



(51) International Patent Classification:

A61K 48/00 (2006.01) C12N 15/63 (2006.01)
C12N 15/113 (2010.01) A61P 35/00 (2006.01)
C07K 19/00 (2006.01)

(21) International Application Number:

PCT/US2018/031674

(22) International Filing Date:

08 May 2018 (08.05.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/503,060 08 May 2017 (08.05.2017) US
62/579,460 31 October 2017 (31.10.2017) US

(71) Applicant: **PRECISION BIOSCIENCES, INC.**
[US/US]; 302 E. Pettigrew Street, Dibrell Building, Suite
A-100, Durham, NC 27701 (US).

(72) Inventors: **HUX, Joann**; 853 Poor Rd, Scotland Neck, NC
27874 (US). **MARTIN, Aaron**; 208 Wild Oak Lane, Car-
rboro, NC 27510 (US). **JANTZ, Derek**; 1011 Gloria Av-
enue, Durham, NC 27701 (US). **BEARD, Clayton**; 3534
Rugby Road, Durham, NC 27707 (US).

(74) Agent: **TWOMEY, Michael, J.**; Wolf, Greenfield &
Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206
(US).

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

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

(54) Title: NUCLEIC ACID MOLECULES ENCODING AN ENGINEERED ANTIGEN RECEPTOR AND AN INHIBITORY NUCLEIC ACID MOLECULE AND METHODS OF USE THEREOF

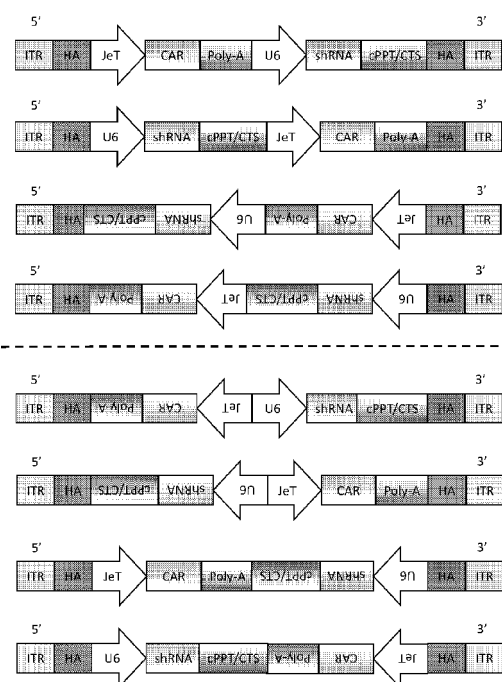


FIGURE 1

(57) Abstract: The present disclosure provides nucleic acid molecules encoding an engineered antigen receptor, such as a chimeric antigen receptor or exogenous T cell receptor, and an inhibitory nucleic acid molecule, such as an RNA interference molecule. The present disclosure further relates to nucleic acids, DNA constructs, vectors, pharmaceutical compositions, genetically-modified cells, and methods of treatment that utilize the nucleic acid molecules of the invention.

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

**NUCLEIC ACID MOLECULES ENCODING AN ENGINEERED ANTIGEN
RECEPTOR AND AN INHIBITORY NUCLEIC ACID MOLECULE AND
METHODS OF USE THEREOF**

FIELD OF THE INVENTION

[0001] The present disclosure relates to the field of molecular biology and recombinant nucleic acid technology. In particular, the present disclosure relates to nucleic acid molecules encoding an engineered antigen receptor, such as a chimeric antigen receptor or exogenous T cell receptor, and an inhibitory nucleic acid molecule, such as an RNA interference molecule. The present disclosure further relates to nucleic acids, DNA constructs, viral vectors, pharmaceutical compositions, genetically-modified cells, and methods of treatment that utilize the nucleic acid molecule of the invention.

REFERENCE TO A SEQUENCE LISTING SUBMITTED AS
A TEXT FILE VIA EFS-WEB

[0002] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on May 8, 2018, is named P109070023WO00-SEQ-MJT.TXT, and is 188579 bytes in size.

BACKGROUND OF THE INVENTION

[0003] T cell adoptive immunotherapy is a promising approach for cancer treatment. This strategy utilizes isolated human T cells that have been genetically-modified to enhance their specificity for a specific tumor associated antigen. Genetic modification may involve the expression of a chimeric antigen receptor (CAR) or an exogenous T cell receptor to graft antigen specificity onto the T cell. By contrast to exogenous T cell receptors, CARs derive their specificity from the variable domains of a monoclonal antibody. Thus, T cells expressing CARs induce tumor immunoreactivity in a major histocompatibility complex (MHC) non-restricted manner. To date, T cell adoptive immunotherapy has been utilized as a clinical therapy for a number of cancers, including B cell malignancies (*e.g.*, acute lymphoblastic leukemia (ALL), B cell non-Hodgkin lymphoma (NHL), and chronic lymphocytic leukemia), multiple myeloma, neuroblastoma, glioblastoma, advanced gliomas, ovarian cancer, mesothelioma, melanoma, and pancreatic cancer.

[0004] Despite its potential usefulness as a cancer treatment, adoptive immunotherapy has been limited, in part, by alloreactivity between host tissues and allogeneic CAR T cells. One cause of alloreactivity arises from the presence of non-host MHC class I molecules on the cell surface of CAR T cells. MHC class I molecules consist of two polypeptide chains, α and β . In humans, the α chain consists of three subunits, $\alpha 1$, $\alpha 2$, and $\alpha 3$, which are encoded by polymorphic human leukocyte antigen (HLA) genes on chromosome 6. The variability of HLA loci, and the encoded α chain subunits, can cause allogeneic CAR T cells to be seen by the host immune system as foreign cells because they bear foreign MHC class I molecules. As a result, CAR T cells administered to a patient can be subject to host versus graft (HvG) rejection, where they are recognized and killed by the host's cytotoxic T cells.

[0005] The β chain of MHC class I molecules consists of beta-2 microglobulin, which is encoded by the non-polymorphic beta-2 microglobulin (B2M) gene on chromosome 15 (SEQ ID NO: 1). Beta-2 microglobulin is non-covalently linked to the $\alpha 3$ subunit and is common to all MHC class I molecules. Furthermore, expression of MHC class I molecules at the cell surface requires its association with beta-2 microglobulin. As such, beta-2 microglobulin represents a logical target for suppressing the expression of MHC class I molecules on CAR T cells, which could render the cells invisible to host cytotoxic T cells and reduce alloreactivity. However, complete knockout of beta-2 microglobulin expression may result in NK cell killing of CAR T cells due to the lack of cell surface MHC class I molecules, which could prompt NK cells to recognize them as non-self and initiate cytotoxic action.

[0006] Another cause of alloreactivity to CAR T cells is the expression of the endogenous T cell receptor on the cell surface. T cell receptors typically consist of variable α and β chains or, in smaller numbers, variable γ and δ chains. The T cell receptor complexes with accessory proteins, including CD3, and functions with cell surface co-receptors (*e.g.*, CD4 and CD8) to recognize antigens bound to MHC molecules on antigen presenting cells. In the case of allogeneic CAR T cells, expression of endogenous T cell receptors may cause the cell to recognize host MHC antigens following administration to a patient, which can lead to the development of graft-versus-host-disease (GVHD).

[0007] To forestall alloreactivity, clinical trials have largely focused on the use of autologous CAR T cells, wherein a donor's T cells are isolated, genetically-modified to incorporate a chimeric antigen receptor, and then re-infused into the same subject. An autologous approach provides immune tolerance to the administered CAR T cells; however,

this approach is constrained by both the time and expense necessary to produce patient-specific CAR T cells after a patient's cancer has been diagnosed.

[0008] Thus, a need exists in the art for the development of allogeneic CAR T cells which exhibit reduced allogenicity but, at the same time, avoid NK cell killing *in vivo*.

SUMMARY OF THE INVENTION

[0009] In one aspect, the invention provides a nucleic acid molecule comprising: (a) a first expression cassette comprising a nucleic acid sequence encoding an engineered antigen receptor; (b) a second expression cassette comprising a nucleic acid sequence encoding an inhibitory nucleic acid molecule; (c) a 5' homology arm; and (d) a 3' homology arm; wherein the 5' homology arm and the 3' homology arm have homology to chromosomal regions flanking a nuclease recognition sequence in a gene of interest.

[0010] In some embodiments, the inhibitory nucleic acid molecule is an RNA interference molecule. In certain embodiments, the RNA interference molecule is a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a hairpin siRNA, a microRNA (miRNA), a precursor miRNA, or an miRNA-adapted shRNA. In particular embodiments, the RNA interference molecule is an shRNA.

[0011] In some embodiments the engineered antigen receptor is a chimeric antigen receptor. In other embodiments, the engineered antigen receptor is an exogenous T cell receptor.

[0012] In some embodiments, the nuclease recognition sequence is an engineered meganuclease recognition sequence, a TALEN recognition sequence, a zinc finger nuclease (ZFN) recognition sequence, a CRISPR/Cas recognition sequence, a compact TALEN recognition sequence, or a megaTAL recognition sequence. In certain embodiments, the nuclease recognition sequence is an engineered meganuclease recognition sequence.

[0013] In some embodiments, the gene of interest is any gene of interest. In certain embodiments, the gene of interest is a human T cell receptor alpha constant region gene. In particular embodiments the nuclease recognition sequence is an engineered meganuclease recognition sequence. In certain embodiments, the engineered meganuclease recognition sequence comprises SEQ ID NO: 1 in a human T cell receptor alpha constant region gene.

[0014] In some embodiments, the first expression cassette further comprises a promoter which drives expression of the engineered antigen receptor. In certain embodiments, the promoter is a JeT promoter.

[0015] In some embodiments, the second expression cassette further comprises a promoter which drives expression of the inhibitory nucleic acid molecule. In certain embodiments, the promoter is a U6 promoter.

[0016] In some embodiments, the first expression cassette comprises a polyadenylation signal to terminate translation of the engineered antigen receptor. In some embodiments, the second expression cassette comprises a central polypurine tract and central terminator sequence (cPPT/CTS) sequence to terminate translation of the inhibitory nucleic acid.

[0017] In some embodiments, the first expression cassette and the second expression cassette are in the same orientation in the nucleic acid molecule. In certain embodiments, the first expression cassette and the second expression cassette are in a 5' to 3' orientation relative to the 5' and 3' homology arms. In some such embodiments, the first expression cassette is 5' upstream of the second expression cassette. In other such embodiments, the second expression cassette is 5' upstream of the first expression cassette.

[0018] In some embodiments, wherein the first expression cassette and the second expression cassette are in the same orientation in the nucleic acid molecule, the first expression cassette and the second expression cassette are in a 3' to 5' orientation relative to the 5' and 3' homology arms. In some such embodiments, the first expression cassette is 5' upstream of the second expression cassette. In other such embodiments, the second expression cassette is 5' upstream of the first expression cassette.

[0019] In some embodiments, the first expression cassette and the second expression cassette are in opposite orientations in the nucleic acid molecule. In some such embodiments, the first expression cassette is in a 3' to 5' orientation and the second expression cassette is in a 5' to 3' orientation relative to the 5' and 3' homology arms. In certain embodiments, the first expression cassette is 5' upstream of the second expression cassette. In other embodiments, the second expression cassette is 5' upstream of the first expression cassette.

[0020] In particular embodiments, wherein the first expression cassette and the second expression cassette are in opposite orientations in the nucleic acid molecule, the first expression cassette is in a 5' to 3' orientation and the second expression cassette is in a 3' to 5' orientation relative to the 5' and 3' homology arms. In some such embodiments, the first expression cassette is 5' upstream of the second expression cassette. In other such embodiments, the second expression cassette is 5' upstream of the first expression cassette.

[0021] In some embodiments, the nucleic acid molecule comprises multiple copies of the second expression cassette. In some such embodiments, the copies are identical. In further embodiments, the copies include a promoter, a coding sequence for the inhibitory nucleic

acid molecule, and a sequence, such as a (cPPT/CTS) sequence, to terminate translation of the inhibitory nucleic acid molecule. In some such embodiments, the copies of the second expression cassette are in tandem in the nucleic acid molecule, and can be in the same orientation, or in opposite orientations. In other such embodiments, the copies may not be in tandem, and can be in the same orientation, or in opposite orientations.

[0022] In some embodiments, the nucleic acid molecule further comprises a 5' inverted terminal repeat and a 3' inverted terminal repeat flanking the first expression cassette and the second expression cassette.

[0023] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin.

[0024] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule. In certain embodiments, the inhibitory molecule is inhibitory against an MHC class I alpha-1 ($\alpha 1$) domain, alpha-2 ($\alpha 2$) domain, alpha-3 ($\alpha 3$) domain, or against beta-2 microglobulin.

[0025] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human CD52.

[0026] In certain embodiments, the inhibitory nucleic acid molecule is an shRNA inhibitory against beta-2 microglobulin, wherein the shRNA has a sequence comprising any one of SEQ ID NOs: 2-4. In particular embodiments, the shRNA has a sequence comprising SEQ ID NO: 2. In some such embodiments, the first expression cassette and the second expression cassette are in a 3' to 5' orientation relative to the 5' and 3' homology arms, and wherein the first expression cassette is 5' upstream of the second expression cassette. In some such embodiments, the first expression cassette comprises: (i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor; (ii) a JeT promoter which drives expression of the chimeric antigen receptor or the exogenous T cell receptor; and (iii) a polyA sequence; and the second expression cassette comprises: (iv) a nucleic acid sequence encoding the shRNA; (v) a U6 promoter which drives expression of the shRNA; and (vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence.

[0027] In certain embodiments, the inhibitory nucleic acid molecule is an shRNA inhibitory against beta-2 microglobulin, wherein the shRNA has a sequence comprising any one of SEQ ID NOs: 2-4. In particular embodiments, the shRNA has a sequence comprising SEQ ID NO: 2. In some such embodiments, the first expression cassette is in a 3' to 5' orientation and the second expression cassette is in a 5' to 3' orientation relative to the 5' and

3' homology arms, and the first expression cassette is 5' upstream of the second expression cassette. In some such embodiments, the first expression cassette comprises: (i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor; (ii) a JeT promoter which drives expression of the chimeric antigen receptor or the exogenous T cell receptor; and (iii) a polyA sequence; and the second expression cassette comprises: (iv) a nucleic acid sequence encoding the shRNA; (v) a U6 promoter which drives expression of the shRNA; and (vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence. In some such embodiments, the nucleic acid molecule comprises a first copy and a second copy of the second expression cassette, wherein the first copy and the second copy are identical, and wherein the first copy and the second copy are in tandem, and further wherein the first copy and the second copy are in the same orientation.

[0028] In another aspect, the invention provides a recombinant DNA construct comprising any nucleic acid molecule of the invention described herein.

[0029] In some embodiments, the recombinant DNA construct encodes a viral vector. In certain embodiments, the viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an adeno-associated viral (AAV) vector. In particular embodiments, the viral vector is a recombinant AAV vector.

[0030] In another aspect, the invention provides a viral vector comprising any nucleic acid molecule of the invention described herein. In certain embodiments, the viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an adeno-associated viral (AAV) vector. In particular embodiments, the viral vector is a recombinant AAV vector.

[0031] In another aspect, the invention provides a method for producing a genetically-modified eukaryotic cell, the method comprising introducing into a cell any nucleic acid molecule of the invention described herein and: (a) a nucleic acid encoding an engineered nuclease having specificity for the nuclease recognition sequence, wherein the engineered nuclease is expressed in the cell; or (b) an engineered nuclease protein having specificity for the nuclease recognition sequence; wherein the engineered nuclease recognizes and cleaves the nuclease recognition sequence in the genome of the cell to generate a cleavage site, and wherein the nucleic acid molecule of the invention is inserted into the genome of the cell at the cleavage site.

[0032] In some embodiments of the method, the genetically-modified eukaryotic cell is a human T cell.

[0033] In some embodiments of the method, the engineered nuclease is an engineered meganuclease, a TALEN, a zinc finger nuclease (ZFN), a CRISPR/Cas, a compact TALEN,

or a megaTAL. In certain embodiments of the method, the engineered nuclease is an engineered meganuclease.

[0034] In some embodiments of the method, the nuclease recognition sequence is in a human T cell receptor alpha constant region gene.

[0035] In certain embodiments of the method, the nuclease recognition sequence is an engineered meganuclease recognition sequence. In particular embodiments, wherein the engineered meganuclease recognition sequence is within a human T cell receptor alpha constant region, the nuclease recognition sequence comprises SEQ ID NO: 1.

[0036] In some embodiments of the method, wherein the nuclease recognition sequence is within a human T cell receptor alpha constant region, cell surface expression of an endogenous T cell receptor is reduced compared to a control cell.

[0037] In some embodiments of the method, the nucleic acid encoding the engineered nuclease is an mRNA. In certain embodiments, the mRNA is a polycistronic mRNA encoding the engineered nuclease and at least one additional polypeptide or nucleic acid molecule.

[0038] In some embodiments of the method, the nucleic acid molecule of the invention described herein is introduced into the cell using a viral vector. In certain embodiments of the method, the viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an AAV vector. In particular embodiments of the method, the viral vector is a recombinant AAV vector, such as a recombinant AAV vector previously described herein.

[0039] In some embodiments of the method, the nucleic acid molecule of the invention described herein is introduced into the cell using a recombinant DNA construct. In certain embodiments of the method, the recombinant DNA construct is a recombinant DNA construct previously described herein.

[0040] In some embodiments of the method, the nucleic acid molecule of the invention described herein is inserted into the genome of the cell at the cleavage site by homologous recombination.

[0041] In some embodiments of the method, the engineered antigen receptor is a chimeric antigen receptor. In other embodiments of the method, the engineered antigen receptor is an exogenous T cell receptor.

[0042] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin. In certain embodiments of the method, cell surface expression of beta-2 microglobulin is between about 1% and about 50% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the

method, cell surface expression of beta-2 microglobulin is between about 1% and about 25% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the method, cell surface expression of beta-2 microglobulin is between about 1% and about 10% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the method, cell surface expression of beta-2 microglobulin is between about 1% and about 5% of cell surface beta-2 microglobulin expression on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

[0043] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin. In certain embodiments of the method, cell surface expression of beta-2 microglobulin is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the method, cell surface expression of beta-2 microglobulin is reduced by 50% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the method, cell surface expression of beta-2 microglobulin is reduced by 75% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In other embodiments of the method, cell surface expression of beta-2 microglobulin is reduced by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

[0044] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule. In certain embodiments of the method, cell surface expression of MHC class I molecules is between about 1% and about 50% of expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is between about 1% and about 25% of expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is between about 1% and about 10% of expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is between about 1% and about 5% of expression of MHC class I molecules on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

[0045] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule. In certain embodiments of the

method, cell surface expression of MHC class I molecules is reduced by 10% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is reduced by 50% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is reduced by 75% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments of the method, cell surface expression of MHC class I molecules is reduced by 90% to 95% compared to expression of MHC class I molecules on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

[0046] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against human CD52. In certain embodiments of the method, cell surface expression of CD52 is between about 1% and about 50% of cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is between about 1% and about 25% of cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is between about 1% and about 10% of cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is between about 1% and about 5% of cell surface CD52 expression on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface expression of CD52.

[0047] In some embodiments of the method, the inhibitory nucleic acid molecule is inhibitory against human CD52. In certain embodiments of the method, cell surface expression of CD52 is reduced by 10% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is reduced by 50% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is reduced by 75% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments of the method, cell surface expression of CD52 is reduced by 90% to 95% compared to cell surface CD52 expression on a control cell. In particular embodiments of the method, a control cell is not genetically-modified to reduce cell surface expression of CD52.

[0048] In another aspect the invention provides a genetically-modified eukaryotic cell made by any of the methods described herein above.

[0049] In another aspect, the invention provides a genetically-modified eukaryotic cell comprising any nucleic acid molecule of the invention described herein, wherein the

engineered antigen receptor and the inhibitory nucleic acid molecule are expressed in the genetically-modified eukaryotic cell.

[0050] In some embodiments, the genetically-modified eukaryotic cell is a genetically-modified human T cell.

[0051] In some embodiments, the nucleic acid molecule of the invention is inserted into the genome of the genetically-modified eukaryotic cell at the nuclease recognition sequence.

[0052] In some embodiments, the gene of interest is a human T cell receptor alpha constant region gene.

[0053] In some embodiments, the nuclease recognition sequence is an engineered meganuclease recognition sequence, a TALEN recognition sequence, a zinc finger nuclease (ZFN) recognition sequence, a CRISPR/Cas recognition sequence, a compact TALEN recognition sequence, or a megaTAL recognition sequence. In certain embodiments, the nuclease recognition sequence is an engineered meganuclease recognition sequence.

[0054] In particular embodiments, wherein the nuclease recognition sequence is within a human T cell receptor alpha constant region gene, the nuclease recognition sequence is an engineered meganuclease recognition sequence comprising SEQ ID NO: 1.

[0055] In some embodiments, cell surface expression of an endogenous T cell receptor is reduced compared to a control cell.

[0056] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin. In certain embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 50% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 25% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 10% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 5% of cell surface beta-2 microglobulin expression on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

[0057] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin. In certain embodiments, cell surface expression of beta-2 microglobulin is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is reduced by 50% to 95% compared to cell surface beta-2 microglobulin

expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is reduced by 75% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is reduced by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

[0058] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule. In certain embodiments, cell surface expression of MHC class I molecules is between about 1% and about 50% of expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is between about 1% and about 25% of expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is between about 1% and about 10% of expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is between about 1% and about 5% of expression of MHC class I molecules on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

[0059] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule. In certain embodiments, cell surface expression of MHC class I molecules is reduced by 10% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is reduced by 50% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is reduced by 75% to 95% compared to expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules is reduced by 90% to 95% compared to expression of MHC class I molecules on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

[0060] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human CD52. In certain embodiments, cell surface expression of CD52 is between about 1% and about 50% of cell surface CD52 expression on a control cell. In other embodiments, cell surface expression of CD52 is between about 1% and about 25% of cell surface CD52 expression on a control cell. In other embodiments, cell surface expression of CD52 is between about 1% and about 10% of cell surface CD52 expression on a control cell. In other

embodiments, cell surface expression of CD52 is between about 1% and about 5% of cell surface CD52 expression on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface expression of CD52.

[0061] In some embodiments, the inhibitory nucleic acid molecule is inhibitory against human CD52. In certain embodiments, cell surface expression of CD52 is reduced by 10% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments, cell surface expression of CD52 is reduced by 50% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments, cell surface expression of CD52 is reduced by 75% to 95% compared to cell surface CD52 expression on a control cell. In other embodiments, cell surface expression of CD52 is reduced by 90% to 95% compared to cell surface CD52 expression on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface expression of CD52.

[0062] In another aspect, the invention provides a genetically-modified eukaryotic cell comprising in its genome a nucleic acid sequence encoding an engineered antigen receptor which is expressed by the genetically-modified eukaryotic cell, wherein cell surface expression of beta-2 microglobulin on the genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell. In certain embodiments, cell surface expression of beta-2 microglobulin on the genetically-modified eukaryotic cell is reduced between 50% and 95% compared to cell surface beta-2 microglobulin expression on a control cell. In certain embodiments, cell surface expression of beta-2 microglobulin on the genetically-modified eukaryotic cell is reduced between 75% and 95% compared to cell surface beta-2 microglobulin expression on a control cell. In certain embodiments, cell surface expression of beta-2 microglobulin on the genetically-modified eukaryotic cell is reduced between 90% and 95% compared to cell surface beta-2 microglobulin expression on a control cell. In particular embodiments, the control cell is not genetically-modified to reduce cell surface expression of beta-2 microglobulin.

[0063] In some embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 50% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 25% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 10% of cell surface beta-2 microglobulin expression on a control cell. In other embodiments, cell surface expression of beta-2 microglobulin is between about 1% and about 5% of cell surface beta-2 microglobulin expression on a control cell. In particular embodiments, a

control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

[0064] In certain embodiments, the genetically-modified eukaryotic cell comprises in its genome a nucleic acid sequence encoding an inhibitory nucleic acid molecule which is inhibitory against beta-2 microglobulin. In particular embodiments, the inhibitory nucleic acid molecule is an RNA interference molecule. In some embodiments, the RNA interference molecule is a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a hairpin siRNA, a microRNA (miRNA), a precursor miRNA, or an miRNA-adapted shRNA. In certain embodiments, the RNA interference molecule is an shRNA. In particular embodiments, the shRNA comprises a sequence of any one of SEQ ID NOs: 2-4. In specific embodiments, the shRNA comprises a sequence of SEQ ID NO: 2.

[0065] In some embodiments, the nucleic acid sequence encoding the engineered antigen receptor is integrated at the same location within the genome as the nucleic acid sequence encoding the inhibitory nucleic acid molecule. In particular embodiments, the genetically-modified eukaryotic cell comprises in its genome the nucleic acid molecule of the invention.

[0066] In other embodiments, the nucleic acid sequence encoding the engineered antigen receptor is integrated at a different location within the genome than the nucleic acid sequence encoding the inhibitory nucleic acid molecule.

[0067] In some embodiments, the genetically-modified eukaryotic cell is less susceptible to endogenous NK cell killing when compared to a control cell, has extended persistence in a subject when compared to a control cell, exhibits enhanced expansion in a subject when compared to a control cell, and/or exhibits reduced allogenicity when compared to a control cell.

[0068] In some embodiments, the engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

[0069] In some embodiments, the genetically-modified eukaryotic cell is a genetically-modified human T cell.

[0070] In particular embodiments, the genetically-modified eukaryotic cell is a genetically-modified human T cell, and the engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

[0071] In another aspect, the invention provides a genetically-modified eukaryotic cell comprising in its genome a nucleic acid sequence encoding an engineered antigen receptor which is expressed by the genetically-modified eukaryotic cell, wherein cell surface expression of MHC class I molecules on the genetically-modified eukaryotic cell is reduced

by 10% to 95% compared to cell surface expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules on the genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules on the genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface expression of MHC class I molecules on a control cell. In certain embodiments, cell surface expression of MHC class I molecules on the genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell. In particular embodiments, the control cell is not genetically-modified to reduce cell surface expression of a component of the MHC class I molecule.

[0072] In some embodiments, cell surface expression of MHC class I molecules is between about 1% and about 50% of cell surface MHC class I molecule expression on a control cell. In other embodiments, cell surface expression of MHC class I molecules is between about 1% and about 25% of cell surface MHC class I molecule expression on a control cell. In other embodiments, cell surface expression of MHC class I molecules is between about 1% and about 10% of cell surface MHC class I molecule expression on a control cell. In other embodiments, cell surface expression of MHC class I molecules is between about 1% and about 5% of cell surface MHC class I molecule expression on a control cell. In particular embodiments, a control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

[0073] In certain embodiments, the genetically-modified eukaryotic cell comprises in its genome a nucleic acid sequence encoding an inhibitory nucleic acid molecule which is inhibitory against a component of the MHC class I molecule. In certain embodiments, the inhibitory molecule is inhibitory against an MHC class I alpha-1 ($\alpha 1$) domain, alpha-2 ($\alpha 2$) domain, alpha-3 ($\alpha 3$) domain, or against beta-2 microglobulin. In a particular embodiment, the inhibitory molecule is inhibitory against beta-2 microglobulin.

[0074] In particular embodiments, the inhibitory nucleic acid molecule is an RNA interference molecule. In some embodiments, the RNA interference molecule is a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a hairpin siRNA, a microRNA (miRNA), a precursor miRNA, or an miRNA-adapted shRNA. In certain embodiments, the RNA interference molecule is an shRNA.

[0075] In certain embodiments, the inhibitory nucleic acid molecule is an shRNA inhibitory against beta-2 microglobulin, wherein the shRNA has a sequence comprising any one of SEQ ID NOs: 2-4. In particular embodiments, the shRNA has a sequence comprising SEQ ID NO: 2. In some such embodiments, the first expression cassette and the second expression cassette are in a 3' to 5' orientation relative to the 5' and 3' homology arms, and wherein the first expression cassette is 5' upstream of the second expression cassette. In some such embodiments, the first expression cassette comprises: (i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor; (ii) a JeT promoter which drives expression of the chimeric antigen receptor or the exogenous T cell receptor; and (iii) a polyA sequence; and the second expression cassette comprises: (iv) a nucleic acid sequence encoding the shRNA; (v) a U6 promoter which drives expression of the shRNA; and (vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence.

[0076] In certain embodiments, the inhibitory nucleic acid molecule is an shRNA inhibitory against beta-2 microglobulin, wherein the shRNA has a sequence comprising any one of SEQ ID NOs: 2-4. In particular embodiments, the shRNA has a sequence comprising SEQ ID NO: 2. In some such embodiments, the first expression cassette is in a 3' to 5' orientation and the second expression cassette is in a 5' to 3' orientation relative to the 5' and 3' homology arms, and the first expression cassette is 5' upstream of the second expression cassette. In some such embodiments, the first expression cassette comprises: (i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor; (ii) a JeT promoter which drives expression of the chimeric antigen receptor or the exogenous T cell receptor; and (iii) a polyA sequence; and the second expression cassette comprises: (iv) a nucleic acid sequence encoding the shRNA; (v) a U6 promoter which drives expression of the shRNA; and (vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence. In some such embodiments, the nucleic acid molecule comprises a first copy and a second copy of the second expression cassette, wherein the first copy and the second copy are identical, and wherein the first copy and the second copy are in tandem, and further wherein the first copy and the second copy are in the same orientation.

[0077] In some embodiments, the nucleic acid sequence encoding the engineered antigen receptor is integrated at the same location within the genome as the nucleic acid sequence encoding the inhibitory nucleic acid molecule. In particular embodiments, the genetically-modified eukaryotic cell comprises in its genome the nucleic acid molecule of the invention.

[0078] In other embodiments, the nucleic acid sequence encoding the engineered antigen receptor is integrated at a different location within the genome than the nucleic acid sequence encoding the inhibitory nucleic acid molecule.

[0079] In some embodiments, the genetically-modified eukaryotic cell is less susceptible to endogenous NK cell killing when compared to a control cell, has extended persistence in a subject when compared to a control cell, exhibits enhanced expansion in a subject when compared to a control cell, and/or exhibits reduced allogenicity when compared to a control cell.

[0080] In some embodiments, the engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

[0081] In some embodiments, the genetically-modified eukaryotic cell is a genetically-modified human T cell.

[0082] In particular embodiments, the genetically-modified eukaryotic cell is a genetically-modified human T cell, and the engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

[0083] In another aspect, the invention provides a pharmaceutical composition comprising a pharmaceutically-acceptable carrier and a therapeutically effective amount of any genetically-modified eukaryotic cell described herein above.

[0084] In some particular embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, the engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and cell surface expression of beta-2 microglobulin is between about 1% and about 50%, about 1% and about 25%, about 1% and about 10%, or about 1% and about 5% of cell surface beta-2 microglobulin expression on a control cell.

[0085] In other particular embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, the engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and cell surface expression of beta-2 microglobulin on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of beta-2 microglobulin on a control cell.

[0086] In some particular embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, the engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and cell surface expression of MHC class I molecules is between about 1% and about 50%, about 1% and

about 25%, about 1% and about 10%, or about 1% and about 5% of cell surface expression of MHC class I molecules on a control cell.

[0087] In other particular embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, the engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and cell surface expression of MHC class I molecules on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

[0088] In other particular embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, and the engineered antigen receptor is a chimeric antigen receptor, and cell surface expression of CD52 is between about 1% and about 50%, about 1% and about 25%, about 1% and about 10%, or about 1% and about 5% of cell surface CD52 expression on a control cell.

[0089] In some embodiments, the genetically-modified eukaryotic cell of the pharmaceutical composition is a genetically-modified human T cell, the engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and cell surface expression of CD52 on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of CD52 on a control cell.

[0090] In certain embodiments, the pharmaceutical composition of the invention is for immunotherapy in the treatment of cancer in a subject in need thereof.

[0091] In another aspect, the invention provides a population of genetically-modified eukaryotic cells comprising a plurality of any genetically-modified eukaryotic cell described herein.

[0092] In some embodiments, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to 100%, of cells in the population are a genetically-modified eukaryotic cell as described herein.

[0093] In particular embodiments, the genetically-modified eukaryotic cells of the population are genetically-modified human T cells, or cells derived therefrom, or genetically-modified NK cells, or cells derived therefrom.

[0094] In certain embodiments, the genetically-modified eukaryotic cells of the population comprise a cell surface chimeric antigen receptor or exogenous T cell receptor. In

some of these embodiments, the chimeric antigen receptor or exogenous T cell receptor comprises an extracellular ligand-binding domain having specificity for a tumor-specific antigen.

[0095] In some embodiments, the genetically-modified eukaryotic cells of the population have no cell surface expression of an endogenous T cell receptor when compared to an unmodified control cell. In some embodiments, the genetically-modified eukaryotic cells of the population have reduced cell surface expression of beta-2 microglobulin, MHC class I molecules, or CD52.

[0096] In another aspect, the invention provides a method of using immunotherapy to treat a disease in a subject in need thereof, the method comprising administering to the subject a genetically-modified eukaryotic cell described herein; wherein the genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor or an exogenous T cell receptor; and wherein cell surface expression of beta-2 microglobulin on the genetically-modified human T cell is between about 1% and about 50%, about 1% and about 25%, about 1% and about 10%, or about 1% and about 5% of cell surface beta-2 microglobulin expression on a control cell.

[0097] In some embodiments of the method, cell surface expression of beta-2 microglobulin on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

[0098] In some embodiments of the method, endogenous NK cell killing of the genetically-modified human T cell is reduced in the subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression.

[0099] In some embodiments of the method, the subject is administered any pharmaceutical composition described herein in which cell surface beta-2 microglobulin expression is reduced on the genetically-modified human T cell when compared to a control cell.

[0100] In some embodiments of the method, the genetically-modified human T cell is allogeneic to the subject.

[0101] In some embodiments of the method, persistence of the genetically-modified human T cell is extended in the subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

[0102] In some embodiments of the method, expansion of the genetically-modified human T cell is enhanced in the subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

[0103] In some embodiments of the method, allogenicity of the genetically-modified human T cell is reduced when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

[0104] In some embodiments of the method, the disease is cancer.

[0105] In some embodiments of the method, the cancer is selected from the group consisting of a cancer of carcinoma, lymphoma, sarcoma, blastomas, and leukemia. In certain embodiments of the method, the cancer is selected from the group consisting of a cancer of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyo sarcoma, leukemia, and Hodgkin's lymphoma. In particular embodiments of the method, the cancer of B-cell origin is selected from the group consisting of B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma.

[0106] In another aspect, the invention provides a method of using immunotherapy to treat a disease in a subject in need thereof, the method comprising administering to the subject a genetically-modified eukaryotic cell described herein; wherein the genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor or an exogenous T cell receptor; and wherein cell surface expression of MHC class I molecules on the genetically-modified human T cell is between about 1% and about 50%, about 1% and about 25%, about 1% and about 10%, or about 1% and about 5% of cell surface expression of MHC class I molecules on a control cell.

[0107] In some embodiments of the method, cell surface expression of MHC class I molecules on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

[0108] In some embodiments of the method, endogenous NK cell killing of the genetically-modified human T cell is reduced in the subject when compared to a genetically-modified human T cell having no cell surface expression of MHC class I molecules.

[0109] In some embodiments of the method, the subject is administered any pharmaceutical composition described herein in which cell surface expression of MHC class I molecules is reduced on the genetically-modified human T cell when compared to a control cell.

[0110] In some embodiments of the method, the genetically-modified human T cell is allogeneic to the subject.

[0111] In some embodiments of the method, persistence of the genetically-modified human T cell is extended in the subject when compared to a genetically-modified human T cell having no cell surface MHC class I molecule expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

[0112] In some embodiments of the method, expansion of the genetically-modified human T cell is enhanced in the subject when compared to a genetically-modified human T cell having no cell surface MHC class I molecule expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

[0113] In some embodiments of the method, allogenicity of the genetically-modified human T cell is reduced when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

[0114] In some embodiments of the method, the disease is cancer.

[0115] In some embodiments of the method, the cancer is selected from the group consisting of a cancer of carcinoma, lymphoma, sarcoma, blastomas, and leukemia. In certain embodiments of the method, the cancer is selected from the group consisting of a cancer of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyo sarcoma, leukemia, and Hodgkin's lymphoma. In particular embodiments of the method, the cancer of B-cell origin is selected from the group consisting of B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma.

[0116] In another aspect, the invention provides a method of using immunotherapy to treat cancer in a subject in need thereof, the method comprising administering to the subject a genetically-modified eukaryotic cell described herein; wherein the genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor or an exogenous T cell receptor; and wherein cell surface expression of CD52 on the

genetically-modified human T cell is between 1% and 50%, 1% and 25%, 1% and 10%, or 1% and 5% of cell surface CD52 expression on a control cell.

[0117] In some embodiments of the method, cell surface expression of CD52 on the genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of CD52 on a control cell.

[0118] In some embodiments of the method, the subject is administered a pharmaceutical composition described herein in which cell surface expression of CD52 is reduced on the genetically-modified human T cell when compared to a control cell.

[0119] In some embodiments of the method, the genetically-modified human T cell is allogeneic to the subject.

[0120] In some embodiments of the method, the cancer is selected from the group consisting of a cancer of carcinoma, lymphoma, sarcoma, blastomas, and leukemia. In certain embodiments of the method, the cancer is selected from the group consisting of a cancer of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyo sarcoma, leukemia, and Hodgkin's lymphoma. In particular embodiments of the method, the cancer of B-cell origin is selected from the group consisting of B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma.

[0121] In another aspect, the invention provides a method for preparing an enriched population of genetically-modified eukaryotic cells comprising an engineered antigen receptor, the method comprising preparing a population of cells comprising a genetically-modified eukaryotic cell described herein and cells expressing a wild-type level of cell surface CD52, wherein cell surface expression of CD52 on the genetically-modified eukaryotic cell is between about 1% and about 50%, about 1% and about 25%, about 1% and about 10%, or about 1% and about 5% when compared to a control cell, the method comprising: (a) contacting the population of cells with beads conjugated to an anti-CD52 binding molecule, wherein cells expressing a wild-type level of cell surface CD52 are bound to the beads and the genetically-modified eukaryotic cell is not bound to the beads; and (b) removing the beads from the population of cells to produce the enriched population of cells; wherein the enriched population of cells is enriched for the genetically-modified eukaryotic cell.

[0122] In some embodiments of the method, cell surface expression of CD52 on the genetically-modified eukaryotic cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% when compared to a control cell.

[0123] In some embodiments of the method, the beads are magnetic beads. In certain embodiments of the method, the magnetic beads are removed from the population of cells by magnetic separation.

[0124] In some embodiments of the method, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, or up to 100% of cells in the enriched population are the genetically-modified eukaryotic cell.

[0125] In some embodiments of the method, the genetically-modified eukaryotic cell expresses a chimeric antigen receptor. In other embodiments of the method, the genetically-modified eukaryotic cell expresses an exogenous T cell receptor.

[0126] In some embodiments of the method, the genetically-modified eukaryotic cell is a genetically-modified human T cell, such as any genetically-modified T cell described herein.

[0127] In another aspect, the present disclosure provides a genetically-modified eukaryotic cell described herein for use as a medicament. The present disclosure further provides the use of a genetically-modified eukaryotic cell described herein in the manufacture of a medicament for treating a disease in a subject in need thereof. In one such embodiment, the medicament is useful in the treatment of cancer.

[0128] The foregoing and other aspects and embodiments of the present invention can be more fully understood by reference to the following detailed description and claims. Certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. All combinations of the embodiments are specifically embraced by the present invention and are disclosed herein just as if each and every combination was individually and explicitly disclosed. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. All sub-combinations of features listed in the embodiments are also specifically embraced by the present invention and are disclosed herein just as if each and every such sub-combination was individually and explicitly disclosed herein. Embodiments of each aspect of the present invention disclosed herein apply to each other aspect of the invention *mutatis mutandis*.

BRIEF DESCRIPTION OF THE FIGURES

[0129] Figure 1 shows diagrams of various embodiments of the nucleic acid molecule of the invention. The JeT promoter is shown as an example of a promoter driving expression of the engineered antigen receptor. A U6 promoter is shown as an example of a promoter driving expression of the inhibitory nucleic acid molecule. A chimeric antigen receptor (CAR) is shown as an example of an engineered antigen receptor. An shRNA is shown as an example of an inhibitory nucleic acid molecule. A poly-A sequence is shown as an example of a sequence which terminates translation of the engineered antigen receptor. A cPPT/CTS sequence is shown as an example of a sequence which terminates translation of the inhibitory nucleic acid molecule. 5' and 3' homology arms are shown flanking the first expression cassette and second expression cassette of each construct. Optional 5' and 3' inverted terminal repeats are further shown in each construct. Constructs above the dashed line have first and second expression cassettes in the same orientation, whereas constructs below the dashed line have first and second expression cassettes in opposite orientations.

[0130] Figure 2A-2L shows flow plots which represent NK cell killing of primary human T cells. The indicated ratios represent the ratio of NK cells to T cells (E:T) in each experiment. Figure 2A shows NK cell killing of B2M⁺ T cells using a 2:1 ratio. Figure 2B shows NK cell killing of B2M⁺ T cells using a 1:1 ratio. Figure 2C shows NK cell killing of B2M⁺ T cells using a 0.5:1 ratio. Figure 2D shows NK cell killing of B2M⁺ T cells using a 0:1 ratio. Figure 2E shows NK cell killing of B2M⁻ T cells using a 2:1 ratio. Figure 2F shows NK cell killing of B2M⁻ T cells using a 1:1 ratio. Figure 2G shows NK cell killing of B2M⁻ T cells using a 0.5:1 ratio. Figure 2H shows NK cell killing of B2M⁻ T cells using a 0:1 ratio. Figure 2I shows NK cell killing of Daudi Class I-negative cells using a 2:1 ratio. Figure 2J shows NK cell killing of Daudi Class I-negative cells using a 1:1 ratio. Figure 2K shows NK cell killing of Daudi Class I-negative cells using a 0.5:1 ratio. Figure 2L shows NK cell killing of Daudi Class I-negative cells using a 0:1 ratio.

[0131] Figure 3 shows a chart summarizing NK cell killing of B2M⁺ and B2M⁻ cells at different ratios.

[0132] Figure 4 shows percentage knockdown of human B2M in primary human T cells by three candidate B2M shRNAs.

[0133] Figures 5A-5D show flow diagrams representing NK cell lysis or allogeneic cell lysis of K562 cells or mock-treated primary human T cells. Figure 5A shows NK cell lysis of K562 cells. Figure 5B shows allogeneic cell lysis of K562 cells. Figure 5C shows NK cell

lysis of mock-treated primary human T cells. Figure 5D shows allogeneic cell lysis of mock-treated primary human T cells.

[0134] Figures 6A-6D show flow diagrams representing NK cell lysis or allogeneic cell lysis of primary human T cells treated with B2M shRNAs. Figure 6A shows NK cell lysis of primary human T cells treated with shRNA254. Figure 6B shows allogeneic cell lysis of primary human T cells treated with shRNA254. Figure 6C shows NK cell lysis of primary human T cells treated with shRNA472. Figure 6D shows allogeneic cell lysis of primary human T cells treated with shRNA472.

[0135] Figures 7A-7F show diagrams of various nucleic acid molecule constructs encoding a chimeric antigen receptor and an shRNA against beta-2 microglobulin. Figure 7A shows construct 7007 (SEQ ID NO: 18). Figure 7B shows construct 7217 (SEQ ID NO: 19). Figure 7C shows construct 7008 (SEQ ID NO: 20). Figure 7D shows construct 7218 (SEQ ID NO: 21). Figure 7E shows construct 7009 (SEQ ID NO: 22). Figure 7F shows construct 7219 (SEQ ID NO: 23).

[0136] Figure 8 shows percentage knockdown of human CD52 in primary human T cells by three different candidate CD52 shRNAs.

[0137] Figure 9 A-C show knockdown of CD52 using shRNA and magnetic enrichment of the knockdown population of primary human T cells by CD52 magnetic depletion. Figure 9A shows T cells that were mock transduced. Figure 9B shows T cells transduced with an shRNA-568 lentivirus. Figure 9C shows lentivirus-shRNA568 transduced cells that have undergone a CD52 magnetic depletion.

[0138] Figures 10A-10H shows diagrams of nucleic acid molecule constructs encoding a chimeric antigen receptor and an shRNA against CD52. Figure 10A shows construct 7005 (SEQ ID NO: 10) which encodes a CAR only. Figure 10B shows construct 7002 (SEQ ID NO: 11) which encodes a CAR only. Figure 10C shows construct 7004 (SEQ ID NO: 12). Figure 10D shows construct 7204 (SEQ ID NO: 13). Figure 10E shows construct 7013 (SEQ ID NO: 14). Figure 10F shows construct 7213 (SEQ ID NO: 15). Figure 10G shows construct 7014 (SEQ ID NO: 16). Figure 10H shows construct 7214 (SEQ ID NO: 17).

[0139] Figure 11A-11D shows CD52 knockdown profiles using CAR/CD52 shRNA constructs with different orientations. Figure 11A shows CD52 expression when a CAR is expressed in the absence of a CD52 shRNA. Figure 11B shows CD52 expression when using the 7013 construct. Figure 11C shows CD52 expression when using the 7004 construct. Figure 11D shows CD52 expression when using the 7014 construct.

[0140] Figure 12A-12C shows B2M knockdown on CAR T cells using CAR/B2M shRNA constructs having one or multiple shRNA cassettes. Figure 12A shows B2M expression in CAR T cells expressing no B2M shRNA (7002 – shaded curve) or a single B2M shRNA cassette (7008 – open curve). Figure 12B shows B2M expression in CAR T cells expressing no B2M shRNA (7002 - shaded) or two B2M shRNA cassettes (7029 - open). Figure 12C shows B2M expression in CAR-/CD3+ (*i.e.* non-edited) populations from cultures electroporated with 7002, 7008, or 7029.

[0141] Figure 13A-13C shows cell surface expression of beta-2 microglobulin on T cells transfected with linearized DNA to express a control CAR-negative construct (7002), CAR constructs expressing a single shRNA472 copy in a 3' to 5' head-to-tail configuration with the CAR (7056), or in a 3' to 5'/5' to 3' tail-to-tail configuration with the CAR (7059), or a CAR construct expressing two shRNA cassette copies in a 3' to 5'/5' to 3' tail-to-tail configuration with the CAR (7060). Figure 13A shows CAR T cells expressing the 7002 and 7056 constructs. Figure 13B shows CAR T cells expressing the 7002 and 7059 constructs. Figure 13C shows CAR T cells expressing the 7002 and 7060 constructs.

[0142] Figure 14 A-D shows beta-2 microglobulin expression or HLA-A, B, and C expression (*i.e.*, MHC class I molecule expression) on T cells transduced with an AAV comprising construct 7056 which expresses a single copy of the shRNA472 in a 3' to 5' head-to-tail configuration with the CAR. Figure 14A shows the B2M surface levels in CD3-/CAR+ cells compared to meganuclease-edited cells expressing no shRNA from a control culture. Figure 14B shows B2M levels on CD3-/CAR+ versus CD3+/CAR- populations in the same culture. Figure 14C shows HLA-ABC (*i.e.*, MHC class I molecule) surface levels in CD3-/CAR+ cells compared to meganuclease-edited cells expressing no shRNA from a control culture. Figure 14D shows HLA-ABC levels on CD3-/CAR+ versus CD3+/CAR- populations in the same culture.

BRIEF DESCRIPTION OF THE SEQUENCES

[0143] SEQ ID NO: 1 sets forth the nucleic acid sequence of the TRC 1-2 recognition sequence.

[0144] SEQ ID NO: 2 sets forth the nucleic acid sequence of the anti-beta-2 microglobulin shRNA472.

[0145] SEQ ID NO: 3 sets forth the nucleic acid sequence of the anti-beta-2 microglobulin shRNA256.

