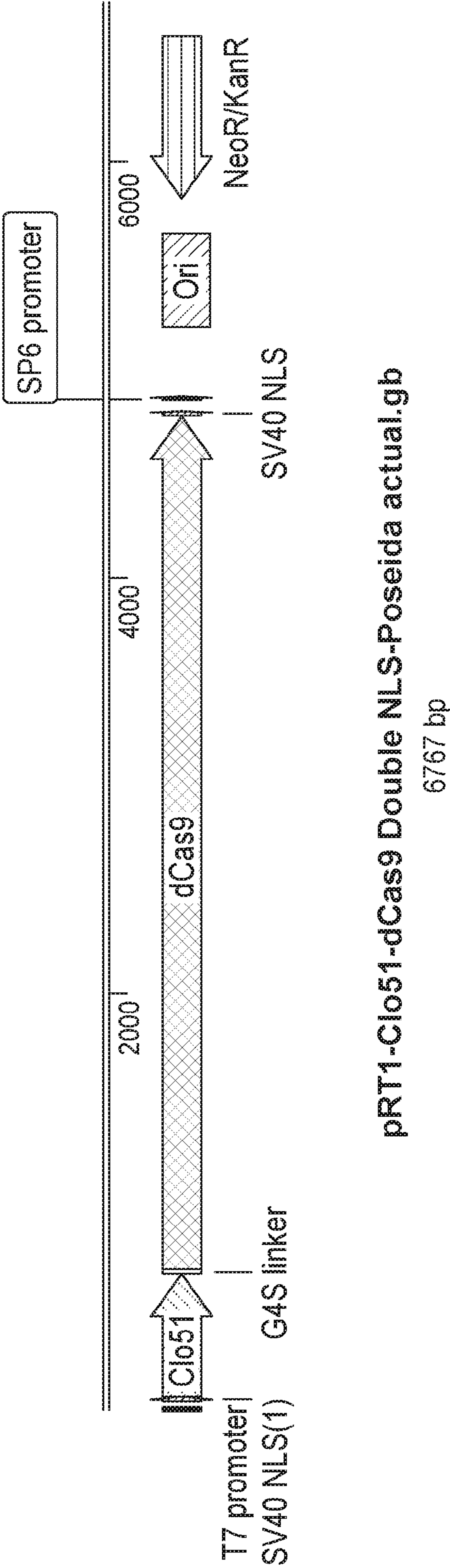


FIGURE 15



pRT1-Clo51-dCas9 Double NLS-Poseida actual.gb

6767 bp

FIGURE 16
Cas-Clover mRNA Comparison