- [0146] SEQ ID NO: 4 sets forth the nucleic acid sequence of the anti-beta-2 microglobulin shRNA254.
- [0147] SEQ ID NO: 5 sets forth the nucleic acid sequence of the anti-CD52 shRNA572.
- [0148] SEQ ID NO: 6 sets forth the nucleic acid sequence of the anti-CD52 shRNA876.
- [0149] SEQ ID NO: 7 sets forth the nucleic acid sequence of the anti-CD52 shRNA568.
- [0150] SEQ ID NO: 8 sets forth the nucleic acid sequence of the anti-CD52 shRNA569.
- [0151] SEQ ID NO: 9 sets forth the nucleic acid sequence of the anti-CD52 shRNA571.
- [0152] SEQ ID NO: 10 sets forth the nucleic acid sequence of the CAR 7005 construct.
- [0153] SEQ ID NO: 11 sets forth the nucleic acid sequence of the CAR 7002 construct.
- [0154] SEQ ID NO: 12 sets forth the nucleic acid sequence of the CAR 7004 construct.
- [0155] SEQ ID NO: 13 sets forth the nucleic acid sequence of the CAR 7204 construct.
- [0156] SEQ ID NO: 14 sets forth the nucleic acid sequence of the CAR 7013 construct.
- [0157] SEQ ID NO: 15 sets forth the nucleic acid sequence of the CAR 7213 construct.
- [0158] SEQ ID NO: 16 sets forth the nucleic acid sequence of the CAR 7014 construct.
- [0159] SEQ ID NO: 17 sets forth the nucleic acid sequence of the CAR 7214 construct.
- [0160] SEQ ID NO: 18 sets forth the nucleic acid sequence of the CAR 7007 construct.
- [0161] SEQ ID NO: 19 sets forth the nucleic acid sequence of the CAR 7217 construct.
- [0162] SEQ ID NO: 20 sets forth the nucleic acid sequence of the CAR 7008 construct.
- [0163] SEQ ID NO: 21 sets forth the nucleic acid sequence of the CAR 7218 construct.
- [0164] SEQ ID NO: 22 sets forth the nucleic acid sequence of the CAR 7009 construct.
- [0165] SEQ ID NO: 23 sets forth the nucleic acid sequence of the CAR 7219 construct.
- [0166] SEQ ID NO: 24 sets forth the nucleic acid sequence of the CAR 7029 construct.
- [0167] SEQ ID NO: 25 sets forth the nucleic acid sequence of the CAR 7056 construct.
- [0168] SEQ ID NO: 26 sets forth the nucleic acid sequence of the CAR 7059 construct.
- [0169] SEQ ID NO: 27 sets forth the nucleic acid sequence of the CAR 7060 construct.

DETAILED DESCRIPTION OF THE INVENTION

1.1 References and Definitions

[0170] The patent and scientific literature referred to herein establishes knowledge that is available to those of skill in the art. The issued US patents, allowed applications, published foreign applications, and references, including GenBank database sequences, which are cited herein are hereby incorporated by reference to the same extent as if each was specifically and individually indicated to be incorporated by reference.

[0171] The present disclosure can be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the present disclosure to those skilled in the art. For example, features illustrated with respect to one embodiment can be incorporated into other embodiments, and features illustrated with respect to a particular embodiment can be deleted from that embodiment. In addition, numerous variations and additions to the embodiments suggested herein will be apparent to those skilled in the art in light of the instant disclosure, which do not depart from the present disclosure.

[0172] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. The terminology used in the description of the present disclosure herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the present disclosure.

[0173] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference herein in their entirety.

[0174] As used herein, “a,” “an,” or “the” can mean one or more than one. For example, “a” cell can mean a single cell or a multiplicity of cells.

[0175] As used herein, unless specifically indicated otherwise, the word “or” is used in the inclusive sense of “and/or” and not the exclusive sense of “either/or.”

[0176] The terms “expression cassette,” “recombinant DNA construct,” “recombinant construct,” “expression construct,” “chimeric construct,” “construct,” and “recombinant DNA fragment” are used interchangeably herein and are nucleic acid fragments. A recombinant construct comprises an artificial combination of nucleic acid fragments, including, without limitation, regulatory and coding sequences that are not found together in nature. For example, a recombinant DNA construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source and arranged in a manner different than that found in nature. Such a construct may be used by itself or may be used in conjunction with a vector.

[0177] As used herein, a “vector” or “recombinant DNA vector” may be a construct that includes a replication system and sequences that are capable of transcription and translation of a polypeptide-encoding sequence in a given host cell. If a vector is used then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. Vectors can include, without limitation, plasmid vectors and

recombinant lentiviral or recombinant AAV vectors, or any other vector known in that art suitable for delivering a gene encoding a co-stimulatory domain of the present disclosure to a target cell. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleotides or nucleic acid sequences of the present disclosure.

[0178] As used herein, a “vector” can also refer to a viral vector. Viral vectors can include, without limitation, retroviral vectors, lentiviral vectors, adenoviral vectors, and adeno-associated viral vectors (AAV).

[0179] As used herein, the term “operably linked” is intended to mean a functional linkage between two or more elements. For example, an operable linkage between a nucleic acid sequence encoding a nuclease as disclosed herein and a regulatory sequence (*e.g.*, a promoter) is a functional link that allows for expression of the nucleic acid sequence encoding the nuclease. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by operably linked is intended that the coding regions are in the same reading frame.

[0180] As used herein, the term “RNA interference” or “RNAi” refers to a phenomenon in which the introduction of double-stranded RNA (dsRNA) into a diverse range of organisms and cell types causes degradation of the complementary mRNA. In the cell, long dsRNAs are cleaved into short 21-25 nucleotide small interfering RNAs, or siRNAs, by a ribonuclease known as Dicer. The siRNAs subsequently assemble with protein components into an RNA-induced silencing complex (RISC), unwinding in the process. Activated RISC then binds to complementary transcript by base pairing interactions between the siRNA antisense strand and the mRNA. The bound mRNA is cleaved and sequence specific degradation of mRNA results in gene silencing. See, for example, U.S. Pat. No. 6,506,559.

[0181] The term “siRNA” as used herein refers to small interfering RNA, also known as short interfering RNA or silencing RNA. siRNAs can be, for example, 18 to 30, 20 to 25, 21 to 23 or 21 nucleotide-long double-stranded RNA molecules. An “shRNA” as used herein is a short hairpin RNA, which is a sequence of RNA that makes a tight hairpin turn that can also be used to silence gene expression via RNA interference. shRNA can be operably linked to the U6 promoter expression. The shRNA hairpin structure is cleaved by the cellular machinery into siRNA. shRNA disclosed herein can comprise a sequence complementary to at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides,

at least 21 nucleotides, at least 22 nucleotides, or 23 nucleotides of the mRNA a target protein.

[0182] As used herein, an “engineered antigen receptor” refers to an exogenous receptor introduced into a cell, such as a chimeric antigen receptor or exogenous T cell receptor, which induces an activating signal in the cell upon stimulation/binding to a ligand or antigen (*e.g.*, a tumor-specific antigen).

[0183] As used herein, a “chimeric antigen receptor” or “CAR” refers to an engineered receptor that grafts specificity for an antigen or other ligand or molecule onto an immune effector cell (*e.g.*, a T cell or NK cell). A chimeric antigen receptor typically comprises at least an extracellular ligand-binding domain or moiety and an intracellular domain that comprises one or more signaling domains and/or co-stimulatory domains.

[0184] In some embodiments, the extracellular ligand-binding domain or moiety is in the form of a single-chain variable fragment (scFv) derived from a monoclonal antibody, which provides specificity for a particular epitope or antigen (*e.g.*, an epitope or antigen preferentially present on the surface of a cell, such as a cancer cell or other disease-causing cell or particle). In some embodiments, the scFv is attached via a linker sequence. In some embodiments, the extracellular ligand-binding domain is specific for any antigen or epitope of interest. In some embodiments, the scFv is humanized. In some embodiments, the extracellular domain of a chimeric antigen receptor comprises an autoantigen (see, Payne *et al.* (2016) *Science*, Vol. 353 (6295): 179-184), which is recognized by autoantigen-specific B cell receptors on B lymphocytes, thus directing T cells to specifically target and kill autoreactive B lymphocytes in antibody-mediated autoimmune diseases. Such CARs can be referred to as chimeric autoantibody receptors (CAARs), and the incorporation of one or more co-stimulatory domains described herein into such CAARs is encompassed by the present disclosure.

[0185] The extracellular domain of a chimeric antigen receptor can also comprise a naturally-occurring ligand for an antigen of interest, or a fragment of a naturally-occurring ligand which retains the ability to bind the antigen of interest.

[0186] Intracellular signaling domains are cytoplasmic domains which transmit an activation signal to the cell following binding of the extracellular domain. An intracellular signaling domain can be any intracellular signaling domain of interest that is known in the art. Such cytoplasmic signaling domains can include, without limitation, CD3 ζ .

[0187] In some embodiments, the intracellular domain also includes one or more intracellular co-stimulatory domains, such as those described herein, which transmit a co-

stimulatory signal which promotes cell proliferation, cell survival, and/or cytokine secretion after binding of the extracellular domain. As used herein, a “co-stimulatory domain” refers to a polypeptide domain which transmits an intracellular proliferative and/or cell-survival signal upon activation. Activation of a co-stimulatory domain may occur following homodimerization of two co-stimulatory domain polypeptides. Activation may also occur, for example, following activation of a construct comprising the co-stimulatory domain (*e.g.*, a chimeric antigen receptor or an inducible regulatory construct). Generally, a co-stimulatory domain can be derived from a transmembrane co-stimulatory receptor, particularly from an intracellular portion of a co-stimulatory receptor. Such intracellular co-stimulatory domains can be any of those known in the art and can include, without limitation, CD27, CD28, CD8, 4-1BB (CD137), OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3 and a ligand that specifically binds with CD83, N1, N6, or any combination thereof.

[0188] As used herein, a “co-stimulatory signal” refers to an intracellular signal induced by a co-stimulatory domain that promotes cell proliferation, expansion of a cell population *in vitro* and/or *in vivo*, promotes cell survival, modulates (*e.g.*, upregulates or downregulates) the secretion of cytokines, and/or modulates the production and/or secretion of other immunomodulatory molecules. In some embodiments, a co-stimulatory signal is induced following homodimerization of two co-stimulatory domain polypeptides. In some embodiments, a co-stimulatory signal is induced following activation of a construct comprising the co-stimulatory domain (*e.g.*, a chimeric antigen receptor or an inducible regulatory construct).

[0189] A chimeric antigen receptor can further include additional structural elements, including a transmembrane domain that is attached to the extracellular ligand-binding domain via a hinge or spacer sequence. The transmembrane domain can be derived from any membrane-bound or transmembrane protein. For example, the transmembrane polypeptide can be a subunit of the T-cell receptor (*i.e.*, an α , β , γ or ζ , polypeptide constituting CD3 complex), IL2 receptor p55 (α chain), p75 (β chain) or γ chain, subunit chain of Fc receptors (*e.g.*, Fc γ receptor III) or CD proteins such as the CD8 alpha chain. Alternatively the transmembrane domain can be synthetic and can comprise predominantly hydrophobic residues such as leucine and valine.

[0190] The hinge region refers to any oligo- or polypeptide that functions to link the transmembrane domain to the extracellular ligand-binding domain. For example, a hinge region may comprise up to 300 amino acids, preferably 10 to 100 amino acids and most

preferably 25 to 50 amino acids. Hinge regions may be derived from all or part of naturally occurring molecules, such as from all or part of the extracellular region of CD8, CD4 or CD28, or from all or part of an antibody constant region. Alternatively, the hinge region may be a synthetic sequence that corresponds to a naturally occurring hinge sequence, or may be an entirely synthetic hinge sequence. In particular examples, a hinge domain can comprise a part of a human CD8 alpha chain, FcγRIIIa receptor or IgG1.

[0191] As used herein, the term “activation” refers to the state of a cell (*e.g.*, a T cell) that has been sufficiently stimulated to induce detectable effector function. In some embodiments, activation is associated with induced cytokine production and/or induced cell proliferation and expansion.

[0192] As used herein, an “exogenous T cell receptor” or “exogenous TCR” refers to a TCR whose sequence is introduced into the genome of an immune effector cell (*e.g.*, a human T cell) that may or may not endogenously express the TCR. Expression of an exogenous TCR on an immune effector cell can confer specificity for a specific epitope or antigen (*e.g.*, an epitope or antigen preferentially present on the surface of a cancer cell or other disease-causing cell or particle). Such exogenous T cell receptors can comprise alpha and beta chains or, alternatively, may comprise gamma and delta chains. Exogenous TCRs useful in the invention may have specificity to any antigen or epitope of interest.

[0193] As used herein, with respect to a protein, the term “engineered” or “recombinant” means having an altered amino acid sequence as a result of the application of genetic engineering techniques to nucleic acids which encode the protein, and cells or organisms which express the protein. With respect to a nucleic acid, the term “recombinant” means having an altered nucleic acid sequence as a result of the application of genetic engineering techniques. Genetic engineering techniques include, but are not limited to, PCR and DNA cloning technologies; transfection, transformation and other gene transfer technologies; homologous recombination; site-directed mutagenesis; and gene fusion. In accordance with this definition, a protein having an amino acid sequence identical to a naturally-occurring protein, but produced by cloning and expression in a heterologous host, is not considered recombinant.

[0194] As used herein, the term “wild-type” refers to the most common naturally occurring polynucleotide or polypeptide sequence responsible for a given phenotype. Whereas a wild-type allele or polypeptide can confer a normal phenotype in an organism, a mutant or variant allele or polypeptide can, in some instances, confer an altered phenotype.

[0195] As used herein with respect to recombinant proteins, the term “modification” means any insertion, deletion or substitution of an amino acid residue in the recombinant sequence relative to a reference sequence (*e.g.*, a wild-type or a native sequence).

[0196] As used herein, the term “meganuclease” refers to an endonuclease that binds double-stranded DNA at a recognition sequence that is greater than 12 base pairs. In some embodiments, the recognition sequence for a meganuclease of the present disclosure is 22 base pairs. A meganuclease can be an endonuclease that is derived from I-CreI, and can refer to an engineered variant of I-CreI that has been modified relative to natural I-CreI with respect to, for example, DNA-binding specificity, DNA cleavage activity, DNA-binding affinity, or dimerization properties. Methods for producing such modified variants of I-CreI are known in the art (*e.g.* WO 2007/047859). A meganuclease as used herein binds to double-stranded DNA as a heterodimer. A meganuclease may also be a “single-chain meganuclease” in which a pair of DNA-binding domains are joined into a single polypeptide using a peptide linker. The term “homing endonuclease” is synonymous with the term “meganuclease.” Meganucleases of the present disclosure are substantially non-toxic when expressed in cells, particularly in human T cells, such that cells can be transfected and maintained at 37°C without observing deleterious effects on cell viability or significant reductions in meganuclease cleavage activity when measured using the methods described herein.

[0197] As used herein, the term “single-chain meganuclease” refers to a polypeptide comprising a pair of nuclease subunits joined by a linker. A single-chain meganuclease has the organization: N-terminal subunit – Linker – C-terminal subunit. The two meganuclease subunits will generally be non-identical in amino acid sequence and will recognize non-identical DNA sequences. Thus, single-chain meganucleases typically cleave pseudo-palindromic or non-palindromic recognition sequences. A single-chain meganuclease may be referred to as a “single-chain heterodimer” or “single-chain heterodimeric meganuclease” although it is not, in fact, dimeric. For clarity, unless otherwise specified, the term “meganuclease” can refer to a dimeric or single-chain meganuclease.

[0198] As used herein, the term “linker” refers to an exogenous peptide sequence used to join two meganuclease subunits into a single polypeptide. A linker may have a sequence that is found in natural proteins, or may be an artificial sequence that is not found in any natural protein. A linker may be flexible and lacking in secondary structure or may have a propensity to form a specific three-dimensional structure under physiological conditions. A linker can

include, without limitation, any of those encompassed by U.S. Patent Nos. 8,445,251 and 9,434,931.

[0199] As used herein, the term “zinc finger nuclease” or “ZFN” refers to a chimeric protein comprising a zinc finger DNA-binding domain fused to a nuclease domain from an endonuclease or exonuclease, including but not limited to a restriction endonuclease, homing endonuclease, S1 nuclease, mung bean nuclease, pancreatic DNase I, micrococcal nuclease, and yeast HO endonuclease. Nuclease domains useful for the design of zinc finger nucleases include those from a Type IIs restriction endonuclease, including but not limited to FokI, FomI, and StsI restriction enzyme. Additional Type IIs restriction endonucleases are described in International Publication No. WO 2007/014275, which is incorporated by reference in its entirety. The structure of a zinc finger domain is stabilized through coordination of a zinc ion. DNA binding proteins comprising one or more zinc finger domains bind DNA in a sequence-specific manner. The zinc finger domain can be a native sequence or can be redesigned through rational or experimental means to produce a protein which binds to a pre-determined DNA sequence ~18 basepairs in length, comprising a pair of nine basepair half-sites separated by 2-10 basepairs. See, for example, U.S. Pat. Nos. 5,789,538, 5,925,523, 6,007,988, 6,013,453, 6,200,759, and International Publication Nos. WO 95/19431, WO 96/06166, WO 98/53057, WO 98/54311, WO 00/27878, WO 01/60970, WO 01/88197, and WO 02/099084, each of which is incorporated by reference in its entirety. By fusing this engineered protein domain to a nuclease domain, such as FokI nuclease, it is possible to target DNA breaks with genome-level specificity. The selection of target sites, zinc finger proteins and methods for design and construction of zinc finger nucleases are known to those of skill in the art and are described in detail in U.S. Publications Nos. 20030232410, 20050208489, 2005064474, 20050026157, 20060188987 and International Publication No. WO 07/014275, each of which is incorporated by reference in its entirety. Cleavage by a zinc finger nuclease can create a blunt end or a 5' overhang of variable length (frequently four basepairs).

[0200] As used herein, the term “TALEN” refers to an endonuclease comprising a DNA-binding domain comprising a plurality of TAL domain repeats fused to a nuclease domain or an active portion thereof from an endonuclease or exonuclease, including but not limited to a restriction endonuclease, homing endonuclease, S1 nuclease, mung bean nuclease, pancreatic DNase I, micrococcal nuclease, and yeast HO endonuclease. See, for example, Christian et al. (2010) *Genetics* 186:757-761, which is incorporated by reference in its entirety. Nuclease domains useful for the design of TALENs include those from a Type IIs restriction

endonuclease, including but not limited to FokI, FoM, StsI, HhaI, HindIII, Nod, BbvCI, EcoRI, BglI, and AlwI. Additional Type IIs restriction endonucleases are described in International Publication No. WO 2007/014275. In some embodiments, the nuclease domain of the TALEN is a FokI nuclease domain or an active portion thereof. TAL domain repeats can be derived from the TALE (transcription activator-like effector) family of proteins used in the infection process by plant pathogens of the *Xanthomonas* genus. TAL domain repeats are 33-34 amino acid sequences with divergent 12th and 13th amino acids. These two positions, referred to as the repeat variable dipeptide (RVD), are highly variable and show a strong correlation with specific nucleotide recognition. Each base pair in the DNA target sequence is contacted by a single TAL repeat, with the specificity resulting from the RVD. In some embodiments, the TALEN comprises 16-22 TAL domain repeats. DNA cleavage by a TALEN requires two DNA recognition regions flanking a nonspecific central region (*i.e.*, the “spacer”). The term “spacer” in reference to a TALEN refers to the nucleic acid sequence that separates the two nucleic acid sequences recognized and bound by each monomer constituting a TALEN. The TAL domain repeats can be native sequences from a naturally-occurring TALE protein or can be redesigned through rational or experimental means to produce a protein which binds to a pre-determined DNA sequence (see, for example, Boch et al. (2009) *Science* 326(5959):1509-1512 and Moscou and Bogdanove (2009) *Science* 326(5959):1501, each of which is incorporated by reference in its entirety). See also, U.S. Publication No. 20110145940 and International Publication No. WO 2010/079430 for methods for engineering a TALEN to recognize a specific sequence and examples of RVDs and their corresponding target nucleotides. In some embodiments, each nuclease (*e.g.*, FokI) monomer can be fused to a TAL effector sequence that recognizes a different DNA sequence, and only when the two recognition sites are in close proximity do the inactive monomers come together to create a functional enzyme.

[0201] As used herein, the term “compact TALEN” refers to an endonuclease comprising a DNA-binding domain with one or more TAL domain repeats fused in any orientation to any portion of the I-TevI homing endonuclease or any of the endonucleases listed in Table 2 in U.S. Application No. 20130117869 (which is incorporated by reference in its entirety), including but not limited to MmeI, EndA, End1, I-BasI, I-TevII, I-TevIII, I-TwoI, MspI, MvaI, NucA, and NucM. Compact TALENs do not require dimerization for DNA processing activity, alleviating the need for dual target sites with intervening DNA spacers. In some embodiments, the compact TALEN comprises 16-22 TAL domain repeats.

[0202] As used herein, the term “CRISPR” refers to a caspase-based endonuclease comprising a caspase, such as Cas9, and a guide RNA that directs DNA cleavage of the caspase by hybridizing to a recognition site in the genomic DNA. The caspase component of a CRISPR is an RNA-guided DNA endonuclease. In certain embodiments, the caspase is a class II Cas enzyme. In some of these embodiments, the caspase is a class II, type II enzyme, such as Cas9. In other embodiments, the caspase is a class II, type V enzyme, such as Cpf1. The guide RNA comprises a direct repeat and a guide sequence (often referred to as a spacer in the context of an endogenous CRISPR system), which is complementary to the target recognition site. In certain embodiments, the CRISPR further comprises a tracrRNA (trans-activating CRISPR RNA) that is complementary (fully or partially) to a direct repeat sequence (sometimes referred to as a tracr-mate sequence) present on the guide RNA. In particular embodiments, the caspase can be mutated with respect to a corresponding wild-type enzyme such that the enzyme lacks the ability to cleave one strand of a target polynucleotide, functioning as a nickase, cleaving only a single strand of the target DNA. Non-limiting examples of caspase enzymes that function as a nickase include Cas9 enzymes with a D10A mutation within the RuvC I catalytic domain, or with a H840A, N854A, or N863A mutation.

[0203] As used herein, the term “megaTAL” refers to a single-chain nuclease comprising a transcription activator-like effector (TALE) DNA binding domain with an engineered, sequence-specific homing endonuclease.

[0204] As used herein, the term “recognition sequence” refers to a DNA sequence that is bound and cleaved by an endonuclease. In the case of a meganuclease, a recognition sequence comprises a pair of inverted, 9 basepair “half sites” which are separated by four basepairs. In the case of a single-chain meganuclease, the N-terminal domain of the protein contacts a first half-site and the C-terminal domain of the protein contacts a second half-site. Cleavage by a meganuclease produces four basepair 3’ “overhangs”. “Overhangs”, or “sticky ends” are short, single-stranded DNA segments that can be produced by endonuclease cleavage of a double-stranded DNA sequence. In the case of meganucleases and single-chain meganucleases derived from I-CreI, the overhang comprises bases 10-13 of the 22 basepair recognition sequence. In the case of a compact TALEN, the recognition sequence comprises a first CNNNGN sequence that is recognized by the I-TevI domain, followed by a non-specific spacer 4-16 basepairs in length, followed by a second sequence 16-22 bp in length that is recognized by the TAL-effector domain (this sequence typically has a 5' T base). Cleavage by a Compact TALEN produces two basepair 3' overhangs. In the case of a

CRISPR, the recognition sequence is the sequence, typically 16-24 basepairs, to which the guide RNA binds to direct cleavage. Full complementarity between the guide sequence and the recognition sequence is not necessarily required to effect cleavage. Cleavage by a CRISPR can produce blunt ends (such as by a class II, type II caspase) or overhanging ends (such as by a class II, type V caspase), depending on the caspase. In those embodiments wherein a CpfI caspase is utilized, cleavage by the CRISPR complex comprising the same will result in 5' overhangs and in certain embodiments, 5 nucleotide 5' overhangs. Each caspase enzyme also requires the recognition of a PAM (protospacer adjacent motif) sequence that is near the recognition sequence complementary to the guide RNA. The precise sequence, length requirements for the PAM, and distance from the target sequence differ depending on the caspase enzyme, but PAMs are typically 2-5 base pair sequences adjacent to the target/recognition sequence. PAM sequences for particular caspase enzymes are known in the art (see, for example, U.S. Patent No. 8,697,359 and U.S. Publication No. 20160208243, each of which is incorporated by reference in its entirety) and PAM sequences for novel or engineered caspase enzymes can be identified using methods known in the art, such as a PAM depletion assay (see, for example, Karvelis et al. (2017) *Methods* 121-122:3-8, which is incorporated herein in its entirety). In the case of a zinc finger, the DNA binding domains typically recognize an 18-bp recognition sequence comprising a pair of nine basepair "half-sites" separated by 2-10 basepairs and cleavage by the nuclease creates a blunt end or a 5' overhang of variable length (frequently four basepairs).

[0205] As used herein, the term "target site" or "target sequence" refers to a region of the chromosomal DNA of a cell comprising a recognition sequence for a nuclease.

[0206] As used herein, the term "homologous recombination" or "HR" refers to the natural, cellular process in which a double-stranded DNA-break is repaired using a homologous DNA sequence as the repair template (see, *e.g.* Cahill *et al.* (2006), *Front. Biosci.* 11:1958-1976). The homologous DNA sequence may be an endogenous chromosomal sequence or an exogenous nucleic acid that was delivered to the cell.

[0207] As used herein, the term "non-homologous end-joining" or "NHEJ" refers to the natural, cellular process in which a double-stranded DNA-break is repaired by the direct joining of two non-homologous DNA segments (see, *e.g.* Cahill *et al.* (2006), *Front. Biosci.* 11:1958-1976).

[0208] As used herein, the term "reduced" refers to any reduction in the symptoms or severity of a disease or any reduction in the proliferation or number of cancerous cells. In either case, such a reduction may be up to 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%,

90%, 95%, or up to 100%. Accordingly, the term “reduced” encompasses both a partial reduction and a complete reduction of a disease state.

[0209] As used herein, the term “reduced” can also refer to a decrease in the cell surface expression of a polypeptide when compared to an appropriate control cell. In the present context, a reduction is distinct from knockout of polypeptide expression, wherein expression is reduced by 100%. Rather, in the present invention, a reduction indicates that expression is decreased but not completely eliminated. Such as a reduction can be, for example, a reduction in cell surface beta-2 microglobulin, MHC class I molecule, or CD52 expression when compared to a control cell which has not been genetically-modified to reduce beta-2 microglobulin, MHC class I molecules, or CD52, respectively. A reduction in expression can be between about 10% and about 99% or any number or range therein. For example, a reduction can be between about 10% and 95%, about 50% and about 95%, about 75% and about 95%, or about 90% and about 95%, when compared to a control cell. A reduction can also be by about 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% when compared to a control cell.

[0210] As used herein, the term “MHC class I molecule” refers to a major histocompatibility complex (MHC) found on the cell surface which displays peptide fragments of non-self proteins. MHC class I molecules consist of two polypeptide chains. The alpha chain consists of 3 polypeptides referred to as the alpha-1 ($\alpha 1$), alpha-2 ($\alpha 2$), and alpha-3 ($\alpha 3$) domains. The alpha chain is linked non-covalently via the $\alpha 3$ domain to a beta chain which consists of beta-2 microglobulin (B2M). The alpha chain is polymorphic and is encoded by the HLA gene (*i.e.*, HLA-A, HLA-B, and HLA-C), whereas beta-2 microglobulin is not polymorphic and it encoded by the B2M gene.

[0211] As used herein, the term “beta-2 microglobulin” refers to the beta chain component of MHC class I molecules. Human beta-2 microglobulin is encoded by the B2M gene (*e.g.*, NCBI Gene ID 567). Expression of beta-2 microglobulin is necessary for assembly and function of MHC class I molecules on the cell surface.

[0212] As used herein, the term “CD52” refers to the polypeptide encoded by the human CD52 gene (*e.g.*, NCBI gene ID 1043) which is also referred to as cluster of differentiation 52.

[0213] As used herein, the term “anti-tumor activity” or “anti-tumor effect” refers to a biological effect which can be manifested by a decrease in tumor volume, a decrease in the

number of tumor cells, a decrease in the number of metastases, an increase in life expectancy, or amelioration of various physiological symptoms associated with the cancerous condition. An “anti-tumor effect” can also be manifested by the ability of the genetically-modified cells of the present disclosure in prevention of the occurrence of tumor in the first place.

[0214] The term “effective amount” or “therapeutically effective amount” refers to an amount sufficient to effect beneficial or desirable biological and/or clinical results. The therapeutically effective amount will vary depending on the therapeutic (*e.g.*, genetically-modified cell, CAR T cell, etc.) formulation or composition, the disease and its severity, and the age, weight, physical condition and responsiveness of the subject to be treated. In specific embodiments, an effective amount of a cell comprising a co-stimulatory domain disclosed herein or pharmaceutical compositions disclosed herein reduces at least one symptom or the progression of a disease.

[0215] As used herein, the term “treat” or “treatment” means to reduce the frequency or severity of at least one sign or symptom of a disease or disorder experienced by a subject.

[0216] As used herein, the term “cancer” should be understood to encompass any neoplastic disease (whether invasive or metastatic) which is characterized by abnormal and uncontrolled cell division causing malignant growth or tumor.

[0217] As used herein, the term “carcinoma” refers to a malignant growth made up of epithelial cells.

[0218] As used herein, the term “leukemia” refers to malignancies of the hematopoietic organs/systems and is generally characterized by an abnormal proliferation and development of leukocytes and their precursors in the blood and bone marrow.

[0219] As used herein, the term “sarcoma” refers to a tumor which is made up of a substance like the embryonic connective tissue and is generally composed of closely packed cells embedded in a fibrillary, heterogeneous, or homogeneous substance.

[0220] As used herein, the term “melanoma” refers to a tumor arising from the melanocytic system of the skin and other organs.

[0221] As used herein, the term “lymphoma” refers to a group of blood cell tumors that develop from lymphocytes.

[0222] As used herein, the term “blastoma” refers to a type of cancer that is caused by malignancies in precursor cells or blasts (immature or embryonic tissue).

[0223] As used herein, “transfected” or “transformed” or “transduced” or “nucleofected” refers to a process by which exogenous nucleic acid is transferred or introduced into the host cell. A “transfected” or “transformed” or “transduced” cell is one which has been transfected,

transformed or transduced with exogenous nucleic acid. The cell includes the primary subject cell and its progeny.

[0224] As used herein, a “human T cell” or “T cell” refers to a T cell isolated from a human donor. Human T cells, and cells derived therefrom, include isolated T cells that have not been passaged in culture, T cells that have been passaged and maintained under cell culture conditions without immortalization, and T cells that have been immortalized and can be maintained under cell culture conditions indefinitely.

[0225] As used herein, a “human natural killer cell” or “human NK cell” or “natural killer cell” or “NK cell” refers to a type of cytotoxic lymphocyte critical to the innate immune system. The role NK cells play is analogous to that of cytotoxic T-cells in the vertebrate adaptive immune response. NK cells provide rapid responses to virally infected cells and respond to tumor formation, acting at around 3 days after infection.

[0226] As used herein, a “control” or “control cell” refers to a cell that provides a reference point for measuring changes in genotype or phenotype of a genetically-modified cell. A control cell may comprise, for example: (a) a wild-type cell, *i.e.*, of the same genotype as the starting material for the genetic alteration which resulted in the genetically-modified cell; (b) a cell of the same genotype as the genetically-modified cell but which has been transformed with a null construct (*i.e.*, with a construct which has no known effect on the trait of interest); or, (c) a cell genetically identical to the genetically-modified cell but which is not exposed to conditions, stimuli, or further genetic modifications that would induce expression of altered genotype or phenotype. In particular embodiments, a control cell is otherwise identical to a genetically-modified cell but has not been genetically-modified to reduce cell surface expression of a particular polypeptide (*e.g.*, beta-2 microglobulin, MHC class I molecules, CD52).

[0227] As used herein with respect to both amino acid sequences and nucleic acid sequences, the terms “percent identity,” “sequence identity,” “percentage similarity,” “sequence similarity,” and the like, refer to a measure of the degree of similarity of two sequences based upon an alignment of the sequences which maximizes similarity between aligned amino acid residues or nucleotides, and which is a function of the number of identical or similar residues or nucleotides, the number of total residues or nucleotides, and the presence and length of gaps in the sequence alignment. A variety of algorithms and computer programs are available for determining sequence similarity using standard parameters. As used herein, sequence similarity is measured using the BLASTp program for amino acid sequences and the BLASTn program for nucleic acid sequences, both of which are available

through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov), and are described in, for example, Altschul *et al.* (1990), *J. Mol. Biol.* 215:403-410; Gish and States (1993), *Nature Genet.* 3:266-272; Madden *et al.* (1996), *Meth. Enzymol.* 266:131-141; Altschul *et al.* (1997), *Nucleic Acids Res.* 25:33 89-3402; Zhang *et al.* (2000), *J. Comput. Biol.* 7(1-2):203-14. As used herein, percent similarity of two amino acid sequences is the score based upon the following parameters for the BLASTp algorithm: word size=3; gap opening penalty=-11; gap extension penalty=-1; and scoring matrix=BLOSUM62. As used herein, percent similarity of two nucleic acid sequences is the score based upon the following parameters for the BLASTn algorithm: word size=11; gap opening penalty=-5; gap extension penalty=-2; match reward=1; and mismatch penalty=-3.

[0228] As used herein with respect to modifications of two proteins or amino acid sequences, the term “corresponding to” is used to indicate that a specified modification in the first protein is a substitution of the same amino acid residue as in the modification in the second protein, and that the amino acid position of the modification in the first proteins corresponds to or aligns with the amino acid position of the modification in the second protein when the two proteins are subjected to standard sequence alignments (*e.g.*, using the BLASTp program). Thus, the modification of residue “X” to amino acid “A” in the first protein will correspond to the modification of residue “Y” to amino acid “A” in the second protein if residues X and Y correspond to each other in a sequence alignment, and despite the fact that X and Y may be different numbers.

[0229] As used herein, the recitation of a numerical range for a variable is intended to convey that the present disclosure may be practiced with the variable equal to any of the values within that range. Thus, for a variable which is inherently discrete, the variable can be equal to any integer value within the numerical range, including the end-points of the range. Similarly, for a variable which is inherently continuous, the variable can be equal to any real value within the numerical range, including the end-points of the range. As an example, and without limitation, a variable which is described as having values between 0 and 2 can take the values 0, 1 or 2 if the variable is inherently discrete, and can take the values 0.0, 0.1, 0.01, 0.001, or any other real values ≥ 0 and ≤ 2 if the variable is inherently continuous.

2.1 Principle of the Invention

[0230] The present disclosure is based, in part, on the observation that knockdown of cell surface beta-2 microglobulin, and consequently MHC class I molecules, can reduce allogenicity of genetically-modified cells, such as CAR T cells. Importantly, the inventors

have discovered that an incomplete knockdown of beta-2 microglobulin and MHC class I molecules (*i.e.*, to a low percentage of wild-type expression, but not complete knockout) not only reduces allogenicity of genetically-modified cells, but also serves to dramatically reduce killing by NK cells, which can recognize cells that are B2M-negative as non-self and induce a cytotoxic action.

[0231] The present invention is also based, in part, on the inventors' discovery that a population of CAR-positive T cells can be enriched by an advantageous negative-selection method when the CAR-encoding construct includes a coding sequence for an RNA interfering molecule against CD52. In this manner, a population of CAR T cells can be contacted with beads conjugated to an anti-CD52 antibody in order to capture CD52-positive cells. Separation of the beads, and thus the CD52-positive cells, results in an enriched population of CAR-positive cells having reduced cell surface expression of CD52.

[0232] Accordingly, a nucleic acid molecule is provided comprising a first expression cassette which encodes an engineered antigen receptor, such as a chimeric antigen receptor, and a second expression cassette which encodes an inhibitory nucleic acid molecule, such as an RNA interfering molecule. Further, the nucleic acid molecule is flanked by 5' and 3' homology arms to promote targeted insertion of the nucleic acid into the genome of a cell at a double-strand break, such as a cleavage site produced by an engineered nuclease. In certain embodiments of the invention, the inhibitory nucleic acid molecule can be against human beta-2 microglobulin, a component of the MHC class I molecule, or CD52.

[0233] Further disclosed herein are recombinant DNA constructs and viral vectors comprising the nucleic acid molecule, genetically-modified cells comprising the nucleic acid molecule, and pharmaceutical compositions comprising such cells. Also disclosed are genetically-modified cells expressing an engineered antigen receptor (*e.g.*, a CAR or exogenous TCR) which have reduced cell-surface expression of beta-2 microglobulin, MHC class I molecules, or CD52, and may or may not express the particular nucleic acid molecule of the invention.

[0234] In some embodiments, administration of genetically-modified cells of the invention reduces the symptoms or severity of diseases, such cancers, which can be targeted by genetically-modified cells of the present disclosure.

[0235] Also disclosed herein are methods of immunotherapy for treating cancer in a subject in need thereof comprising administering to the subject a pharmaceutical composition comprising a genetically-modified cell disclosed herein and a pharmaceutically acceptable carrier. In such methods, wherein a CAR is expressed and cell surface beta-2 microglobulin

and/or MHC class I molecules is reduced, incomplete knockout leads to a reduction in both allogenicity of the cells and killing of the cells by NK cells.

[0236] Further disclosed are methods for producing an enriched population of genetically-modified cells, wherein a CAR is expressed and cell surface CD52 is reduced by RNA interference, allowing for negative selection of CAR-positive cells having reduced CD52 expression.

2.2 Nucleic Acid Molecules

[0237] In certain embodiments, the invention provides a nucleic acid molecule comprising: (a) a first expression cassette comprising a nucleic acid sequence encoding an engineered antigen receptor; (b) a second expression cassette comprising a nucleic acid sequence encoding an inhibitory nucleic acid molecule; (c) a 5' homology arm; and (d) a 3' homology arm. The 5' homology arm and the 3' homology arm can be engineered at any suitable length to have homology to chromosomal regions flanking a nuclease recognition sequence in a gene of interest, which can be any desired gene in a target cell in which a suitable recognition sequence is present.

[0238] The nucleic acid molecule of the invention can have any number of orientations. Some non-limiting examples illustrated in Figure 1. In particular embodiments, the first and second expression cassettes can be in the same orientation. This orientation can be either 5' to 3' relative to the homology arms or, alternatively, 3' to 5'. In either case, the first expression cassette may be 5' to the second cassette, or the second cassette may be 5' to the first cassette. In other embodiments, the first and second expression cassettes can be in different orientations in the nucleic acid molecule. For example, the first expression cassette may be oriented 5' to 3', whereas the second expression cassette may be oriented 3' to 5'. Alternatively, the first expression cassette may be oriented 3' to 5' and the second expression cassette may be oriented 5' to 3'.

[0239] In embodiments wherein the expression cassettes are in opposite orientations, they may be oriented in a “tail-to-tail” configuration, such that the first expression cassette is oriented 3' to 5' and is positioned 5' to the second expression cassette, which is oriented 5' to 3'. In a similar tail-to-tail embodiment, the second expression cassette is oriented 3' to 5' and is positioned 5' to the first expression cassette, which is oriented 5' to 3'.

[0240] In other embodiments wherein the expression cassettes are in opposite orientations, they may be oriented in a “head-to-head” configuration, such that the first expression cassette is oriented 5' to 3' and is positioned 5' to the second expression cassette,

which is oriented 3' to 5'. In a similar head-to-head embodiment, the second expression cassette is oriented 5' to 3' and is positioned 5' to the first expression cassette, which is oriented 3' to 5'.

[0241] In some embodiments, the nucleic acid molecule can comprise multiple copies of the second expression cassette. The copies of the second expression cassette can be identical or vary from one another. In some cases, the copies can include a promoter, a coding sequence for the inhibitory nucleic acid molecule, and a sequence, such as a (cPPT/CTS) sequence, to terminate translation of the inhibitory nucleic acid molecule. The copies of the second expression cassette can be in tandem to one another in the nucleic acid molecule, and can be in the same orientation, or in opposite orientations. Alternatively, the copies may not be in tandem, and can be in the same orientation, or in opposite orientations.

[0242] The expression cassettes of the nucleic acid molecule can include various promoters which drive expression of the engineered antigen receptor and/or the inhibitory nucleic acid molecule. One example of a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. This promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. Another example of a suitable promoter is Elongation Growth Factor-1 α (EF-1 α). However, other constitutive promoter sequences may also be used, including, but not limited to the simian virus 40 (SV40) early promoter, mouse mammary tumor virus (MMTV), human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcoma virus promoter, as well as human gene promoters such as, but not limited to, the actin promoter, the myosin promoter, the hemoglobin promoter, and the creatine kinase promoter. Further, the present disclosure should not be limited to the use of constitutive promoters. Inducible promoters are also contemplated as part of the present disclosure. The use of an inducible promoter provides a molecular switch capable of turning on expression of the polynucleotide sequence which it is operatively linked when such expression is desired, or turning off the expression when expression is not desired. Examples of inducible promoters include, but are not limited to a metallothionein promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter.

[0243] Synthetic promoters are also contemplated as part of the present disclosure. For example, in particular embodiments, the promoter driving expression of the engineered antigen receptor is a JeT promoter (see, WO/2002/012514).

[0244] In some embodiments, the promoters are selected based on the desired outcome. It is recognized that different applications can be enhanced by the use of different promoters in the expression cassettes to modulate the timing, location and/or level of expression of the polynucleotides disclosed herein. Such expression constructs may also contain, if desired, a promoter regulatory region (*e.g.*, one conferring inducible, constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific/selective expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

[0245] Promoters particularly useful for driving expression of an RNA interference molecule are well known in the art and can include, without limitation, pol III promoters, such as U6 or H1.

[0246] The 5' and 3' homology arms of the nucleic acid molecule have sequence homology to corresponding sequences 5' upstream and 3' downstream of the nuclease recognition sequence in the genome. The homology arms promote insertion of the nucleic acid molecule into the cleavage site generated by the nuclease. In general, homology arms can have a length of at least 50 base pairs, preferably at least 100 base pairs, and up to 2000 base pairs or more, and can have at least 90%, preferably at least 95%, or more, sequence homology to their corresponding sequences in the genome.

[0247] In order to assess the expression of an engineered antigen receptor (*e.g.* a CAR or exogenous T cell receptor) in a genetically-modified cell, the nucleic acid molecule of the invention can optionally comprise an epitope which can be used to detect the presence of the encoded cell surface protein. In some examples described herein, a CAR coding sequence may include a QBend10 epitope which allows for detection using an anti-CD34 antibody (see, WO2013/153391).

[0248] In other examples, an expression cassette can also contain either a selectable marker gene or a reporter gene, or both, to facilitate identification and selection of expressing cells from the population of cells sought to be transfected or infected through viral vectors. In other aspects, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers include, for example, antibiotic-resistance genes and fluorescent marker genes.

[0249] Expression may also be assessed by determining protein expression of the polypeptide targeted by the inhibitory nucleic acid sequence. For example, expression of beta-2 microglobulin and CD52 can be detected on the cell surface by a number of techniques

known in the art. Expression can also be determined by positive or negative selection procedures which purify particular populations of cells expressing, or lacking expression, of the cell surface polypeptides.

[0250] Also provided herein are vectors comprising the nucleic acid molecules of the present disclosure. In some embodiments, the nucleic acid molecule is cloned into a vector including, but not limited to a plasmid, a phagemid, a phage derivative, an animal virus, or a cosmid. Vectors of particular interest include expression vectors, replication vectors, probe generation vectors, and sequencing vectors.

[0251] In other embodiments, nucleic acid molecules of the invention are provided on viral vectors, such as retroviral vectors, lentiviral vectors, adenoviral vectors, and adeno-associated viral (AAV) vectors. Viral vector technology is well known in the art and is described, for example, in Sambrook *et al.* (2001, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York), and in other virology and molecular biology manuals. Viruses, which are useful as vectors include, but are not limited to, retroviruses, adenoviruses, adeno-associated viruses, herpes viruses, and lentiviruses. In general, a suitable vector contains an origin of replication functional in at least one organism, a promoter sequence, convenient restriction endonuclease sites, and one or more selectable markers, (*e.g.*, WO 01/96584; WO 01/29058; and U.S. Pat. No. 6,326,193). Where the nucleic acid of the invention is provided in a viral vector that promotes random integration into the genome, and does not require the presence of 5' and 3' homology arms for homologous recombination, the nucleic acid of the invention can be provided without 5' and 3' homology arms.

2.3 Chimeric Antigen Receptors (CARs)

[0252] Provided herein are genetically-modified cells expressing an engineered antigen receptor. In some embodiments, the engineered antigen receptor is a chimeric antigen receptor (CAR). Generally, a CAR of the present disclosure will comprise at least an extracellular domain and an intracellular domain. In some embodiments, the extracellular domain comprises a target-specific binding element otherwise referred to as a ligand-binding domain or moiety. In some embodiments, the intracellular domain, or cytoplasmic domain, comprises at least one co-stimulatory domain and one or more signaling domains. In other embodiments, the CAR may only comprise a signaling domain, such as CD3 ζ , and the cell may comprise one or more co-stimulatory domains on another construct expressed in the cell.

[0253] In some embodiments, a CAR comprises an extracellular, target-specific binding element otherwise referred to as a ligand-binding domain or moiety. The choice of ligand-binding domain depends upon the type and number of ligands that define the surface of a target cell. For example, the ligand-binding domain may be chosen to recognize a ligand that acts as a cell surface marker on target cells associated with a particular disease state. Thus, examples of cell surface markers that may act as ligands for the ligand-binding domain in the CAR of the present disclosure can include those associated with viral, bacterial and parasitic infections, autoimmune disease, and cancer cells. In some embodiments, the CAR of the present disclosure is engineered to target a tumor antigen of interest by way of engineering a desired ligand-binding moiety that specifically binds to an antigen on a tumor cell. In the context of the present disclosure, “tumor antigen” refers to antigens that are common to specific hyperproliferative disorders such as cancer.

[0254] In some embodiments, the extracellular ligand-binding domain of the CAR is specific for any antigen or epitope of interest, particularly any tumor antigen or epitope of interest. As non-limiting examples, in some embodiments the antigen of the target is a tumor-associated surface antigen, such as ErbB2 (HER2/neu), carcinoembryonic antigen (CEA), epithelial cell adhesion molecule (EpCAM), epidermal growth factor receptor (EGFR), EGFR variant III (EGFRvIII), CD19, CD20, CD22, CD30, CD40, CLL-1, disialoganglioside GD2, ductal-epithelial mucine, gp36, TAG-72, glycosphingolipids, glioma-associated antigen, B-human chorionic gonadotropin, alphafetoprotein (AFP), lectin-reactive AFP, thyroglobulin, RAGE-1, MN-CA IX, human telomerase reverse transcriptase, RU1, RU2 (AS), intestinal carboxyl esterase, mut hsp70-2, M-CSF, prostase, prostase specific antigen (PSA), PAP, NY-ESO-1, LAGA-la, p53, prostein, PSMA, surviving and telomerase, prostate-carcinoma tumor antigen-1 (PCTA-1), MAGE, ELF2M, neutrophil elastase, ephrin B2, insulin growth factor (IGFI)-I, IGF-II, IGFI receptor, mesothelin, a major histocompatibility complex (MHC) molecule presenting a tumor-specific peptide epitope, 5T4, ROR1, Nkp30, NKG2D, tumor stromal antigens, the extra domain A (EDA) and extra domain B (EDB) of fibronectin and the A1 domain of tenascin-C (TnC A1) and fibroblast associated protein (fap); a lineage-specific or tissue specific antigen such as CD3, CD4, CD8, CD24, CD25, CD33, CD34, CD38, CD123, CD133, CD138, CTLA-4, B7- 1 (CD80), B7-2 (CD86), endoglin, a major histocompatibility complex (MHC) molecule, BCMA (CD269, TNFRSF 17), CS1, or a virus-specific surface antigen such as an HIV-specific antigen (such as HIV gp120); an EBV-specific antigen, a CMV-specific antigen, a HPV-specific antigen such as the E6 or E7 oncoproteins, a Lasse Virus-specific antigen, an Influenza Virus-specific

antigen, as well as any derivate or variant of these surface markers. In a particular embodiment of the present disclosure, the ligand-binding domain is specific for CD19.

[0255] In some embodiments, the extracellular domain of a chimeric antigen receptor further comprises an autoantigen (see, Payne *et al.* (2016) *Science*, Vol. 353 (6295): 179-184), which can be recognized by autoantigen-specific B cell receptors on B lymphocytes, thus directing T cells to specifically target and kill autoreactive B lymphocytes in antibody-mediated autoimmune diseases. Such CARs can be referred to as chimeric autoantibody receptors (CAARs), and the incorporation of one or more co-stimulatory domains described herein into such CAARs is encompassed by the present disclosure.

[0256] In some embodiments, the extracellular domain of a chimeric antigen receptor can comprise a naturally-occurring ligand for an antigen of interest, or a fragment of a naturally-occurring ligand which retains the ability to bind the antigen of interest.

[0257] In some embodiments, a CAR comprises a transmembrane domain which links the extracellular ligand-binding domain or autoantigen with the intracellular signaling and co-stimulatory domains via a hinge or spacer sequence. The transmembrane domain can be derived from any membrane-bound or transmembrane protein. For example, the transmembrane polypeptide can be a subunit of the T-cell receptor (*i.e.*, an α , β , γ or ζ , polypeptide constituting CD3 complex), IL2 receptor p55 (α chain), p75 (β chain) or γ chain, subunit chain of Fc receptors (*e.g.*, Fc γ receptor III) or CD proteins such as the CD8 α chain. Alternatively the transmembrane domain can be synthetic and can comprise predominantly hydrophobic residues such as leucine and valine. In particular examples, the transmembrane domain is a CD8 α transmembrane polypeptide.

[0258] The hinge region refers to any oligo- or polypeptide that functions to link the transmembrane domain to the extracellular ligand-binding domain. For example, a hinge region may comprise up to 300 amino acids, preferably 10 to 100 amino acids and most preferably 25 to 50 amino acids. Hinge regions may be derived from all or part of naturally occurring molecules, such as from all or part of the extracellular region of CD8, CD4 or CD28, or from all or part of an antibody constant region. Alternatively, the hinge region may be a synthetic sequence that corresponds to a naturally occurring hinge sequence, or may be an entirely synthetic hinge sequence. In particular examples, a hinge domain can comprise a part of a human CD8 α chain, Fc γ R11a receptor or IgG1.

[0259] The intracellular signaling domain of a CAR of the present disclosure is responsible for activation of at least one of the normal effector functions of the cell in which

the CAR has been placed and/or activation of proliferative and cell survival pathways. The term “effector function” refers to a specialized function of a cell. Effector function of a T cell, for example, may be cytolytic activity or helper activity including the secretion of cytokines. An intracellular signaling domain, such as CD3 ζ , can provide an activation signal to the cell in response to binding of the extracellular domain. As discussed, the activation signal can induce an effector function of the cell such as, for example, cytolytic activity or cytokine secretion.

[0260] The intracellular domain of the CAR can include one or more intracellular co-stimulatory domains which transmit a co-stimulatory signal to promote cell proliferation, cell survival, and/or cytokine secretion after binding of the extracellular domain. Such intracellular co-stimulatory domains include those known in the art such as, without limitation, CD27, CD28, CD8, 4-1BB (CD137), OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3 and a ligand that specifically binds with CD83, N1, or N6.

[0261] The CAR can be specific for any type of cancer cell. Such cancers can include, without limitation, carcinoma, lymphoma, sarcoma, blastomas, leukemia, cancers of B cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyosarcoma, leukemia, and Hodgkin’s lymphoma. In certain embodiments, cancers of B cell origin include, without limitation, B lineage acute lymphoblastic leukemia, B cell chronic lymphocytic leukemia, B cell non-Hodgkin’s lymphoma, and multiple myeloma.

2.4 Methods for Producing Recombinant Viral Vectors

[0262] In some embodiments, the present disclosure provides recombinant AAV vectors for use in the methods of the present disclosure. Recombinant AAV vectors are typically produced in mammalian cell lines such as HEK-293. Because the viral *cap* and *rep* genes are removed from the vector to prevent its self-replication to make room for the therapeutic gene(s) to be delivered (*e.g.* the endonuclease gene), it is necessary to provide these *in trans* in the packaging cell line. In addition, it is necessary to provide the “helper” (*e.g.* adenoviral) components necessary to support replication (Cots D, Bosch A, Chillon M (2013) *Curr. Gene Ther.* 13(5): 370-81). Frequently, recombinant AAV vectors are produced using a triple-transfection in which a cell line is transfected with a first plasmid encoding the “helper” components, a second plasmid comprising the *cap* and *rep* genes, and a third plasmid

comprising the viral ITRs containing the intervening DNA sequence to be packaged into the virus. Viral particles comprising a genome (ITRs and intervening gene(s) of interest) encased in a capsid are then isolated from cells by freeze-thaw cycles, sonication, detergent, or other means known in the art. Particles are then purified using cesium-chloride density gradient centrifugation or affinity chromatography and subsequently delivered to the gene(s) of interest to cells, tissues, or an organism such as a human patient. Accordingly, methods are provided herein for producing recombinant AAV vectors comprising the nucleic acid molecules of the invention described herein.

[0263] In some embodiments, genetic transfer is accomplished via lentiviral vectors. Lentiviruses, in contrast to other retroviruses, in some contexts may be used for transducing certain non-dividing cells. Non-limiting examples of lentiviral vectors include those derived from a lentivirus, such as Human Immunodeficiency Virus 1 (HIV-1), HIV-2, an Simian Immunodeficiency Virus (SIV), Human T-lymphotropic virus 1 (HTLV-1), HTLV-2 or equine infection anemia virus (EIAV). For example, lentiviral vectors have been generated by multiply attenuating the HIV virulence genes, for example, the genes *env*, *vif*, *vpr*, *vpu* and *nef* are deleted, making the vector safer for therapeutic purposes. Lentiviral vectors are known in the art, see Naldini *et al.*, (1996 and 1998); Zufferey *et al.*, (1997); Dull *et al.*, 1998, U.S. Pat. Nos. 6,013,516; and 5,994,136). In some embodiments, these viral vectors are plasmid-based or virus-based, and are configured to carry the essential sequences for incorporating foreign nucleic acid, for selection, and for transfer of the nucleic acid into a host cell. Known lentiviruses can be readily obtained from depositories or collections such as the American Type Culture Collection (“ATCC”; 10801 University Blvd., Manassas, Va. 20110-2209), or isolated from known sources using commonly available techniques.

[0264] In specific embodiments, lentiviral vectors are prepared using a first plasmid encoding the *gag*, *pol*, *tat*, and *rev* genes cloned from human immunodeficiency virus (HIV) and a second plasmid encoding the envelope protein from vesicular stomatitis virus (VSV-G) used to pseudotype viral particles. A transfer vector, such as the pCDH-EF1-MCS vector, can be used with a suitable promoter. All three plasmids can then be transfected into lentivirus cells, such as the Lenti-X-293T cells, and lentivirus can then be harvested, concentrated and screened after a suitable incubation time. Accordingly, methods are provided herein for producing recombinant lentiviral vectors comprising the nucleic acid molecule of the invention described herein.

2.5 Genetically-Modified Cells

[0265] Provided herein are cells genetically-modified to comprise the nucleic acid molecule of the invention described herein. Further provided are genetically-modified cells (*e.g.*, human T cells expressing a CAR or exogenous TCR) with reduced cell-surface expression of beta-2 microglobulin, MHC class I molecules, and or CD52, which do not necessarily comprise the particular nucleic acid molecule of the invention.

[0266] In different variations of the present disclosure, a nucleic acid molecule of the invention is present within the genome of the genetically-modified cell or, alternatively, is not integrated into the genome of the cell. In particular embodiments, the nucleic acid molecule of the invention is inserted into the genome of a cell by targeted insertion at a cleavage site produced by a double-strand break, such as that produced by an engineered nuclease. The presence of 5' and 3' homology arms flanking the first and second expression cassettes of the nucleic acid molecule promote homologous recombination of the nucleic acid molecule into the cleavage site, resulting in targeted insertion.

[0267] In some embodiments where the nucleic acid molecule is not integrated into the genome, the nucleic acid molecule can be present in the genetically-modified cell in a recombinant DNA construct, in an mRNA, in a viral genome, or other nucleic acid which is not integrated into the genome of the cell.

[0268] In specific embodiments, the cells comprising the nucleic acid molecule of the invention, and other genetically-modified cells of the invention, are eukaryotic cells. In particular embodiments, such cells are T cells or NK cells, particularly human T cells or NK cells. In some embodiments, the cells are primary T cells or primary NK cells.

[0269] T cells and NK cells can be obtained from a number of sources, including peripheral blood mononuclear cells, bone marrow, lymph node tissue, cord blood, thymus tissue, tissue from a site of infection, ascites, pleural effusion, spleen tissue, and tumors. In certain embodiments of the present disclosure, any number of T cell and NK cell lines available in the art may be used. In some embodiments of the present disclosure, T cells and NK cells are obtained from a unit of blood collected from a subject using any number of techniques known to the skilled artisan. In one embodiment, cells from the circulating blood of an individual are obtained by apheresis.

[0270] Genetically-modified cells comprising the nucleic acid molecule disclosed herein, and other genetically-modified cells of the invention, can exhibit a number of functional properties dependent upon which polypeptide is reduced in the cell and/or targeted by the inhibitory nucleic acid molecule. For example, in some genetically-modified cells of the

invention, beta-2 microglobulin is reduced, or the inhibitory nucleic acid molecule is against human beta-2 microglobulin, and cell surface beta-2 microglobulin expression is reduced, to a small percentage of wild-type expression. Such genetically-modified cells can be less susceptible to endogenous NK cell killing, have extended persistence time in a subject, exhibit enhanced expansion in a subject, and/or have reduced allogenicity than cells with wild-type levels of B2M or cells which are completely B2M-negative. Reductions in beta-2 microglobulin consequently result in a reduction in cell surface expression of MHC class I molecules, because beta-2 microglobulin is necessary for their assembly and function. Therefore, the same properties are also applicable to genetically-modified cells of the invention which have reduced cell surface expression of MHC class I molecules.

[0271] Susceptibility to NK cell killing can be determined by methods known in the art such as those described further herein. Reductions in NK cell killing can be by about 5%, 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or up to 100% when compared to a control cell.

[0272] The genetically-modified cells of the invention are capable of expansion in a subject following administration. Here, expansion is considered an increase in cell number resulting from proliferation and division *in vivo*. The degree of expansion depends, in part, on the subject's response to the cells; for example, if the cells are identified as allogeneic and/or non-self, the subject's immune system may reduce the ability of the cells to expand and further reduce persistence of the cells post-administration. Thus, in some examples, genetically-modified cells of the invention can exhibit an increase in expansion in a subject that is about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450%, 500%, up to 1000%, or more, when compared to a control cell. Expansion *in vivo* can be determined post-administration by any method known in the art. Persistence time of a genetically-modified cell in a subject can be considered, for example, as the amount of time post-administration of the cell that it can be detected in the subject by any method known in the art. In some examples, a genetically-modified cell of the invention will have an increase in persistence time that is up to about 1 week, 2 weeks, 3 weeks, 4 weeks, 5 weeks, 6 weeks, 7 weeks, 2 months, 3 months, 4 months, 5 months, 6 months, 1 year, or more, longer than a control cell.

[0273] Allogenicity can be determined by any method known in the art, such as those methods described further herein. The genetically-modified cells of the invention can exhibit a reduction in allogenicity of about 5%, 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%,

75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96, 97%, 98%, 99%, or up to 100% when compared to a control cell.

2.6 Methods for Producing Genetically-Modified Cells

[0274] The present disclosure provides methods for producing genetically-modified cells comprising the nucleic acid molecule of the invention described herein. In specific embodiments, methods are provided for modifying the cell to comprise the nucleic acid molecule. In different aspects of the present disclosure, the nucleic acid molecule is integrated into the genome of the cell or, alternatively, is not integrated into the genome of the cell.

[0275] In some embodiments, the nucleic acid molecule is introduced into a cell using any technology known in the art. In specific embodiments, vectors or expression cassettes comprising the nucleic acid molecule disclosed herein is introduced into a cell using a viral vector. Such vectors are known in the art and include lentiviral vectors, adenoviral vectors, and adeno-associated virus (AAV) vectors (reviewed in Vannucci, *et al.* (2013) *New Microbiol.* 36:1-22). Recombinant AAV vectors useful in the present disclosure can have any serotype that allows for transduction of the virus into the cell and insertion of the nucleic acid gene into the cell and, in particular embodiments, into the cell genome. In particular embodiments, recombinant AAV vectors have a serotype of AAV2, AAV6, or AAV8. Recombinant AAV vectors can also be self-complementary such that they do not require second-strand DNA synthesis in the host cell (McCarty, *et al.* (2001) *Gene Ther.* 8:1248-54).

[0276] In some embodiments, nucleic acid molecules disclosed herein are delivered into a cell in the form of DNA (*e.g.*, circular or linearized plasmid DNA or PCR products) and/or via a viral vector. In some embodiments, the nucleic acid molecule disclosed herein is coupled covalently or non-covalently to a nanoparticle or encapsulated within such a nanoparticle using methods known in the art (Sharma, *et al.* (2014) *Biomed Res Int.* 2014). A nanoparticle is a nanoscale delivery system whose length scale is <1 μm , preferably <100 nm. Such nanoparticles may be designed using a core composed of metal, lipid, polymer, or biological macromolecule, and multiple copies of the nucleic acid molecules can be attached to or encapsulated with the nanoparticle core. This increases the copy number of the DNA that is delivered to each cell and, so, increases the intracellular expression to maximize the likelihood that the encoded products will be expressed. The surface of such nanoparticles may be further modified with polymers or lipids (*e.g.*, chitosan, cationic polymers, or cationic

lipids) to form a core-shell nanoparticle whose surface confers additional functionalities to enhance cellular delivery and uptake of the payload (Jian *et al.* (2012) *Biomaterials*. 33(30): 7621-30). Nanoparticles may additionally be advantageously coupled to targeting molecules to direct the nanoparticle to the appropriate cell type and/or increase the likelihood of cellular uptake. Examples of such targeting molecules include antibodies specific for cell surface receptors and the natural ligands (or portions of the natural ligands) for cell surface receptors.

[0277] In some embodiments, the nucleic acid molecule disclosed herein can be encapsulated within liposomes or complexed using cationic lipids (see, *e.g.*, LIPOFECTAMINE, Life Technologies Corp., Carlsbad, CA; Zuris *et al.* (2015) *Nat Biotechnol.* 33: 73-80; Mishra *et al.* (2011) *J Drug Deliv.* 2011:863734). The liposome and lipoplex formulations can protect the payload from degradation, and facilitate cellular uptake and delivery efficiency through fusion with and/or disruption of the cellular membranes of the cells.

[0278] In some embodiments, the nucleic acid molecule disclosed herein can be encapsulated within polymeric scaffolds (*e.g.*, PLGA) or complexed using cationic polymers (*e.g.*, PEI, PLL) (Tamboli *et al.* (2011) *Ther Deliv.* 2(4): 523-536). In some embodiments, the nucleic acid molecule disclosed herein can be combined with amphiphilic molecules that self-assemble into micelles (Tong *et al.* (2007) *J Gene Med.* 9(11): 956-66). Polymeric micelles may include a micellar shell formed with a hydrophilic polymer (*e.g.*, polyethyleneglycol) that can prevent aggregation, mask charge interactions, and reduce nonspecific interactions outside of the cell.

[0279] In some embodiments, the nucleic acid molecule disclosed herein can be formulated as an emulsion for delivery to the cell. The term “emulsion” refers to, without limitation, any oil-in-water, water-in-oil, water-in-oil-in-water, or oil-in-water-in-oil dispersions or droplets, including lipid structures that can form as a result of hydrophobic forces that drive apolar residues (*e.g.*, long hydrocarbon chains) away from water and polar head groups toward water, when a water immiscible phase is mixed with an aqueous phase. These other lipid structures include, but are not limited to, unilamellar, paucilamellar, and multilamellar lipid vesicles, micelles, and lamellar phases. Emulsions are composed of an aqueous phase and a lipophilic phase (typically containing an oil and an organic solvent). Emulsions also frequently contain one or more surfactants. Nanoemulsion formulations are well known, *e.g.*, as described in US Patent Application Nos. 2002/0045667 and 2004/0043041, and US Pat. Nos. 6,015,832, 6,506,803, 6,635,676, and 6,559,189, each of which is incorporated herein by reference in its entirety.

[0280] In some embodiments, the nucleic acid molecule disclosed herein can be covalently attached to, or non-covalently associated with, multifunctional polymer conjugates, DNA dendrimers, and polymeric dendrimers (Mastorakos *et al.* (2015) *Nanoscale*. 7(9): 3845-56; Cheng *et al.* (2008) *J Pharm Sci.* 97(1): 123-43). The dendrimer generation can control the payload capacity and size, and can provide a high payload capacity. Moreover, display of multiple surface groups can be leveraged to improve stability and reduce nonspecific interactions.

[0281] Methods of introducing and expressing genes into a cell are known in the art. In the context of an expression vector, the vector can be readily introduced into a host cell, *e.g.*, mammalian, bacterial, yeast, or insect cell by any method in the art. For example, the expression vector can be transferred into a host cell by physical, chemical, or biological means. Physical methods for introducing a polynucleotide into a host cell include calcium phosphate precipitation, lipofection, particle bombardment, microinjection, electroporation, and the like. Methods for producing cells comprising vectors and/or exogenous nucleic acids are well-known in the art. See, for example, Sambrook *et al.* (2001, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York). A preferred method for the introduction of a polynucleotide into a host cell is calcium phosphate transfection. Biological methods for introducing a polynucleotide of interest into a host cell include the use of DNA and RNA vectors. Viral vectors, and especially retroviral vectors, have become the most widely used method for inserting genes into mammalian, *e.g.*, human cells. Other viral vectors can be derived from lentivirus, poxviruses, herpes simplex virus I, adenoviruses and adeno-associated viruses, and the like. See, for example, U.S. Pat. Nos. 5,350,674 and 5,585,362. Chemical means for introducing a polynucleotide into a host cell include colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. An exemplary colloidal system for use as a delivery vehicle *in vitro* and *in vivo* is a liposome (*e.g.*, an artificial membrane vesicle).

2.7 Pharmaceutical Compositions

[0282] In some embodiments, the present disclosure provides a pharmaceutical composition comprising a genetically-modified cell, or a population of genetically-modified cells, of the present disclosure and a pharmaceutically-acceptable carrier. Such pharmaceutical compositions can be prepared in accordance with known techniques. See, *e.g.*, Remington, *The Science And Practice of Pharmacy* (21st ed. 2005). In the manufacture

of a pharmaceutical formulation according to the present disclosure, cells are typically admixed with a pharmaceutically acceptable carrier and the resulting composition is administered to a subject. The carrier must, of course, be acceptable in the sense of being compatible with any other ingredients in the formulation and must not be deleterious to the subject. In some embodiments, pharmaceutical compositions of the present disclosure further comprises one or more additional agents useful in the treatment of a disease in the subject. In additional embodiments, where the genetically-modified cell is a genetically-modified human T cell or NK cell (or a cell derived therefrom), pharmaceutical compositions of the present disclosure further include biological molecules, such as cytokines (*e.g.*, IL-2, IL-7, IL-15, and/or IL-21), which promote *in vivo* cell proliferation and engraftment. Pharmaceutical compositions comprising genetically-modified cells of the present disclosure can be administered in the same composition as an additional agent or biological molecule or, alternatively, can be co-administered in separate compositions.

[0283] In some embodiments, the pharmaceutical compositions of the present disclosure are useful for treating any disease state that can be targeted by T cell adoptive immunotherapy. In a particular embodiment, the pharmaceutical compositions of the present disclosure are useful as immunotherapy in the treatment of cancer. Such cancers can include, without limitation, carcinoma, lymphoma, sarcoma, blastomas, leukemia, cancers of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyo sarcoma, leukemia, and Hodgkin's lymphoma. In certain embodiments, cancers of B-cell origin include, without limitation, B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma. Non-limiting examples of cancer which may be treated with the pharmaceutical compositions and medicaments of the present disclosure are carcinomas, lymphomas, sarcomas, melanomas, blastomas, leukemias, and germ cell tumors, including but not limited to cancers of B-cell origin, neuroblastoma, osteosarcoma, prostate cancer, renal cell carcinoma, rhabdomyosarcoma, liver cancer, gastric cancer, bone cancer, pancreatic cancer, skin cancer, cancer of the head or neck, breast cancer, lung cancer, cutaneous or intraocular malignant melanoma, renal cancer, uterine cancer, ovarian cancer, colorectal cancer, colon cancer, rectal cancer, cancer of the anal region, stomach cancer, testicular cancer, uterine cancer, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, non-Hodgkin's lymphoma, cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland,

cancer of the adrenal gland, sarcoma of soft tissue, cancer of the urethra, cancer of the penis, solid tumors of childhood, lymphocytic lymphoma, cancer of the bladder, cancer of the kidney or ureter, carcinoma of the renal pelvis, neoplasm of the central nervous system (CNS), primary CNS lymphoma, tumor angiogenesis, spinal axis tumor, brain stem glioma, pituitary adenoma, Kaposi's sarcoma, epidermoid cancer, squamous cell cancer, environmentally induced cancers including those induced by asbestos, multiple myeloma, Hodgkin lymphoma, non-Hodgkin's lymphomas, acute myeloid lymphoma, chronic myelogenous leukemia, chronic lymphoid leukemia, immunoblastic large cell lymphoma, acute lymphoblastic leukemia, mycosis fungoides, anaplastic large cell lymphoma, and T-cell lymphoma, and any combinations of said cancers. In certain embodiments, cancers of B-cell origin include, without limitation, B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, B-cell lymphoma, diffuse large B cell lymphoma, pre-B ALL (pediatric indication), mantle cell lymphoma, follicular lymphoma, marginal zone lymphoma, Burkitt's lymphoma, and B-cell non-Hodgkin's lymphoma.

[0284] In some of these embodiments wherein cancer is treated with the presently disclosed genetically-modified cells, the subject administered the genetically-modified cells is further administered an additional therapeutic, such as radiation, surgery, or a chemotherapeutic agent.

[0285] The invention further provides a population of genetically-modified cells comprising a plurality of genetically-modified cells described herein. Such genetically-modified cells can comprise in their genome a nucleic acid molecule encoding an engineered antigen receptor, such as a chimeric antigen receptor or exogenous T cell receptor, and an inhibitory nucleic acid molecule, such as an RNA interference molecule. Such genetically-modified cells can also comprise in their genome a nucleic acid molecule encoding an engineered antigen receptor, such as a chimeric antigen receptor or exogenous T cell receptor, and have reduced cell surface expression of beta-2 microglobulin, MHC class I molecules, or CD52, without necessarily comprising the particular nucleic acid molecule of the invention. Thus, in various embodiments of the invention, a population of genetically-modified cells is provided wherein at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to 100%, of cells in the population are a genetically-modified cell described herein.

2.8 Methods of Administering Genetically-Modified Cells

[0286] Another aspect disclosed herein is the administration of the genetically-modified cells of the present disclosure to a subject in need thereof. In particular embodiments, the pharmaceutical compositions described herein are administered to a subject in need thereof. For example, an effective amount of a genetically-modified cell or population of genetically-modified cells of the invention which express a cell surface chimeric antigen receptor or exogenous T cell receptor, can be administered to a subject having a disease. In particular embodiments, the disease can be cancer, such as a cancer of B-cell origin. Thus, the present disclosure also provides a method for providing a T cell-mediated immune response to a target cell population or tissue in a mammal, comprising the step of administering to the mammal a CAR T cell, wherein the CAR comprises an extracellular ligand-binding domain that specifically interacts with a predetermined target, such as a tumor antigen, and an intracellular domain that comprises at least one signaling domain, such as CD3 ζ , and optionally one or more co-stimulatory signaling domains. The administered CAR T cells are able to reduce the proliferation, reduce the number, or kill target cells in the recipient. Unlike antibody therapies, genetically-modified cells of the present disclosure are able to replicate and expand *in vivo*, resulting in long-term persistence that can lead to sustained control of a disease.

[0287] In examples wherein the inhibitory nucleic acid molecule is against human beta-2 microglobulin or a component of the MHC class I molecule, or wherein beta-2 microglobulin or MHC class I molecules are otherwise reduced, expansion and/or persistence of such CAR T cells can be enhanced in the subject when compared to a CAR T cell having wild-type levels of beta-2 microglobulin or MHC class I molecules, or no cell surface beta-2 microglobulin or MHC class I expression. Further, allogenicity of the CAR T cell can be reduced when compared to a CAR T cell having a wild-type level of cell surface expression of beta-2 microglobulin and MHC class I molecules. These advantageous characteristics result from the incomplete reduction of cell surface beta-2 microglobulin (and consequently MHC class I molecules) to a small percentage of wild-type expression, which allows for reduced allogenicity but avoidance of NK cells which would otherwise target a beta-2 microglobulin-negative cell.

[0288] Examples of possible routes of administration include parenteral, (*e.g.*, intravenous (IV), intramuscular (IM), intradermal, subcutaneous (SC), or infusion) administration. Moreover, the administration may be by continuous infusion or by single or multiple boluses. In specific embodiments, one or both of the agents is infused over a period

of less than about 12 hours, 6 hours, 4 hours, 3 hours, 2 hours, or 1 hour. In still other embodiments, the infusion occurs slowly at first and then is increased over time.

[0289] In some embodiments, a genetically-modified eukaryotic cell or population thereof of the present disclosure targets a tumor antigen for the purposes of treating cancer. Such cancers can include, without limitation, carcinoma, lymphoma, sarcoma, blastomas, leukemia, cancers of B cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyosarcoma, leukemia, and Hodgkin's lymphoma. In specific embodiments, cancers and disorders include but are not limited to pre-B ALL (pediatric indication), adult ALL, mantle cell lymphoma, diffuse large B cell lymphoma, salvage post allogenic bone marrow transplantation, and the like. These cancers can be treated using a combination of CARs that target, for example, CD19, CD20, CD22, and/or ROR1. In some non-limiting examples, a genetically-modified eukaryotic cell or population thereof of the present disclosure targets carcinomas, lymphomas, sarcomas, melanomas, blastomas, leukemias, and germ cell tumors, including but not limited to cancers of B-cell origin, neuroblastoma, osteosarcoma, prostate cancer, renal cell carcinoma, rhabdomyosarcoma, liver cancer, gastric cancer, bone cancer, pancreatic cancer, skin cancer, cancer of the head or neck, breast cancer, lung cancer, cutaneous or intraocular malignant melanoma, renal cancer, uterine cancer, ovarian cancer, colorectal cancer, colon cancer, rectal cancer, cancer of the anal region, stomach cancer, testicular cancer, uterine cancer, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, non-Hodgkin's lymphoma, cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of soft tissue, cancer of the urethra, cancer of the penis, solid tumors of childhood, lymphocytic lymphoma, cancer of the bladder, cancer of the kidney or ureter, carcinoma of the renal pelvis, neoplasm of the central nervous system (CNS), primary CNS lymphoma, tumor angiogenesis, spinal axis tumor, brain stem glioma, pituitary adenoma, Kaposi's sarcoma, epidermoid cancer, squamous cell cancer, environmentally induced cancers including those induced by asbestos, multiple myeloma, Hodgkin lymphoma, non-Hodgkin's lymphomas, acute myeloid lymphoma, chronic myelogenous leukemia, chronic lymphoid leukemia, immunoblastic large cell lymphoma, acute lymphoblastic leukemia, mycosis fungoides, anaplastic large cell lymphoma, and T-cell lymphoma, and any combinations of said cancers. In certain embodiments, cancers of B-cell origin include, without limitation, B-lineage acute lymphoblastic leukemia, B-cell chronic

lymphocytic leukemia, B-cell lymphoma, diffuse large B cell lymphoma, pre-B ALL (pediatric indication), mantle cell lymphoma, follicular lymphoma, marginal zone lymphoma, Burkitt's lymphoma, multiple myeloma, and B-cell non-Hodgkin's lymphoma.

[0290] When an “effective amount” or “therapeutic amount” is indicated, the precise amount of the compositions of the present disclosure to be administered can be determined by a physician with consideration of individual differences in age, weight, tumor size (if present), extent of infection or metastasis, and condition of the patient (subject). In some embodiments, a pharmaceutical composition comprising the genetically-modified cells described herein is administered at a dosage of 10^4 to 10^9 cells/kg body weight, including all integer values within those ranges. In further embodiments, the dosage is 10^5 to 10^6 cells/kg body weight, including all integer values within those ranges. In some embodiments, cell compositions are administered multiple times at these dosages. The cells can be administered by using infusion techniques that are commonly known in immunotherapy (see, *e.g.*, Rosenberg *et al.*, New Eng. J. of Med. 319:1676, 1988). The optimal dosage and treatment regime for a particular patient can readily be determined by one skilled in the art of medicine by monitoring the patient for signs of disease and adjusting the treatment accordingly.

[0291] In some embodiments, administration of genetically-modified cells of the present disclosure reduce at least one symptom of a target disease or condition. For example, administration of genetically-modified cells of the present disclosure can reduce at least one symptom of a cancer, such as cancers of B-cell origin. Symptoms of cancers, such as cancers of B-cell origin, are well known in the art and can be determined by known techniques.

EXPERIMENTAL

[0292] This disclosure is further illustrated by the following examples, which should not be construed as limiting. Those skilled in the art will recognize, or be able to ascertain, using no more than routine experimentation, numerous equivalents to the specific substances and procedures described herein. Such equivalents are intended to be encompassed in the scope of the claims that follow the examples below.

EXAMPLE 1

NK cell killing of B2M knockout primary human T cells

1. Methods and Materials

[0293] Primary human T cells were stimulated for 3 days using ImmunoCult anti-CD2/CD3/CD28 (StemCell Technologies) in the presence of IL-2 (Gibco) in XVIVO-15 medium (Lonza) supplemented with 5% fetal bovine serum. RNA encoding B2M13-14x479 nuclease was introduced into the T cells using the 4D Nucleofector (Lonza). Cells were cultured in the presence of IL-2 for 6 days before magnetic depletion of remaining B2M⁺ cells using biotinylated anti-human B2M (BioLegend) and a Biotin Selection Kit (StemCell Technologies). NK cells were isolated from PBMC samples of the same donor using a CD56 positive selection kit (StemCell Technologies). Daudi cells were purchased from ATCC. Daudi cells are naturally B2M⁻ and are reported to be highly sensitive to NK cytotoxicity. All target cells were labeled with 1μM CellTrace Violet (LifeTechnologies) to distinguish them from effectors in mixed cultures. Isolated NK cells were mixed with either autologous B2M⁺ T cell targets (negative control for NK cytotoxicity), Daudi targets (positive control for NK cytotoxicity), or autologous B2M KO T cell targets (experimental sample) at effector:target ratios of 2:1, 1:1, 0.5:1, and 0:1. Killing was assessed after 2h of co-culture. Killing by NK cells was measured by staining with CaspACE - VAD-FMK (Promega).

2. Results

[0294] NK cells elicited only dim VAD-FMK signals in autologous B2M⁺ targets, indicating low levels of apoptosis induction by active caspases (Figures 2A-2D). In comparison, high VAD-FMK signals were induced in large percentages of Class I- Daudi cells (71-83%, Figures 2I-2L), indicating extensive caspase cascade activation. Similarly, B2M⁻ autologous T cells returned high VAD-FMK signals in response to NK encounter (19-47%, Figure 2 E-H), indicative of caspase-mediated apoptosis induction. A graphical summary of these results appears in Figure 3.

3. Conclusions

[0295] Complete knockout of cell surface B2M using engineered meganucleases sensitizes primary human T cells to NK cell attack and killing.

EXAMPLE 2

Characterization of candidate shRNAs against B2M and effect of B2M knockdown on NK cell killing of primary human T cells

1. Materials and Methods

[0296] Five Mission-shRNA lentiviral transfer plasmids encoding different B2M targeting sequences were purchased from Sigma-Aldrich. Second-generation lentiviral vectors were produced in-house using Lenti-X 293T cells (ClonTech) and a triple transfection method (Lipofectamine 2000 – Thermo-Fisher). T cells were prepared for lentiviral transduction by stimulating for 3 days with ImmunoCult anti-CD2/CD3/CD28 as in Example 1. Transduction was carried out in the presence of 5uM polybrene (Sigma-Aldrich) and transduced cells were selected with puromycin (InVivoGen) beginning at 48h post-transduction and concluding 72h following drug addition. Selected cells were expanded for 5 days in IL-2 supplemented medium before a flow cytometric analysis of B2M surface expression to determine the extent of knockdown. Cultures receiving B2M shRNAs were used as targets in NK and CTL killing assays. The NK killing assays were carried out as described in Example 1, but the K562 cell line was used as the positive control for NK cytolysis. For the CTL killing assay, CD8+ T cells from a donor unrelated to the donor of the target cells were isolated and used as effectors. The NK killing assay was carried out for 2h and the CTL assay was carried out for 6h. For both assays, target cells were labeled with 1uM CellTrace Violet (Life Technologies), and killing was measured using CaspACE-VAD-FMK (Promega).

2. Results

[0297] Five shRNAs were screened in human T cells for interference with B2M expression. Two sequences did not reduce the mean fluorescence intensity of B2M in a cytometric analysis (not shown). Three shRNA sequences did reduce the MFI of B2M expression, with sequence 254 and 255 reducing MFI by approximately 50% and sequence 472 reducing the MFI by approximately 95% (Figure 4).

[0298] CTL and NK killing of targets exhibiting altered B2M expression was next measured. NK cells, but not CTLs induced caspase activation (measured by VAD-FMK signal) in Class I deficient K562 cells (46% vs. 5% - Figure 5A and 5B). Conversely, CTLs induced caspase activation (32%) in mismatched B2M⁺ T cells while NK cells induced a signal in a lower frequency of mismatched T cells (14%) (Figure 5C and 5D). In T cell targets exhibiting a 50% reduction in B2M antigen density, NK cells elicited caspase activity in 17% of targets while mismatched CTLs did so in 36% of targets (Figure 6A and 6B). In T cell targets exhibiting a 95% knockdown of B2M levels, NK cells elicited caspase activation in 16% of targets, while mismatched CTLs did so in 20.8% of targets (Figure 5 C and D).

3. Conclusions

[0299] B2M expression can be effectively knocked down using shRNA delivered by a viral vector. Using caspase (VAD-FMK) activity to measure apoptosis induction in target cells by NK cells or CTLs, it was determined that B2M knockdown does not alter a target's susceptibility to NK cytotoxicity, as both B2M knockdown targets exhibited the same VAD-FMK frequency as un-manipulated targets, and less VAD-FMK signal than K562 targets. In addition, B2M knockdown confers some protection against CTL cytotoxicity, as the frequency of VAD-FMK+ targets in the shRNA 472 group was approximately half the frequency observed in the positive control. In fact, there was a direct relationship between the degree of knockdown and the degree of protection against CTL activity from NK cells.

EXAMPLE 3

Production and characterization of CAR T cells utilizing shRNA to reduce cell surface expression of B2M

1. Materials and methods

[0300] A number of constructs were prepared comprising an anti-CD19 CAR coding sequence and an shRNA against B2M. These are illustrated in Figure 7A-7F and are provided in SEQ ID NOs: 18-23. CAR constructs 7007 and 7217 (SEQ ID NOs: 18 and 19) comprise the CAR coding sequence and the shRNA472 sequence in the same 5' to 3' orientation. CAR constructs 7008 and 7218 (SEQ ID NOs: 20 and 21) comprise the CAR coding sequence in the 3' to 5' orientation, and shRNA472 sequence in the 5' to 3' orientation (*i.e.*, tail-to-tail). CAR constructs 7009 and 7219 (SEQ ID NOs: 22 and 23) comprise both the CAR coding sequence and the shRNA472 sequence in the 3' to 5' orientation. The 5' and 3' homology arms flanking the CAR coding sequence and the shRNA472 sequence have homology to regions upstream and downstream of the TRC 1-2 recognition sequence in the TRAC locus.

[0301] CAR T cells will be prepared using primary donor human T cells transduced with recombinant AAV vectors comprising one of the CAR/shRNA constructs above, with simultaneous nucleofection of mRNA encoding the TRC 1-2x.87EE to induce a double-strand break at the TRC 1-2 recognition sequence and promote targeted insertion of the construct into the genome of the T cells. Beta-2 microglobulin expression will be determined as described above to determine which orientation of the first and second expression cassettes

will result in the highest and/or the most consistent CAR expression, along with the most consistent level of B2M knockdown on the cell surface.

[0302] CAR T cells produced with certain constructs will be evaluated in both the allogenicity and NK cell killing assays previously described above. Further, CAR T cells produced using the disclosed constructs will be evaluated in various stress tests, in which the CAR T cells are repeatedly exposed to antigen in order to determine changes in cell proliferation/expansion and cytotoxic potential. CAR T cells produced using the disclosed constructs will also be utilized with *in vivo* tumor models to determine their ability to clear tumor cells in an animal and to evaluate their ability to persist *in vivo*. It is expected, based on the Examples described herein, that CAR T cells having a reduced but incomplete knockdown of cell surface beta-2 microglobulin will have greater persistence and/or enhanced expansion *in vivo* when compared to CAR T cells which are completely B2M-negative and may be susceptible to NK cell killing.

[0303] In a particular study, CAR T cells were prepared that are TCR-negative, CAR-positive, and have reduced cell surface expression of B2M. CAR T cells were prepared using donor templates that comprise a promoter-driven CAR coding sequence, a T2A element, and one or multiple promoter-driven B2M shRNA cassettes. In this study, an apheresis sample was drawn from a healthy, informed, and compensated donor, and the T cells were enriched using the CD3 positive selection kit II in accord with the manufacturer's instructions (Stem Cell Technologies). T cells were activated using ImmunoCult T cell stimulator (anti-CD2/CD3/CD28 – Stem Cell Technologies) in X-VIVO 15 medium (Lonza) supplemented with 5% fetal bovine serum and 10ng/ml IL-2 (Gibco). After 3 days of stimulation, cells were collected and samples of 1e6 cells were electroporated with the following mixture of nucleic acid species using the Lonza 4D NucleoFector.

- 1µg mRNA encoding the TRC 1-2x.87EE meganuclease which produces a double-strand break in Exon 1 of the T cell receptor alpha constant region gene
- 1.5 µl of 100mM siRNA specific for TMEM173 (STING)
- 1µg of linearized plasmid DNA comprising a donor template

In this experiment, three different CAR constructs were analyzed for their ability to knock down B2M surface expression. All three constructs use homology to genomic regions flanking the TRC 1-2x.87EE binding site (referred to as the TRC 1-2 recognition site) to direct targeted insertion into the T cell receptor alpha constant region locus, and they all express a CAR that comprises a CD34 epitope tag (for detection). Construct 7002 (SEQ ID

NO: 11) does not encode an shRNA gene. Construct 7008 (SEQ ID NO: 20) encodes one copy of shRNA472. Construct 7029 (SEQ ID NO: 24) encodes two copies of this shRNA cassette. Expression from each shRNA cassette is driven by a U6 promoter.

[0304] Cell cultures were maintained for 10 additional days in X-VIVO15 medium supplemented with 5% FBS and 30ng/ml of IL-2. On d4, 7, and 10 post-nucleofection, the cultures were sampled and analyzed for surface expression of CD3 (anti-CD3-BV711, BioLegend), CD34 (anti-CD34-PE, LifeTechnologies), and B2M (anti-B2M-APC, BioLegend). Flow cytometry data were acquired on a Beckman-Coulter CytoFLEX-LX.

2. Results

[0305] B2M surface levels were measured in samples nucleofected with a control CAR construct (7002) or with CAR constructs expressing one (7008) or two (7029) copies of B2M-specific shRNA (Figure 12). When comparing the CD3-/CD34+ populations in 7002 (control) and 7008 (single shRNA) expressing cells, 7008 expressing cells were observed to display slightly lower levels of surface B2M (Figure 12A). Notably, cells nucleofected with construct 7029 (two shRNA copies) displayed approximately half of the amount of B2M displayed on the surface of control cells (7002) (Figure 12B). This observation was specific to the CD3-/CD34+ population, but was not observed in the CD3-/CD34- population (Figure 12C).

3. Conclusions

[0306] A pre-screened B2M-targeting shRNA can knock down B2M expression levels on the surface of cells into which the construct has been delivered (via targeted insertion into the T cell receptor alpha constant region locus). Due to the high abundance of B2M transcripts, these data suggest that a single shRNA copy can be sufficient for low levels of B2M knockdown, whereas multiple copies of the shRNA cassette may be required to achieve more significant knockdown.

EXAMPLE 4

Production and characterization of CAR T cells utilizing shRNA to reduce cell surface expression of B2M1. Materials and methods

[0307] A number of constructs were prepared comprising an anti-CD19 CAR coding sequence and an shRNA against B2M. These are illustrated in Figure 7A-7F and are provided in SEQ ID NOs: 18-23. CAR constructs 7007 and 7217 (SEQ ID NOs: 18 and 19) comprise the CAR coding sequence and the shRNA472 sequence in the same 5' to 3' orientation. CAR constructs 7008 and 7218 (SEQ ID NOs: 20 and 21) comprise the CAR coding sequence in the 3' to 5' orientation, and shRNA472 sequence in the 5' to 3' orientation (*i.e.*, tail-to-tail). CAR constructs 7009 and 7219 (SEQ ID NOs: 22 and 23) comprise both the CAR coding sequence and the shRNA472 sequence in the 3' to 5' orientation. CAR constructs 7056, 7059, and 7060 contain modified versions of the U6-shRNA gene cassette. A cloning site that was located between the U6 promoter and the hairpin sequence in constructs 7007-7009, and in 7217-7219 was removed. Construct 7056 comprises the CAR coding sequence and the shRNA472 sequence in the 3' to 5' orientation. Construct 7056 comprises the CAR coding sequence in 3' to 5' orientation, and the shRNA472 sequence in the 5' to 3' orientation (tail-to-tail). Construct 7060 comprises the CAR coding sequence in 3' to 5' orientation and two copies of the U6-shRNA472 sequence in 5' to 3' orientation. The 5' and 3' homology arms flanking the CAR coding sequence and the shRNA472 sequence have homology to regions upstream and downstream of the TRC 1-2 recognition sequence in the T cell receptor alpha constant locus.

[0308] In a particular study, CAR T cells were prepared that are TCR-negative, CAR-positive, and have reduced cell surface expression of B2M. CAR T cells were prepared using donor templates that comprise a promoter-driven CAR coding sequence, and one or multiple promoter-driven B2M shRNA cassettes. In this study, an apheresis sample was drawn from a healthy, informed, and compensated donor, and the T cells were enriched using the CD3 positive selection kit II in accord with the manufacturer's instructions (Stem Cell Technologies). T cells were activated using ImmunoCult T cell stimulator (anti-CD2/CD3/CD28 – Stem Cell Technologies) in X-VIVO 15 medium (Lonza) supplemented with 5% fetal bovine serum and 10ng/ml IL-2 (Gibco). After 3 days of stimulation, cells were collected and samples of 1e6 cells were electroporated with the following mixture of nucleic acid species using the Lonza 4D NucleoFector.

- 1µg mRNA encoding the TRC 1-2x.87EE meganuclease which produces a double-strand break in Exon 1 of the T cell receptor alpha constant region gene
- 1.5 µl of 100mM siRNA specific for TMEM173 (STING)
- 1µg of linearized plasmid DNA comprising a donor template

In this experiment, four different CAR constructs were analyzed for their ability to knock down B2M surface expression. All four constructs use homology to genomic regions flanking the TRC 1-2x.87EE recognition site to direct targeted insertion into the T cell receptor alpha constant region locus, and they all express a CAR that comprises a CD34 epitope tag (for detection). Construct 7002 (SEQ ID NO: 11) does not encode an shRNA gene. Construct 7056 (SEQ ID NO: 25) encodes one copy of shRNA472, and both cassettes are in the 3' to 5' orientation (head-to-tail). Construct 7059 (SEQ ID NO: 26) encodes one copy of this shRNA cassette, with the CAR expression cassette in a 3' to 5' orientation, and the shRNA472 cassette in a 5' to 3' orientation (tail-to-tail). Construct 7060 (SEQ ID NO: 27) is in the same orientation as construct 7059 but encodes two copies of the shRNA472 cassette (tail-to-tail). Expression from each shRNA cassette is driven by a U6 promoter. Cell cultures were maintained for up to 10 additional days in X-VIVO15 medium supplemented with 5% FBS and 30ng/ml of IL-2. On d4, 7, and/or 10 post-nucleofection, the cultures were sampled and analyzed for surface expression of CD3 (anti-CD3-BV711, BioLegend), CD34 (anti-CD34-PE, or APC, LifeTechnologies), B2M (anti-B2M-APC, or PE, BioLegend), and/or HLA-A, B, and C (clone W6/32, BV605). Flow cytometry data were acquired on a Beckman-Coulter CytoFLEX-LX.

2. Results

[0309] B2M surface levels were measured in samples nucleofected with a control CAR construct (7002) or with CAR constructs expressing one (7056 or 7059) or two (7060) copies of B2M-specific shRNA in either head-to-tail (7056) or tail-to-tail (7059, 7060) configurations. A restriction digest site that was present in previous constructs between the U6 promoter and the shRNA sequence was been removed from these shRNA472 vectors. It was hypothesized that the palindromic restriction digest site interfered with the efficacy of the constructs and the ability of the shRNA to knock down B2M.

[0310] When comparing the CD3-/CD34+ populations in 7002 (control) and 7056 (single shRNA) expressing cells, 7056 expressing cells were observed to display lower levels of surface B2M (77% knockdown) (Figure 13A). Cells nucleofected with 7059 (single copy,

tail-to-tail) displayed a 90.1% knockdown of B2M relative to 7002 control cells (Figure 13B), while 7060 nucleofection (two copies, tail-to-tail) resulted in a 92% knockdown relative to 7002 controls (Figure 13C).

3. Conclusions

[0311] A pre-screened B2M-targeting shRNA can knock down B2M expression levels on the surface of cells into which the construct has been delivered (via targeted insertion into the T cell receptor alpha constant region locus). Removing the cloning site between the U6 promoter and the hairpin sequence improves the efficiency with which B2M is knocked down. 7008 (tail-to-tail, one shRNA472 cassette - Figure 12A) supports minimal B2M knockdown while 7059 (one cassette, tail-to-tail – Figure 13B) supports greater than 90% knockdown. As was observed using a CD52-specific shRNA, superior knockdown was observed when the CAR promoter and the shRNA promoter were oriented in different directions (tail-to-tail configuration). Adding a second shRNA sequence did not provide any noticeable benefit (92% versus 90.1% knockdown).

EXAMPLE 5

Production and characterization of CAR T cells utilizing shRNA to reduce cell surface expression of B2M

1. Materials and methods

[0312] In this study, an apheresis sample was drawn from a healthy, informed, and compensated donor, and the T cells were enriched using the CD3 positive selection kit II in accord with the manufacturer's instructions (Stem Cell Technologies). T cells were activated using ImmunoCult T cell stimulator (anti-CD2/CD3/CD28 – Stem Cell Technologies) in X-VIVO 15 medium (Lonza) supplemented with 5% fetal bovine serum and 10ng/ml IL-2 (Gibco). After 3 days of stimulation, cells were collected and samples of 1e6 cells were electroporated with 1ug of RNA encoding the TRC 1-2L.1592 meganuclease, which recognizes and cleaves the TRC 1-2 recognition sequence in the T cell receptor alpha constant locus, and were transduced with AAV packaged with construct 7056 at an MOI of 25000 viral genomes/cell.

[0313] Cell cultures were maintained for up to 10 additional days in X-VIVO15 medium supplemented with 5% FBS and 30ng/ml of IL-2. On day 4, 7, and/or 10 post-nucleofection, the cultures were sampled and analyzed for surface expression of CD3 (anti-CD3-PE,

BioLegend), (anti-FMC63 anti-CAR clone VM16 conjugated to AlexaFluor488), B2M (anti-B2M-APC, or PE, BioLegend), and HLA-A, B, and C (clone W6/32, BV605). Flow cytometry data were acquired on a Beckman-Coulter CytoFLEX-LX.

2. Results

[0314] B2M and HLA-ABC levels were measured in samples expressing construct 7056 and control populations. Figure 14A shows the B2M surface levels in CD3-/CAR+ cells compared to TRAC-edited cells expressing no shRNA from a control culture. Figure 14B shows B2M levels on CD3-/CAR+ versus CD3+/CAR- populations in the same culture. Figure 14C and 14D make the same respective comparisons in displays of HLA-ABC surface levels. The CD3-/CAR+ fraction of cells transduced with AAV-7056 displayed levels of B2M and HLA-ABC that are reduced by greater than 90% compared to control populations.

3. Conclusions

[0315] A pre-screened B2M-targeting shRNA can knock down B2M expression levels on the surface of cells into which the construct has been delivered (via targeted insertion into the T cell receptor alpha constant region locus). This effect is specific to CAR+ populations (*i.e.*, cells in which targeted integration into the TRAC locus has occurred). This experiment demonstrates that B2M can be efficiently knocked down using a single copy of shRNA472 co-delivered to the TRAC locus with the CAR gene on the same AAV template.

EXAMPLE 6

Characterization of candidate shRNAs against CD52 in primary human T cells

1. Materials and Methods

[0316] Five Mission-shRNA lentiviral transfer plasmids encoding different CD52 targeting sequences were purchased from Sigma-Aldrich. Second-generation lentiviral vectors were produced in-house using Lenti-X 293T cells (ClonTech) and a triple transfection method (Lipofectamine 2000 – Thermo-Fisher). T cells were prepared for lentiviral transduction by stimulating for 3 days with ImmunoCult anti-CD2/CD3/CD28 as in Example 1. Transduction was carried out in the presence of 5uM polybrene (Sigma-Aldrich) and transduced cells were expanded for 5 days in IL-2 supplemented medium before a flow cytometric analysis of CD52 surface levels. Cells were not selected with puromycin because a heterogeneous population was desired for downstream attempts at magnetic depletion of

CD52Hi cells. Cells transduced with a lentivirus encoding shRNA 568 were labeled with biotinylated anti-CD52 (Miltenyi Biotec), and magnetic separation was performed using a Biotin Positive Selection Kit (StemCell Technologies). A post-separation analysis of surface CD52 was performed.

2. Results

[0317] Of the 5 shRNA sequences screened, 3 (shRNA568, shRNA572, and shRNA876) interfered with CD52 expression. CD52 surface expression profiles are displayed in Figure 8. Levels of CD52 displayed on the surface of T cells are shown in Figure 9 for mock transduced T cells (9A), T cells transduced with an shRNA-568 lentivirus (9B), and LV-shRNA568 transduced cells that have undergone a CD52 magnetic depletion (9C).

3. Conclusions

[0318] CD52 antigen density on the surface of cells can be reduced using shRNA delivered by a viral vector. Sequence 568 exhibited the highest degree of CD52 knockdown. Knockdown of CD52 using this shRNA sequence was sufficient to allow for magnetic depletion of non-transduced CD52 Hi cells.