Embodiment 1 vs Embodiment 2

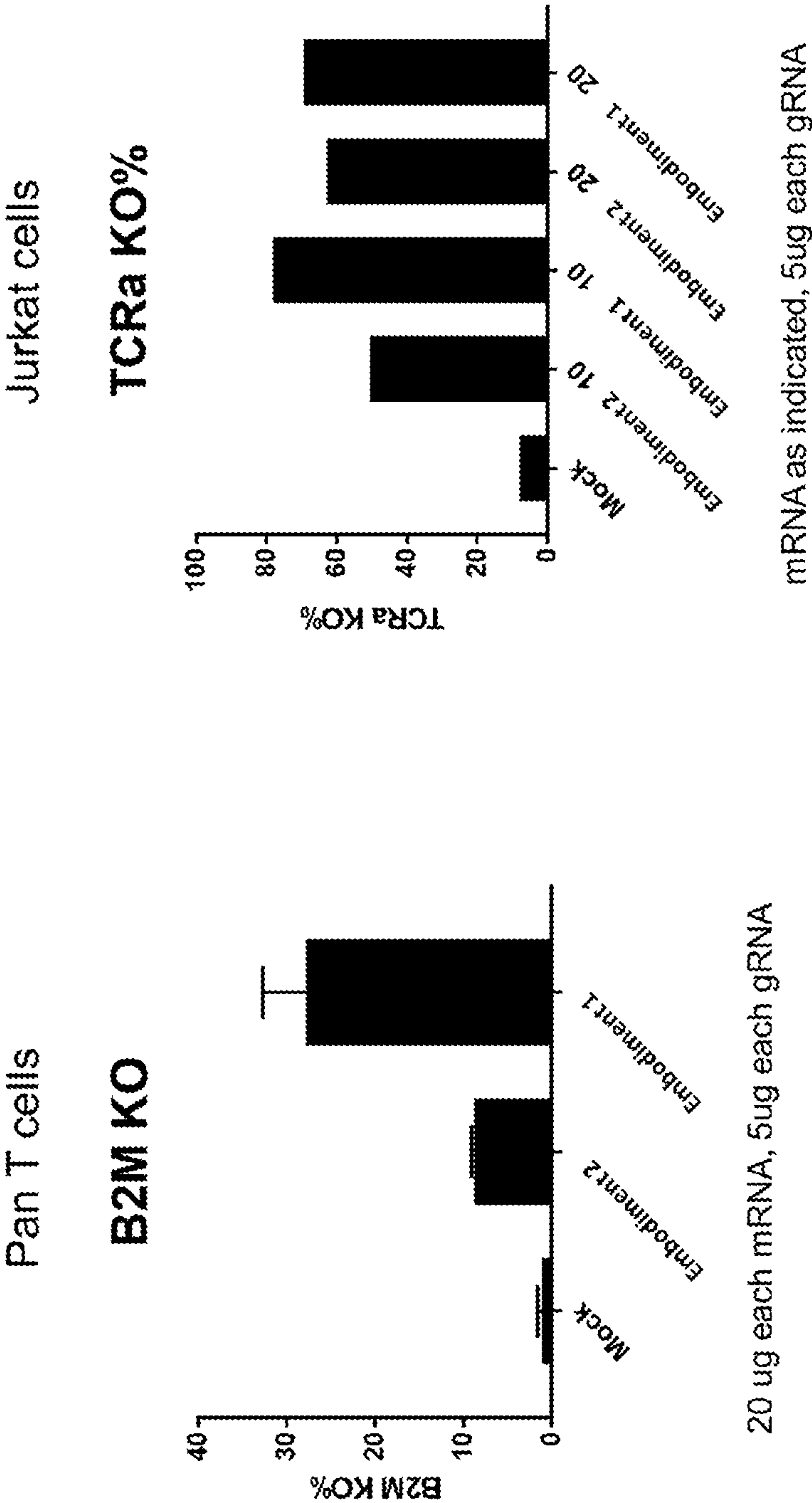


FIGURE 17

Cas-Clover mRNA Comparison Embodiment 1 v. Embodiment 2