EXAMPLE 7

CD52 knockdown profiles using CAR/CD52 constructs with different orientations

1. Materials and Methods

[0319] T cells were stimulated for 3 days using ImmunoCult anti-CD2/CD3/CD28 as described in EXAMPLE 1. After 3 days, TRC 1-2x.87EE mRNA, STING siRNA, and linearized AAV transfer vector encoding different CAR constructs (Figure 10) were delivered to the T cells using the 4-D Nucleofector (Lonza). Cultures of nucleofected T cells were carried for 10 days in medium supplemented with IL-2 prior to flow cytometric analyses of CD3, CAR (CD34 epitope-tagged), and CD52.

2. Results

[0320] To demonstrate un-manipulated levels of CD52 surface display on TRAC-edited CAR T cells, a TRC 1-2x.87EE nuclease and a CD34-tagged CAR construct encoding no shRNA sequence were delivered. CD52 levels on TCR KO cells, TCR KO CAR+ cells and nonedited cells are overlaid in the histogram in Figure 11A. Three CAR constructs encoding a U6 promoter-controlled CD52 shRNA were evaluated for ability to knock down CD52 when integrated into the TRAC locus. When the CAR gene and the shRNA cassette are both in forward orientation, CD52 antigen density is reduced by approximately 50% (Figure 11C; construct 7004). Reversing the transcriptional orientation of the CAR gene alone (*i.e.*, a tail-to-tail configuration) reduces the amount of CD52 displayed on the surface by approximately 95% (Figure 11B; construct 7013), while reversing the orientation of both the CAR gene and the U6-shRNA element reduces the CD52 signal by approximately 90% (Figure 11D; construct 7014).

3. Conclusions

[0321] The CD52 specific shRNA sequence 568 can interfere with CD52 expression when only one copy is delivered by targeted insertion into the TRAC locus. Altering the transcriptional orientation of either the CAR gene only (*i.e.*, tail-to-tail configuration) or both the CAR and shRNA genes can influence the efficiency of target gene knockdown. Reversing only the CAR gene's orientation resulted in the most efficient knockdown.

4. Further Studies

[0322] A number of constructs were prepared comprising an anti-CD19 CAR coding sequence and an shRNA against CD52. These are illustrated in Figure 10A-10H and are provided in SEQ ID NOs: 10-17. As described above, CAR constructs 7004 (SEQ ID NO: 12), 7013 (SEQ ID NO: 14), and 7014 (SEQ ID NO: 16) were previously evaluated for their ability to reduce CD52 expression while expressing a CAR. The 5' and 3' homology arms flanking the CAR coding sequence and the shRNA sequence have homology to regions upstream and downstream of the TRC 1-2 recognition sequence in the TRAC locus.

[0323] In additional studies, CAR T cells will be prepared using primary donor human T cells transduced with recombinant AAV vectors comprising one of the CAR/shRNA constructs above, with simultaneous nucleofection of mRNA encoding the TRC 1-2x.87EE to induce a double-strand break at the TRC 1-2 recognition sequence and promote targeted

insertion of the construct into the genome of the T cells. CD52 expression will be determined as described above to determine which orientation of the first and second expression cassettes will result in the highest and/or the most consistent CAR expression, along with the most consistent level of CD52 knockdown on the cell surface.

[0324] CAR T cells produced with certain constructs will be evaluated in both the allogenicity and NK cell killing assays previously described above. Further, CAR T cells produced using the disclosed constructs will be evaluated in various stress tests, in which the CAR T cells are repeatedly exposed to antigen in order to determine changes in cell proliferation/expansion and cytotoxic potential. CAR T cells produced using the disclosed constructs will also be utilized with *in vivo* tumor models to determine their ability to clear tumor cells in an animal and to evaluate their ability to persist *in vivo*. It is expected, based on the Examples described herein, that enriched populations of CAR T cells can be produced for *in vivo* use by an advantageous negative-selection for CAR T cells having reduced cell surface expression of CD52.

CLAIMS

1. A nucleic acid molecule comprising:
 - (a) a first expression cassette comprising a nucleic acid sequence encoding an engineered antigen receptor;
 - (b) a second expression cassette comprising a nucleic acid sequence encoding an inhibitory nucleic acid molecule;
 - (c) a 5' homology arm; and
 - (d) a 3' homology arm;wherein said 5' homology arm and said 3' homology arm have homology to chromosomal regions flanking a nuclease recognition sequence in a gene of interest.
2. The nucleic acid molecule of claim 1, wherein said inhibitory nucleic acid molecule is an RNA interference molecule.
3. The nucleic acid molecule of claim 2, wherein said RNA interference molecule is a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a hairpin siRNA, a microRNA (miRNA), a precursor miRNA, or an miRNA-adapted shRNA.
4. The nucleic acid molecule of claim 2 or claim 3, wherein said RNA interference molecule is an shRNA.
5. The nucleic acid molecule of any one of claims 1-4, wherein said engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.
6. The nucleic acid molecule of any one of claims 1-5, wherein said nuclease recognition sequence is an engineered meganuclease recognition sequence, a TALEN recognition sequence, a zinc finger nuclease (ZFN) recognition sequence, a CRISPR/Cas recognition sequence, a compact TALEN recognition sequence, or a megaTAL recognition sequence.
7. The nucleic acid molecule of any one of claims 1-6, wherein said nuclease recognition sequence is an engineered meganuclease recognition sequence.

8. The nucleic acid molecule of any one of claims 1-7, wherein said gene of interest is a human T cell receptor alpha constant region gene.
9. The nucleic acid molecule of any one of claims 1-8, wherein said nuclease recognition sequence comprises SEQ ID NO: 1.
10. The nucleic acid molecule of any one of claims 1-9, wherein said first expression cassette further comprises a promoter which drives expression of said engineered antigen receptor.
11. The nucleic acid molecule of claim 10, wherein said promoter is a JeT promoter.
12. The nucleic acid molecule of any one of claims 1-11, wherein said second expression cassette further comprises a promoter which drives expression of said inhibitory nucleic acid molecule.
13. The nucleic acid molecule of claim 12, wherein said promoter is a U6 promoter.
14. The nucleic acid molecule of any one of claims 1-13, wherein said first expression cassette comprises a polyadenylation signal to terminate translation of said engineered antigen receptor.
15. The nucleic acid molecule of any one of claims 1-14, wherein said second expression cassette comprises a central polypurine tract and central terminator sequence (cPPT/CTS) sequence to terminate translation of said inhibitory nucleic acid.
16. The nucleic acid molecule of any one of claims 1-15, wherein said first expression cassette and said second expression cassette are in the same orientation in said nucleic acid molecule.

17. The nucleic acid molecule of claim 16, wherein said first expression cassette and said second expression cassette are in a 5' to 3' orientation relative to said 5' and 3' homology arms.

18. The nucleic acid molecule of claim 17, wherein said first expression cassette is 5' upstream of said second expression cassette.

19. The nucleic acid molecule of claim 17, wherein said second expression cassette is 5' upstream of said first expression cassette.

20. The nucleic acid molecule of claim 16, wherein said first expression cassette and said second expression cassette are in a 3' to 5' orientation relative to said 5' and 3' homology arms.

21. The nucleic acid molecule of claim 20, wherein said first expression cassette is 5' upstream of said second expression cassette.

22. The nucleic acid molecule of claim 20, wherein said second expression cassette is 5' upstream of said first expression cassette.

23. The nucleic acid molecule of any one of claims 1-15, wherein said first expression cassette and said second expression cassette are in opposite orientations in said nucleic acid molecule.

24. The nucleic acid molecule of claim 23, wherein said first expression cassette is in a 3' to 5' orientation and said second expression cassette is in a 5' to 3' orientation relative to said 5' and 3' homology arms.

25. The nucleic acid molecule of claim 24, wherein said first expression cassette is 5' upstream of said second expression cassette.

26. The nucleic acid molecule of claim 24, wherein said second expression cassette is 5' upstream of said first expression cassette.

27. The nucleic acid molecule of claim 23, wherein said first expression cassette is in a 5' to 3' orientation and said second expression cassette is in a 3' to 5' orientation relative to said 5' and 3' homology arms.

28. The nucleic acid molecule of claim 27, wherein said first expression cassette is 5' upstream of said second expression cassette.

29. The nucleic acid molecule of claim 27, wherein said second expression cassette is 5' upstream of said first expression cassette.

30. The nucleic acid molecule of any one of claims 1-29, wherein said nucleic acid molecule further comprises a 5' inverted terminal repeat and a 3' inverted terminal repeat flanking said first expression cassette and said second expression cassette.

31. The nucleic acid molecule of any one of claims 1-30, wherein said inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin, a component of the MHC class I molecule, or CD52.

32. The nucleic acid molecule of any one of claims 1-31, wherein said inhibitory nucleic acid molecule is an shRNA inhibitory against beta-2 microglobulin, wherein said shRNA has a sequence comprising any one of SEQ ID NOs: 2-4.

33. The nucleic acid molecule of claim 32, wherein said shRNA has a sequence comprising SEQ ID NO: 2.

34. The nucleic acid molecule of claim 32 or claim 33, wherein said first expression cassette and said second expression cassette are in a 3' to 5' orientation relative to said 5' and 3' homology arms, and wherein said first expression cassette is 5' upstream of said second expression cassette.

35. The nucleic acid molecule of claim 34, wherein said first expression cassette comprises:

(i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor;

(ii) a JeT promoter which drives expression of said chimeric antigen receptor or said exogenous T cell receptor; and

(iii) a polyA sequence;

and wherein said second expression cassette comprises:

(iv) a nucleic acid sequence encoding said shRNA;

(v) a U6 promoter which drives expression of said shRNA; and

(vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence.

36. The nucleic acid molecule of claim 32 or claim 33, wherein said first expression cassette is in a 3' to 5' orientation and said second expression cassette is in a 5' to 3' orientation relative to said 5' and 3' homology arms, and wherein said first expression cassette is 5' upstream of said second expression cassette.

37. The nucleic acid molecule of claim 36, wherein said first expression cassette comprises:

(i) a nucleic acid sequence encoding a chimeric antigen receptor or an exogenous T cell receptor;

(ii) a JeT promoter which drives expression of said chimeric antigen receptor or said exogenous T cell receptor; and

(iii) a polyA sequence;

and wherein said second expression cassette comprises:

(iv) a nucleic acid sequence encoding said shRNA;

(v) a U6 promoter which drives expression of said shRNA; and

(vi) a central polypurine tract and central terminator sequence (cPPT/CTS) sequence.

38. The nucleic acid molecule of claim 37, wherein said nucleic acid molecule comprises a first copy and a second copy of said second expression cassette, wherein said first copy and said second copy are identical, and wherein said first copy and said second copy are in tandem, and wherein said first copy and said second copy are in the same orientation.

39. A recombinant DNA construct comprising said nucleic acid molecule of any one of claims 1-38.

40. The recombinant DNA construct of claim 39, wherein said recombinant DNA construct encodes a viral vector.

41. The recombinant DNA construct of claim 40, wherein said viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an adeno-associated viral (AAV) vector.

42. The recombinant DNA construct of claim 41, wherein said viral vector is a recombinant AAV vector.

43. A viral vector comprising said nucleic acid molecule of any one of claims 1-38.

44. The viral vector of claim 43, wherein said viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an AAV vector.

45. The viral vector of claim 44, wherein said viral vector is a recombinant AAV vector.

46. A genetically-modified eukaryotic cell comprising said nucleic acid molecule of any one of claims 1-31, wherein said engineered antigen receptor and said inhibitory nucleic acid molecule are expressed in said genetically-modified eukaryotic cell.

47. The genetically-modified eukaryotic cell of claim 46, wherein said engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

48. The genetically-modified eukaryotic cell of claim 46 or claim 47, wherein said nucleic acid molecule is inserted into the genome of said genetically-modified eukaryotic cell at said nuclease recognition sequence.

49. The genetically-modified eukaryotic cell of any one of claims 46-48, wherein said nuclease recognition sequence is an engineered meganuclease recognition sequence, a TALEN recognition sequence, a zinc finger nuclease (ZFN) recognition sequence, a CRISPR/Cas recognition sequence, a compact TALEN recognition sequence, or a megaTAL recognition sequence.

50. The genetically-modified eukaryotic cell of claim 49, wherein said nuclease recognition sequence is an engineered meganuclease recognition sequence.

51. The genetically-modified eukaryotic cell of any one of claims 46-50, wherein said gene of interest is a human T cell receptor alpha constant region gene.

52. The genetically-modified eukaryotic cell of any one of claims 46-51, wherein said nuclease recognition sequence comprises SEQ ID NO: 1.

53. The genetically-modified eukaryotic cell of claim 51 or claim 52, wherein cell surface expression of an endogenous T cell receptor is reduced compared to a control cell.

54. The genetically-modified eukaryotic cell of any one of claims 46-53, wherein said inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin.

55. The genetically-modified eukaryotic cell of claim 54, wherein said genetically-modified eukaryotic cell comprises said nucleic acid molecule of any one of claims 32-38.

56. The genetically-modified eukaryotic cell of claim 54 or claim 55, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

57. The genetically-modified eukaryotic cell of any one of claims 54-56, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

58. The genetically-modified eukaryotic cell of any one of claims 54-57, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

59. The genetically-modified eukaryotic cell of any one of claims 54-58, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

60. The genetically-modified eukaryotic cell of any one of claims 56-59, wherein said control cell is not genetically-modified to reduce cell surface beta-2 microglobulin expression.

61. The genetically-modified eukaryotic cell of any one of claims 46-53, wherein said inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule.

62. The genetically-modified eukaryotic cell of claim 61, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

63. The genetically-modified eukaryotic cell of claim 62 or claim 63, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

64. The genetically-modified eukaryotic cell of any one of claims 61-63, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

65. The genetically-modified eukaryotic cell of any one of claims 61-64, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

66. The genetically-modified eukaryotic cell of any one of claims 62-65, wherein said control cell is not genetically-modified to reduce cell surface expression of MHC class I molecules.

67. The genetically-modified eukaryotic cell of any one of claims 54-66, wherein said genetically-modified eukaryotic cell is less susceptible to endogenous NK cell killing when compared to a control cell, wherein said genetically-modified eukaryotic cell has extended persistence in a subject when compared to a control cell, wherein said genetically-modified eukaryotic cell exhibits enhanced expansion in a subject when compared to a control cell, and/or wherein allogenicity of said genetically-modified eukaryotic cell is reduced when compared to a control cell.

68. The genetically-modified eukaryotic cell of any one of claims 46-67, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell.

69. The genetically-modified eukaryotic cell of any one of claims 54-67, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell, and wherein said engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

70. The genetically-modified eukaryotic cell of any one of claims 46-53, wherein said inhibitory nucleic acid molecule is inhibitory against human CD52.

71. A genetically-modified eukaryotic cell comprising in its genome a nucleic acid sequence encoding an engineered antigen receptor which is expressed by said genetically-modified eukaryotic cell, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

72. The genetically-modified eukaryotic cell of claim 71, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

73. The genetically-modified eukaryotic cell of claim 71 or claim 72, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

74. The genetically-modified eukaryotic cell of any one of claims 71-73, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

75. The genetically-modified eukaryotic cell of any one of claims 71-74, wherein said control cell is not genetically-modified to reduce cell surface expression of beta-2 microglobulin.

76. The genetically-modified eukaryotic cell of any one of claims 71-75, wherein said genetically-modified eukaryotic cell comprises in its genome a nucleic acid sequence encoding an inhibitory nucleic acid molecule which is inhibitory against beta-2 microglobulin.

77. A genetically-modified eukaryotic cell comprising in its genome a nucleic acid sequence encoding an engineered antigen receptor which is expressed by said genetically-modified eukaryotic cell, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

78. The genetically-modified eukaryotic cell of claim 77, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

79. The genetically-modified eukaryotic cell of claim 77 or claim 78, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

80. The genetically-modified eukaryotic cell of any one of claims 77-79, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

81. The genetically-modified eukaryotic cell of any one of claims 77-80, wherein said control cell is not genetically-modified to reduce cell surface expression of a component of the MHC class I molecule.

82. The genetically-modified eukaryotic cell of claim 77-81, wherein said genetically-modified eukaryotic cell comprises in its genome a nucleic acid sequence encoding an inhibitory nucleic acid molecule which is inhibitory against a component of the MHC class I molecule.

83. The genetically-modified eukaryotic cell of claim 82, wherein said component of the MHC class I molecule is beta-2 microglobulin.

84. The genetically-modified eukaryotic cell of any one of claims 76, 82, or 83, wherein said inhibitory nucleic acid molecule is an RNA interference molecule.

85. The genetically-modified eukaryotic cell of claim 84, wherein said RNA interference molecule is a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a hairpin siRNA, a microRNA (miRNA), a precursor miRNA, or an miRNA-adapted shRNA.

86. The genetically-modified eukaryotic cell of claim 84 or claim 85, wherein said RNA interference molecule is an shRNA.

87. The genetically-modified eukaryotic cell of any one of claims 76 or 82-86, wherein said nucleic acid sequence encoding said engineered antigen receptor is integrated at

the same location within the genome as said nucleic acid sequence encoding said inhibitory nucleic acid molecule.

88. The genetically-modified eukaryotic cell of claim 87, wherein said genetically-modified eukaryotic cell comprises in its genome said nucleic acid molecule of any one of claims 1-31.

89. The genetically-modified eukaryotic cell of claim 87, wherein said genetically-modified eukaryotic cell comprises said nucleic acid molecule of any one of claims 32-38.

90. The genetically-modified eukaryotic cell of any one of claims 76 or 82-86, wherein said nucleic acid sequence encoding said engineered antigen receptor is integrated at a different location within the genome than said nucleic acid sequence encoding said inhibitory nucleic acid molecule.

91. The genetically-modified eukaryotic cell of any one of claims 71-90, wherein said genetically-modified eukaryotic cell is less susceptible to endogenous NK cell killing when compared to a control cell, wherein said genetically-modified eukaryotic cell has extended persistence in a subject when compared to a control cell, wherein said genetically-modified eukaryotic cell exhibits enhanced expansion in a subject when compared to a control cell, and/or wherein allogenicity of said genetically-modified eukaryotic cell is reduced when compared to a control cell.

92. The genetically-modified eukaryotic cell of any one of claims 71-91, wherein said engineered antigen receptor is a chimeric antigen receptor (CAR) or an exogenous T cell receptor (TCR).

93. The genetically-modified eukaryotic cell of any one of claims 71-92, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell.

94. The genetically-modified eukaryotic cell of claim 93, wherein said genetically-modified human T cell expresses a chimeric antigen receptor or an exogenous T cell receptor.

95. A method for producing a genetically-modified eukaryotic cell, said method comprising introducing into a cell said nucleic acid molecule of any one of claims 1-38 and:

(a) a nucleic acid encoding an engineered nuclease having specificity for said nuclease recognition sequence, wherein said engineered nuclease is expressed in said cell; or

(b) an engineered nuclease protein having specificity for said nuclease recognition sequence;

wherein said engineered nuclease recognizes and cleaves said nuclease recognition sequence in the genome of said cell to generate a cleavage site,

and wherein said nucleic acid molecule of any one of claims 1-38 is inserted into the genome of said cell at said cleavage site.

96. The method of claim 95, wherein said engineered nuclease is an engineered meganuclease, a TALEN, a zinc finger nuclease (ZFN), a CRISPR/Cas, a compact TALEN, or a megaTAL.

97. The method of claim 96, wherein said engineered nuclease is an engineered meganuclease.

98. The method of any one of claims 95-97, wherein said nuclease recognition sequence is in a human T cell receptor alpha constant region gene.

99. The method of any one of claims 95-98, wherein said nuclease recognition sequence comprises SEQ ID NO: 1.

100. The method of any one of claim 98 or claim 99, wherein cell surface expression of an endogenous T cell receptor is reduced compared to a control cell.

101. The method of any one of claims 95-100, wherein said nucleic acid encoding said engineered nuclease is an mRNA.

102. The method of claim 101, wherein said mRNA is a polycistronic mRNA encoding said engineered nuclease and at least one additional polypeptide or nucleic acid molecule.

103. The method of any one of claims 95-102, wherein said nucleic acid molecule of any one of claims 1-38 is introduced into said cell using a viral vector.

104. The method of claim 103, wherein said viral vector is an adenoviral vector, a lentiviral vector, a retroviral vector, or an AAV vector.

105. The method of claim 104, wherein said viral vector is a recombinant AAV vector.

106. The method of claim 104, wherein said viral vector is said viral vector of any one of claims 43-45.

107. The method of any one of claims 95-102, wherein said nucleic acid molecule of any one of claims 1-38 is introduced into said cell using a recombinant DNA construct.

108. The method of claim 107, wherein said recombinant DNA construct is said recombinant DNA construct of any one of claims 39-42.

109. The method of any one of claims 95-108, wherein said nucleic acid molecule of any one of claims 1-38 is inserted into the genome of said cell at said cleavage site by homologous recombination.

110. The method of any one of claims 95-109, wherein said inhibitory nucleic acid molecule is inhibitory against human beta-2 microglobulin.

111. The method of claim 110, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

112. The method of claim 110 or claim 111, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

113. The method of any one of claims 110-112, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

114. The method of any one of claims 110-113, wherein cell surface expression of beta-2 microglobulin on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

115. The method of any one of claims 111-114, wherein said control cell has not been genetically-modified to reduce cell surface beta-2 microglobulin expression.

116. The method of any one of claims 95-109, wherein said inhibitory nucleic acid molecule is inhibitory against a component of the MHC class I molecule.

117. The method of claim 116, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 10% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

118. The method of claim 116 or claim 117, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 50% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

119. The method of any one of claims 116-118, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 75% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

120. The method of any one of claims 116-119, wherein cell surface expression of MHC class I molecules on said genetically-modified eukaryotic cell is reduced by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

121. The method of any one of claims 117-120, wherein said control cell has not been genetically-modified to reduce cell surface expression of MHC class I molecules.

122. The method of any one of claims 95-121, wherein said engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

123. The method of any one of claims 95-122, wherein said genetically-modified eukaryotic cell is a human T cell.

124. The method of any one of claims 110-123, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell, and wherein said engineered antigen receptor is a chimeric antigen receptor or an exogenous T cell receptor.

125. The method of any one of claims 95-109, wherein said inhibitory nucleic acid molecule is inhibitory against human CD52.

126. A genetically-modified eukaryotic cell made by the method of any one of claims 95-125.

127. A genetically-modified eukaryotic cell made by the method of claim 124.

128. A pharmaceutical composition comprising a pharmaceutically-acceptable carrier and a therapeutically effective amount of said genetically-modified eukaryotic cell of any one of claims 46-94, 126, or 127.

129. A pharmaceutical composition comprising a pharmaceutically-acceptable carrier and a therapeutically effect amount of said genetically-modified eukaryotic cell of any one of claims 69, 94, or 127, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell, and wherein said engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and wherein cell surface expression of beta-2 microglobulin on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

130. A pharmaceutical composition comprising a pharmaceutically-acceptable carrier and a therapeutically effective amount of said genetically-modified eukaryotic cell of any one of claims 69, 94, or 127, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell, and wherein said engineered antigen receptor is a chimeric antigen receptor or exogenous T cell receptor, and wherein cell surface expression of MHC class I molecules on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface expression of MHC class I molecules on a control cell.

131. A pharmaceutical composition comprising a pharmaceutically-acceptable carrier and a therapeutically effective amount of said genetically-modified eukaryotic cell of claim 70 or claim 126, wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell, and wherein said engineered antigen receptor is a chimeric antigen receptor, and wherein cell surface expression of CD52 on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface CD52 expression on a control cell.

132. The pharmaceutical composition of any one of claims 128-131, wherein said pharmaceutical composition is for immunotherapy in the treatment of cancer in a subject in need thereof.

133. A method of using immunotherapy to treat a disease in a subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of said genetically-modified eukaryotic cell of any one of claims 69, 94, or 127; wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor or an exogenous T cell receptor; and wherein cell surface expression of beta-2 microglobulin on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface beta-2 microglobulin expression on a control cell.

134. The method of claim 133, wherein endogenous NK cell killing of said genetically-modified human T cell is reduced in said subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression.

135. The method of claim 133 or claim 134, wherein said subject is administered said pharmaceutical composition of claim 129.

136. The method of any one of claims 133-135, wherein said genetically-modified human T cell is allogeneic to said subject.

137. The method of any one of claims 133-136, wherein persistence of said genetically-modified human T cell is extended in said subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

138. The method of any one of claims 133-137, wherein expansion of said genetically-modified human T cell is enhanced in said subject when compared to a genetically-modified human T cell having no cell surface beta-2 microglobulin expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

139. The method of any one of claims 133-138, wherein allogenicity of said genetically-modified human T cell is reduced when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of beta-2 microglobulin.

140. A method of using immunotherapy to treat a disease in a subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of said genetically-modified eukaryotic cell of any one of claims 69, 94, or 127;

wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor or an exogenous T cell receptor;

and wherein cell surface expression of MHC class I molecules on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to expression of MHC class I molecules on a control cell.

141. The method of claim 140, wherein endogenous NK cell killing of said genetically-modified human T cell is reduced in said subject when compared to a genetically-modified human T cell having no cell surface MHC class I molecule expression.

142. The method of claim 140 or claim 141, wherein said subject is administered said pharmaceutical composition of claim 130.

143. The method of any one of claims 140-142, wherein said genetically-modified human T cell is allogeneic to said subject.

144. The method of any one of claims 140-143, wherein persistence of said genetically-modified human T cell is extended in said subject when compared to a genetically-modified human T cell having no cell surface MHC class I molecule expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

145. The method of any one of claims 140-144, wherein expansion of said genetically-modified human T cell is enhanced in said subject when compared to a genetically-modified human T cell having no cell surface MHC class I molecule expression, or when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

146. The method of any one of claims 140-145, wherein allogenicity of said genetically-modified human T cell is reduced when compared to a genetically-modified human T cell having a wild-type level of cell surface expression of MHC class I molecules.

147. A method of using immunotherapy to treat a disease in a subject in need thereof, said method comprising administering to said subject a therapeutically effective amount of said genetically-modified eukaryotic cell of claim 70 or claim 126;

wherein said genetically-modified eukaryotic cell is a genetically-modified human T cell expressing a chimeric antigen receptor and an inhibitory nucleic acid against CD52;

and wherein cell surface expression of CD52 on said genetically-modified human T cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface CD52 expression on a control cell.

148. The method of claim 147, wherein said subject is administered said pharmaceutical composition of claim 131.

149. The method of any one of claims 133-148, wherein said disease is cancer.

150. The method of claim 149, wherein said cancer is selected from the group consisting of a cancer of carcinoma, lymphoma, sarcoma, blastomas, and leukemia.

151. The method of claim 149 or claim 150, wherein said cancer is selected from the group consisting of a cancer of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, melanoma, prostate cancer, colon cancer, renal cell carcinoma, ovarian cancer, rhabdomyo sarcoma, leukemia, and Hodgkin's lymphoma.

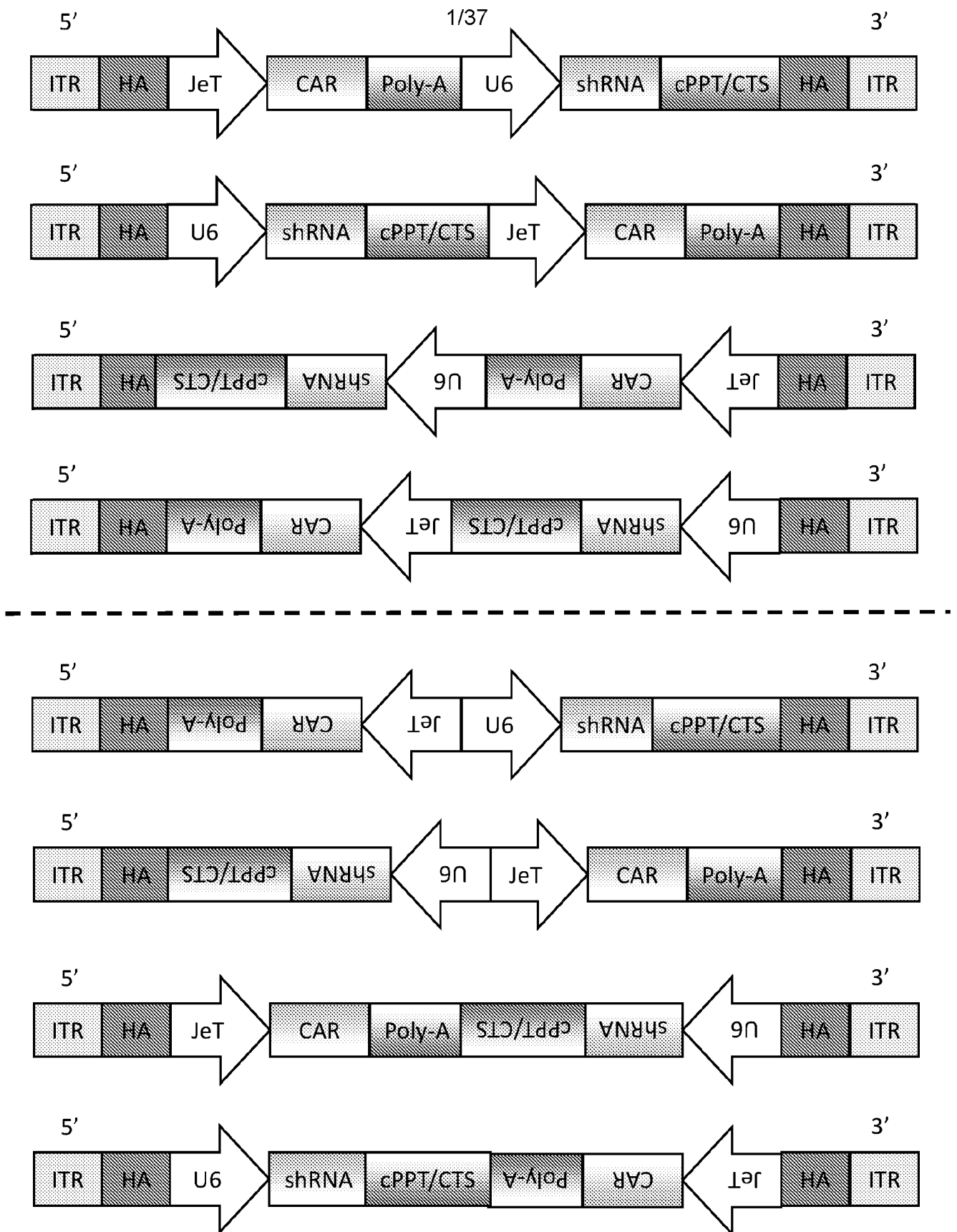
152. The method of claim 151, wherein said cancer of B-cell origin is selected from the group consisting of B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma.

153. A method for preparing an enriched population of genetically-modified eukaryotic cells comprising an engineered antigen receptor, said method comprising preparing a population of cells comprising said genetically-modified eukaryotic cell of claim 70 or claim 126, and cells expressing a wild-type level of cell surface CD52, wherein cell surface expression of CD52 on said genetically-modified cell is reduced by 10% to 95%, by 50% to 95%, by 75% to 95%, or by 90% to 95% compared to cell surface CD52 expression on a control cell, said method comprising:

(a) contacting said population of cells with beads conjugated to an anti-CD52-binding molecule, wherein cells expressing a wild-type level of cell surface CD52 are bound to said beads and said genetically-modified eukaryotic cell is not bound to said beads; and

(b) removing said beads from said population of cells to produce said enriched population of cells;

wherein said enriched population of cells is enriched for said genetically-modified eukaryotic cell.

**FIGURE 1**

2/37

FIGURE 2A

B2m+ T cells
2:1

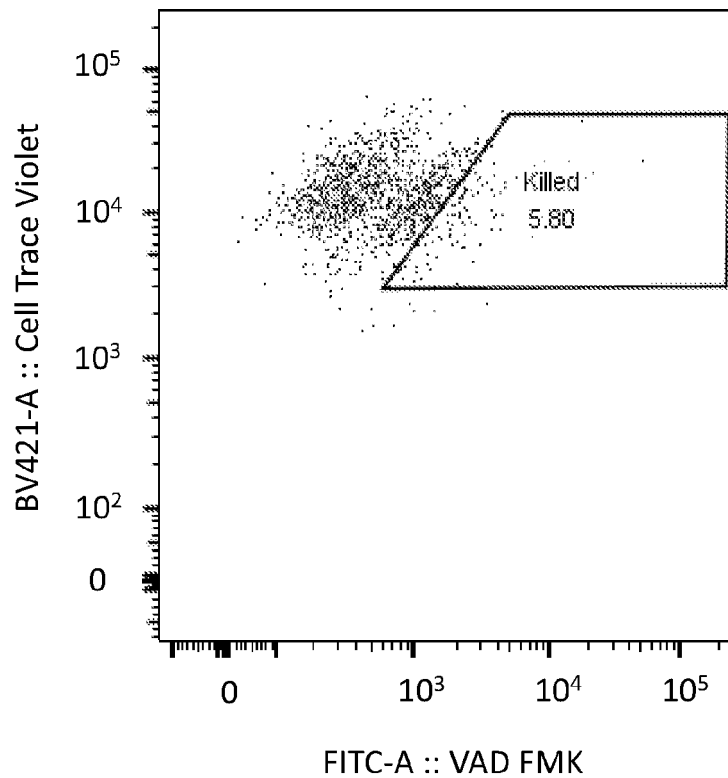
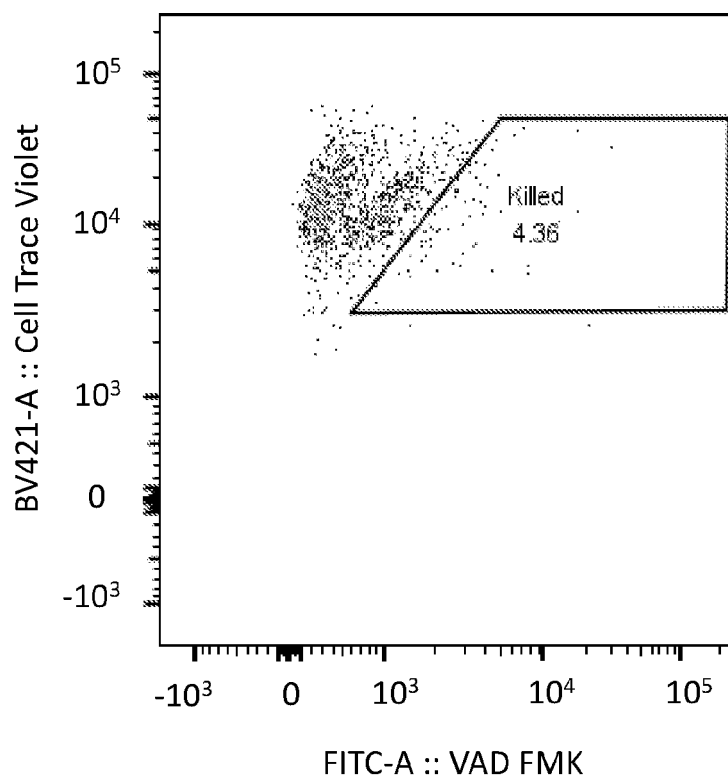


FIGURE 2B

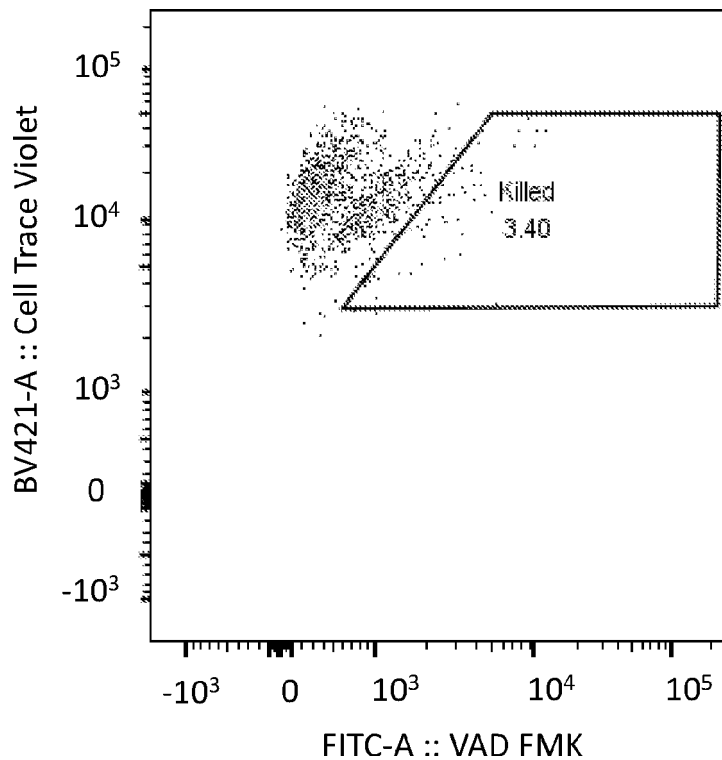
1:1



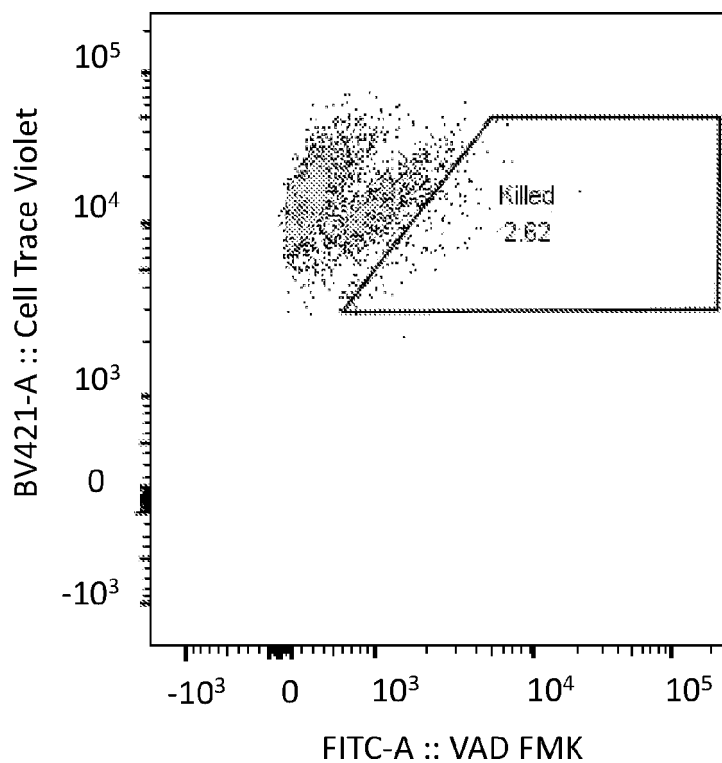
3/37

FIGURE 2CB2m+ T cells

0.5:1

**FIGURE 2D**

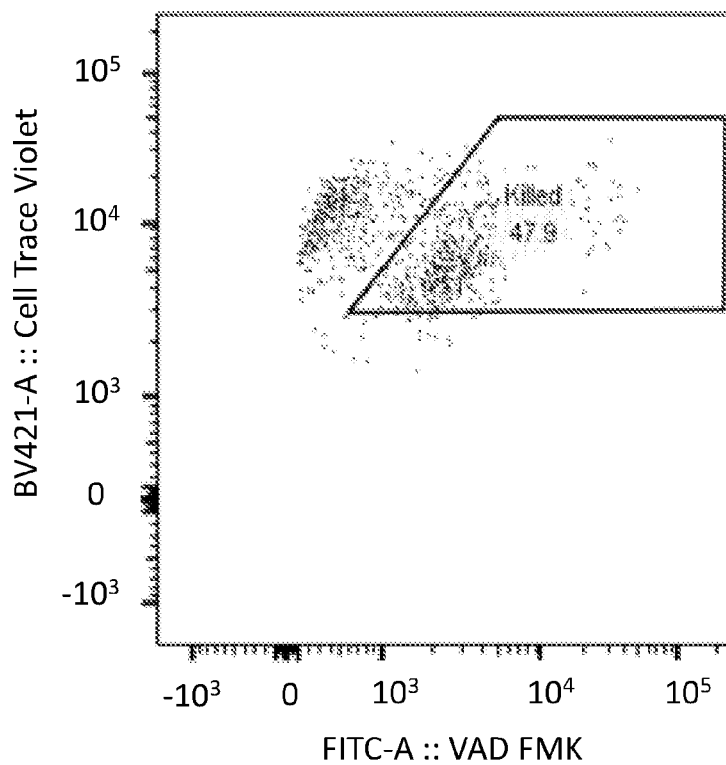
0:1



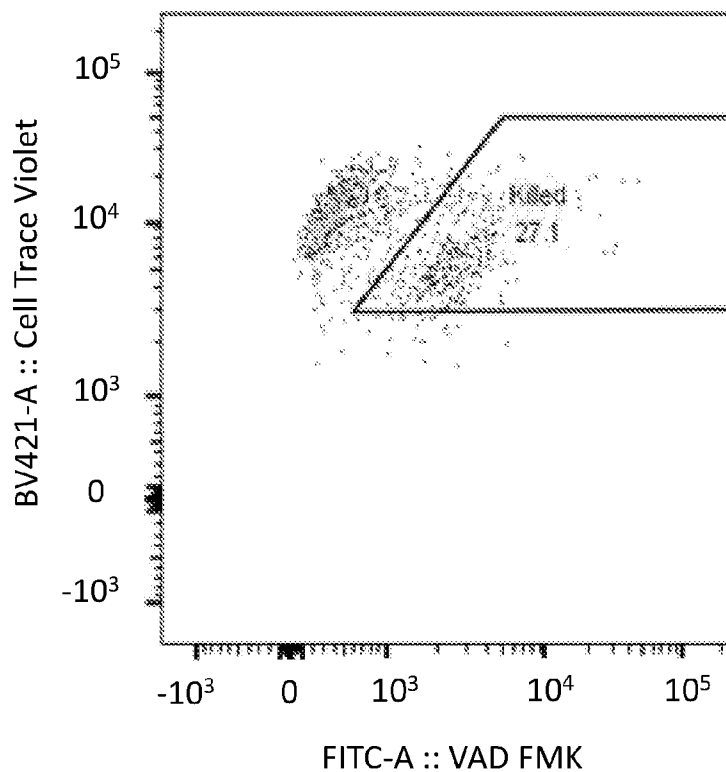
4/37

FIGURE 2EB2m- T cells

2:1

**FIGURE 2F**

1:1



5/37

FIGURE 2G B2m- T cells
0.5:1

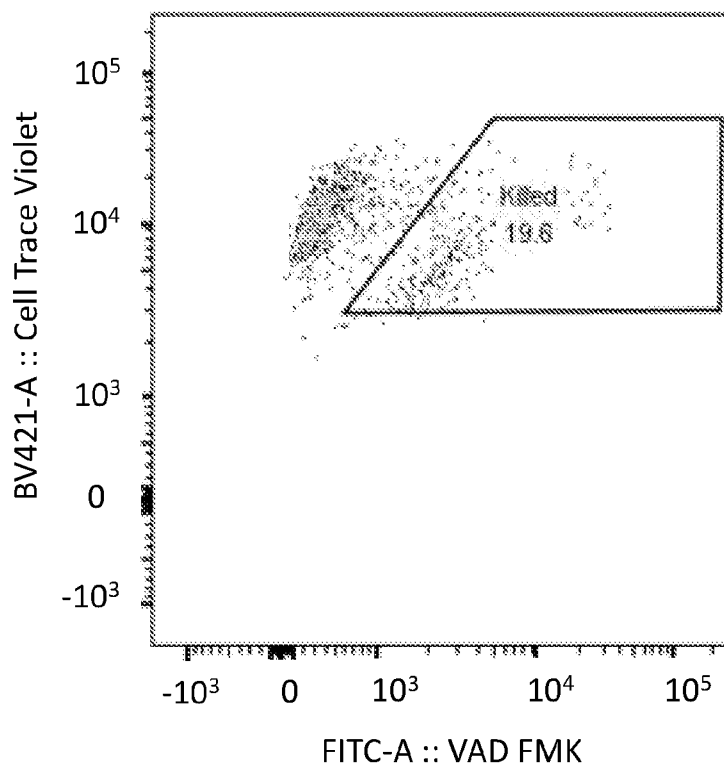
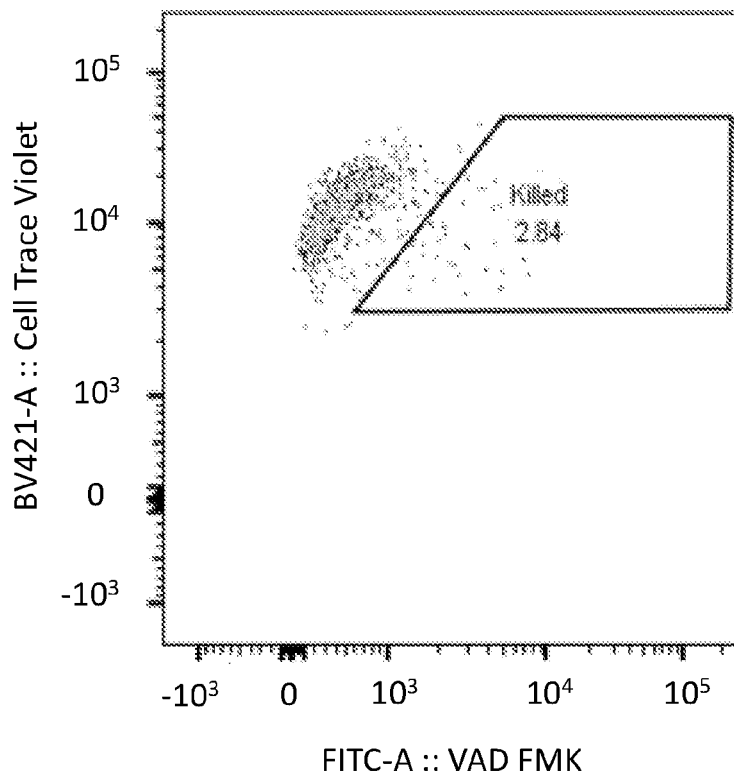


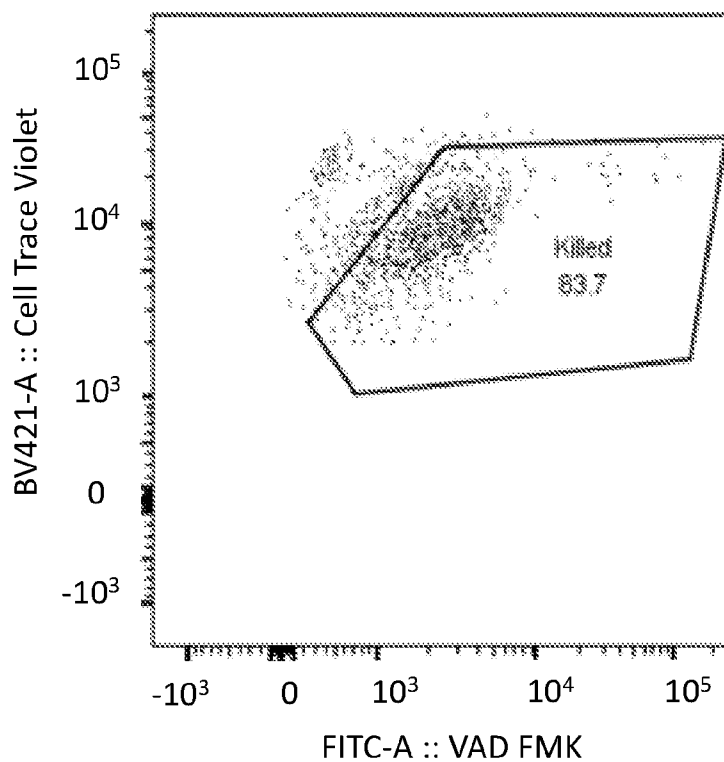
FIGURE 2H 0:1



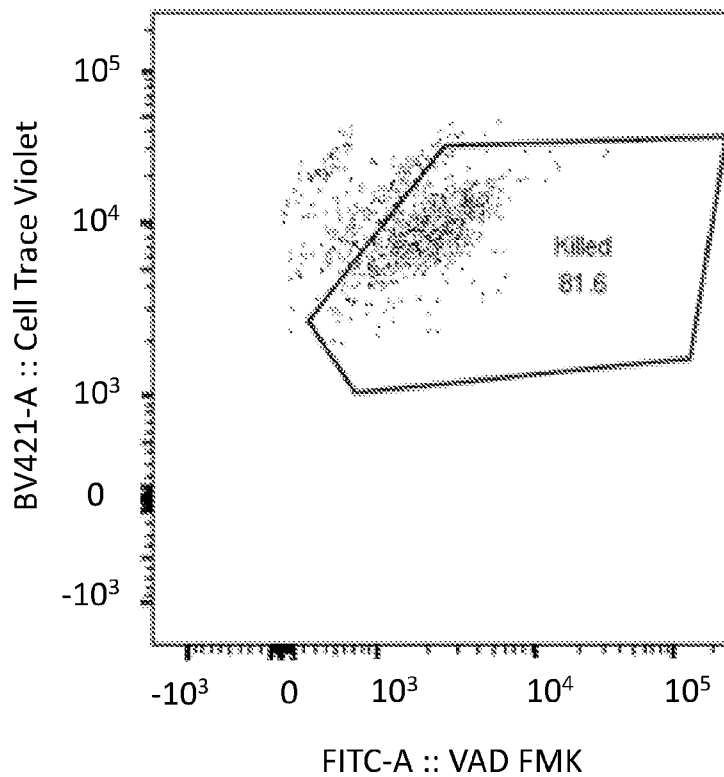
6/37

FIGURE 2I Daudi Class I Neg

2:1

**FIGURE 2J**

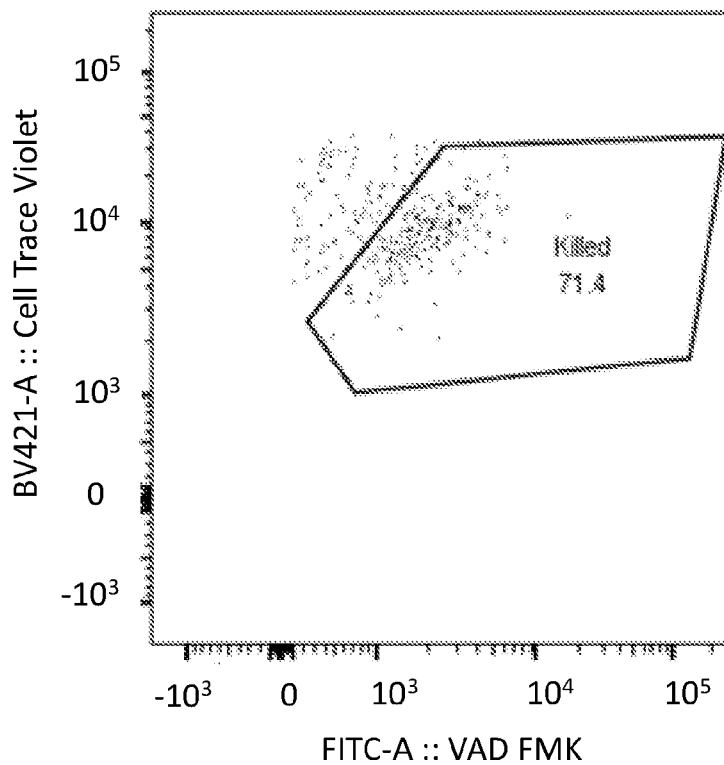
1:1



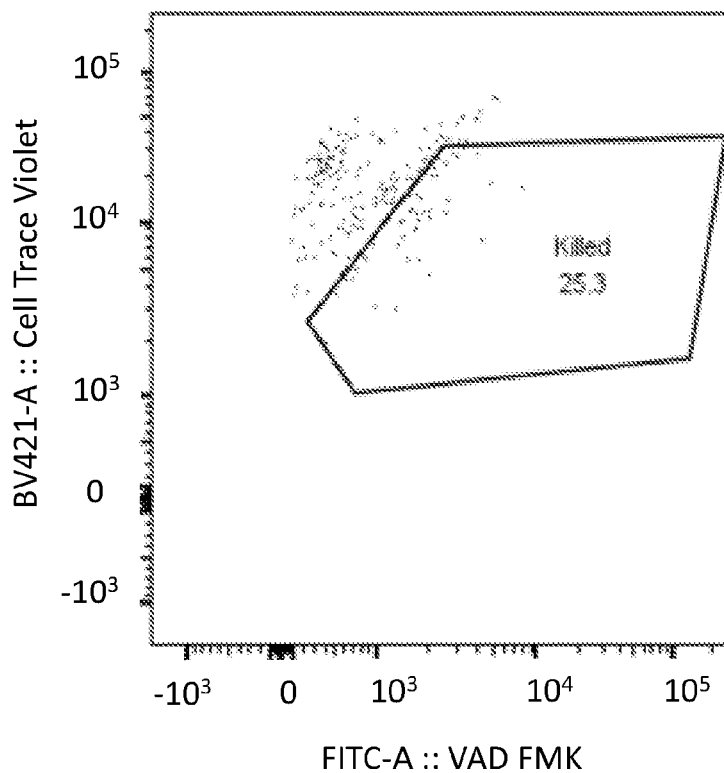
7/37

FIGURE 2K Daudi Class I Neg

0.5:1

**FIGURE 2L**

0:1



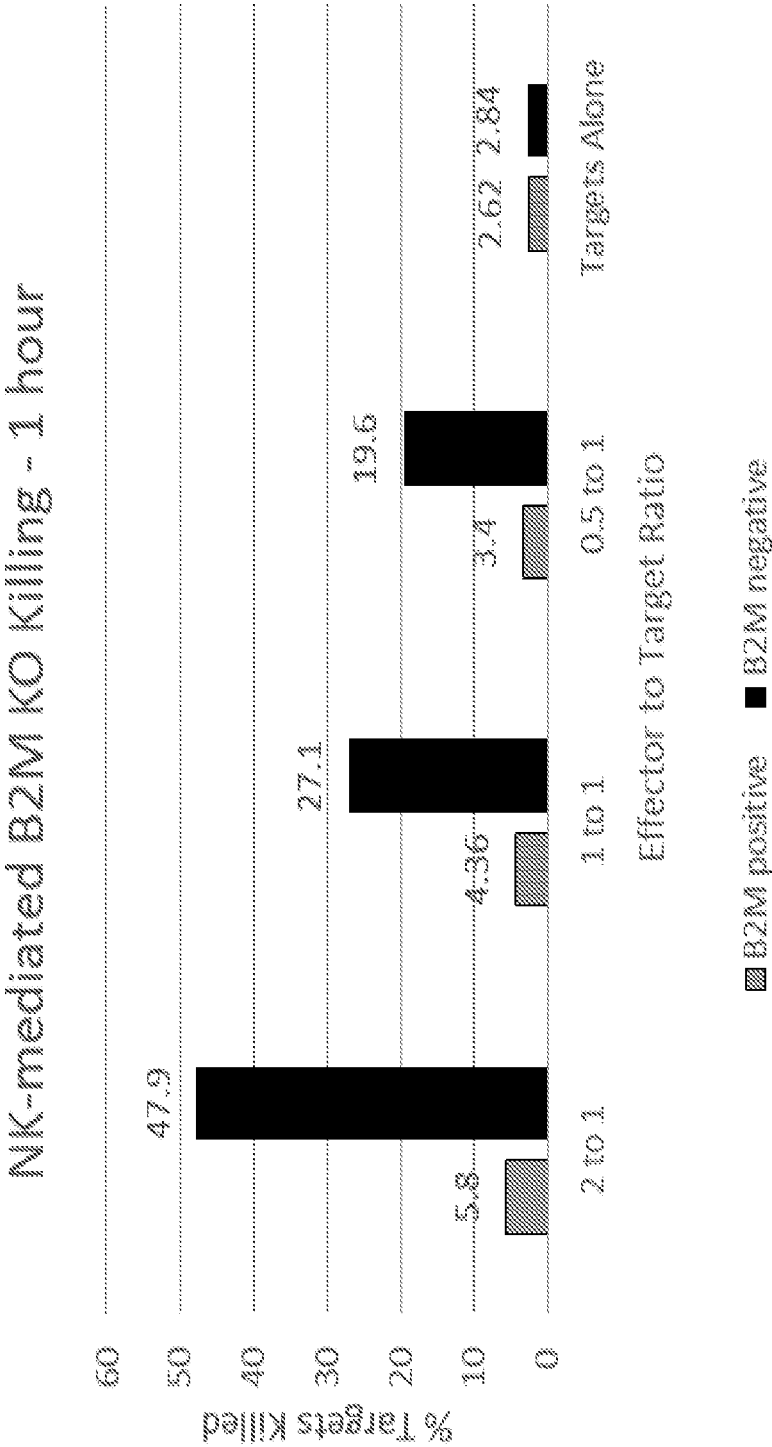
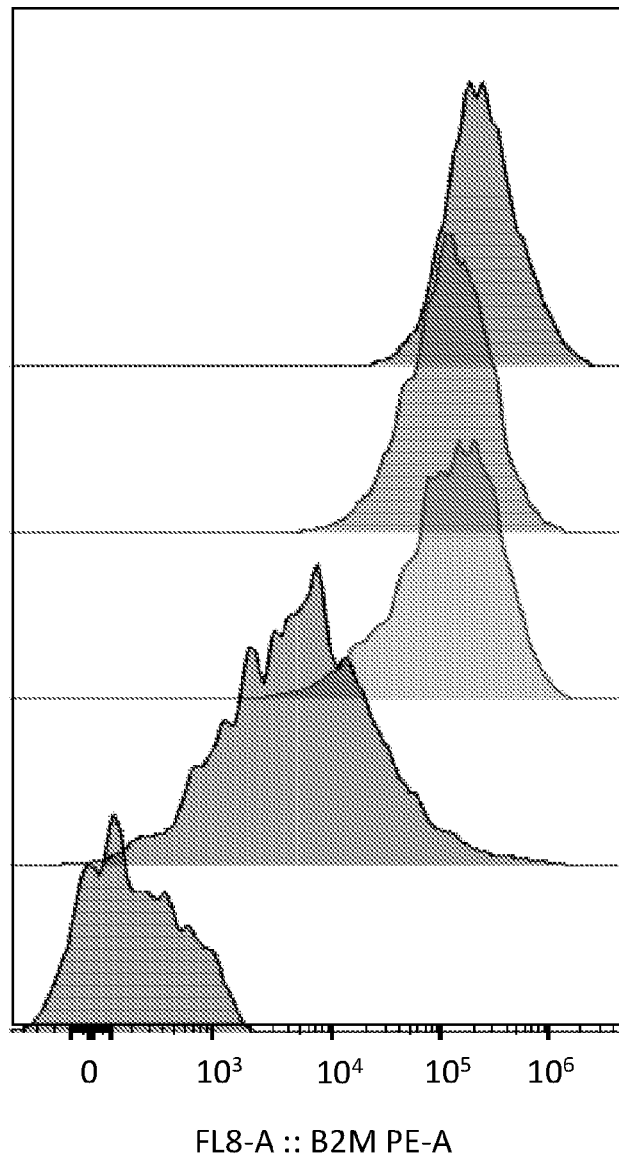


FIGURE 3

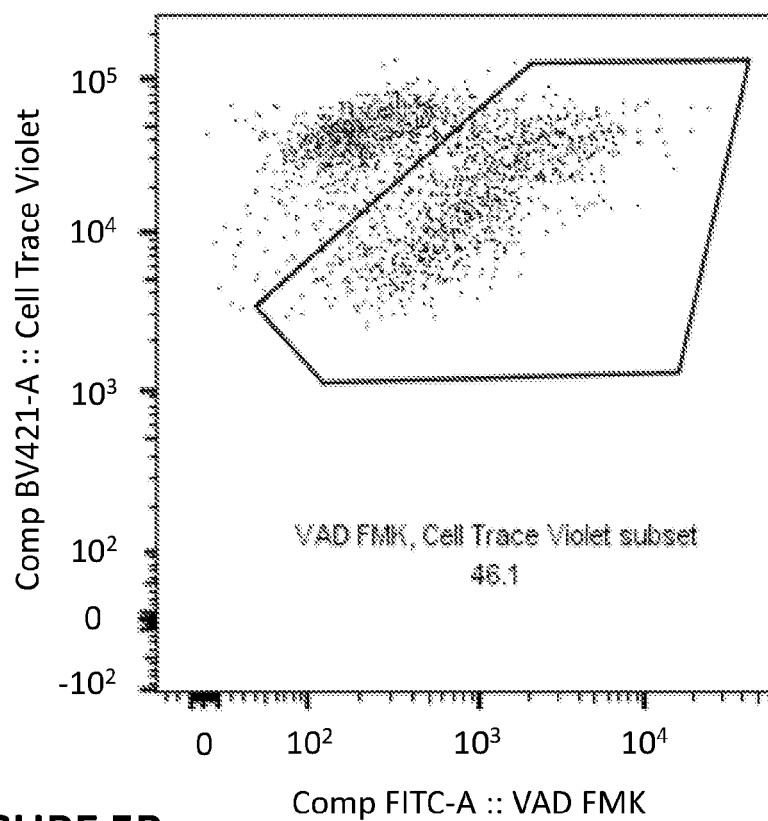
9/37



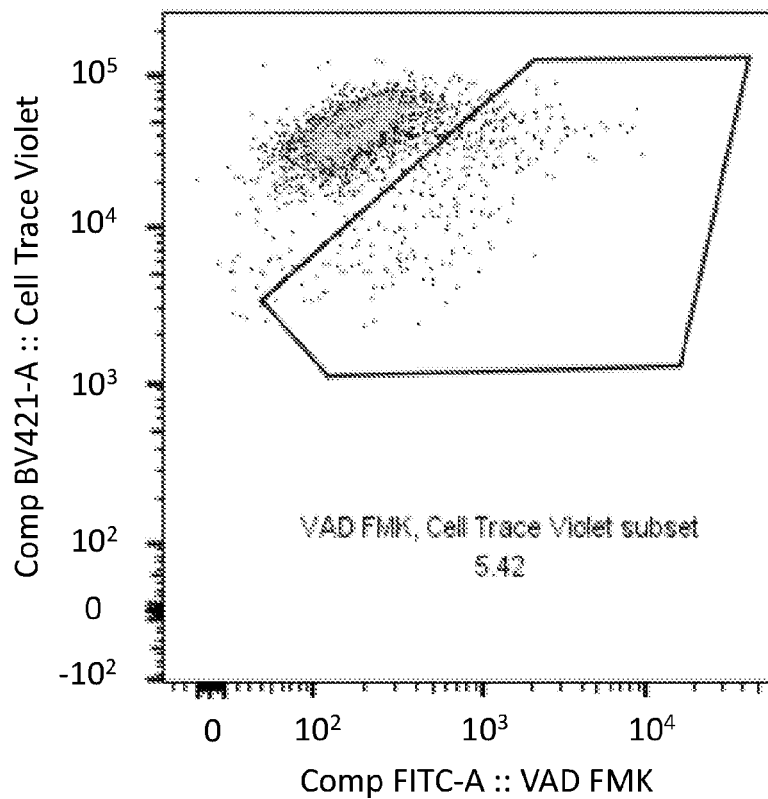
Sample Name	Median: FL8-A	Mean : FL8-A
Mock.fcs	233673	325119
256.Fcs	120555	163628
254.Fcs	131049	185393
472.Fcs	5538	23540
Unstained.fcs	186	292

FIGURE 4

10/37

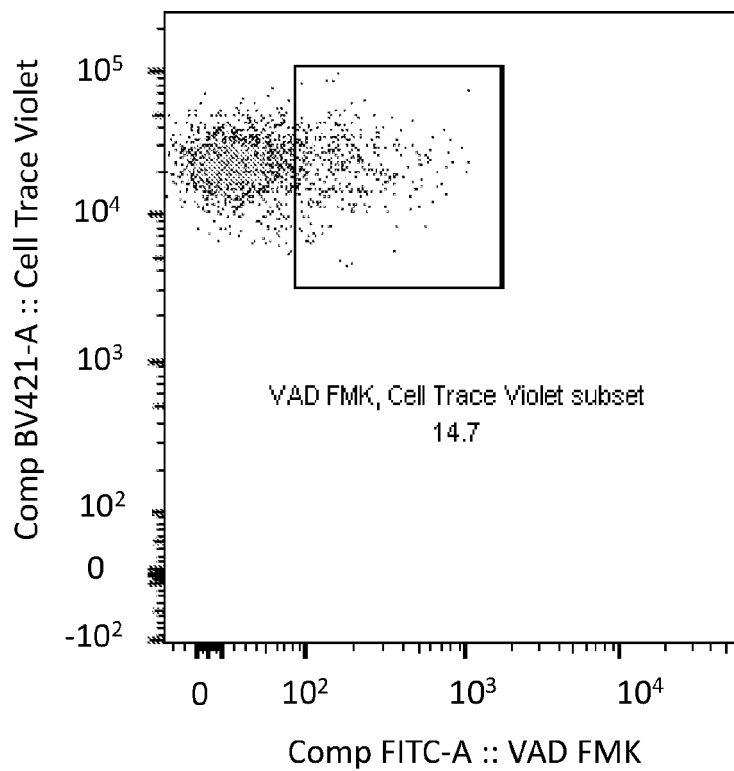
FIGURE 5AK562

NK Cell Lysis

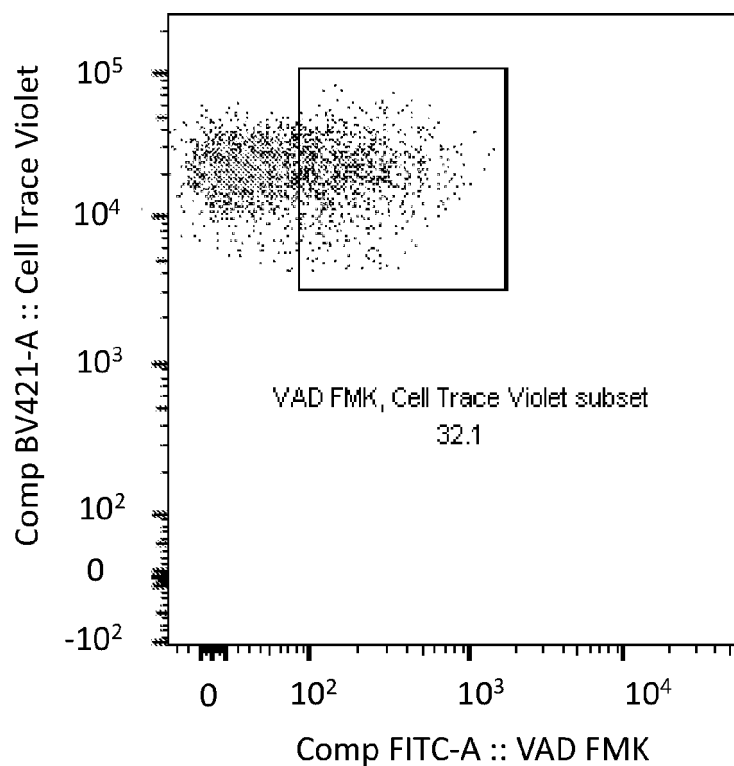
FIGURE 5B

Allo Cell Lysis

11/37

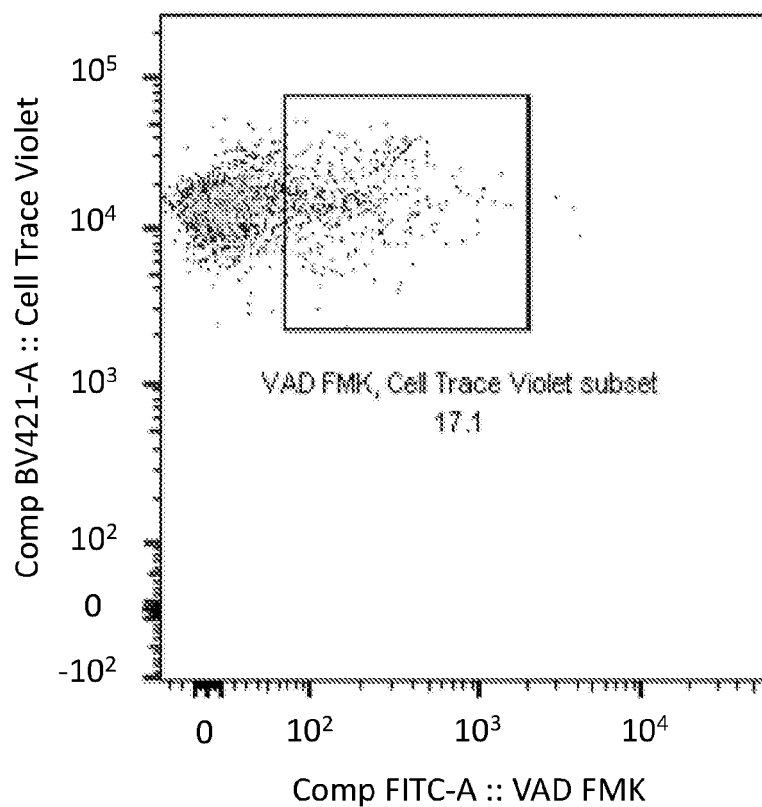
FIGURE 5CMock T cells

NK Cell Lysis

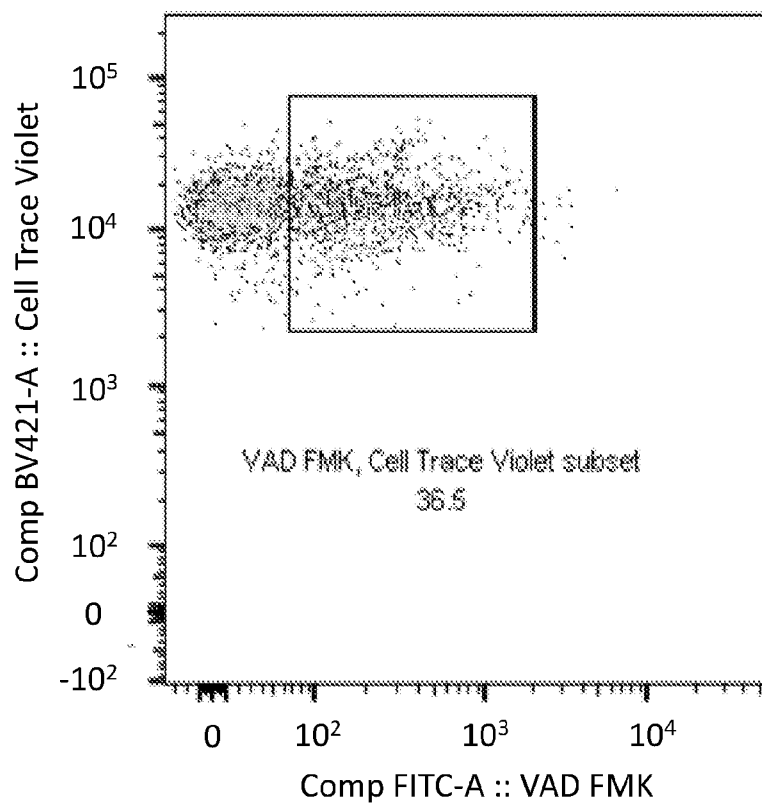
FIGURE 5D

Allo Cell Lysis

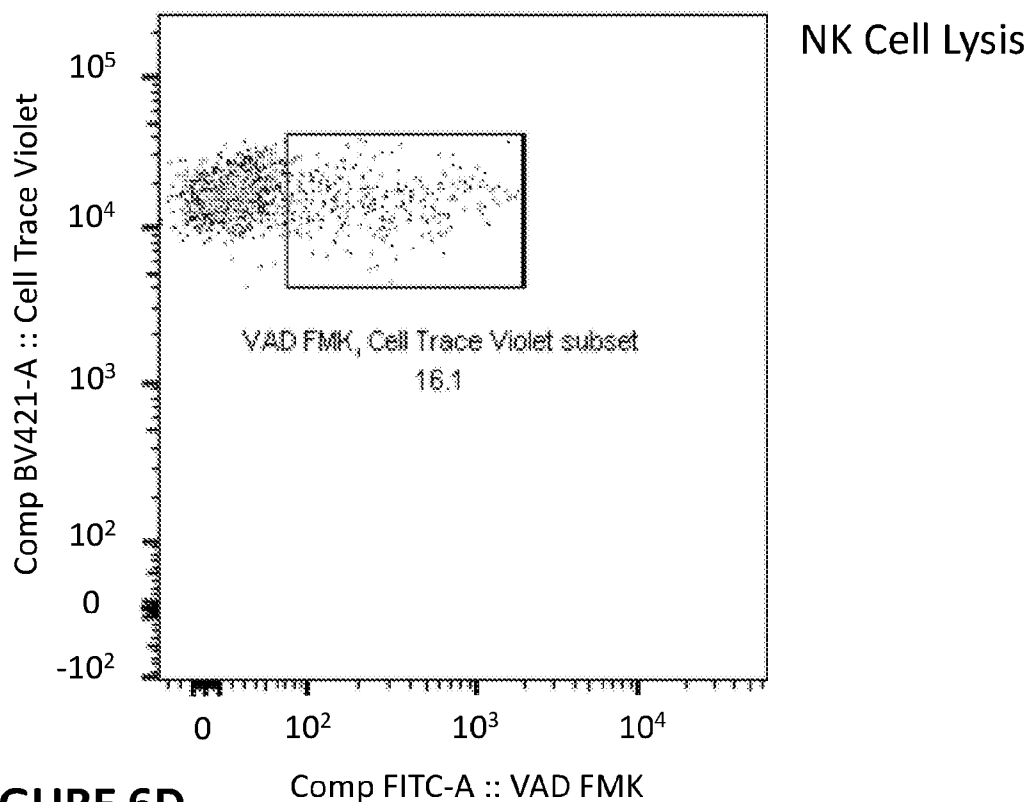
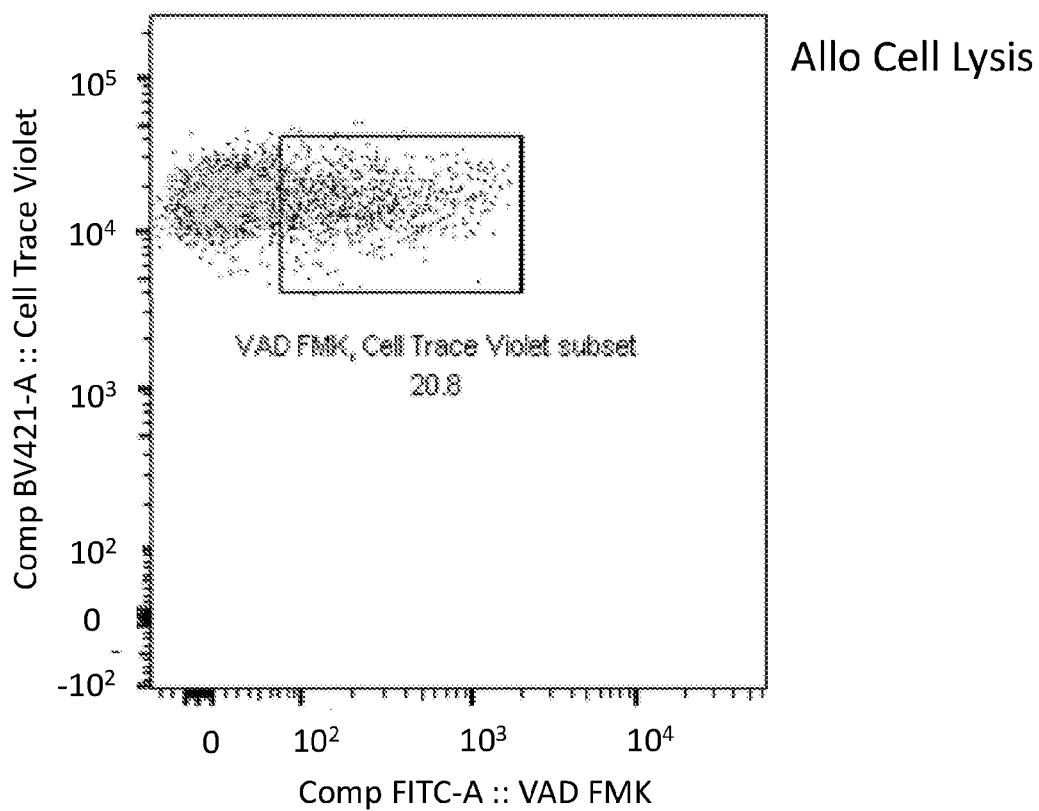
12/37

FIGURE 6AB2M shRNA 254

NK Cell Lysis

FIGURE 6B

Allo Cell Lysis

FIGURE 6CB2M shRNA 472**FIGURE 6D**

7007 (SEQ ID NO: 18)

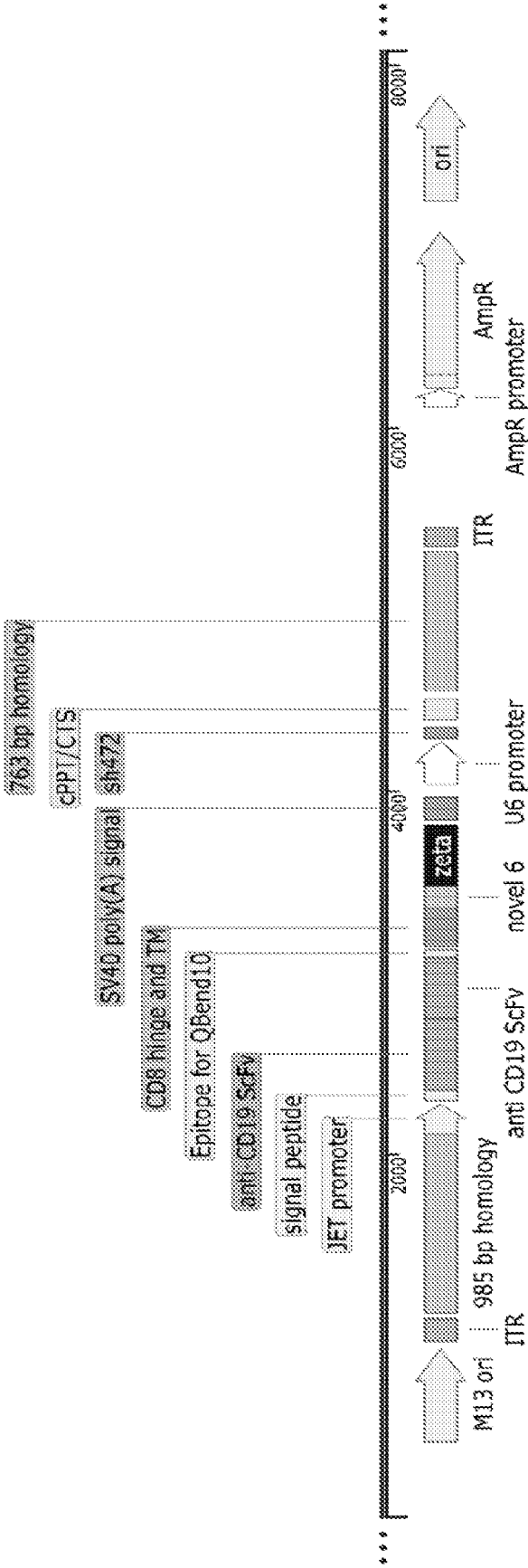


FIGURE 7A

7217 (SEQ ID NO: 19)

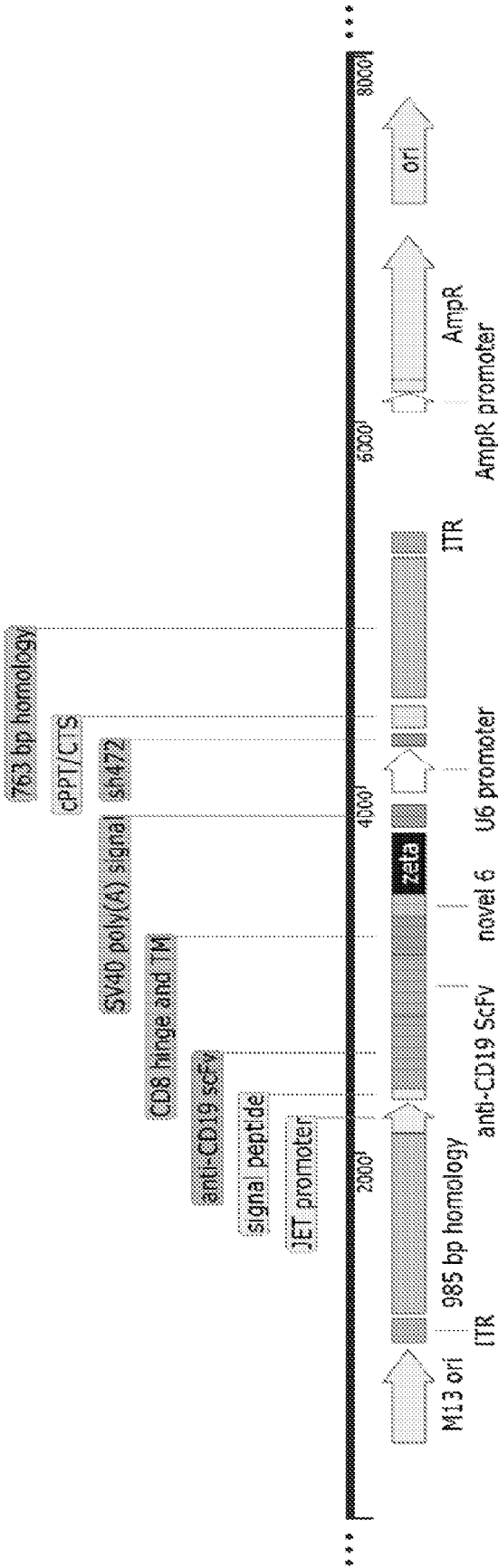


FIGURE 7B

7008 (SEQ ID NO: 20)

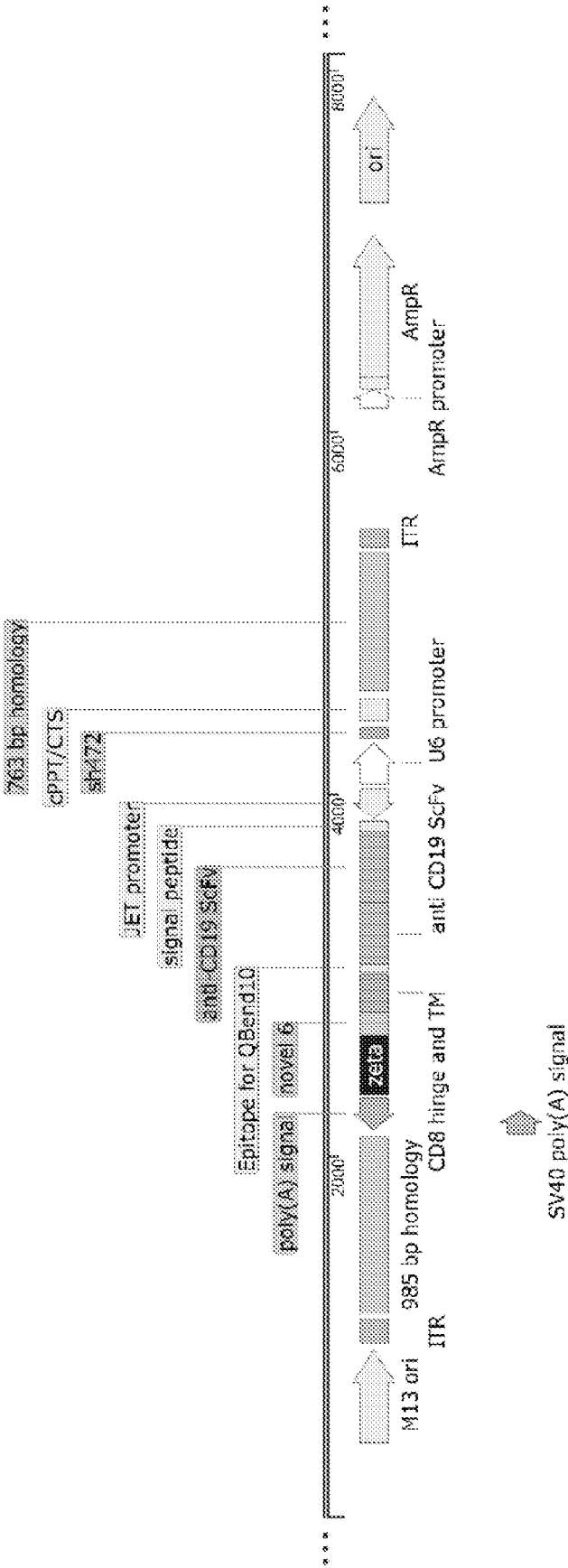


FIGURE 7C

7218 (SEQ ID NO: 21)

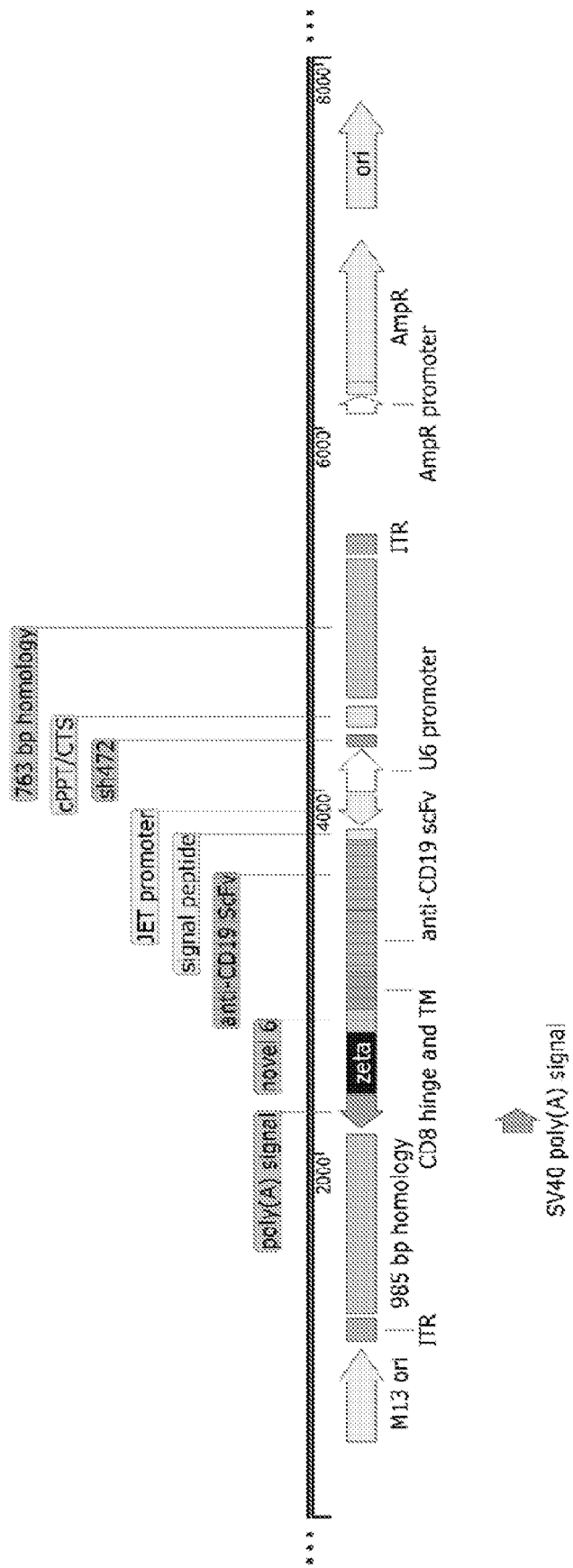


FIGURE 7D

7009 (SEQ ID NO: 22)

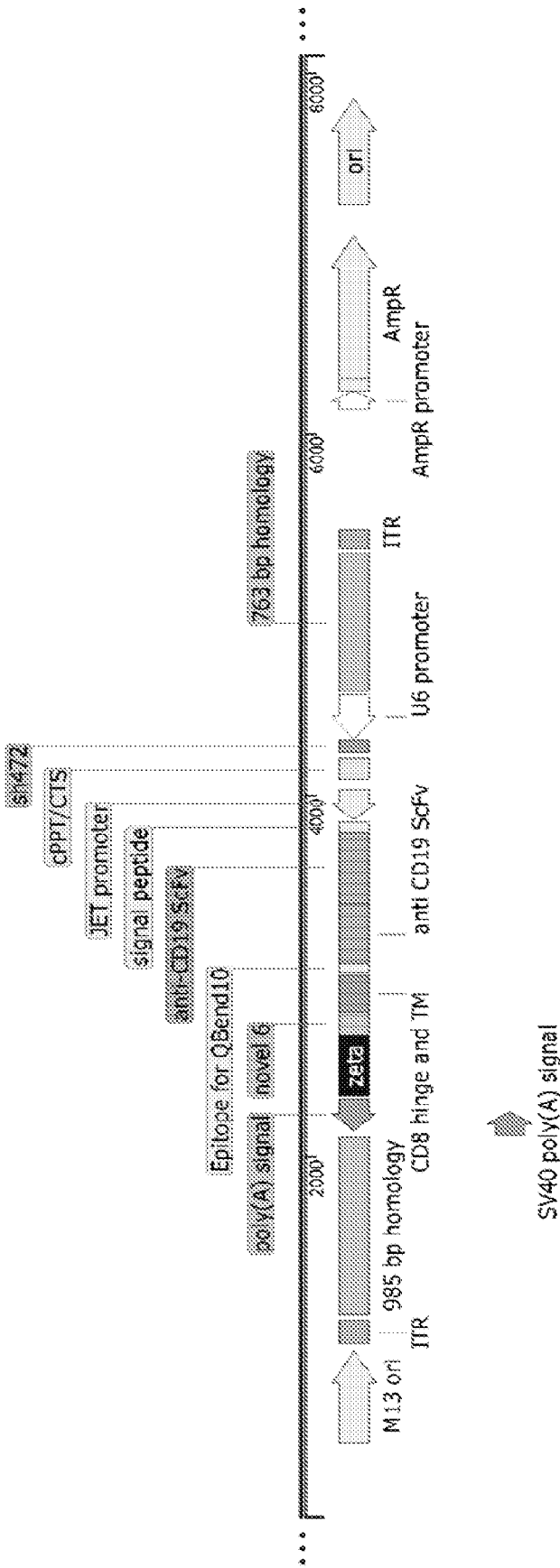


FIGURE 7E

7219 (SEQ ID NO: 23)

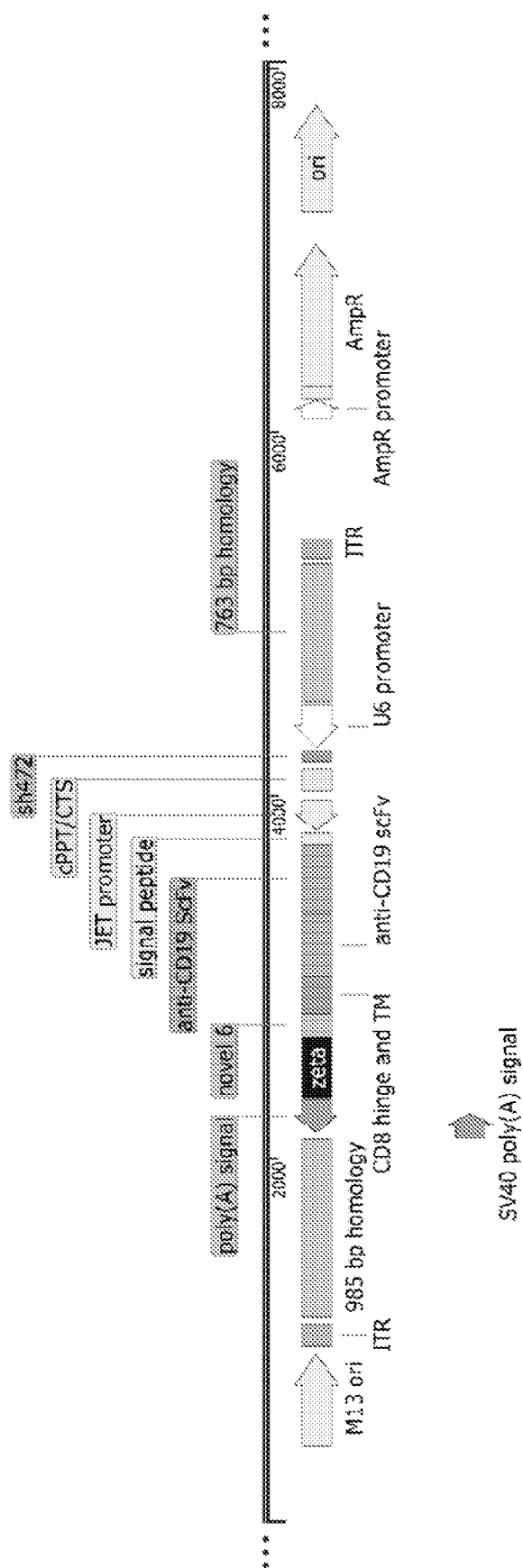
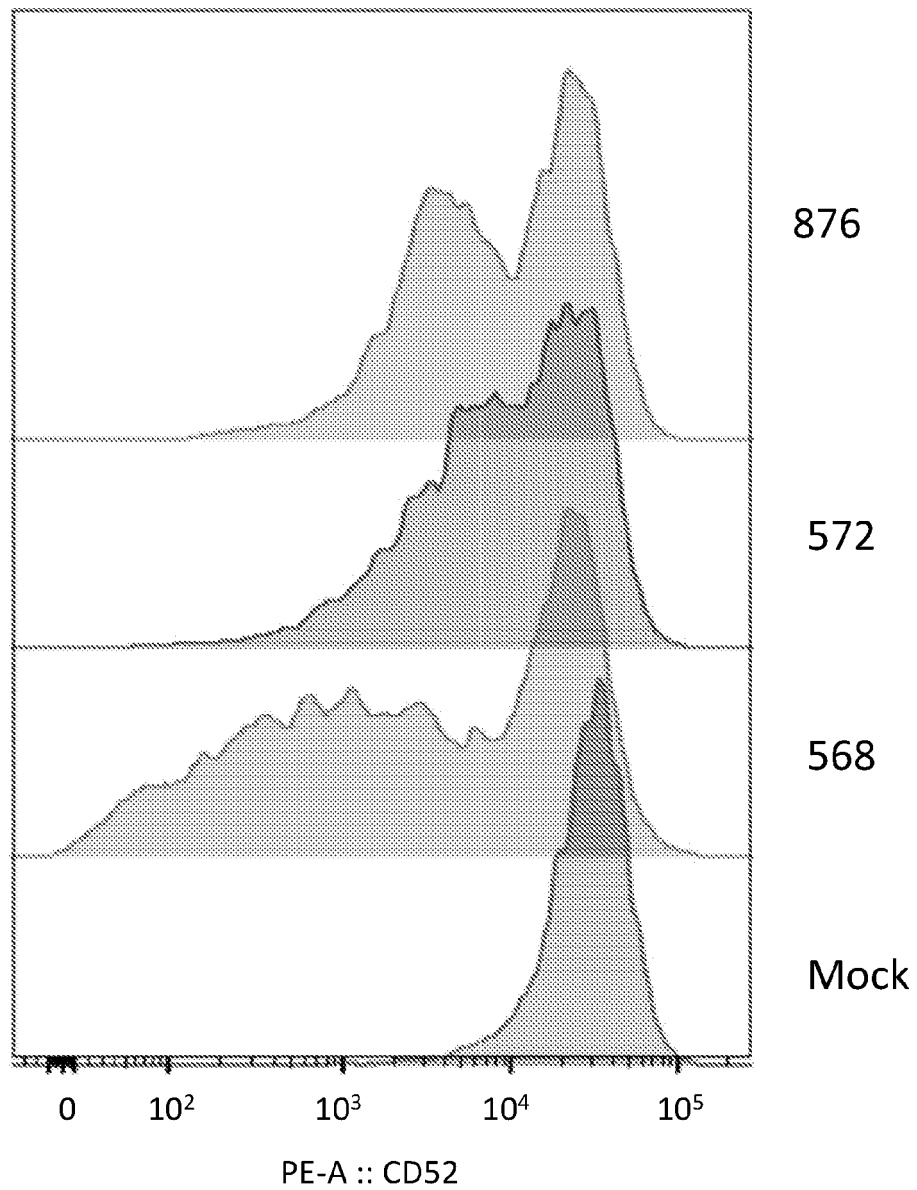