mRNA

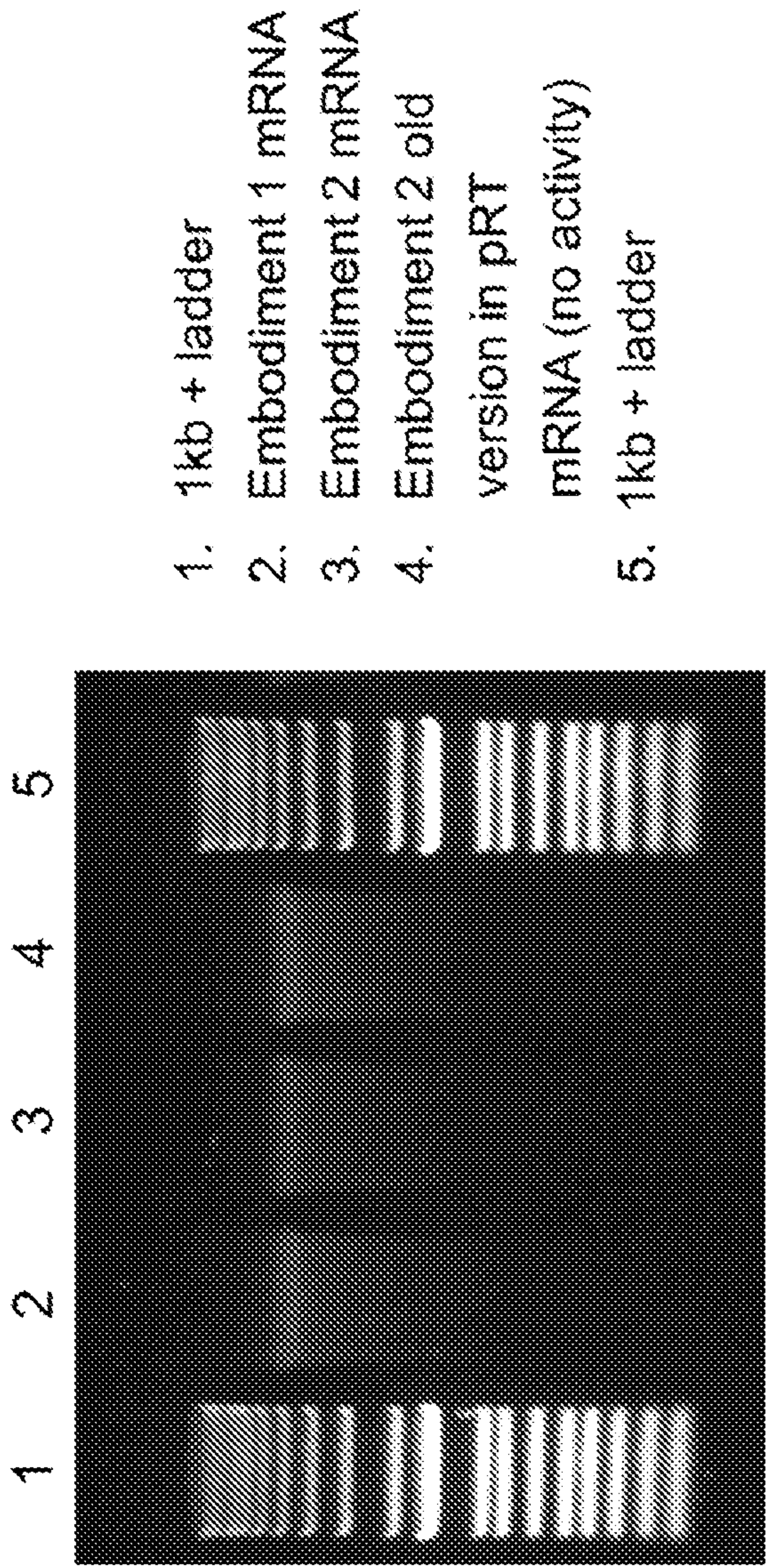
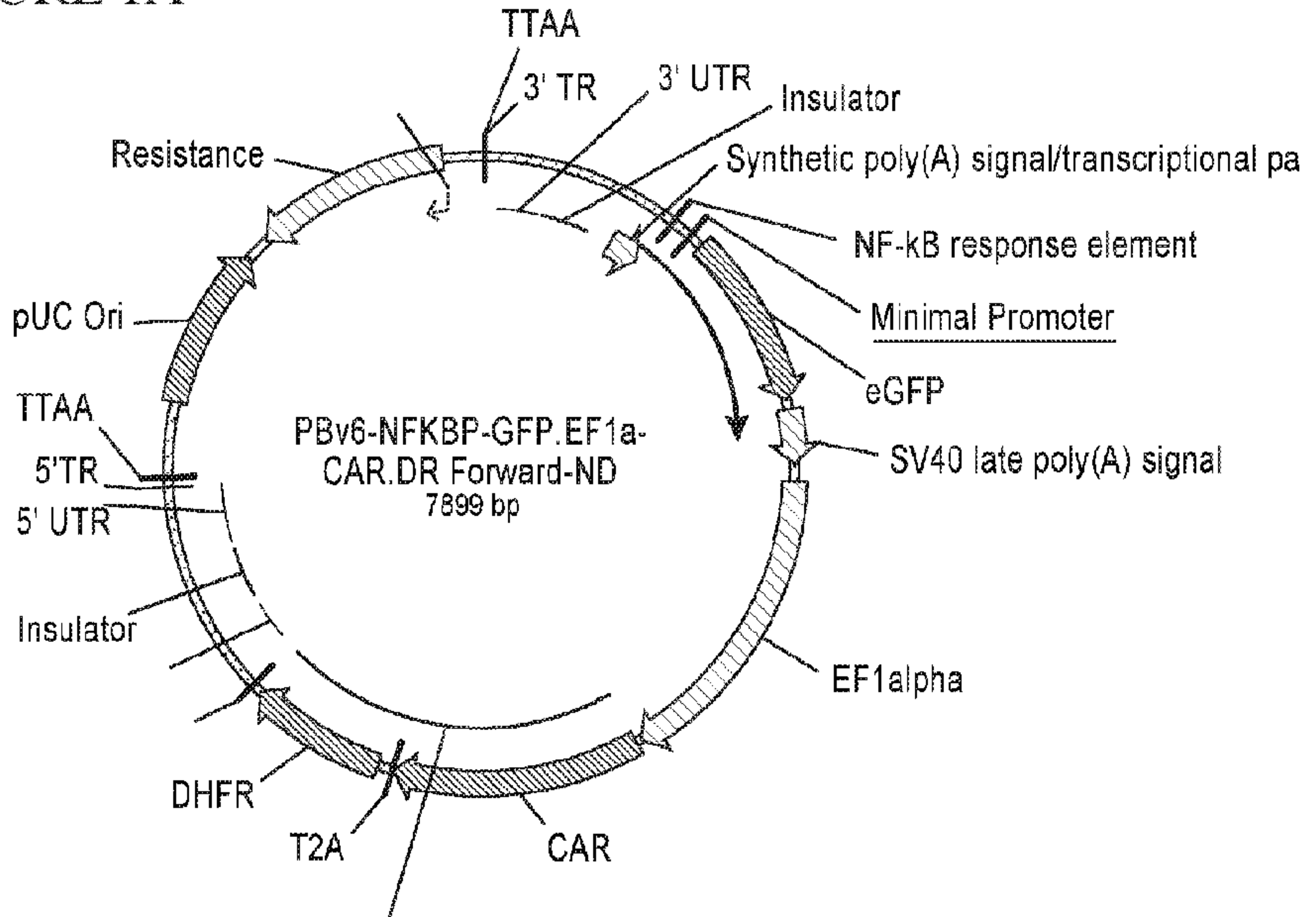


FIGURE 1A



pNFKB-GFP & EF1a-CAR in the same direction