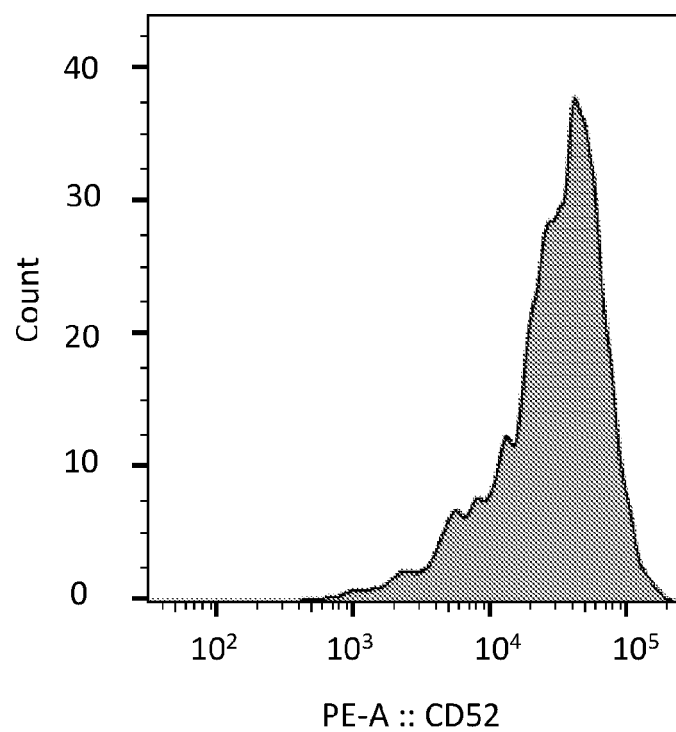
FIGURE 7F

**FIGURE 8**

21/37

FIGURE 9A

Mock

**FIGURE 9B**

Transduced w/ shRNA encoding LV

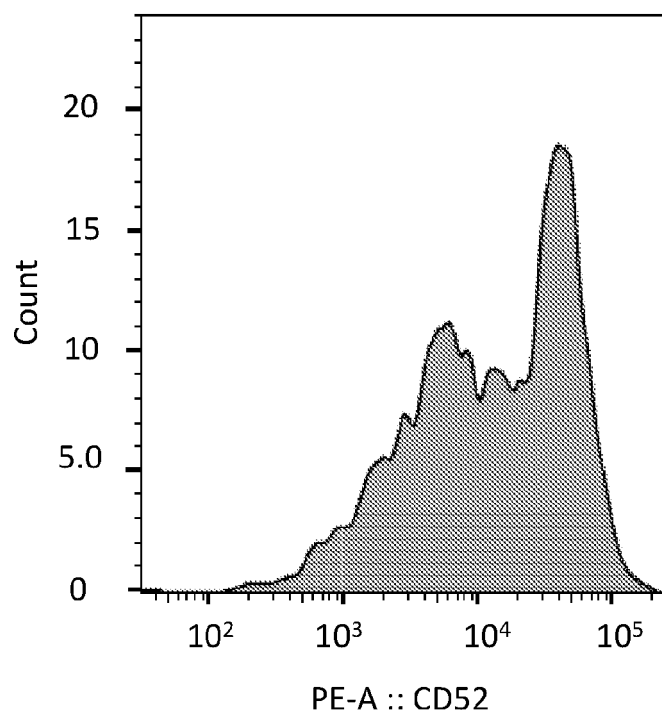
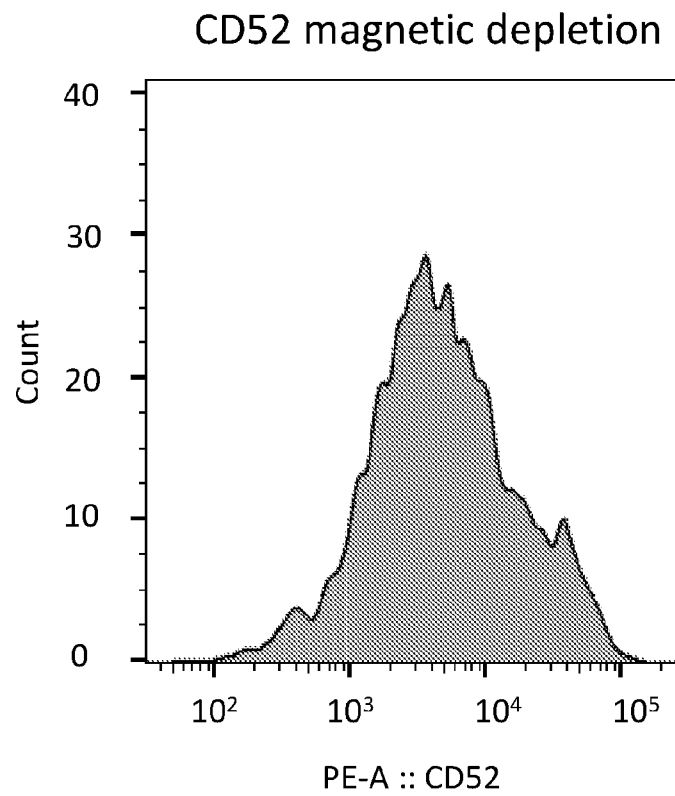


FIGURE 9C

7005 (SEQ ID NO: 10)

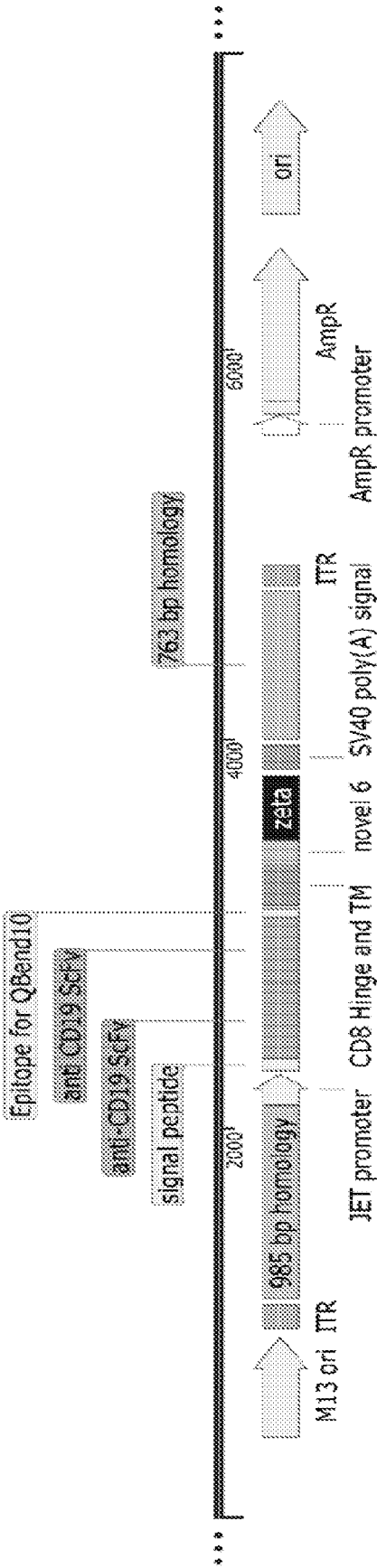


FIGURE 10A

7002 (SEQ ID NO: 11)

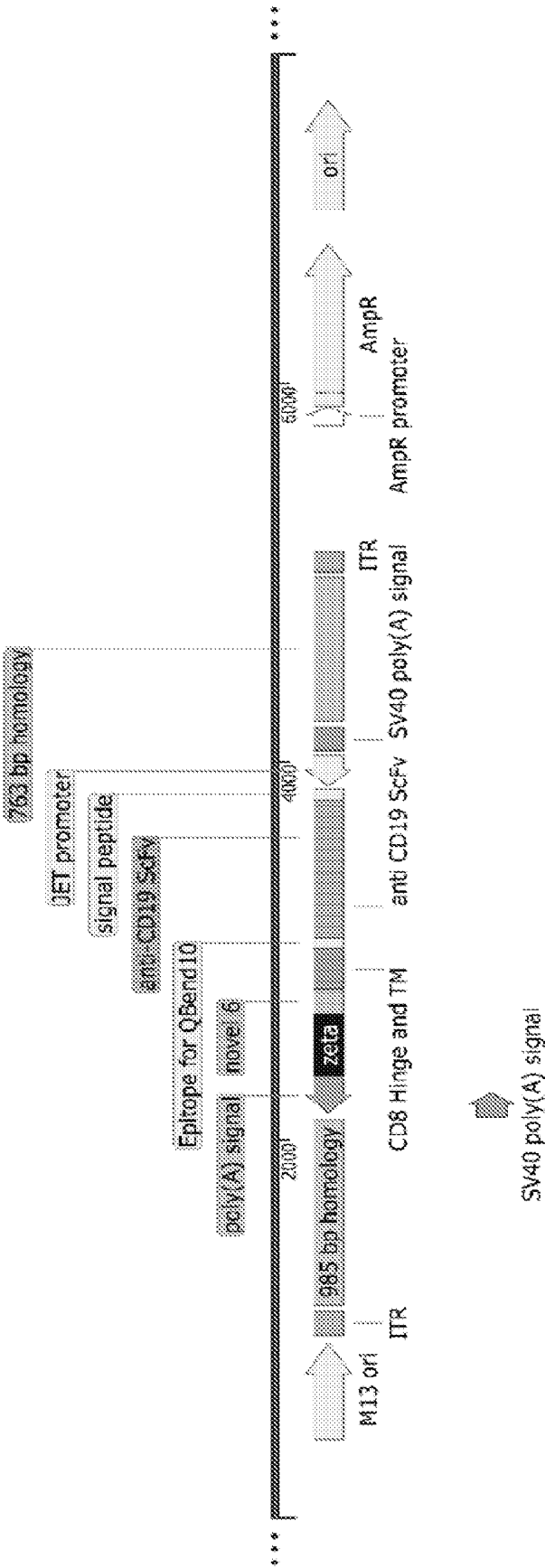


FIGURE 10B

7004 (SEQ ID NO: 12)

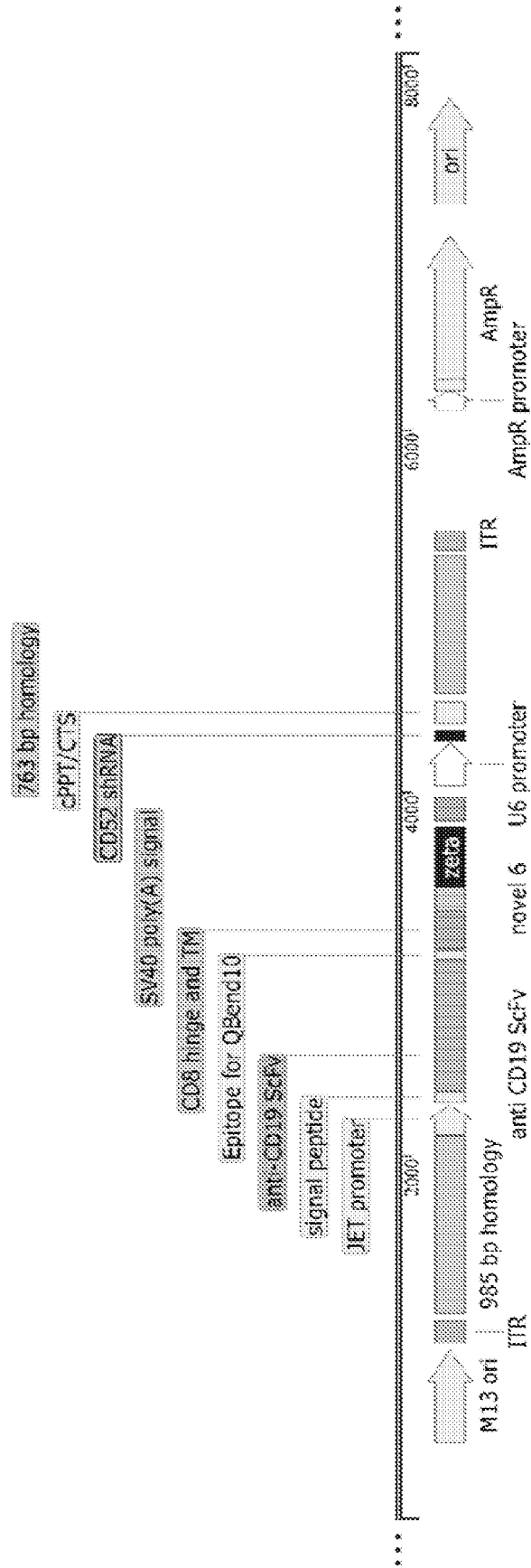


FIGURE 10C

7204 (SEQ ID NO: 13)

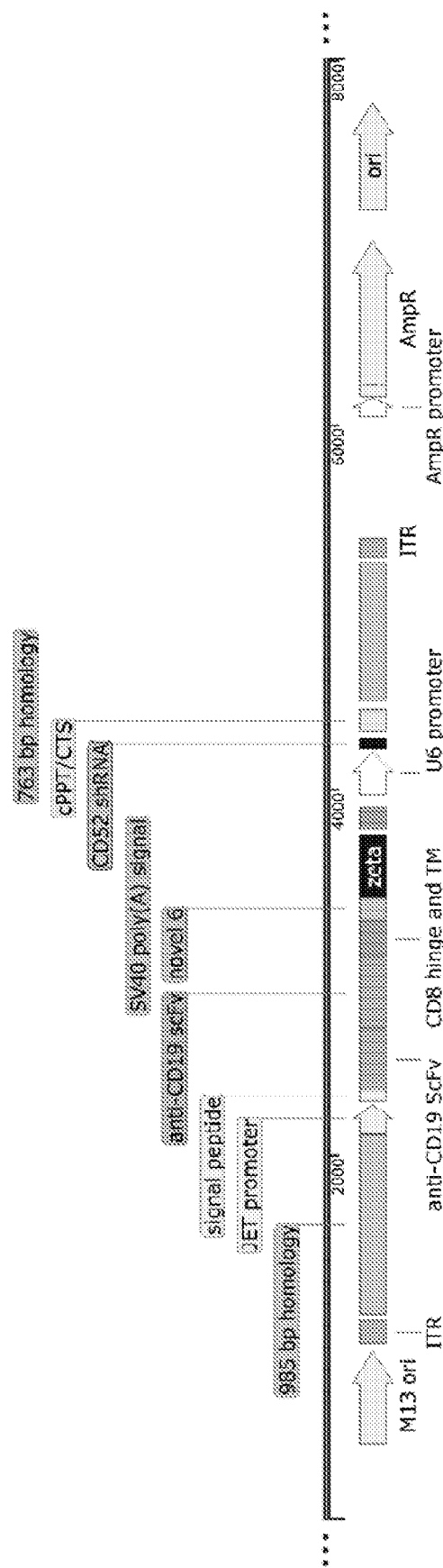


FIGURE 10D

7013 (SEQ ID NO: 14)

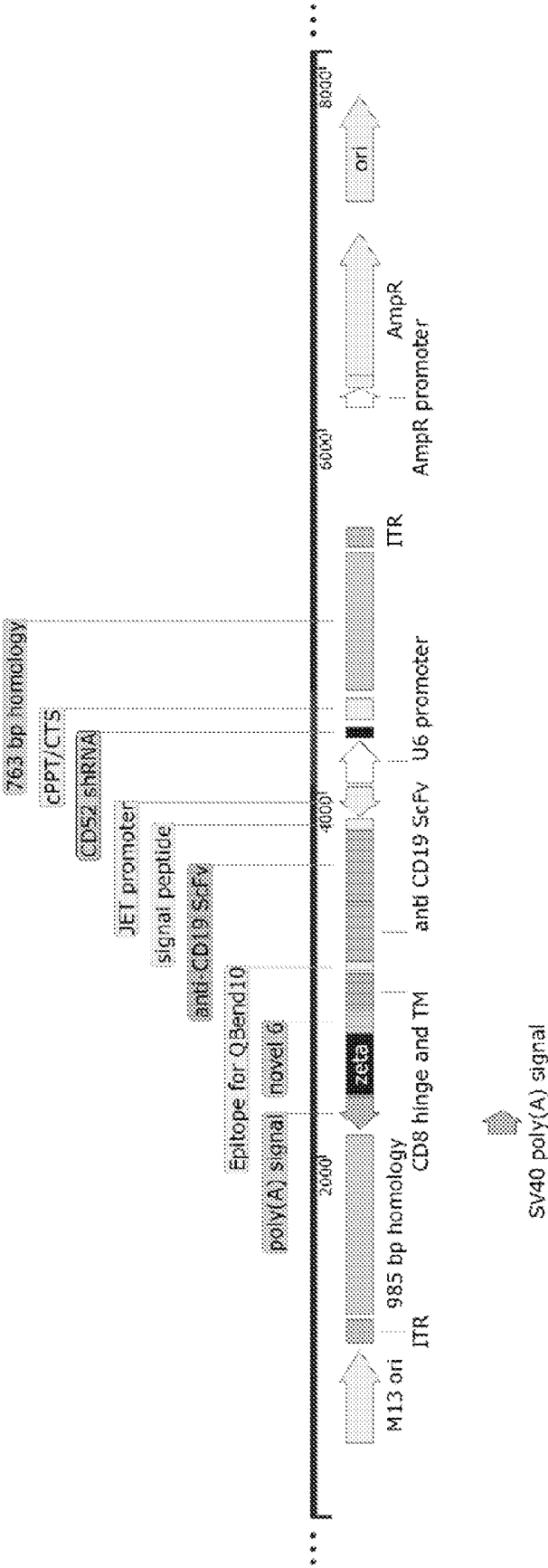


FIGURE 10E

7213 (SEQ ID NO: 15)

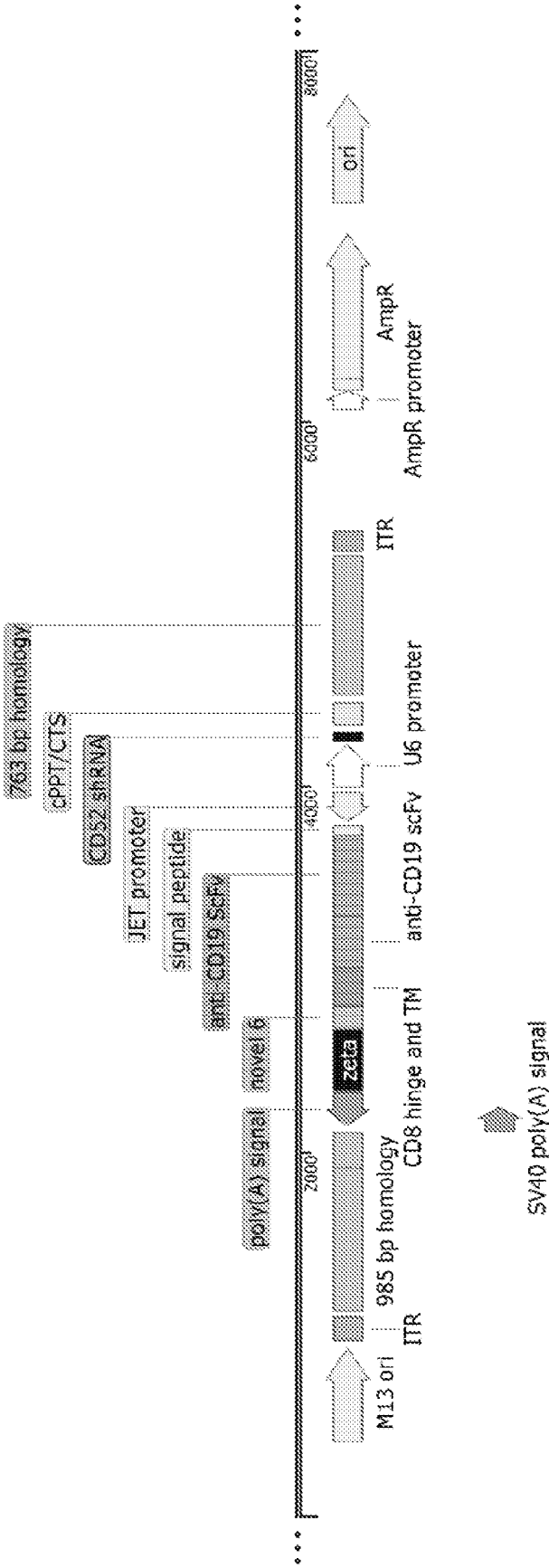


FIGURE 10F

7014 (SEQ ID NO: 16)

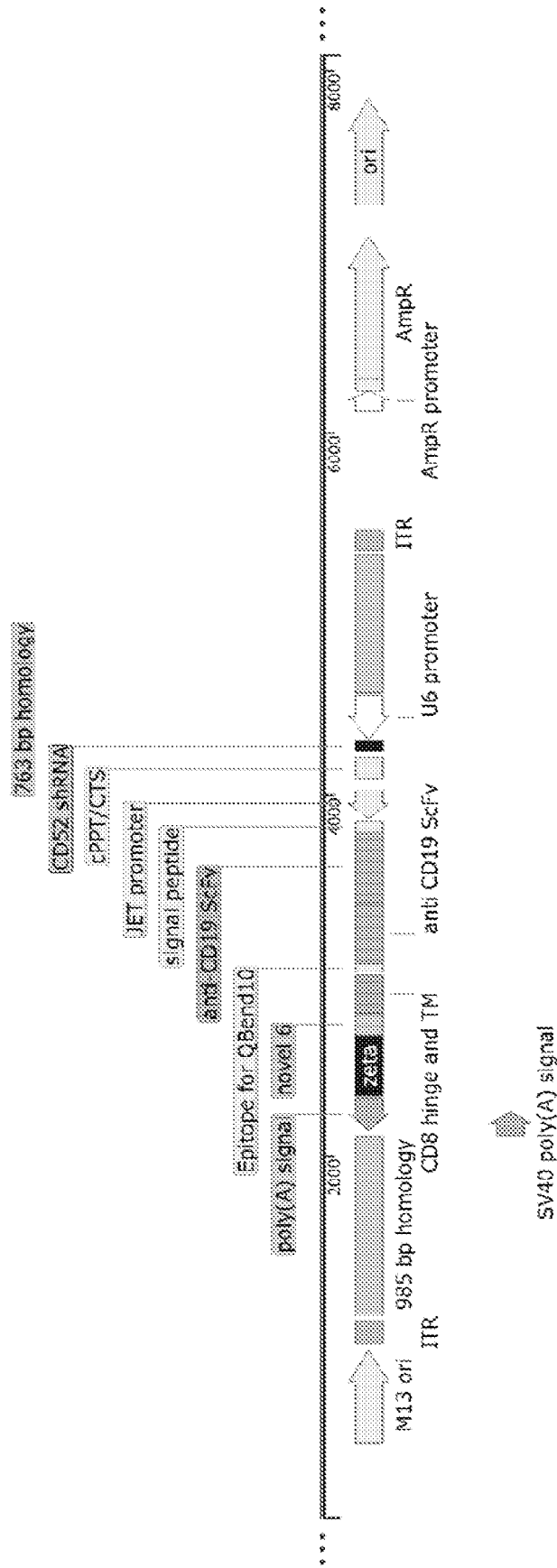


FIGURE 10G

7214 (SEQ ID NO: 17)

6. Chimeric anti-CD19 scFv

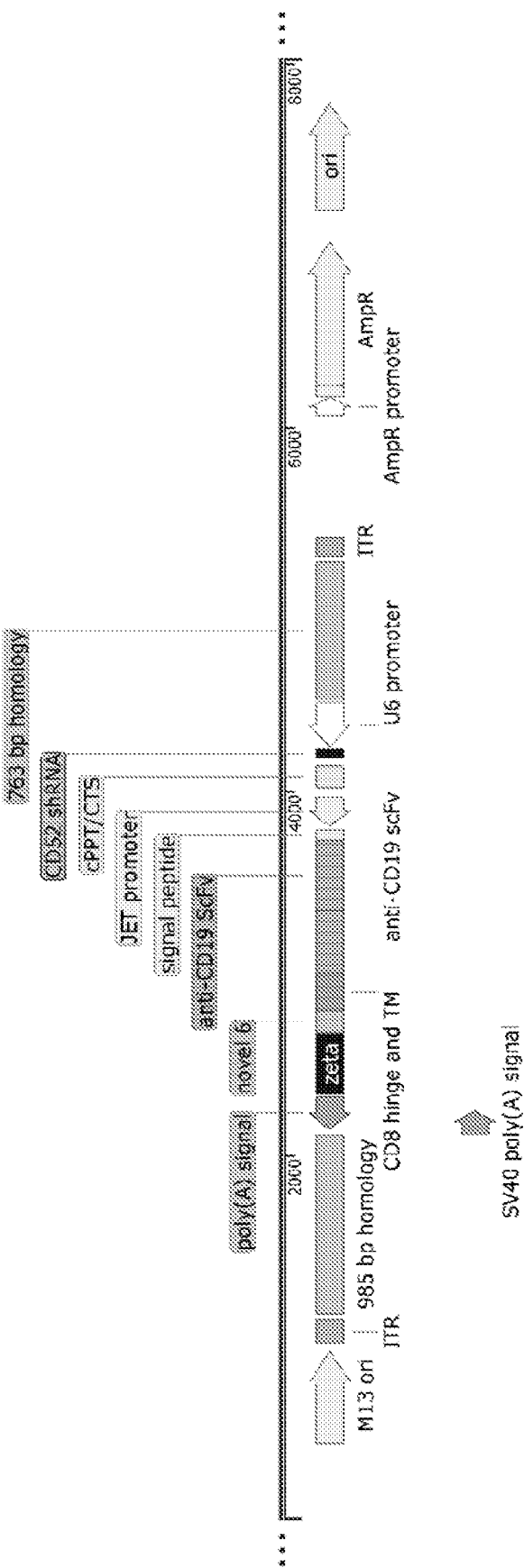
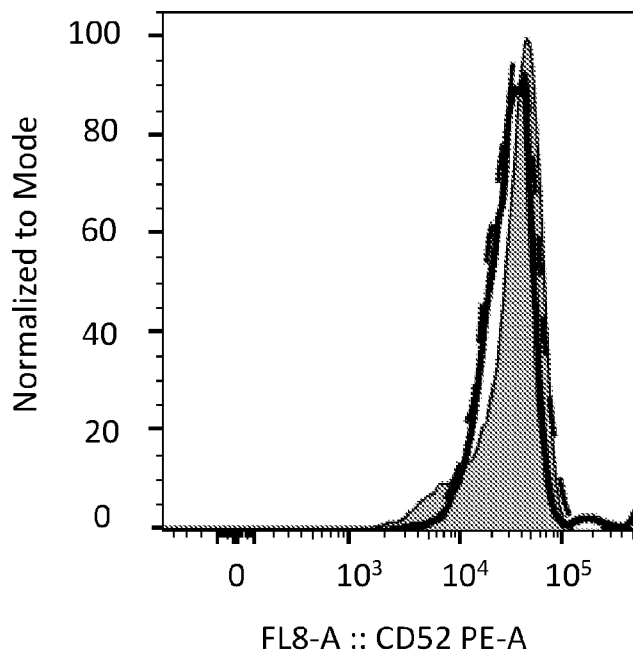


FIGURE 10H

FIGURE 11A

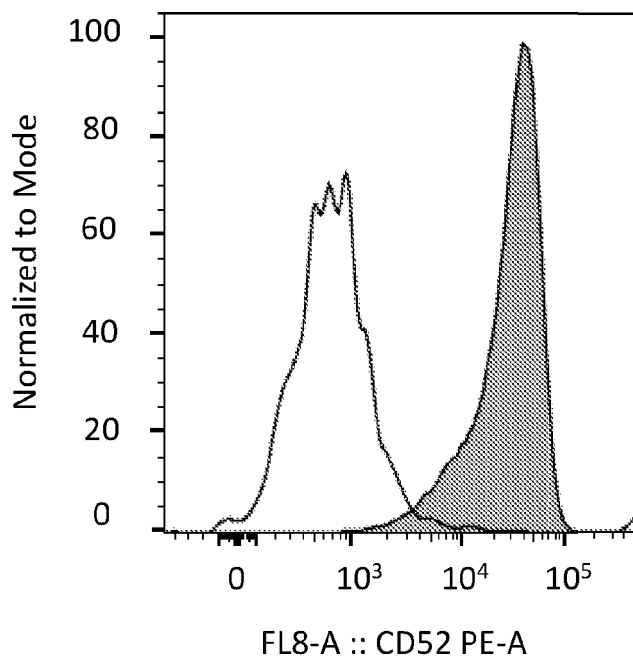
34N6



	Sample Name	Subset Name	Geometric Mean : FL8-A
	34N6.fcs	CAR-CD3-	29721
	34N6.fcs	CAR+CD3-	29846
	34N6.fcs	CAR+CD3+	29966

FIGURE 11B

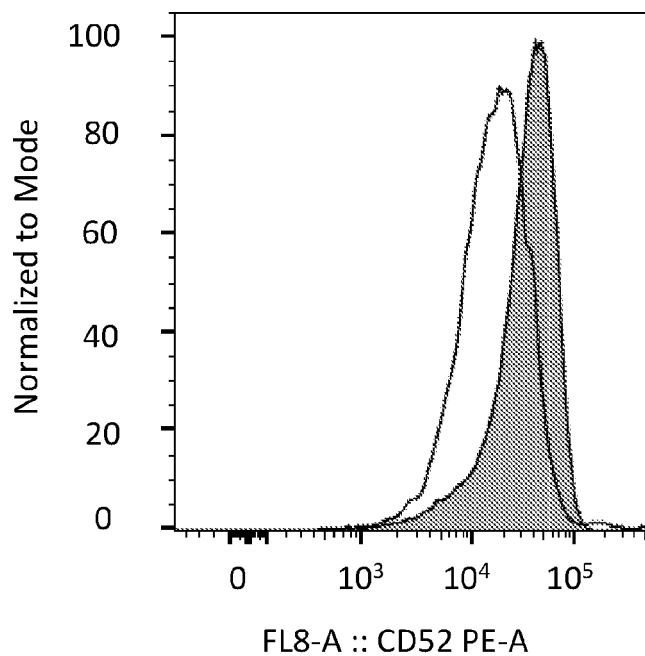
7013



	Sample Name	Subset Name	Geometric Mean : FL8-A
	7013.fcs	CAR-CD3-	25822
	7013.fcs	CAR+CD3-	725

FIGURE 11C

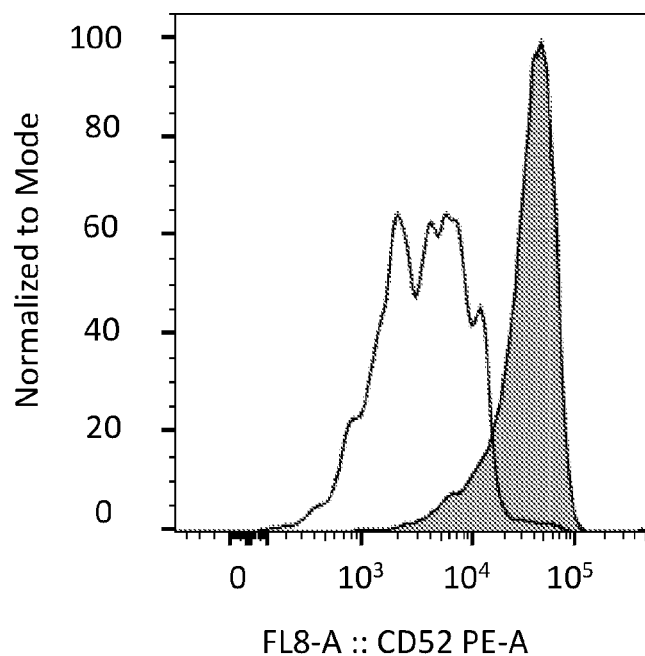
7004



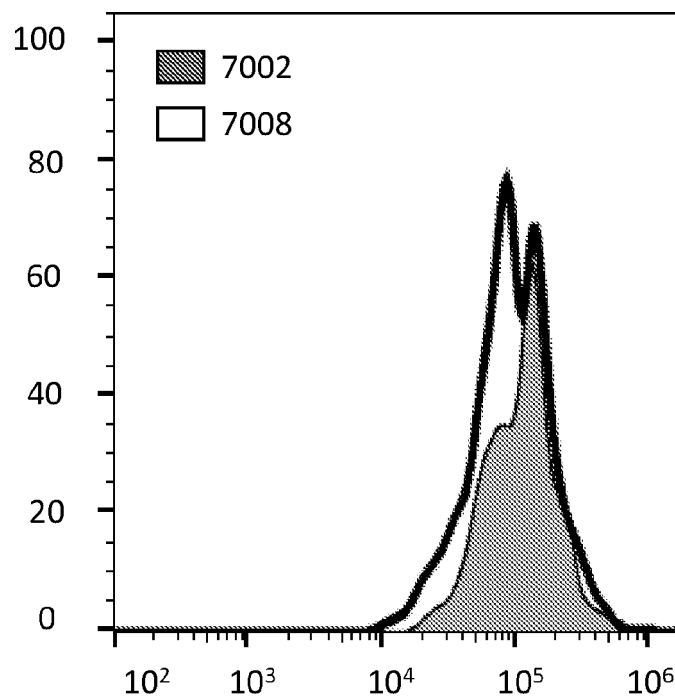
Sample Name	Subset Name	Geometric Mean : FL8-A
7004.fcs	CAR-CD3-	29454
7004.fcs	CAR+CD3-	15864

FIGURE 11D

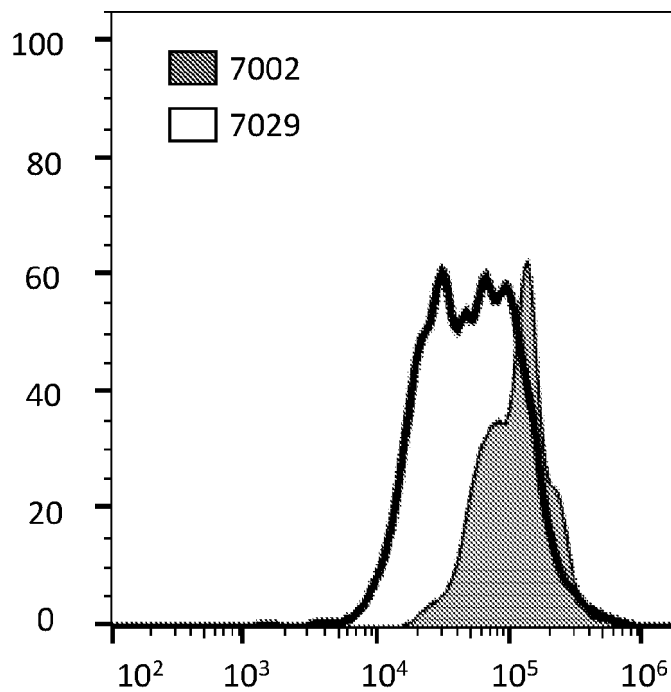
7014



Sample Name	Subset Name	Geometric Mean : FL8-A
7014.fcs	CAR-CD3-	30684
7014.fcs	CAR+CD3-	3322

FIGURE 12A

FL9-A :: B2m R660-APC-A

FIGURE 12B

FL9-A :: B2m R660-APC-A

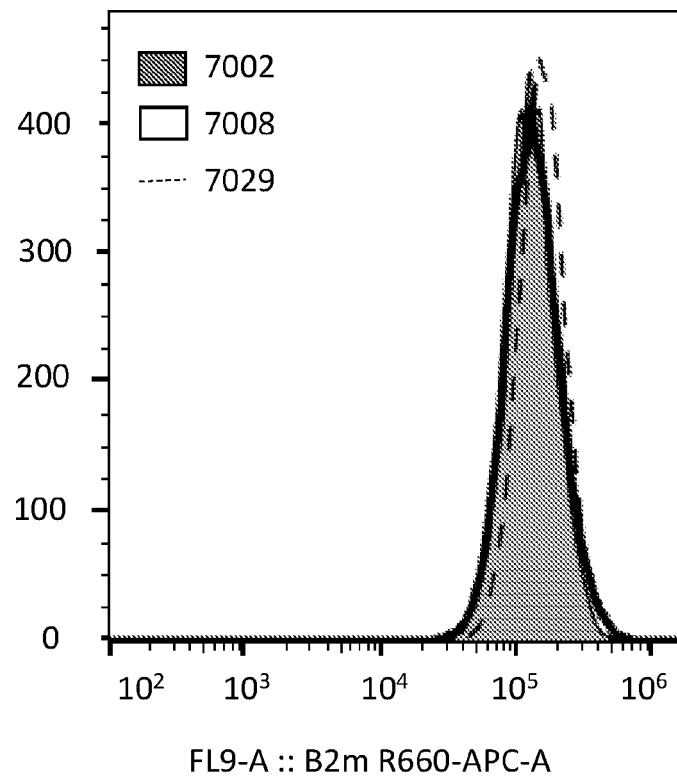
FIGURE 12C

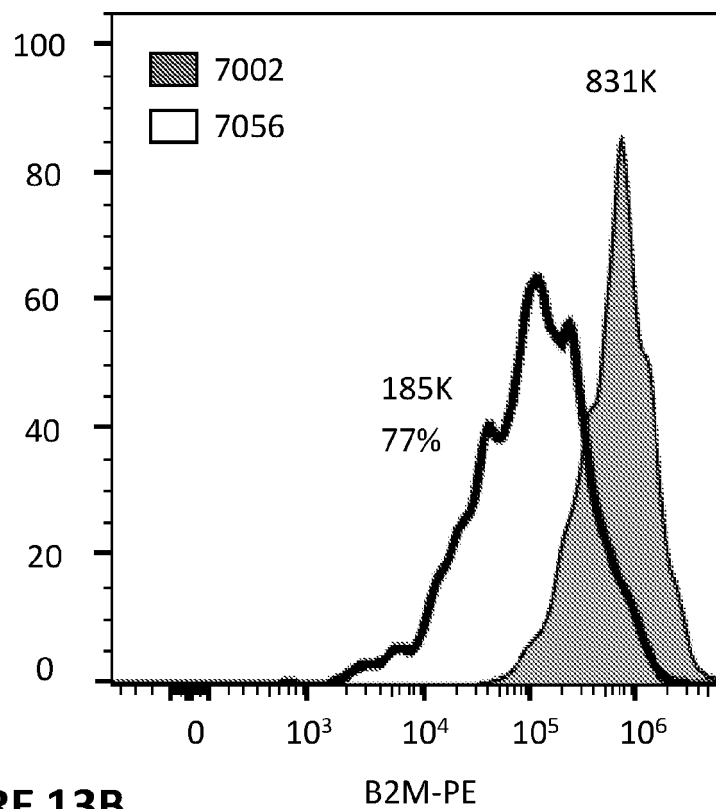
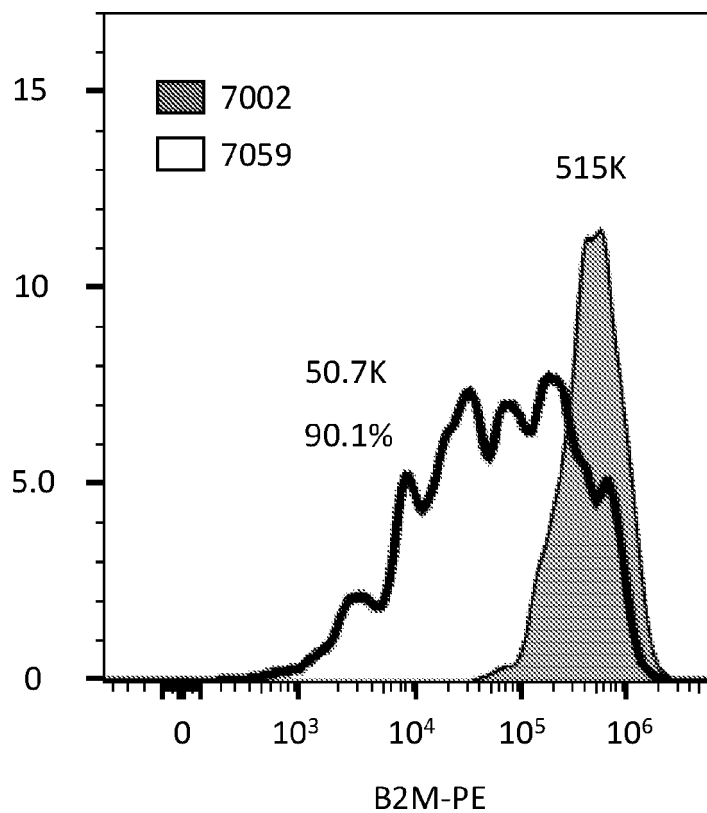
FIGURE 13A**FIGURE 13B**

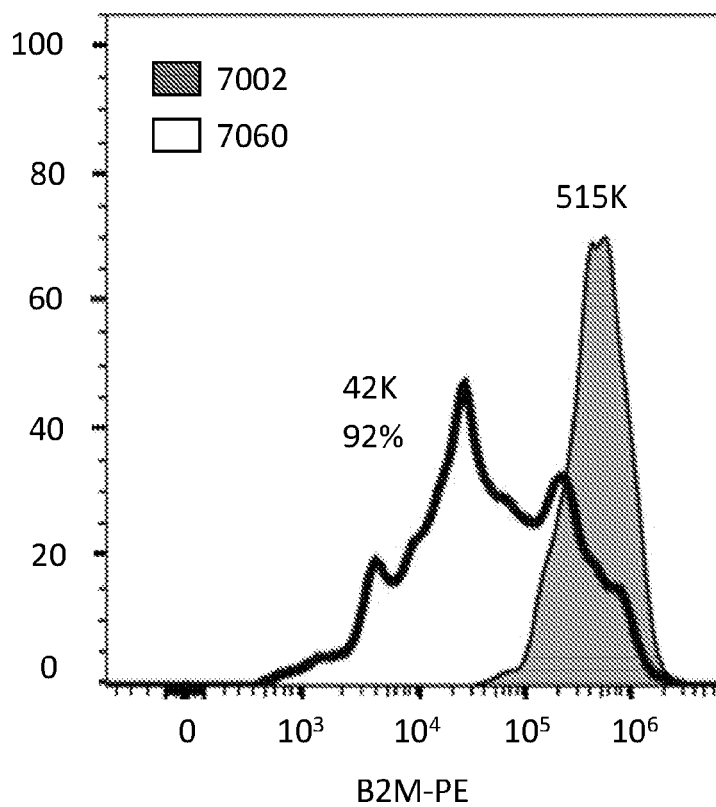
FIGURE 13C

FIGURE 14A

shRNA vs no shRNA

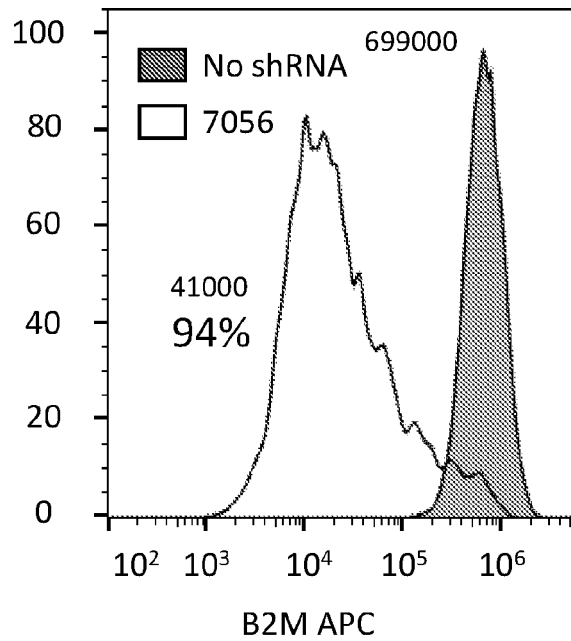


FIGURE 14B

CAR+ vs CAR- reference

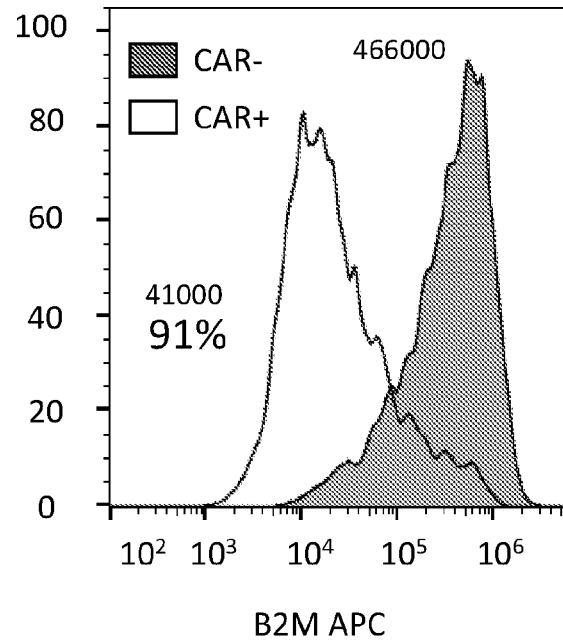


FIGURE 14C

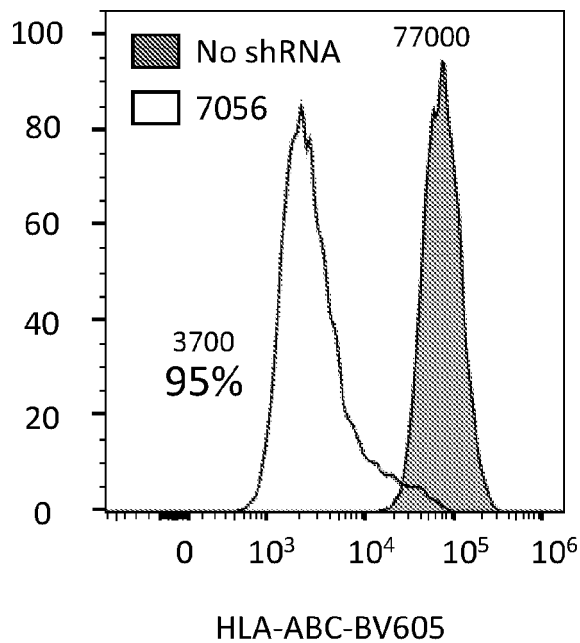
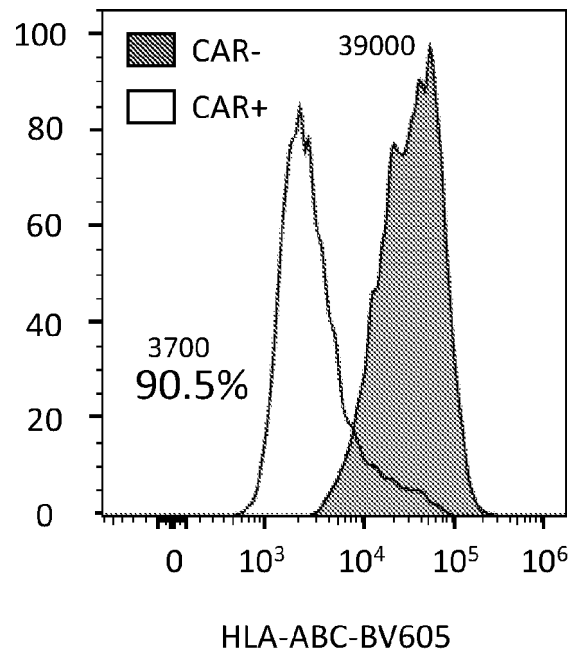


FIGURE 14D



	Subset Name	Mean : FL9-A	Mean : FL13-A
	no shRNA	6.99E5	77582
	no edits	4.66E5	39276
	KD KI	41240	3699

P109070023W000-SEQ-MJT.txt
SEQUENCE LISTING

<110> Precision BioSciences, Inc.

<120> NUCLEIC ACID MOLECULES ENCODING AN ENGINEERED ANTIGEN RECEPTOR
AND AN INHIBITORY NUCLEIC ACID MOLECULE AND METHODS OF USE
THEREOF

<130> P1090.70023W000

<140> Not Yet Assigned

<141> Concurrently Herewith

<150> US 62/503,060

<151> 2017-05-08

<150> US 62/579,460

<151> 2017-10-31

<160> 27

<170> PatentIn version 3.5

<210> 1

<211> 22

<212> DNA

<213> Homo sapiens

<400> 1

tggcctggag caacaaatct ga

22

<210> 2

<211> 62

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 2

gtaccggagg tttgaagatg ccgcatttct cgagaaatgc ggcattctca aacctttttt

60

tg

62

<210> 3

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 3
ccggctggtc tttctatctc ttgtactcga gtacaagaga tagaaagacc agtttttg 58

<210> 4
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic polynucleotide

<400> 4
ccggcagcag agaatggaaa gtcaactcga gttgactttc cattctctgc tgtttttg 58

<210> 5
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic polynucleotide

<400> 5
ccggcagata caaactggac tctcactcga gtgagagtcc agtttgatc tgtttttg 58

<210> 6
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic polynucleotide

<400> 6
ccggccacca tcactcgcaa gagaactcga gttctcttgc gagtgatggt ggttttttg 59

<210> 7
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic polynucleotide

<400> 7
ccggcctcct ggttatggta cagatctcga gatctgtacc ataaccagga ggtttttg 58

<210> 8
 <211> 58
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 8
 ccggcattgc cataatccac ctcttctcga gaagaggtgg attatggcat tgtttttg 58

<210> 9
 <211> 58
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 9
 ccgggcagca tgagcggagg catttctcga gaaatgcctc cgctcatgct gctttttg 58

<210> 10
 <211> 7524
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 10
 cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgag 60
 cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc 120
 gtttttcttg ttgcaatggc tggcggtaat attgttcttg atattaccag caaggccgat 180
 agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca 240
 acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac 300
 acttctcagg attctggcgt accgttcttg tctaaaatcc ctttaatcgg cctcctgttt 360
 agtccccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata 420
 gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 480
 cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cctttctcgc 540

cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gaggccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagttagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtactta cagtttatta aatagatggt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttccc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggcttat ggacttcaag agcaacagtg	2100

P109070023W000-SEQ-MJT.txt

ctgtggcctg gagcaactag tgggcggagt tagggcggag ccaatcagcg tgcgccgttc	2160
cgaaagttgc cttttatggc tgggcggaga atgggcgggtg aacgccgatg attatataag	2220
gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg gatttgggtc gcggttcttg	2280
tttgttccgg aaagccacca tggcgctccc agtgacagcc ttacttttac ctctggcggtt	2340
attattgcac gcggctcgtc ctgacataca gatgactcag actacctctt ccctatctgc	2400
ttcttttaggc gaccgagtaa caatatcttg ccgggccagc caggacatct caaaatactt	2460
aaactgggat cagcagaagc cggacggaac agttaagttg ctcatattacc acacgtcgag	2520
attacactca ggcgttccta gccgattttc gggttccggt tccggtacgg actacagcct	2580
gacaatcagt aaccttgagc aggaggacat cgccacctac ttctgtcagc agggcaacac	2640
gctcccgtac acattcgggtg ggggaactaa gctggagatt accggaggcg gtggcagcgg	2700
tggcggcggc agcgggggtg gcggctcgga ggtcaagtta caggagagcg gaccgggctt	2760
ggtcgcacct agccagagcc tctcagtcac gtgcactgtg tctggagtca gtctcccaga	2820
ctacggggta tcatggatac gacagccgcc tagaaagggc ttagagtggc tgggggttat	2880
ctggggaagt gaaaccacat actacaactc agctctcaag agccgcctca ccatcattaa	2940
ggacaacagt aagtcgcagg ttttcttaaa gatgaactct ctccagactg acgacaccgc	3000
tatttactac tgcgcggaagc actactacta cggcgggagt tacgcaatgg actactgggg	3060
tcagggcact tctgtgaccg tatccagcga gttacctacc cagggaacat tttcaaagt	3120
ttctacaaat gtatccccag cgaagcccac tactaccca gccccacgtc cccccacgcc	3180
agctccaacg atagcaagtc agcccttatc tcttcgccct gaggcttgca ggccccgcggc	3240
gggcggcgcc gttcacacgc gaggactaga cttcgccgtg gacatctaca tctgggcacc	3300
actagccggg acttgcgag tggtgtgtgt gagcttggtg ataacgctct actgcaaagc	3360
gagccgcaaa aaagcggcgg cggcgggctaa aagcccgttt gcgagcccgg cgagcagcgc	3420
gcaggaagaa gatgcgagca gctgccgcgc gccgagcgaa gaagaaggca gctgcgaact	3480
gagagtgaag ttctctcgct ccgcggacgc acccgcttac cagcagggtc agaaccagct	3540
atacaacgag ttaaacctgg ggcgccggga ggagtacgac gtgttagaca agcgtagagg	3600
tagggacccg gagatgggag gcaagcctcg gagaaagaac ccccaggagg gcctgtacaa	3660

cgaactccag aaggacaaga tggctgaggc gtactcggag attggtatga agggcgagag	3720
acgtcgcgga aagggaacac acggcttata ccaggggctt tccaccgcga ccaaggacac	3780
atacgacgcg ctgcacatgc aagccttacc acctcgatga ggtaccagcg gccgcttcga	3840
gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa	3900
aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc	3960
aataaacaag ttaacaacaa caattcgaat ttaaatcgga tccgcaacaa atctgacttt	4020
gcatgtgcaa acgccttcaa caacagcatt attccagaag acaccttctt ccccgagcca	4080
ggtaagggca gctttggtgc cttcgcaggc tgtttccttg cttcaggaat ggccaggttc	4140
tgcccagagc tctggtcaat gatgtctaaa actcctctga ttggtggtct cggccttatc	4200
cattgccacc aaaaccctct ttttactaag aaacagtgag ccttgttctg gcagtccaga	4260
gaatgacacg ggaaaaaagc agatgaagag aaggtggcag gagagggcac gtggcccagc	4320
ctcagtctct ccaactgagt tcctgcctgc ctgcctttgc tcagactgtt tgccccttac	4380
tgctcttcta ggcctcattc taagcccctt ctccaagttg cctctcctta tttctccctg	4440
tctgcaaaaa aatctttccc agctcactaa gtcagtctca cgcagtcact cattaacca	4500
ccaatcactg attgtgccgg cacatgaatg caccaggtgt tgaagtggag gaattaaaaa	4560
gtcagatgag ggggtgtgcc agaggaagca ccattctagt tgggggagcc catctgtcag	4620
ctgggaaaag tccaaataac ttcagattgg aatgtgtttt aactcagggt tgagaaaaca	4680
gccaccttca ggacaaaagt cagggaaggg ctctctgaag aaatgctact tgaagatacc	4740
agccctacca agggcagggg gaggaccaat tgatggagtt ggccactccc tctctgcgcg	4800
ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc	4860
cggcctcagt gagcgagcga gcgcgcagag agggagtggc caacggcgcg cctgcaggtc	4920
tcaaaaatag ctaccctctc cgcatgaat ttatcagcta gaacggttga atatcatatt	4980
gatggtgatt tgactgtctc cggcctttct caccgtttg aatctttacc tacacattac	5040
tcaggcattg catttaaaat atatgagggt tctaaaaatt tttatccttg cgttgaaata	5100
aaggcttctc ccgcaaaagt attacagggt cataatgttt ttggtacaac cgatttagct	5160
ttatgctctg aggctttatt gcttaatttt gctaattctt tgccttgcct gtatgattta	5220

ttggatgttg gaattcctga tgcggtatTT tctccttacg catctgtgcg gtatttcaca	5280
ccgcatatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg	5340
acacccgccca acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta	5400
cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc	5460
gaaacgcgcg agacgaaagg gcctcgtgat acgcctatTT ttataggTTa atgtcatgat	5520
aataatggTT tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaacccctat	5580
ttgtttatTT ttctaaatac attcaaatat gtatccgctc atgagacaat aacctgata	5640
aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct	5700
tattcccttt tttgcggcat tttgccttcc tgtttttgct caccagaaa cgctggtgaa	5760
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa	5820
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt	5880
taaagtTctg ctatgtggcg cggTattatc ccgtattgac gccgggcaag agcaactcgg	5940
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca	6000
tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtataa	6060
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt	6120
gcacaacatg ggggatcatg taactcgcct tgatcgTTgg gaaccggagc tgaatgaagc	6180
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa	6240
actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga	6300
ggcggataaa gttgcaggac cacttctgcg ctcgccctt ccggctggct ggtttattgc	6360
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga	6420
tggtaaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga	6480
acgaaataga cagatcgctg agataggtgc ctactgatt aagcattggt aactgtcaga	6540
ccaagtttac tcatatatac tttagattga tttaaaactt catTTTTaat ttaaaaggat	6600
ctaggtgaag atcctTTTTg ataatctcat gacaaaaatc ccttaacgtg agttttcgtt	6660
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct	6720
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggTgg tttgtttgcc	6780

ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagataacc	6840
aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc	6900
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc	6960
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg	7020
aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata	7080
cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta	7140
tccggtaaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc	7200
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg	7260
atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt	7320
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt	7380
ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga	7440
gcgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac cgcctctccc	7500
cgcgcgttgg ccgattcatt aatg	7524

<210> 11

<211> 7739

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 11

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttcttg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcggcg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480

cgctacactt gccagcgccc tagcgccgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtagac atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgaacgc cgggctttgc ccgggcggcc tcagtgaacg agcgaacgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtactta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040

P109070023W000-SEQ-MJT.txt

tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaag tggtacctca	2340
tcgagggtgg aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaaag	2400
cccctgggat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccattctccg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgcg gcgggcctgc aagcctcagg	2940
gcgaagagat aagggctgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaaca tttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgacct cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatacg gtgtcgtag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggtcttt	3240
gagagctgag ttgtagtatg tggtttctact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac ccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccgggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac ccccgctgcc gccgccaccg ctgccaccgc ctccggtaat	3480
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600

P109070023W000-SEQ-MJT.txt

ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aaatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggcccggcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780
agtcactctgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgac ccaaatccc	3900
gctgacgagg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcggtcac	3960
cgccattct ccgcccagcc ataaaaggca actttcggaa cggcgacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc ttcgagcaga catgataaga tacattgatg	4080
agtttgaca aaccacaact agaatgcagt gaaaaaatg ctttatttgt gaaatttgtg	4140
atgctattgc tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt	4200
cgaatttaaa tcggatccgc aacaaatctg actttgcatg tgcaaacgcc ttcaacaaca	4260
gcattattcc agaagacacc ttcttccca gccaggtaa gggcagctt ggtgccttcg	4320
caggctgtt cttgcttca ggaatggcca ggttctgccc agagctctgg tcaatgatgt	4380
ctaaaactcc tctgattggt ggtctcgcc ttatccattg ccacaaaac cctcttttta	4440
ctaagaaaca gtgagccttg ttctggcagt ccagagaatg acacgggaaa aaagcagatg	4500
aagagaaggt ggcaggagag ggcacgtggc ccagcctcag tctctccaac tgagttcctg	4560
cctgcctgcc tttgctcaga ctgtttgcc cttactgctc ttctaggcct cattctaagc	4620
cccttctcca agttgcctct cttatttct ccctgtctgc caaaaaatct ttcccagctc	4680
actaagtcag tctcacgcag tctactatta accaccaat cactgattgt gccggcacat	4740
gaatgcacca ggtgttgaag tggaggaatt aaaaagtcag atgaggggtg tgcccagagg	4800
aagcaccatt ctagttgggg gagcccatct gtcagctggg aaaagtccaa ataacttcag	4860
attggaatgt gttttaactc agggttgaga aaacagccac cttcaggaca aaagtcaggg	4920
aagggtctc tgaagaaatg ctacttgaag ataccagccc taccaagggc agggagagga	4980
ccaattgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgccc	5040
ggcaaagccc gggcgtcggg cgaccttgg tcgcccggcc tcagtgcgc agcgagcgc	5100
cagagagga gtggccaacg gcgcgcctgc aggtctcaaa aatagctacc ctctccggca	5160

tgaatttatc agctagaacg gttgaatatc atattgatgg tgatttgact gtctccggcc	5220
tttctcaccg gtttgaatct ttacctacac attactcagg cattgcattt aaaatatatg	5280
agggttctaa aaatTTTTat ccttgcggtg aaataaaggc ttctcccgca aaagtattac	5340
agggtcataa tgTTTTtggt acaaccgatt tagctttatg ctctgaggct ttattgctta	5400
atTTtgctaa ttctttgcct tgcctgtatg atttattgga tgTTggaatt cctgatgcgg	5460
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatggtgcac tctcagtaca	5520
atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg	5580
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg	5640
agctgcatgt gtcagagggt ttcaccgtca tcaccgaaac gcgcgagacg aaagggcctc	5700
gtgatacgcc tatttttata ggttaatgtc atgataataa tggtttctta gacgtcagggt	5760
ggcacttttc ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca	5820
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg	5880
aagagtatga gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc	5940
cttcctgttt ttgctcaccg agaaacgctg gtgaaagtaa aagatgctga agatcagttg	6000
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt	6060
cgccccgaag aacgttttcc aatgatgagc actttttaag ttctgctatg tggcgcggta	6120
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat	6180
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga	6240
gaattatgca gtgctgcat aaccatgagt gataaactg cggccaactt acttctgaca	6300
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact	6360
cgcttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc	6420
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact	6480
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt	6540
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt	6600
gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt	6660
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata	6720

ggtgcctcac tgattaagca ttggttaactg tcagaccaag tttactcata tatacttttag	6780
attgatttaa aacttcattt ttaattttaa aggatctagg tgaagatcct ttttgataat	6840
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa	6900
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca	6960
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt	7020
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg	7080
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaate	7140
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga	7200
cgatagttag cgataaggc gcagcggctg ggctgaacgg ggggttcgtg cacacagccc	7260
agcttgagac gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc	7320
gccacgcttc ccgaaggag aaaggcggac aggtatccgg taagcggcag ggtcgggaaca	7380
ggagagcgca cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg	7440
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta	7500
tggaaaaacg ccagcaacgc ggctttttta cggttcctgg ccttttgctg gccttttgct	7560
cacatgttct ttctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag	7620
tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa	7680
gcggaagagc gcccaatacg caaacgcct ctccccgcgc gttggccgat tcattaatg	7739

<210> 12
 <211> 8073
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 12	
cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240

acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgtt ttcttccctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtagat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggct gcccgacgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgcgcgcg	1080
cagagaggga gtggccaact ccactcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt ttggtacttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgcgtggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800

acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agaggggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgggcggagt tagggcggag ccaatcagcg tgcgccgttc	2160
cgaaagttgc cttttatggc tgggcggaga atgggcgggtg aacgccgatg attatataag	2220
gacgcgccgg gtgtggcaca gctagtccg tcgcagccgg gatttgggtc gcggttcttg	2280
tttgttccgg aaagccacca tggcgctccc agtgacagcc ttacttttac ctctggcggt	2340
attattgcac gcggctcgtc ctgacataca gatgactcag actacctctt ccctatctgc	2400
ttctttaggc gaccgagtaa caatatcttg ccgggccagc caggacatct caaaatactt	2460
aaactggtat cagcagaagc cggacggaac agttaagttg ctcatctacc acacgtcgag	2520
attacactca ggcgttccta gccgattttc gggttccggt tccggtacgg actacagcct	2580
gacaatcagt aaccttgagc aggaggacat cgccacctac ttctgtcagc agggcaacac	2640
gctcccgtac acattcggtg ggggaactaa gctggagatt accggaggcg gtggcagcgg	2700
tggcggcggc agcgggggtg gcggctcggg ggtcaagtta caggagagcg gaccgggctt	2760
ggtcgcacct agccagagcc tctcagtcac gtgcactgtg tctggagtca gtctcccaga	2820
ctacggggta tcatggatac gacagccgcc tagaaagggc ttagagtggc tgggggttat	2880
ctggggaagt gaaaccacat actacaactc agctctcaag agccgcctca ccatcattaa	2940
ggacaacagt aagtcgcagg ttttcttaaa gatgaactct ctccagactg acgacaccgc	3000
tatttactac tgcgcgaagc actactacta cggcgggagt tacgcaatgg actactgggg	3060
tcagggcact tctgtgaccg tatccagcga gttacctacc cagggaacat tttcaaatgt	3120
ttctacaaat gtatccccag cgaagcccac tactacccca gccccacgtc cccccacgcc	3180
agctccaacg atagcaagtc agcccttatc tcttcgccct gaggcttgca ggcccgcggc	3240
gggcggcgcc gttcacacgc gaggactaga cttcgcctgc gacatctaca tctgggcacc	3300
actagccggg acttgcgag tgttgttgtt gagcttggtg ataacgctct actgcaaagc	3360

P109070023W000-SEQ-MJT.txt

gagccgcaaa aaagcggcgg cggcgggctaa aagcccgttt gcgagcccgg cgagcagcgc	3420
gcaggaagaa gatgcgagca gctgccgcgc gccgagcgaa gaagaaggca gctgcgaact	3480
gagagtgaag ttctctcgct ccgcggacgc acccgcttac cagcagggtc agaaccagct	3540
atacaacgag ttaaacctgg ggcgccggga ggagtacgac gtgttagaca agcgtagagg	3600
tagggacccg gagatgggag gcaagcctcg gagaaagaac ccccaggagg gcctgtacaa	3660
cgaactccag aaggacaaga tggctgaggc gtactcggag attggtatga agggcgagag	3720
acgtcgcgga aagggaacg acggcttata ccaggggctt tccaccgcga ccaaggacac	3780
atacgacgcg ctgcacatgc aagccttacc acctcgatga ggtaccagcg gccgcttcga	3840
gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa	3900
aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc	3960
aataaacaag ttaacaacaa caattcgaag gatctcgacg gtatcgatca cgagactagc	4020
ctcgagcggc cgcccccttc accgagggcc tatttcccat gattccttca tatttgcata	4080
tacgatacaa ggctgttaga gagataattg gaattaattt gactgtaaac acaaagatat	4140
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaaat	4200
tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg	4260
ctttatatat cttgtggaaa ggacgaaaca ccggcctcct ggttatggta cagatctcga	4320
gatctgtacc ataaccagga ggtttttgaa ttctcgacct cgagacaaat ggacgtattc	4380
atccacaatt ttaaaagaaa aggggggatt ggggggtaca gtgcagggga aagaatagta	4440
gacataatag caacagacat acaaactaaa gaattacaaa aacaaattac aaaaattcaa	4500
aattttcggg tttattacag ggacagcaga gatccacttt ggccgcggat ccgcaacaaa	4560
tctgactttg catgtgcaaa cgccttcaac aacagcatta ttccagaaga caccttcttc	4620
cccagcccag gtaagggcag ctttggtgcc ttgcgaggct gtttccttgc ttcaggaatg	4680
gccaggttct gccagagct ctggtcaatg atgtctaaaa ctctctgat tgggtggtctc	4740
ggccttatcc attgccacca aaaccctctt ttactaaga aacagtgagc cttgttctgg	4800
cagtccagag aatgacacgg gaaaaaagca gatgaagaga aggtggcagg agagggcacg	4860
tggcccagcc tcagtctctc caactgagtt cctgcctgcc tgcctttgct cagactgttt	4920

gccccttact gctcttctag gcctcattct aagcccccttccccaagttgc ctctccttat	4980
ttctccctgt ctgccaaaaa atctttccca gctcactaag tcagtctcac gcagtcactc	5040
attaaccac caatcactga ttgtgccggc acatgaatgc accaggtggt gaagtggagg	5100
aattaaaaag tcagatgagg ggtgtgcca gaggaagcac cattctagtt gggggagccc	5160
atctgtcagc tgggaaaagt ccaaataact tcagattgga atgtgtttta actcagggtt	5220
gagaaaacag ccaccttcag gacaaaagtc aggggaagggc tctctgaaga aatgctactt	5280
gaagatacca gccctaccaa gggcagggag aggaccaatt gatggagttg gccactccct	5340
ctctgcgcgc tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct	5400
ttggtcgccc ggcctcagt agcgagcgag cgcgagaga gggagtggcc aacggcgcg	5460
ctgcaggtct caaaaatagc taccctctcc ggcatgaatt tatcagctag aacggttgaa	5520
tatcatattg atggtgattt gactgtctcc ggcctttctc acccgtttga atctttacct	5580
acacattact caggcattgc atttaaaata tatgagggtt ctaaaaattt ttatccttgc	5640
gttgaaataa aggcttctcc cgcaaaagta ttacagggtc ataatgtttt tgggtacaacc	5700
gatttagctt tatgctctga ggctttattg cttaattttg ctaattcttt gccttgctg	5760
tatgatttat tggatgttg aattcctgat gcggtatttt ctccttacgc atctgtgcgg	5820
tatttcacac cgcatatggt gactctcag tacaatctgc tctgatgccg catagttaag	5880
ccagccccga caccgcga caccgctga cgcgccctga cgggcttgct tgctcccggc	5940
atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc	6000
gtcatcaccg aaacgcgcga gacgaaagg cctcgtgata cgcctatttt tataggttaa	6060
tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg	6120
aaccctatt tgtttatttt tctaaataca ttcaaataatg tatccgctca tgagacaata	6180
accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg	6240
tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac	6300
gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact	6360
ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat	6420
gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga	6480

gcaactcggg	cgccgcatac	actatttctca	gaatgacttg	gttgagtact	caccagtcac	6540
agaaaagcat	cttacggatg	gcatgacagt	aagagaatta	tgcagtgtg	ccataacat	6600
gagtataac	actgcggcca	acttacttct	gacaacgatc	ggaggaccga	aggagctaac	6660
cgcttttttg	cacaacatgg	gggatcatgt	aactcgcctt	gatcgttggg	aaccggagct	6720
gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	cctgtagcaa	tggcaacaac	6780
gttgcgcaaa	ctattaactg	gcgaactact	tactctagct	tcccggcaac	aattaataga	6840
ctggatggag	gcggataaag	ttgcaggacc	acttctgcgc	tcggcccttc	cggctggctg	6900
gtttattgct	gataaatctg	gagccgggtga	gcgtgggtct	cgcggtatca	ttgcagcact	6960
ggggccagat	ggtaagccct	cccgtatcgt	agttatctac	acgacgggga	gtcaggcaac	7020
tatggatgaa	cgaaatagac	agatcgctga	gatagggtgcc	tcactgatta	agcattggta	7080
actgtcagac	caagtttact	catatatact	ttagattgat	ttaaaacttc	atttttaatt	7140
taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acaaaaatcc	cttaacgtga	7200
gttttcgttc	cactgagcgt	cagaccccgt	agaaaagatc	aaaggatctt	cttgagatcc	7260
tttttttctg	cgcgtaatct	gctgcttgca	aacaaaaaaaa	ccaccgctac	cagcggtggt	7320
ttgtttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	tcagcagagc	7380
gcagatacca	aatactgtcc	ttctagtgtg	gccgtagtta	ggccaccact	tcaagaactc	7440
tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	ctgccagtgg	7500
cgataagtcg	tgtcttaccg	ggttggactc	aagacgatag	ttaccggata	aggcgcagcg	7560
gtcgggctga	acgggggggtt	cgtgcacaca	gcccagcttg	gagcgaacga	cctacaccga	7620
actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	ggagaaaggc	7680
ggacaggtat	ccggtaagcg	gcagggtcgg	aacaggagag	cgcacgaggg	agcttccagg	7740
gggaaacgcc	tggtatcttt	atagtcctgt	cgggtttcgc	cacctctgac	ttgagcgtcg	7800
atttttgtga	tgctcgtcag	gggggaggag	cctatggaaa	aacgccagca	acgcggcctt	7860
tttacgggtc	ctggcctttt	gctggccttt	tgctcacatg	ttctttcctg	cgttatcccc	7920
tgattctgtg	gataaccgta	ttaccgcctt	tgagttagct	gataccgctc	gccgcagccg	7980
aacgaccgag	cgcagcgagt	cagttagcga	ggaagcggaa	gagcgcccaa	tacgcaaacc	8040

gcctctcccc gcgcgttggc cgattcatta atg 8073

<210> 13

<211> 8013

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 13

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccggggc	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatgtt	1200

P109070023W000-SEQ-MJT.txt

tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat tttgcagggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agaggggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgagggtctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgggcggagt tagggcggag ccaatcagcg tgcgccgttc	2160
cgaaagttgc cttttatggc tgggcggaga atgggcgggtg aacgccgatg attatataag	2220
gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg gatttgggtc gcggttcttg	2280
ttgtttccgg aaagccacca tggcgctccc agtgacagcc ttacttttac ctctggcgtt	2340
attattgcac gcggctcgtc ctgacataca gatgactcag actacctctt ccctatctgc	2400
ttcttttaggc gaccgagtaa caatatcttg ccgggccagc caggacatct caaaatactt	2460
aaactgggat cagcagaagc cggacggaac agttaagttg ctcatTTacc acacgtcgag	2520
attacactca ggcgttccta gccgattttc gggttccggt tccggtacgg actacagcct	2580
gacaatcagt aaccttgagc aggaggacat cgccacctac ttctgtcagc agggcaacac	2640
gctcccgtac acattcggtg ggggaactaa gctggagatt accggaggcg gtggcagcgg	2700
tggcggcggc agcgggggtg gcggctcgga ggtcaagtta caggagagcg gaccgggctt	2760

ggtcgcacct agccagagcc tctcagtcac gtgcactgtg tctggagtca gtctcccaga	2820
ctacggggta tcatggatac gacagccgcc tagaaagggc ttagagtggc tgggggttat	2880
ctggggaagt gaaaccacat actacaactc agctctcaag agccgcctca ccatcattaa	2940
ggacaacagt aagtcgcagg ttttcttaaa gatgaactct ctccagactg acgacaccgc	3000
tatttactac tgcgcgaagc actactacta cggcgggagt tacgcaatgg actactgggg	3060
tcagggcact tctgtgaccg tatccagcac tactacccca gccccacgtc cccccacgcc	3120
agctccaacg atagcaagtc agcccttatc tcttcgccct gaggcttgca ggcccgcggc	3180
gggcggcgcc gttcacacgc gaggactaga cttcgcctgc gacatctaca tctgggcacc	3240
actagccggg acttgccggag tgttggtgtt gagcttggtg ataacgctct actgcaaagc	3300
gagccgcaaa aaagcggcgg cggcggcgaa aagcccgttt gcgagcccgg cgagcagcgc	3360
gcaggaagaa gatgcgagca gctgccgcgc gccgagcgaa gaagaaggca gctgcgaact	3420
gagagtgaag ttctctcgct ccgcggacgc acccgcttac cagcagggtc agaaccagct	3480
atacaacgag ttaaacctgg ggccgccgga ggagtacgac gtgttagaca agcgtagagg	3540
tagggacccg gagatgggag gcaagcctcg gagaaagaac ccccaggagg gcctgtacaa	3600
cgaactccag aaggacaaga tggctgaggc gtactcggag attggtatga agggcgagag	3660
acgtcgcgga aagggacacg acggcttata ccaggggctt tccaccgcga ccaaggacac	3720
atacgacgcg ctgcacatgc aagccttacc acctcgatga ggtaccagcg gccgcttcga	3780
gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa	3840
aatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc	3900
aataaacaag ttaacaacaa caattcgaag gatctcgacg gtatcgatca cgagactagc	3960
ctcgagcggc cgcccccttc accgagggcc tatttcccat gattccttca tatttgcata	4020
tacgatacaa ggctgttaga gagataattg gaattaattt gactgtaaac acaaagatat	4080
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaaat	4140
tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg	4200
ctttatatat cttgtggaag ggacgaaaca ccggcctcct ggttatggta cagatctcga	4260
gatctgtacc ataaccagga ggtttttgaa ttctcgacct cgagacaaat ggcagtattc	4320

atccacaatt ttaaaagaaa aggggggatt ggggggtaca gtgcagggga aagaatagta	4380
gacataatag caacagacat acaaactaaa gaattacaaa aacaaattac aaaaattcaa	4440
aattttcggg tttattacag ggacagcaga gatccacttt ggccgcggat ccgcaacaaa	4500
tctgactttg catgtgcaaa cgccttcaac aacagcatta ttccagaaga caccttcttc	4560
cccagcccag gtaagggcag ctttggtgcc ttgcaggtt gtttccttgc ttcaggaatg	4620
gccaggttct gccagagct ctggtcaatg atgtctaaaa ctcctctgat tgggtggtctc	4680
ggccttatcc attgccacca aaacctctt tttactaaga aacagtgagc cttgttctgg	4740
cagtccagag aatgacacgg gaaaaaagca gatgaagaga aggtggcagg agagggcacg	4800
tggcccagcc tcagtctctc caactgagtt cctgcctgcc tgcctttgct cagactgttt	4860
gccccttact gctcttctag gcctcattct aagccccctt tccaagttgc ctctccttat	4920
ttctccctgt ctgccaaaa atctttcca gctcactaag tcagtctcac gcagtcactc	4980
attaaccac caatcactga ttgtgccggc acatgaatgc accaggtgtt gaagtggagg	5040
aattaaaaag tcagatgagg ggtgtgcca gaggaagcac cattctagtt gggggagccc	5100
atctgtcagc tgggaaaagt ccaaataact tcagattgga atgtgtttta actcagggtt	5160
gagaaaacag ccaccttcag gacaaaagtc aggggaagggc tctctgaaga aatgctactt	5220
gaagatacca gccctacca gggcagggag aggaccaatt gatggagttg gccactccct	5280
ctctgcgcgc tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct	5340
ttggtcgccc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aacggcgcg	5400
ctgcaggtct caaaaatagc taccctctcc ggcatgaatt tatcagctag aacggttgaa	5460
tatcatattg atggtgattt gactgtctcc ggcctttctc acccgtttga atctttacct	5520
acacattact caggcattgc atttaaaata tatgaggggt ctaaaaattt ttatccttgc	5580
gttgaaataa aggcttctcc cgcaaaagta ttacagggtc ataatgtttt tgggtacaacc	5640
gatttagctt tatgctctga ggctttattg cttaattttg ctaattcttt gccttgctg	5700
tatgatttat tggatgttgg aattcctgat gcggtatttt ctccttacgc atctgtgcgg	5760
tatttcacac cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag	5820
ccagccccga caccgccaa caccgctga cgcgccctga cgggcttgtc tgctcccg	5880

atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc	5940
gtcatcaccg aaacgcgcga gacgaaaggg cctcgtgata cgcctatatt tataggttaa	6000
tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg	6060
aaccctatt tgtttatttt tctaaatata ttcaaataatg tatccgctca tgagacaata	6120
accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg	6180
tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac	6240
gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact	6300
ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat	6360
gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga	6420
gcaactcggc cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac	6480
agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgtg ccataaccat	6540
gagtgataac actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac	6600
cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct	6660
gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac	6720
gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac aattaataga	6780
ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg	6840
gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact	6900
ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac	6960
tatggatgaa cgaaatagac agatcgctga gatagggtgcc tctactgatta agcattggta	7020
actgtcagac caagtttact catatatact ttagattgat ttaaaacttc atttttaatt	7080
taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga	7140
gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc	7200
tttttttctg cgcgtaatct gctgcttgca aacaaaaaaaa ccaccgctac cagcggtggt	7260
ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc	7320
gcagatacca aatactgtcc ttctagtgtg gccgtagtta ggccaccact tcaagaactc	7380
tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg	7440

```

cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg      7500
gtcgggctga acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga      7560
actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc      7620
ggacaggtat ccggttaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg      7680
gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg      7740
atTTTTgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt      7800
tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc      7860
tgattctgtg gataaccgta ttaccgcctt tgagttagct gataccgctc gccgcagccg      7920
aacgaccgag cgcagcgagt cagttagcga ggaagcggaa gagcgcccaa tacgcaaacc      7980
gcctctcccc gcgcgttggc cgattcatta atg                                     8013

```

```

<210> 14
<211> 8090
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic polynucleotide

```

```

<400> 14
cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgag      60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc      120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat      180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca      240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac      300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt      360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata      420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac      480
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cttttctcgc      540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt      600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg      660

```

gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgcgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgcgcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tggacttta cagtttatta aatagatggt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgtggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggcttat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttggt aaatttgtga tgctattgct	2220

ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaac tggtagctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cggtaggaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggtg	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaaca tttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgacct cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatacg gtgtcgtcag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggtg ggaggctctt	3240
gagagctgag ttgtagtatg tggtttcact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac ccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag ccggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac cccgctgcc gccgccaccg ctgccaccgc ctccggtaat	3480
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggccccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780

agtcacatgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaattccc	3900
gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcggtcac	3960
cgccatttct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc ccccttcacc gagggcctat ttcccatgat	4080
tccttcatat ttgcatatac gatacaaggc tgtagagag ataattggaa ttaatttgac	4140
tgtaaacaca aagatattag tacaaaatac gtgacgtaga aagtaataat ttcttggtga	4200
gtttgcagtt ttaaaattat gttttaaaat ggactatcat atgcttaccg taacttgaaa	4260
gtatttcgat ttcttggtt tatatatctt gtggaaagga cgaaacaccg gcctcctggt	4320
tatggtacag atctcgagat ctgtaccata accaggagggt ttttgaattc tcgacctga	4380
gacaaatggc agtattcatc cacaatttta aaagaaaagg ggggattggg ggttacagt	4440
caggggaaag aatagtagac ataatagcaa cagacataca aactaaagaa ttacaaaaac	4500
aaattacaaa aattcaaaat tttcgggttt attacaggga cagcagagat ccactttggc	4560
cgcggtaccg caacaaatct gactttgcat gtgcaaacgc cttcaacaac agcattattc	4620
cagaagacac cttcttcccc agcccaggta agggcagctt tggcgccttc gcaggctgtt	4680
tccttgcttc aggaatggcc aggttctgcc cagagctctg gtcaatgatg tctaaaactc	4740
ctctgattgg tggctcggc cttatccatt gccacaaaaa ccctcttttt actaagaaac	4800
agtgagcctt gttctggcag tccagagaat gacacgggaa aaaagcagat gaagagaagg	4860
tggcaggaga gggcacgtgg ccagcctca gtctctccaa ctgagttcct gcctgcctgc	4920
ctttgctcag actgtttgcc cttactgct cttctaggcc tcattctaag ccccttctcc	4980
aagttgcctc tccttatttc tccctgtctg caaaaaatc tttcccagct cactaagtca	5040
gtctcacgca gtcactcatt aaccaccaa tctgattg tgccggcaca tgaatgcacc	5100
aggtgttgaa gtggaggaat taaaaagtca gatgaggggt gtgcccagag gaagcaccat	5160
tctagttggg ggagccatc tgtcagctgg gaaaagtcca aataacttca gattggaatg	5220
tgttttaact cagggttgag aaaacagcca cttcaggac aaaagtcagg gaagggtct	5280
ctgaagaaat gctacttgaa gataccagcc ctaccaaggg caggagagg accaattgat	5340

ggagttggcc actccctctc tgcgcgctcg ctgctcact gaggccgccc gggcaaagcc	5400
cgggcgtcgg gcgacctttg gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg	5460
agtggccaac ggcgcgcctg caggtctcaa aaatagctac cctctccggc atgaatttat	5520
cagctagaac ggttgaatat catattgatg gtgatttgac tgtctccggc ctttctcacc	5580
cgtttgaatc ttacctaca cattactcag gcattgcatt taaaatatat gagggttcta	5640
aaaattttta tccttgcgtt gaaataaagg cttctcccg c aaaagtatta cagggtcata	5700
atgttttttg tacaaccgat ttagctttat gctctgaggc tttattgctt aattttgcta	5760
attctttgcc ttgcctgtat gatttattgg atgttggaaat tcctgatgcg gtattttctc	5820
cttacgcac tgtgcggtat ttcacaccgc atatggtgca ctctcagtac aatctgctct	5880
gatgccgcat agttaagcca gccccgacac ccgccaacac ccgctgacgc gccctgacgg	5940
gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg	6000
tgtcagaggt tttcacgctc atcacgaaa cgcgcgagac gaaagggcct cgtgatacgc	6060
ctatttttat aggttaatgt catgataata atggtttctt agacgtcagg tggcactttt	6120
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat	6180
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg	6240
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttctgtt	6300
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt ggggtgcacga	6360
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa	6420
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt	6480
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt	6540
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc	6600
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga	6660
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat	6720
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct	6780
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc	6840
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg	6900

```

gcccttccgg ctggctgggt tattgctgat aaatctggag ccggtgagcg tgggtctcgc      6960
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg      7020
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca      7080
ctgattaagc attggttaact gtcagaccaa gtttactcat atatacttta gattgattta      7140
aaacttcatt ttttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc      7200
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa      7260
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca      7320
ccgctaccag cgggtggttg tttgccgat caagagctac caactctttt tccgaaggta      7380
actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc      7440
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaata cctgttacca      7500
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta      7560
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag      7620
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaaag cgccacgctt      7680
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc      7740
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac      7800
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac      7860
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgct tcacatgttc      7920
tttcctgcgt tatccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat      7980
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggaagag      8040
cgcccaatac gcaaaccgcc tctccccgcg cgttggccga ttcattaatg      8090

```

<210> 15

<211> 8030

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 15

```

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttccaac agttgcgcag      60

```

cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcggcc tagcgccgc tcctttcgct ttcttccctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgaacgc cgggctttgc ccgggcggcc tcagtgaacg agcgaacgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620

ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttccc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggcttat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaa gcaagtaaac tggtagctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccctc ccattctccg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggcttttc gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgccgccg gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3060
cccgccgtag tagtagtgct tcgcgagta gtaaatacg gtgtcgtag tctggagaga	3120
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggctctt	3180

gagagctgag ttgtagtatg tggtttctact tccccagata acccccagcc actctaagcc	3240
ctttctaggc ggctgtcgtg tccatgatac cccgtagtct gggagactga ctccagacac	3300
agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccgggtccgc tctcctgtaa	3360
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3420
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3480
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3540
ggaacccgaa aatcggttag gaacgcctga gtgtaatctc gacgtgtggt aaatgagcaa	3600
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3660
ggccccggcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3720
agtcactgtg atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3780
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccg	3840
gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcgttcac	3900
cgccccattct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	3960
cgccctaact ccgcccacta gtgcggccgc ccccttcacc gagggcctat ttcccatgat	4020
tccttcatat ttgcatatac gatacaaggc tgttagagag ataattggaa ttaatttgac	4080
tgtaaacaca aagatattag tacaaaatac gtgacgtaga aagtaataat ttcttgggta	4140
gtttgcagtt ttaaaattat gttttaaaat ggactatcat atgcttaccg taacttgaaa	4200
gtatttcgat ttcttggctt tatatatctt gtggaaagga cgaaacaccg gcctcctggt	4260
tatggtacag atctcgagat ctgtaccata accaggaggt ttttgaattc tcgacctga	4320
gacaaatggc agtattcatc cacaatttta aaagaaaagg ggggattggg gggtagctg	4380
caggggaaag aatagtagac ataatagcaa cagacataca aactaaagaa ttacaaaaac	4440
aaattacaaa aattcaaaat tttcgggttt attacaggga cagcagagat ccactttggc	4500
cgcggtaccg caacaaatct gactttgcat gtgcaaacgc cttcaacaac agcattattc	4560
cagaagacac cttcttcccc agcccaggta agggcagctt tggcgccttc gcaggctgtt	4620
tccttgcttc aggaatggcc aggttctgcc cagagctctg gtcaatgatg tctaaaactc	4680
ctctgattgg tggctctggc cttatccatt gccacaaaaa ccctcttttt actaagaaac	4740

agtgagcctt gttctggcag tccagagaat gacacgggaa aaaagcagat gaagagaagg	4800
tggcaggaga gggcacgtgg cccagcctca gtctctccaa ctgagttcct gcctgcctgc	4860
ctttgctcag actgtttgccc ctttactgct cttctaggcc tcatttctaag ccccttctcc	4920
aagttgcctc tccttatttc tccctgtctg ccaaaaaatc tttcccagct cactaagtca	4980
gtctcacgca gtcactcatt aaccaccaa tctactgattg tgccggcaca tgaatgcacc	5040
aggtgttgaa gtggaggaat taaaaagtca gatgaggggt gtgcccagag gaagcaccat	5100
tctagttggg ggagcccatc tgtcagctgg gaaaagtcca aataacttca gattggaatg	5160
tgttttaact cagggttgag aaaacagcca cttcaggac aaaagtcagg gaagggctct	5220
ctgaagaaat gctacttgaa gataccagcc ctaccaaggg caggagagg accaattgat	5280
ggagttggcc actccctctc tgcgcgctcg ctgcctcact gaggccgccc gggcaaagcc	5340
cgggcgtcgg gcgacctttg gtcgcccggc ctgagtgagc gagcgagcgc gcagagaggg	5400
agtggccaac ggcgcgctg caggtctcaa aaatagctac cctctccggc atgaatttat	5460
cagctagaac ggttgaatat catattgatg gtgatttgac tgtctccggc ctttctcacc	5520
cgtttgaatc ttacctaca cattactcag gcattgcatt taaaatatat gagggttcta	5580
aaaattttta tccttgcggt gaaataaagg cttctccgc aaaagtatta cagggtcata	5640
atgttttttg tacaaccgat ttagctttat gctctgaggc tttattgctt aattttgcta	5700
attctttgcc ttgcctgtat gatttattgg atgttggaat tcctgatgcg gtattttctc	5760
cttacgcac tgtgcggtat ttacaccgc atatggtgca ctctcagtac aatctgctct	5820
gatgccgcat agttaagcca gccccgacac ccgccaacac ccgctgacgc gccctgacgg	5880
gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg	5940
tgtcagaggt tttcacgctc atcacgaaa cgcgcgagac gaaagggcct cgtgatacgc	6000
ctatttttat aggttaatgt catgataata atggtttctt agacgtcagg tggcactttt	6060
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat	6120
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg	6180
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctggt	6240
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt ggggtgcacga	6300

gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa	6360
gaacgttttc caatgatgag cactttttaa gttctgctat gtggcgcggt attatcccgt	6420
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt	6480
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc	6540
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga	6600
ggaccgaagg agctaaccgc ttttttgac aacatggggg atcatgtaac tcgccttgat	6660
cgttgggaa cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct	6720
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc	6780
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg	6840
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc	6900
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg	6960
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca	7020
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta	7080
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc	7140
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa	7200
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaacca	7260
ccgctaccag cggtggtttg tttgccgat caagagctac caactctttt tccgaaggta	7320
actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc	7380
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaata cctgttacca	7440
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta	7500
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttgag	7560
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaaag cgccacgctt	7620
cccgaaggga gaaaggcggg caggatatccg gtaagcggca gggtcggaac aggagagcgc	7680
acgaggggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac	7740
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcgagcct atggaaaaac	7800
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc	7860

tttcctgcgt tatcccccta ttctgtggat aaccgtatta ccgcctttga gtgagctgat	7920
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggaagag	7980
cgccaatac gcaaaccgcc tctccccgcg cgttggccga ttcattaatg	8030

<210> 16

<211> 8089

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 16

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttctctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg acccaaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggct gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080

cagagagggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tggtacttta cagtttatta aatagatggt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttcactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttggt aaatttggtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaag tggtacctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtgaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640

agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaaca ttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgacct cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatacg gtgtcgtcag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggtga ggcggtcttt	3240
gagagctgag ttgtagtatg tggtttcact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac ccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag ccggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac ccccgctgcc gccgccaccg ctgccaccgc ctccggtaat	3480
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggccccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780
agtcactctgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccc	3900
gctgcgacgg aactagctgt gccacaccg gcgcgtcctt atataatcat cggcgttcac	3960
cgcccattct ccgcccagcc ataaaaggca actttcgga cggcgcacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc gcggccaaag tggatctctg ctgtccctgt	4080
aataaaccg aaaattttga atttttgtaa tttgtttttg taattcttta gtttgtatgt	4140
ctgttgctat tatgtctact attctttccc ctgcactgta ccccccaatc cccccttttc	4200

ttttaaaatt gtggatgaat actgccatth gtctcgaggt cgagaattca aaaacctcct	4260
ggttatggta cagatctcga gatctgtacc ataaccagga ggccgggtgtt tcgtcctttc	4320
cacaagatat ataaagccaa gaaatcgaaa tactttcaag ttacggtaag catatgatag	4380
tccattttta aacataatth taaaactgca aactacccaa gaaattatta ctttctacgt	4440
cacgtatthh gtactaatat ctttgtgtth acagtcaaat taattccaat tatctctcta	4500
acagccttgt atcgtatatg caaatatgaa ggaatcatgg gaaataggcc ctcgggtgaag	4560
ggggatccgc aacaaatctg actttgcatg tgcaaacgcc ttcaacaaca gcattattcc	4620
agaagacacc ttcttcccca gccaggttaa gggcagctth ggtgccttcg caggctgtth	4680
ccttgcttca ggaatggcca ggttctgccc agagctctgg tcaatgatgt ctaaaactcc	4740
tctgattgggt ggtctcggcc ttatccattg ccaccaaah cctctthhha ctaagaaaca	4800
gtgagccttg ttctggcagt ccagagaatg acacgggaaa aaagcagatg aagagaaggt	4860
ggcaggagag ggcacgtggc ccagcctcag tctctccaac tgagttcctg cctgcctgcc	4920
tttgctcaga ctgtttgccc cttactgctc ttctaggcct cattctaagc cccttctcca	4980
agttgcctct ccttatttht ccctgtctgc caaaaaatct ttcccagctc actaagtcag	5040
tctcacgcag tcactcatta acccaccaat cactgattgt gccggcacat gaatgcacca	5100
ggtgttgaag tggaggaatt aaaaagtcag atgaggggtg tgcccagagg aagcaccatt	5160
ctagttgggg gagcccatct gtcagctggg aaaagtccaa ataacttcag attggaatgt	5220
gtthtaactc agggttgaga aaacagccac cttcaggaca aaagtcaggg aagggtctc	5280
tgaagaaatg ctacttgaag ataccagccc taccaagggc agggagagga ccaattgatg	5340
gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc	5400
gggcgtcggg cgacctthg tcgcccggcc tcagtgaagc agcagcgcg cagagaggga	5460
gtggccaacg gcgcgcctgc aggtctcaaa aatagctacc ctctccggca tgaatttatc	5520
agctagaacg gttgaatatc atattgatgg tgatttgact gtctccggcc tttctcacc	5580
gtttgaatct ttacctacac attactcagg cattgcattt aaaaatatg agggttctaa	5640
aaatthhtht ccttgcgthg aaataaaggc ttctcccgca aaagtattac agggtcataa	5700
tgthhthhgt acaaccgatt tagctthhth ctctgaggct ttattgctta atthhthctaa	5760

ttctttgcct tgcctgtatg atttattgga tgttgggaatt cctgatgcgg tattttctcc	5820
ttacgcatct gtgcggtatt tcacaccgca tatggtgcac tctcagtaca atctgctctg	5880
atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg ccctgacggg	5940
cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg agctgcatgt	6000
gtcagagggtt ttcaccgtca tcaccgaaac gcgcgagacg aaagggcctc gtgatacgcc	6060
tatTTTTata ggTtaatgtc atgataataa tggTTTTctta gacgtcaggt ggcacttttc	6120
ggggaaatgt gcgcggaacc cctatttggtt tatttttcta aatacattca aatatgtatc	6180
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga	6240
gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc cttcctgttt	6300
ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag	6360
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag	6420
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta	6480
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg	6540
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca	6600
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag	6660
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc	6720
gttggggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg	6780
tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc	6840
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg	6900
cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggctctcgcg	6960
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga	7020
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac	7080
tgattaagca ttggtaactg tcagaccaag tttactcata tatacttttag attgatttaa	7140
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca	7200
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag	7260
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac	7320

cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa	7380
ctggcttcag cagagcgagc ataccaaata ctgtccttct agtgtagccg tagttaggcc	7440
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag	7500
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac	7560
cggataaggc gcagcggctg ggctgaacgg ggggttcgtg cacacagccc agcttggagc	7620
gaacgacctt caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc	7680
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcgggaaca ggagagcgca	7740
cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc	7800
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg	7860
ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct cacatgttct	7920
ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata	7980
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc	8040
gccaataacg caaacgcct ctccccgcgc gttggccgat tcattaatg	8089

<210> 17

<211> 8029

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 17

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttcttg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggt gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480

cgctacactt gccagcgccc tagcgccgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgcgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatggt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040

tgtatatcac agacaaaact gtgctagaca tgaggtctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaag tggtacctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccctc ccattctcgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggcttttc gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgcc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3060
cccgccgtag tagtagtgct tcgcgcagta gtaaatacg gtgtcgtag tctggagaga	3120
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggctctt	3180
gagagctgag ttgtagtatg tggtttctact tccccagata acccccagcc actctaagcc	3240
ctttctaggc ggctgtcgta tccatgatac cccgtagtct gggagactga ctccagacac	3300
agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccggtccgc tctcctgtaa	3360
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3420
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3480
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3540
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aaatgagcaa	3600

cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3660
ggcccggcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3720
agtcactctgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3780
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccg	3840
gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcgttcac	3900
cgcccattct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	3960
cgccctaact ccgcccacta gtgcggccgc gcggccaaag tggatctctg ctgtccctgt	4020
aataaacccg aaaattttga atttttgtaa tttgtttttg taattcttta gtttgtatgt	4080
ctgttgctat tatgtctact attctttccc ctgactgta ccccccaatc cccccttttc	4140
ttttaaaatt gtggatgaat actgccattt gtctcgaggt cgagaattca aaaacctcct	4200
ggttatggta cagatctcga gatctgtacc ataaccagga ggccgggtgtt tcgtcctttc	4260
cacaagatat ataaagccaa gaaatcgaag tactttcaag ttacggtaag catatgatag	4320
tccattttta aacataatth taaaactgca aactacccaa gaaattatta ctttctacgt	4380
cacgtatttt gtactaatat ctttgtgttt acagtcaaht taattccaat tatctctcta	4440
acagccttgt atcgtatatg caaatatgaa ggaatcatgg gaaataggcc ctcggtgaag	4500
ggggatccgc aacaaatctg actttgcatg tgcaaacgcc ttcaacaaca gcattattcc	4560
agaagacacc ttcttcccca gccaggtaa gggcagcttt ggtgccttcg caggctgttt	4620
ccttgcttca ggaatggcca ggttctgccc agagctctgg tcaatgatgt ctaaaactcc	4680
tctgattggg ggtctcggcc ttatccattg ccacaaaaac cctcttttta ctaagaaaca	4740
gtgagccttg ttctggcagt ccagagaatg acacgggaaa aaagcagatg aagagaaggt	4800
ggcaggagag ggcacgtggc ccagcctcag tctctccaac tgagttcctg cctgcctgcc	4860
tttgctcaga ctgtttgccc cttactgctc ttctaggcct cattctaagc ccttctcca	4920
agttgcctct ccttatttct ccctgtctgc caaaaaatct ttcccagctc actaagtcag	4980
tctcacgcag tcaactatta acccaccaat cactgattgt gccggcacat gaatgcacca	5040
ggtgttgaag tggaggaatt aaaaagtcag atgaggggtg tgcccagagg aagcaccatt	5100
ctagttgggg gagcccatct gtcagctggg aaaagtccaa ataacttcag attggaatgt	5160

gttttaactc agggttgaga aaacagccac cttcaggaca aaagtcaggg aagggtctctc	5220
tgaagaaatg ctacttgaag ataccagccc taccaagggc agggagagga ccaattgatg	5280
gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc	5340
gggcgtcggg cgacctttgg tcgcccggcc tcagtgaagc agcgaagcgc cagagaggga	5400
gtggccaacg gcgcgcctgc aggtctcaaa aatagctacc ctctccggca tgaatttatc	5460
agctagaacg gttgaatatc atattgatgg tgatttgact gtctccggcc tttctcacc	5520
gtttgaatct ttacctacac attactcagg cattgcattt aaaatatatg agggttctaa	5580
aaatTTTTat ctttgcgttg aaataaaggc ttctcccgca aaagtattac agggtcataa	5640
tgTTTTtggt acaaccgatt tagctttatg ctctgaggct ttattgctta attttgctaa	5700
ttctttgcct tgcctgtatg atttattgga tgttggaatt cctgatgcgg tatTTTctcc	5760
ttacgcatct gtgcggtatt tcacaccgca tatggtgcac tctcagtaca atctgctctg	5820
atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg ccctgacggg	5880
cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg agctgcatgt	5940
gtcagagggtt ttcaccgtca tcaccgaaac gcgcgagacg aaagggcctc gtgatacgcc	6000
tatTTTTata ggTaatgtc atgataataa tggTTTctta gacgtcaggt ggcaTTTTc	6060
ggggaaatgt gcgcggaacc cctatttTgtt tatTTTTctta aatacattca aatatgtatc	6120
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga	6180
gtattcaaca tttccgtgtc gcccttattc cTTTTTTgc ggcatTTTgc cttcctgttt	6240
ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag	6300
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag	6360
aacgtTTTcc aatgatgagc actTTTaaag ttctgctatg tggcgcggtta ttatcccgta	6420
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg	6480
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca	6540
gtgctgccat aaccatgagt gataaactg cggccaactt acttctgaca acgatcggag	6600
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc	6660
gttggggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg	6720

tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc	6780
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg	6840
cccttccggc tggctggttt attgctgata aatctggagc cggtagcgt gggctctcgcg	6900
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga	6960
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac	7020
tgattaagca ttggttaactg tcagaccaag tttactcata tatactttag attgatttaa	7080
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca	7140
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag	7200
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac	7260
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa	7320
ctggcttcag cagagcgcag ataccaaata ctgtccttct agttagccg tagttaggcc	7380
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatac ctgttaccag	7440
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac	7500
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttgagc	7560
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaaagc gccacgcttc	7620
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca	7680
cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc	7740
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg	7800
ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct cacatgttct	7860
ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata	7920
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc	7980
gccaataacg caaaccgcct ctccccgcgc gttggccgat tcattaatg	8029

<210> 18

<211> 8082

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 18

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg acccaaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aaaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgtggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500

cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaaccc tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggcttat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgggcggagt tagggcggag ccaatcagcg tgcgccgttc	2160
cgaaagttgc cttttatggc tgggcggaga atgggcgggtg aacgccgatg attatataag	2220
gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg gatttgggtc gcggttcttg	2280
tttgttccgg aaagccacca tggcgctccc agtgacagcc ttacttttac ctctggcggt	2340
attattgcac gcggctcgtc ctgacataca gatgactcag actacctctt ccctatctgc	2400
ttctttaggc gaccgagtaa caatatcttg ccgggccagc caggacatct caaaatactt	2460
aaactggtat cagcagaagc cggacggaac agttaagttg ctcatattacc acacgtcgag	2520
attacactca ggcgttccta gccgattttc gggttccggt tccggtacgg actacagcct	2580
gacaatcagt aaccttgagc aggaggacat cgccacctac ttctgtcagc agggcaacac	2640
gctcccgtac acattcggtg ggggaactaa gctggagatt accggaggcg gtggcagcgg	2700
tggcggcggc agcgggggtg gcggctcgga ggtcaagtta caggagagcg gaccgggctt	2760
ggtcgcacct agccagagcc tctcagtcac gtgcactgtg tctggagtca gtctcccaga	2820
ctacggggta tcatggatac gacagccgcc tagaaagggc ttagagtggc tgggggttat	2880
ctggggaagt gaaaccacat actacaactc agctctcaag agccgcctca ccatcattaa	2940
ggacaacagt aagtcgcagg ttttcttaaa gatgaactct ctccagactg acgacaccgc	3000
tatttactac tgcgcgaagc actactacta cggcgggagt tacgcaatgg actactgggg	3060

tcagggcact tctgtgaccg tatccagcga gttacctacc cagggaacat tttcaaagt	3120
ttctacaaat gtatccccag cgaagcccac tactaccca gcccacgtc ccccacgcc	3180
agctccaacg atagcaagtc agcccttata tcttcgccct gaggcttgca ggcccgcggc	3240
gggcggcgcc gttcacacgc gaggactaga cttcgccctg gacatctaca tctgggcacc	3300
actagccggg acttgccggag tgttggtgtt gagcttggtataaacgctct actgcaaagc	3360
gagccgcaaa aaagcggcgg cggcggctaa aagcccgttt gcgagcccgg cgagcagcgc	3420
gcaggaagaa gatgcgagca gctgccgcgc gccgagcgaa gaagaaggca gctgcgaact	3480
gagagtgaag ttctctcgct ccgcggacgc acccgcttac cagcagggtc agaaccagct	3540
atacaacgag ttaaacttg ggccgggga ggagtacgac gtgttagaca agcgtagagg	3600
tagggacccg gagatgggag gcaagcctcg gagaaagaac ccccaggagg gcctgtacaa	3660
cgaactccag aaggacaaga tggctgaggc gtactcggag attggtatga agggcgagag	3720
acgtcgcgga aaggacacg acggcttata ccaggggctt tccaccgcga ccaaggacac	3780
atacgcgcg ctgcacatgc aagccttacc acctcgatga ggtaccagcg gccgcttcga	3840
gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat gcagtgtaaa	3900
aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc	3960
aataaacaag ttaacaaca caattcgaag gatctcgacg gtatcgatca cgagactagc	4020
ctcgagcggc cgccccctt accgagggcc tatttcccat gattccttca tatttgcata	4080
tacgatacaa ggctgttaga gagataattg gaattaattt gactgtaaac acaaagatat	4140
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaat	4200
tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg	4260
ctttatatat cttgtggaaa ggacgaaaca ccggtgtacc ggaggtttga agatgccgca	4320
tttctcgaga aatgcggcat cttcaaact ttttttgaat tctcgaccta gggacaaatg	4380
gcagtattca tccacaattt taaaagaaaa ggggggattg gggggtacag tgcaggggaa	4440
agaatagtag acataatagc aacagacata caaactaaag aattacaaaa acaaattaca	4500
aaaattcaaa attttcgggt ttattacagg gacagcagag atccactttg gccgcggatc	4560
cgcaacaaat ctgactttgc atgtgcaaac gccttcaaca acagcattat tccagaagac	4620

accttcttcc ccagcccagg taagggcagc tttgggtgcct tcgcaggctg tttccttgct	4680
tcaggaatgg ccaggttctg cccagagctc tgggtcaatga tgtctaaaac tcctctgatt	4740
gggtggtctcg gccttatcca ttgccaccaa aaccctcttt ttactaagaa acagtgagcc	4800
ttgttctggc agtccagaga atgacacggg aaaaaagcag atgaagagaa ggtggcagga	4860
gagggcacgt ggcccagcct cagtctctcc aactgagttc ctgcctgcct gcctttgctc	4920
agactgtttg ccccttactg ctcttctagg cctcattcta agcccccttct ccaagttgcc	4980
tctccttatt tctccctgtc tgccaaaaaa tctttcccag ctactaagt cagtctcacg	5040
cagtcaactca ttaaccacc aatcactgat tgtgccggca catgaatgca ccaggtgttg	5100
aagtggagga attaaaaagt cagatgaggg gtgtgcccag aggaagcacc attctagttg	5160
ggggagccca tctgtcagct gggaaaagtc caaataactt cagattggaa tgtgttttaa	5220
ctcagggttg agaaaacagc caccttcagg acaaaagtca gggaagggt ctctgaagaa	5280
atgctacttg aagataccag ccctaccaag ggcagggaga ggaccaattg atggagttgg	5340
ccactccctc tctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	5400
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	5460
acggcgcgcc tgcaggtctc aaaaatagct accctctccg gcatgaattt atcagctaga	5520
acggttgaat atcatattga tgggtgatttg actgtctccg gcctttctca cccgtttgaa	5580
tctttaccta cacattactc aggcattgca tttaaaatat atgagggttc taaaaatttt	5640
tatccttgcg ttgaaataaa ggcttctccc gcaaaagtat tacagggtca taatgttttt	5700
ggtacaaccg atttagcttt atgctctgag gctttattgc ttaattttgc taattctttg	5760
ccttgctgt atgatttatt ggatgttgga attcctgatg cggtattttc tccttacgca	5820
tctgtgcggt atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc	5880
atagttaagc cagccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct	5940
gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag	6000
gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac gcctatTTTT	6060
ataggttaat gtcataataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa	6120
tgtgcgcgga acccctatTT gtttattttt ctaaatacat tcaaatatgt atccgctcat	6180

gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca	6240
acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg tttttgctca	6300
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta	6360
catcgaactg gatctcaaca gcggttaagat ccttgagagt tttcgccccg aagaacgttt	6420
tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc	6480
cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc	6540
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgcctgc	6600
cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa	6660
ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga	6720
accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat	6780
ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca	6840
attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc	6900
ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat	6960
tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag	7020
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa	7080
gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca	7140
tttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaaatccc	7200
ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc	7260
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc	7320
agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt	7380
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt	7440
caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc	7500
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa	7560
ggcgcagcgg tcgggctgaa cgggggggtt gtgcacacag cccagcttgg agcgaacgac	7620
ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg	7680
gagaaaggcg gacaggtatc cggtaaagcg cagggtcgga acaggagagc gcacgagggg	7740

gcttccaggg ggaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact	7800
tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa	7860
cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc	7920
gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg	7980
ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcccaat	8040
acgcaaaccg cctctccccg cgcgttggcc gattcattaa tg	8082

<210> 19

<211> 8022

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 19

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttcttg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctcccttttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gaggccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900

P109070023W000-SEQ-MJT.txt

cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgcgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgagcgcg	1080
cagagagga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttgtt tgggtacttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttccc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggtctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgggcggagt tagggcggag ccaatcagcg tgcgccgttc	2160
cgaaagttgc cttttatggc tgggcggaga atgggcggtg aacgccgatg attatataag	2220
gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg gatttgggtc gcggttcttg	2280
tttgttccgg aaagccacca tggcgctccc agtgacagcc ttacttttac ctctggcggt	2340
attattgcac gcggctcgtc ctgacataca gatgactcag actacctctt ccctatctgc	2400
ttctttaggc gaccgagtaa caatatcttg ccgggccagc caggacatct caaaatactt	2460

aaactggtat cagcagaagc cggacggaac agttaagttg ctcatttacc acacgtcgag	2520
attacactca ggcgttccta gccgattttc gggttccggt tccggtacgg actacagcct	2580
gacaatcagt aaccttgagc aggaggacat cgccacctac ttctgtcagc agggcaacac	2640
gctcccgtac acattcggtg ggggaactaa gctggagatt accggaggcg gtggcagcgg	2700
tggcggcggc agcgggggtg gcggtctgga ggtcaagtta caggagagcg gaccgggctt	2760
ggtcgcacct agccagagcc tctcagtcac gtgcactgtg tctggagtca gtctcccaga	2820
ctacggggta tcatggatac gacagccgcc tagaaagggc ttagagtggc tgggggttat	2880
ctggggaagt gaaaccacat actacaactc agctctcaag agccgcctca ccatcattaa	2940
ggacaacagt aagtcgcagg ttttcttaaa gatgaactct ctccagactg acgacaccgc	3000
tatttactac tgcgcgaagc actactacta cggcgggagt tacgcaatgg actactgggg	3060
tcagggcact tctgtgaccg tatccagcac tactacccca gccccacgtc cccccacgcc	3120
agctccaacg atagcaagtc agcccttata tcttcgccct gaggcttgca ggccccgaggc	3180
gggcggcgcc gttcacacgc gaggactaga cttcgctgc gacatctaca tctgggcacc	3240
actagccggg acttgcgag tggtgtgtgt gagcttggtataaacgctct actgcaaagc	3300
gagccgcaaa aaagcggcgg cggcggcgaa aagcccgttt gcgagcccg cgagcagcgc	3360
gcaggaagaa gatgcgagca gctgccgcgc gccgagcgaa gaagaaggca gctgcgaact	3420
gagagtgaag ttctctcgct ccgcggacgc acccgcttac cagcagggtc agaaccagct	3480
atacaacgag ttaaacctgg ggcgccggga ggagtacgac gtgttagaca agcgtagagg	3540
tagggacccg gagatgggag gcaagcctcg gagaaagaac ccccaggagg gcctgtacaa	3600
cgaactccag aaggacaaga tggctgaggc gtactcggag attggtatga agggcgagag	3660
acgtcgcgga aaggacacg acggcttata ccaggggctt tccaccgcga ccaaggacac	3720
atacgacgcg ctgcacatgc aagccttacc acctgatga ggtaccagcg gccgcttcga	3780
gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa	3840
aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc	3900
aataaacaag ttaacaacaa caattcgaag gatctcgacg gtatcgatca cgagactagc	3960
ctcgagcggc cgcccccttc accgagggcc tatttcccat gattccttca tatttgcata	4020

tacgatacaa ggctgttaga gagataattg gaattaatTT gactgtaaac acaaagatat	4080
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaaat	4140
tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg	4200
ctttatatat cttgtggaaa ggacgaaaca ccggtgtacc ggaggtttga agatgccgca	4260
tttctcgaga aatgcggcat cttcaaactt ttttttgaat tctcgacctt gggacaaatg	4320
gcagtattca tccacaatTT taaaagaaaa ggggggattg ggggggtacag tgcaggggaa	4380
agaatagtag acataatagc aacagacata caaactaaag aattacaaaa acaaattaca	4440
aaaattcaaa attttcgggt ttattacagg gacagcagag atccactttg gccgcggatc	4500
cgcaacaaat ctgactttgc atgtgcaaac gccttcaaca acagcattat tccagaagac	4560
accttcttcc ccagcccagg taagggcagc tttggtgcct tcgcaggctg tttccttgct	4620
tcaggaatgg ccaggttctg cccagagctc tggatcaatga tgtctaaaaac tcctctgatt	4680
ggtggtctcg gccttatcca ttgccaccaa aacctctttt ttactaagaa acagtgagcc	4740
ttgttctggc agtccagaga atgacacggg aaaaaagcag atgaagagaa ggtggcagga	4800
gagggcacgt ggcccagcct cagtctctcc aactgagttc ctgcctgcct gcctttgctc	4860
agactgtttg ccccttactg ctcttctagg cctcattcta agcccccttct ccaagttgcc	4920
tctccttatt tctccctgtc tgccaaaaaa tctttcccag ctactaagt cagtctcacg	4980
cagtactca ttaaccacc aatcactgat tgtgccggca catgaatgca ccaggtgttg	5040
aagtggagga attaaaaagt cagatgaggg gtgtgcccag aggaagcacc attctagttg	5100
ggggagccca tctgtcagct gggaaaagtc caaataactt cagattggaa tgtgttttaa	5160
ctcagggttg agaaaacagc caccttcagg acaaaagtca gggaagggtc ctctgaagaa	5220
atgctacttg aagataccag ccctaccaag ggcagggaga ggaccaattg atggagtggg	5280
ccactccctc tctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	5340
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	5400
acggcgcgcc tgcaggtctc aaaaatagct accctctccg gcatgaattt atcagctaga	5460
acggttgaat atcatattga tgggtgattg actgtctccg gcctttctca cccgtttgaa	5520
tctttaccta cacattactc aggcattgca tttaaaatat atgagggttc taaaaatTTT	5580

tatccttgcg ttgaaataaa ggcttctccc gcaaaagtat tacagggtca taatgttttt	5640
ggtacaaccg atttagcttt atgctctgag gctttattgc ttaattttgc taattctttg	5700
ccttgccctgt atgatttatt ggatgttgga attcctgatg cggatatttc tccttacgca	5760
tctgtgcggt atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc	5820
atagttaagc cagccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct	5880
gctcccgga tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag	5940
gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac gcctattttt	6000
ataggttaat gtcataataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa	6060
tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaataatgt atccgctcat	6120
gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca	6180
acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg tttttgctca	6240
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagggggta	6300
catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg aagaacgttt	6360
tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc	6420
cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc	6480
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgctgc	6540
cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa	6600
ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga	6660
accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat	6720
ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca	6780
attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc	6840
ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat	6900
tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag	6960
tcaggcaact atggatgaac gaaatagaca gatcgctgag atagggtgcct cactgattaa	7020
gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca	7080
ttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaaatccc	7140

```

ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc      7200
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc      7260
agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt      7320
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt      7380
caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc      7440
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa      7500
ggcgagcgg tcgggctgaa cgggggggtt gtgcacacag cccagcttgg agcgaacgac      7560
ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc tttccgaagg      7620
gagaaaggcg gacaggtatc cggtaagcgg cagggtcggg acaggagagc gcacgaggga      7680
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact      7740
tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa      7800
cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc      7860
gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg      7920
ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcccaat      7980
acgcaaaccg cctctccccg cgcgttggcc gattcattaa tg                        8022

```

```

<210> 20
<211> 8099
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic polynucleotide

```

```

<400> 20
cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttccaac agttgcgcag      60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc      120
gtttttcctg ttgcaatggc tggcggtaat attgttcttg atattaccag caaggccgat      180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca      240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac      300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt      360

```

agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgccc tagcgcccg ctcctttcgct ttctttccctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctcccttttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaacaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cgggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggct gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat ttgacaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920

tccagaaccc tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaac tggtagctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgctc ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttctgctcgc ctgctcgccg ggctcgcaaa	2760
cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gcccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgccgccg gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtgggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaaca tttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatagcg gtgtcgtcag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatggatga ggcggtcttt	3240
gagagctgag ttgtagtatg tggtttcact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac cccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccgggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3480

ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtagcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aaatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggccccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780
agtcactctgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccg	3900
gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcgttcac	3960
cgcccattct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc ccccttcacc gagggcctat ttcccatgat	4080
tccttcatat ttgcatatac gatacaaggc tgttagagag ataattggaa ttaatttgac	4140
tgtaaacaca aagatattag tacaaaatac gtgacgtaga aagtaataat ttcttgggta	4200
gtttgcagtt ttaaaattat gttttaaaat ggactatcat atgcttaccg taacttgaaa	4260
gtatttcgat ttcttggctt tatatatctt gtggaaagga cgaaacaccg gtgtaccgga	4320
ggtttgaaga tgccgcattt ctcgagaaat gcggcatctt caaacctttt tttgaattct	4380
cgacctaggg acaaattggca gtattcatcc acaatttttaa aagaaaaggg gggattgggg	4440
ggtacagtgc aggggaaaga atagtagaca taatagcaac agacatacaa actaaagaat	4500
tacaaaaaca aattacaaaa attcaaaatt ttcgggttta ttacagggac agcagagatc	4560
cactttggcc gcggatccgc aacaaatctg actttgcatg tgcaaacgcc ttcaacaaca	4620
gcattattcc agaagacacc ttcttcccca gcccaggtaa gggcagcttt ggtgccttcg	4680
caggctgttt ccttgcttca ggaatggcca ggttctgccc agagctctgg tcaatgatgt	4740
ctaaaactcc tctgattggt ggtctcggcc ttatccattg ccaccaaaac cctcttttta	4800
ctaagaaaca gtgagccttg ttctggcagt ccagagaatg acacgggaaa aaagcagatg	4860
aagagaaggt ggcaggagag ggcacgtggc ccagcctcag tctctccaac tgagttcctg	4920
cctgcctgcc tttgctcaga ctgtttgccc cttactgctc ttctaggcct cattctaagc	4980
cccttctcca agttgcctct ccttatttct ccctgtctgc caaaaaatct ttcccagctc	5040

actaagtcag tctcacgcag tcactcatta acccaccaat cactgattgt gccggcacat	5100
gaatgcacca ggtgttgaag tggaggaatt aaaaagtcag atgaggggtg tgcccagagg	5160
aagcaccatt ctagttgggg gagcccatct gtcagctggg aaaagtccaa ataacttcag	5220
attggaatgt gttttaactc agggttgaga aaacagccac cttcaggaca aaagtcaggg	5280
aagggctctc tgaagaaatg ctacttgaag ataccagccc taccaagggc agggagagga	5340
ccaattgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgcccg	5400
ggcaaagccc gggcgtcggg cgaccttttg tcgcccggcc tcagtgagcg agcgagcgcg	5460
cagagagggg gtggccaacg gcgcgcctgc aggtctcaaa aatagctacc ctctccggca	5520
tgaatttatc agctagaacg gttgaatatc atattgatgg tgatttgact gtctccggcc	5580
tttctcaccg gtttgaatct ttacctacac attactcagg cattgcattt aaaatatatg	5640
agggttctaa aaatTTTTat ctttgcgttg aaataaaggc ttctcccgca aaagtattac	5700
agggtcataa tgTTTTTggT acaaccgatt tagctttatg ctctgaggct ttattgctta	5760
atTTtgctaa ttctttgcct tgcctgtatg atttattgga tgTTTgaatt cctgatgcgg	5820
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatggtgcac tctcagtaca	5880
atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg	5940
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg	6000
agctgcatgt gtcagagggt ttaccgtca tcaccgaaac gcgcgagacg aaagggcctc	6060
gtgatacgcc tatttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt	6120
ggcacttttc ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca	6180
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg	6240
aagagtatga gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc	6300
cttcctgttt ttgctcaccg agaaacgctg gtgaaagtaa aagatgctga agatcagttg	6360
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt	6420
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggtg	6480
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat	6540
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga	6600

gaattatgca gtgctgcat aaccatgagt gataacactg cggccaactt acttctgaca	6660
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact	6720
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc	6780
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact	6840
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagtgc aggaccactt	6900
ctgcgctcgg cccttcggc tggctggttt attgctgata aatctggagc cggtagcgt	6960
gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt	7020
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata	7080
ggtgcctcac tgattaagca ttggtaactg tcagaccaag ttactcata tatactttag	7140
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat	7200
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa	7260
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca	7320
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt	7380
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agttagaccg	7440
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaadc	7500
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga	7560
cgatagttag cggataaggc gcagcggctg ggctgaacgg ggggttcgtg cacacagccc	7620
agcttgagac gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc	7680
gccacgcttc ccgaaggag aaaggcggac aggtatccgg taagcggcag ggtcgggaaca	7740
ggagagcgca cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg	7800
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta	7860
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ctttttgctg gccttttgct	7920
cacatgttct ttctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag	7980
tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa	8040
gcggaagagc gcccaatacg caaacgcct ctccccgcgc gttggccgat tcattaatg	8099

<210> 21
 <211> 8039
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 21
 cagcagctgg cgtaatagcg aagaggcccg caccgatcg ccttcccaac agttgcgag 60
 cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc 120
 gtttttctg ttgcaatggc tggcggtaat attgttctg atattaccag caaggccgat 180
 agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca 240
 acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac 300
 acttctcagg attctggcgt accgttctg tctaaaatcc ctttaatcgg cctcctgttt 360
 agtccccgct ctgattctaa cgaggaaagc acgttatagc tgctcgtcaa agcaaccata 420
 gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 480
 cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cttttctcgc 540
 cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 600
 tagtgcttta cggcacctcg acccaaaaaa acttgattag ggtgatgggt cacgtagtgg 660
 gccatcgccc tgatagacgg tttttcgccc ttgacgttg gaggccacgt tctttaatag 720
 tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 780
 ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt 840
 taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt 900
 cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt 960
 acggcgcgcc ggggttgcca ctccctctct gcgcgctcgc tcgctcactg aggccgggag 1020
 accaaaggct gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg 1080
 cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg 1140
 cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatgtt 1200
 tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca 1260
 ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag 1320

gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggtctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttattttgtg aaattttgtga tgctattgct	2220
ttattttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaac tggtagctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtgaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgctc ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggcttttc gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880

ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgcc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggctgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3060
cccgccgtag tagtagtgct tcgcgacagta gtaaatacg gtgtcgtcag tctggagaga	3120
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggtcttt	3180
gagagctgag ttgtagtatg tggtttcact tccccagata acccccagcc actctaagcc	3240
ctttctaggc ggctgtcgta tccatgatac ccgtagtctt gggagactga ctccagacac	3300
agtgcacgtg actgagaggc tctggctagg tgcgaccaag ccggtccgc tctcctgtaa	3360
cttgacctcc gagccgccac ccccgctgcc gccgccaccg ctgccaccgc ctccggtaat	3420
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3480
ggtggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3540
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggt aaatgagcaa	3600
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3660
ggccccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3720
agtcactctgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3780
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaattccg	3840
gctgcgacgg aactagctgt gccacaccg gcgcgtcctt atataatcat cggcgttcac	3900
cgcccattct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	3960
cgccctaact ccgcccacta gtgcggccgc ccccttcacc gagggcctat ttcccatgat	4020
tccttcatat ttgcatatac gatacaaggc tgttagagag ataattggaa ttaatttgac	4080
tgtaaacaca aagatattag tacaaaatac gtgacgtaga aagtaataat ttcttgggta	4140
gtttgcagtt ttaaaattat gttttaaaat ggactatcat atgcttaccg taacttgaaa	4200
gtatttcgat ttcttggctt tatatatctt gtggaaagga cgaaacaccg gtgtaccgga	4260
ggtttgaaga tgccgcattt ctcgagaaat gcggcatctt caaacctttt tttgaattct	4320
cgacctaggg acaaatggca gtattcatcc acaattttaa aagaaaaggg gggattgggg	4380
ggtacagtgc aggggaaaga atagtagaca taatagcaac agacatacaa actaaagaat	4440

tacaaaaaca aattacaaaa attcaaaatt ttcgggttta ttacaggac agcagagatc	4500
cactttggcc gcggatccgc aacaaatctg actttgcatg tgcaaacgcc ttcaacaaca	4560
gcattattcc agaagacacc ttcttcccca gcccaggtaa gggcagcttt ggtgccttcg	4620
caggctgttt ccttgcttca ggaatggcca ggttctgccc agagctctgg tcaatgatgt	4680
ctaaaactcc tctgattggg ggtctcggcc ttatccattg ccaccaaaac cctcttttta	4740
ctaagaaaca gtgagccttg ttctggcagt ccagagaatg acacgggaaa aaagcagatg	4800
aagagaaggt ggcaggagag ggcacgtggc ccagcctcag tctctccaac tgagtctctg	4860
cctgcctgcc tttgctcaga ctgtttgccc cttactgctc ttctaggcct cattctaagc	4920
cccttctcca agttgcctct ccttatttct ccctgtctgc caaaaaatct ttcccagctc	4980
actaagtcag tctcacgcag tcactcatta acccaccaat cactgattgt gccggcacat	5040
gaatgcacca ggtgttgaag tggaggaatt aaaaagtcag atgaggggtg tgcccagagg	5100
aagcaccatt ctagttgggg gagcccatct gtcagctggg aaaagtccaa ataacttcag	5160
attggaatgt gttttaactc agggttgaga aaacagccac cttcaggaca aaagtcaggg	5220
aagggtcttc tgaagaaatg ctacttgaag ataccagccc taccaagggc agggagagga	5280
ccaattgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgcccg	5340
ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg	5400
cagagaggga gtggccaacg gcgcgcctgc aggtctcaaa aatagctacc ctctccggca	5460
tgaatttatc agctagaacg gttgaatatc atattgatgg tgatttgact gtctccggcc	5520
tttctcaccg gtttgaatct ttacctacac attactcagg cattgcattt aaaatatatg	5580
agggttctaa aaatttttat ccttgcgttg aaataaaggc ttctcccgca aaagtattac	5640
agggtcataa tgtttttggg acaaccgatt tagctttatg ctctgaggct ttattgctta	5700
attttgctaa ttctttgcct tgctgtatg atttattgga tgttggaatt cctgatgcgg	5760
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatggtgcac tctcagtaca	5820
atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg	5880
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg	5940
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgagacg aaagggcctc	6000

gtgatacgcc tatttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt	6060
ggcacttttc ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca	6120
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg	6180
aagagtatga gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc	6240
cttcctgttt ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg	6300
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt	6360
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta	6420
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat	6480
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga	6540
gaattatgca gtgctgccat aaccatgagt gataaactg cggccaaactt acttctgaca	6600
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact	6660
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc	6720
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact	6780
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt	6840
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt	6900
gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt	6960
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata	7020
ggtgcctcac tgattaagca ttggtaactg tcagaccaag ttactcata tatactttag	7080
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat	7140
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa	7200
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca	7260
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt	7320
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg	7380
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaadc	7440
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga	7500
cgatagttac cggataaggc gcagcggctg ggctgaacgg ggggttcgtg cacacagccc	7560

agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc	7620
gccacgcttc ccgaaggagg aaaggcggac aggtatccgg taagcggcag ggtcgggaaca	7680
ggagagcgca cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg	7740
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta	7800
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct	7860
cacatgttct ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag	7920
tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa	7980
gcggaagagc gcccaatacg caaacgcct ctccccgcgc gttggccgat tcattaatg	8039

<210> 22

<211> 8098

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 22

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttcctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780

P109070023W000-SEQ-MJT.txt

ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaacaaaa tattaacggt tacaatttaa atatttgctt atacaatctt	900
cctgttttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggct gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tggtagttta cagtttatta aatagatgtt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat tttgcagggt aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagt	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaac tggtagctca	2340

tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaaag	2400
cccctggtat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcttgccgc ctgctcgccg ggctcgcaaa	2760
cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gcccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgccgccg gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtgggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaaca tttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatagcg gtgtcgtcag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggtcttt	3240
gagagctgag ttgtagtatg tggttttact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac cccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3480
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtaggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600
ggaacccgaa aatcggtag gaacgcctga gtgtaatctc gacgtgtggg aaatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggcccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780
agtcacttgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgacgac ccaaatccc	3900

gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcgttcac	3960
cgcccattct ccgcccagcc ataaaaggca acttttcggaa cggcgcacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc gcggccaaag tggatctctg ctgtccctgt	4080
aataaacccg aaaattttga atttttgtaa tttgtttttg taattcttta gtttgtatgt	4140
ctgttgctat tatgtctact attctttccc ctgcactgta ccccccaatc cccccttttc	4200
ttttaaaatt gtggatgaat actgccatth gtccttaggt cgagaattca aaaaaagggt	4260
tgaagatgcc gcatttctcg agaaatgcgg catcttcaaa cctccggtac accggtgttt	4320
cgtcctttcc acaagatata taaagccaag aaatcgaaat actttcaagt tacggtaagc	4380
atatgatagt ccattttaaa acataatttt aaaactgcaa actaccaag aaattattac	4440
tttctacgtc acgtattttg tactaatatc tttgtgttta cagtcaaatt aattccaatt	4500
atctctctaa cagccttgta tcgtatatgc aaatatgaag gaatcatggg aaataggccc	4560
tcggtgaagg gggatccgca acaaactctga ctttgcattg gcaaacgcct tcaacaacag	4620
cattattcca gaagacacct tcttcccag cccaggtaag ggcagctttg gtgccttcgc	4680
aggctgtttc cttgcttcag gaatggccag gttctgcccc gagctctggg caatgatgtc	4740
taaaactcct ctgattgggt gtctcggcct tatccattgc caccaaaacc ctctttttac	4800
taagaaacag tgagccttgt tctggcagtc cagagaatga cacgggaaaa aagcagatga	4860
agagaagggt gcaggagagg gcacgtggcc cagcctcagt ctctccaact gagttcctgc	4920
ctgcctgcct ttgctcagac tgtttgcccc ttactgctct tctaggcctc attctaagcc	4980
ccttctccaa gttgcctctc cttattttct cctgtctgcc aaaaaatctt tcccagctca	5040
ctaagtcagt ctcacgcagt cactcattaa cccaccaatc actgattgtg ccggcacatg	5100
aatgcaccag gtgttgaagt ggaggaatta aaaagtcaga tgaggggtgt gccagagga	5160
agcaccattc tagttggggg agcccatctg tcagctggga aaagtccaaa taacttcaga	5220
ttggaatgtg ttttaactca gggttgagaa aacagccacc ttcaggacaa aagtcaggga	5280
agggtctctt gaagaaatgc tacttgaaga taccagccct accaaggga gggagaggac	5340
caattgatgg agttggccac tccctctctg gcgctcgtc cgctcactga ggccgcccg	5400
gcaaagcccc ggcgtcgggc gacctttggt cggccggcct cagtgagcga gcgagcgcgc	5460

agagagggag tggccaacgg cgcgctgca ggtctcaaaa atagctaccc tctccggcat	5520
gaatttatca gctagaacgg ttgaatatca tattgatggg gatttgactg tctccggcct	5580
ttctcaccgg tttgaatctt tacctacaca ttactcaggc attgcattta aaatatatga	5640
gggttctaaa aattttttatc cttgcgttga aataaaggct tctcccgcaa aagtattaca	5700
gggtcataat gtttttggtg caaccgattt agctttatgc tctgaggctt tattgcttaa	5760
ttttgctaata tctttgcctt gcctgtatga tttattggat gttggaattc ctgatgcggg	5820
attttctcct tacgcatctg tgcgggtattt cacaccgcat atgggtgcact ctgagtacaa	5880
tctgctctga tgccgcatag ttaagccagc cccgacaccc gccaacaccc gctgacgcgc	5940
cctgacggggc ttgtctgctc ccggcatccg cttacagaca agctgtgacc gtctccggga	6000
gctgcatgtg tcagagggtt tcaccgtcat caccgaaacg cgcgagacga aagggcctcg	6060
tgatacgctt atttttatag gttaatgtca tgataataat ggtttcttag acgtcagggtg	6120
gcacttttcg gggaaatgtg cgcggaaccc ctatttggtt atttttctaa atacattcaa	6180
atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga	6240
agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg gcattttgcc	6300
ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg	6360
gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc	6420
gccccgaaga acgtttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat	6480
tatcccgtat tgacgccggg caagagcaac tcggctgccg catacactat tctcagaatg	6540
acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	6600
aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa	6660
cgatcggagg accgaaggag ctaaccgctt ttttgacaaa catgggggat catgtaactc	6720
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	6780
cgatgcctgt agcaatggca acaacgttgc gcaaaactatt aactggcgaa ctacttactc	6840
tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	6900
tgcgctcggc ctttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg	6960
ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctccgt atcgtagtta	7020

tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag	7080
gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga	7140
ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc	7200
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa	7260
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacia	7320
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc	7380
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt	7440
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaattc	7500
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	7560
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca	7620
gcttggagcg aacgacctac accgaactga gataacctaca gcgtgagcta tgagaaagcg	7680
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	7740
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt	7800
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat	7860
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc	7920
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt	7980
gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag	8040
cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatg	8098

<210> 23

<211> 8038

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 23

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgcag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180

agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatatg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgtt ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gaggccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cgggggtacat atgattgaca tgctagtttt	960
acggcgcgcc ggggttgcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg	1140
cagggtcagt ggctccaact aacatttggt tgggtacttta cagtttatta aatagatggt	1200
tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaagggt ggatgaggca	1260
ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgtcggg gcttagacgc aggtgttctg atttatagtt	1380
caaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cttttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttcactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtcct	1740

agtccatcac gagcagctgg tttctaagat gctattttccc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agagggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttattttgtg aaatttgtga tgctattgct	2220
ttattttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaag tggtagctca	2340
tcgaggtggt aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cggtaggaaag	2400
cccctgggat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcgttg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccatctccgg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760
cgggcttttc gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgcc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggctgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ctggatacgg tcacagaagt gccctgacct cagtagtcca ttgcgtaact	3060
cccgccgtag tagtagtgct tcgcgcagta gtaaatagcg gtgtcgtcag tctggagaga	3120
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatggatga ggcggctctt	3180
gagagctgag ttgtagtatg tggtttcact tccccagata acccccagcc actctaagcc	3240
ctttctaggc ggctgtcgta tccatgatac cccgtagtct gggagactga ctccagacac	3300

agtgcacgtg actgagaggc tctggctagg tgcgaccaag cccgggtccgc tctcctgtaa	3360
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3420
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3480
ggtagcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3540
ggaacccgaa aatcggctag gaacgcctga gtgtaatctc gacgtgtggg aaatgagcaa	3600
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3660
ggcccggcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3720
agtcactctg atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3780
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccg	3840
gctgcgacgg aactagctgt gccacaccg gcgcgtcctt atataatcat cggcggtcac	3900
cgccattctt ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	3960
cgccctaact ccgcccacta gtgcgggccgc gcggccaaag tggatctctg ctgtccctgt	4020
aataaacccg aaaattttga atttttgtaa tttgtttttg taattcttta gtttgtatgt	4080
ctgttgctat tatgtctact attctttccc ctgcactgta cccccaatc ccccttttc	4140
ttttaaaatt gtggatgaat actgccatth gtccctagggt cgagaattca aaaaaagggt	4200
tgaagatgcc gcatttctcg agaaatgcgg catcttcaaa cctccgggtac accgggtgtt	4260
cgtcctttcc acaagatata taaagccaag aaatcgaaat actttcaagt tacggtaagc	4320
atatgatagt ccattttaaa acataattht aaaactgcaa actaccaag aaattattac	4380
tttctacgtc acgtattttg tactaatatc tttgtgttta cagtcaaatt aattccaatt	4440
atctctctaa cagccttgta tcgtatatgc aaatatgaag gaatcatggg aaataggccc	4500
tcggtgaagg gggatccgca acaaacttga ctttgcatgt gcaaacgcct tcaacaacag	4560
cattattcca gaagacacct tcttccccag cccaggtaag ggcagctttg gtgccttcgc	4620
aggctgtttc cttgcttcag gaatggccag gttctgccc gagctctggg caatgatgtc	4680
taaaactcct ctgattgggt gtctcggcct tatccattgc caccaaaacc ctctttttac	4740
taagaaacag tgagccttgt tctggcagtc cagagaatga cacgggaaaa aagcagatga	4800
agagaagggt gcaggagagg gcacgtggcc cagcctcagt ctctccaact gagttcctgc	4860

ctgcctgcct ttgctcagac tgtttgcccc ttactgctct tctaggcctc attctaagcc	4920
ccttctccaa gttgcctctc cttatttctc cctgtctgcc aaaaaatctt tcccagctca	4980
ctaagtcagt ctcacgcagt cactcattaa cccaccaatc actgattgtg ccggcacatg	5040
aatgcaccag gtgttgaagt ggaggaatta aaaagtcaga tgaggggtgt gcccagagga	5100
agcaccattc tagttggggg agcccatctg tcagctggga aaagtccaaa taacttcaga	5160
ttggaatgtg ttttaactca gggttgagaa aacagccacc ttcaggacaa aagtcaggga	5220
agggctctct gaagaaatgc tacttgaaga taccagccct accaaggga gggagaggac	5280
caattgatgg agttggccac tccctctctg cgcgctcgct cgctcactga ggccgcccgg	5340
gcaaagcccg ggcgtcgggc gacctttggt cgcccggcct cagtgagcga gcgagcgcgc	5400
agagaggag tggccaacgg cgcgcctgca ggtctcaaaa atagctaccc tctccggcat	5460
gaatttatca gctagaacgg ttgaatatca tattgatggt gatttgactg tctccggcct	5520
ttctcaccg tttgaatctt tacctacaca ttactcaggc attgcattta aaatatatga	5580
gggttctaaa aatttttatc cttgcgttga aataaaggct tctcccga aagtattaca	5640
gggtcataat gtttttggt caaccgattt agctttatgc tctgaggctt tattgcttaa	5700
ttttgcta tctttgcctt gcctgtatga ttattggat gttggaattc ctgatgcggt	5760
attttctcct tacgcatctg tgcggtatct cacaccgcat atggtgcact ctcagtacaa	5820
tctgctctga tgccgcatag ttaagccagc cccgacaccc gccaacaccc gctgacgcgc	5880
cctgacgggc ttgtctgctc ccggcatccg cttacagaca agctgtgacc gtctccggga	5940
gctgcatgtg tcagaggttt tcaccgtcat caccgaaacg cgcgagacga aagggcctcg	6000
tgatacgcct atttttatag gttaatgtca tgataataat ggtttcttag acgtcaggtg	6060
gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa	6120
atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga	6180
agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg gcattttgcc	6240
ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg	6300
gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc	6360
gccccgaaga acgtttttcca atgatgagca cttttaaaagt tctgctatgt ggcgcggtat	6420

tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg	6480
acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	6540
aattatgcag tgctgccata accatgagtg ataactctgc ggccaactta cttctgacaa	6600
cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc	6660
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	6720
cgatgcctgt agcaatggca acaacgttgc gcaactatt aactggcgaa ctacttactc	6780
tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	6840
tgcgctcggc ctttcgggt ggctggttta ttgctgataa atctggagcc ggtgagcgtg	6900
ggtctcgcgg tatcattgca gcaactgggc cagatggtaa gccctcccgt atcgtagtta	6960
tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag	7020
gtgcctcact gattaagcat tggttaactgt cagaccaagt ttactcatat atactttaga	7080
ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc	7140
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa	7200
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacia	7260
aaaaaccacc gctaccagcg gtggtttggt tgccggatca agagctacca actctttttc	7320
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt	7380
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaattc	7440
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	7500
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca	7560
gcttggagcg aacgacctac accgaactga gatactaca gcgtgagcta tgagaaagcg	7620
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	7680
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt	7740
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat	7800
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ctttttgctc	7860
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt	7920
gagctgatac cgctcggcg agccgaacga ccgagcgcag cgagtcagt agcgaggaag	7980

cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatg 8038

<210> 24

<211> 8626

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 24

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag 60

cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc 120

gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat 180

agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca 240

acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaaac 300

acttctcagg attctggcgt accgttctctg tctaaaatcc ctttaatcgg cctcctgttt 360

agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata 420

gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 480

cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cctttctcgc 540

cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 600

tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg 660

gccatcgcgc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag 720

tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 780

ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt 840

taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt 900

cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt 960

acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccggggc 1020

accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg 1080

cagagaggga gtggccaact ccatcactag gggttcctac gcgtagatct catattctgg 1140

cagggtcagt ggctccaact aacatttggt tggtacttta cagtttatta aatagatggt 1200

tatatggaga agctctcatt tctttctcag aagagcctgg ctaggaaggt ggatgaggca	1260
ccatattcat tttgcaggtg aaattcctga gatgtaagga gctgctgtga cttgctcaag	1320
gccttatatc gagtaaacgg tagcgctggg gcttagacgc aggtgttctg atttatagtt	1380
caaaacctct atcaatgaga gagcaatctc ctggtaatgt gatagatttc ccaacttaat	1440
gccaacatac cataaacctc ccattctgct aatgcccagc ctaagttggg gagaccactc	1500
cagattccaa gatgtacagt ttgctttgct gggccttttt cccatgcctg cctttactct	1560
gccagagtta tattgctggg gttttgaaga agatcctatt aaataaaaaga ataagcagta	1620
ttattaagta gccctgcatt tcaggtttcc ttgagtggca ggccaggcct ggccgtgaac	1680
gttactgaa atcatggcct cttggccaag attgatagct tgtgcctgtc cctgagtccc	1740
agtccatcac gagcagctgg tttctaagat gctattttcc gtataaagca tgagaccgtg	1800
acttgccagc cccacagagc cccgcccttg tccatcactg gcatctggac tccagcctgg	1860
gttggggcaa agaggggaaat gagatcatgt cctaaccctg atcctcttgt cccacagata	1920
tccagaacct tgaccctgcc gtgtaccagc tgagagactc taaatccagt gacaagtctg	1980
tctgcctatt caccgatttt gattctcaaa caaatgtgtc acaaagtaag gattctgatg	2040
tgtatatcac agacaaaact gtgctagaca tgaggctctat ggacttcaag agcaacagtg	2100
ctgtggcctg gagcaactag tgatccagac atgataagat acattgatga gtttggacaa	2160
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct	2220
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt	2280
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaac tggtagctca	2340
tcgaggtggg aaggcttgca tgtgcagcgc gtcgtatgtg tccttggtcg cgggtggaag	2400
cccctgggat aagccgtcgt gtccctttcc gcgacgtctc tcgcccttca taccaatctc	2460
cgagtacgcc tcagccatct tgtccttctg gagttcggtg tacaggccct cctggggggt	2520
ctttctccga ggcttgccct ccattctccg gtccctacct ctacgcttgt ctaacacgtc	2580
gtactcctcc cggcgcccca ggtttaactc gttgtatagc tggttctgac cctgctggta	2640
agcgggtgcg tccgcggagc gagagaactt cactctcagt tcgcagctgc cttcttcttc	2700
gctcggcgcg cggcagctgc tcgcatcttc ttcctgcgcg ctgctcgccg ggctcgcaaa	2760

cgggctttta gccgccgccg ccgctttttt gcggctcgct ttgcagtaga gcgttattac	2820
caagctcaac aacaacactc cgcaagtccc ggctagtggg gcccagatgt agatgtcgca	2880
ggcgaagtct agtcctcgcg tgtgaacggc gccgcccgcc gcgggcctgc aagcctcagg	2940
gcgaagagat aagggtgac ttgctatcgt tggagctggc gtggggggac gtggggctgg	3000
ggtagtagtg ggcttcgctg gggatacatt tgtagaaca tttgaaaatg ttccctgggt	3060
aggtaactcg ctggatacgg tcacagaagt gccctgaccc cagtagtcca ttgcgtaact	3120
cccgccgtag tagtagtgct tcgcgcagta gtaaatacg gtgtcgtag tctggagaga	3180
gttcatcttt aagaaaacct gcgacttact gttgtcctta atgatgggta ggcggctctt	3240
gagagctgag ttgtagtatg tggtttact tccccagata acccccagcc actctaagcc	3300
ctttctaggc ggctgtcgta tccatgatac ccgtagtct gggagactga ctccagacac	3360
agtgcacgtg actgagaggc tctggctagg tgcgaccaag ccggtccgc tctcctgtaa	3420
cttgacctcc gagccgccac ccccgtgcc gccgccaccg ctgccaccgc ctccggtaat	3480
ctccagctta gttccccac cgaatgtgta cgggagcgtg ttgccctgct gacagaagta	3540
ggtaggcgatg tcctcctgct caaggttact gattgtcagg ctgtagtccg taccggaacc	3600
ggaacccgaa aatcggttag gaacgcctga gtgtaatctc gacgtgtggg aaatgagcaa	3660
cttaactgtt ccgtccggct tctgctgata ccagtttaag tattttgaga tgtcctggct	3720
ggcccgcaa gatattgtta ctcggtcgcc taaagaagca gatagggaag aggtagtctg	3780
agtcactgt atgtcaggac gagccgcgtg caataataac gccagaggta aaagtaaggc	3840
tgtcactggg agcgccatgg tggctttccg gaacaaacaa gaaccgcgac ccaaatcccg	3900
gctgcgacgg aactagctgt gccacacccg gcgcgtcctt atataatcat cggcgttcac	3960
cgccattct ccgcccagcc ataaaaggca actttcggaa cggcgcacgc tgattggctc	4020
cgccctaact ccgcccacta gtgcggccgc ccccttcacc gagggcctat ttcccatgat	4080
tccttcatat ttgcatatac gatacaaggc tgtagagag ataattggaa ttaatttgac	4140
tgtaaacaca aagatattag tacaaaatac gtgacgtaga aagtaataat ttcttgggta	4200
gtttgcagtt ttaaaattat gttttaaaat ggactatcat atgcttaccg taacttgaaa	4260
gtatttcgat ttcttggctt tatatatctt gtggaaagga cgaaacaccg gtgtaccgga	4320

ggtttgaaga tgccgcattt ctcgagaaat gcggcatctt caaacctttt tttgaattct	4380
cgacctaggg acaaattggca gtattcatcc acaattttta aagaaaagg gggattgggg	4440
ggtacagtgc aggggaaaga atagtagaca taatagcaac agacatacaa actaaagaat	4500
tacaaaaaca aattacaaaa attcaaaatt ttcgggttta ttacagggaac agcagagatc	4560
cactttggcc gcggatcccc cttcaccgag ggcctatttc ccatgattcc ttcataattg	4620
catatacgat acaaggctgt tagagagata attggaatta atttgactgt aaacacaaag	4680
atattagtac aaaatacgtg acgtagaaag taataatttc ttgggtagtt tgcagtttta	4740
aaattatgtt ttaaaatgga ctatcatatg cttaccgtaa cttgaaagta tttcgatttc	4800
ttggctttat atatcttggt gaaaggacga aacaccggtg taccggaggt ttgaagatgc	4860
cgcatcttc gagaaatgcg gcatcttcaa accttttttt gaattctcga cctagggaca	4920
aatggcagta ttcattccaca attttaaaag aaaagggggg attggggggg acagtgcagg	4980
ggaaagaata gtagacataa tagcaacaga catacaaact aaagaattac aaaaacaaat	5040
tacaaaaatt caaaattttc gggtttatta cagggaacagc agagatccac tttggccgcg	5100
gatccgcaac aaatctgact ttgcatgtgc aaacgccttc aacaacagca ttattccaga	5160
agacaccttc ttccccagcc caggtaaggg cagctttggt gccttcgcag gctgtttcct	5220
tgcttcagga atggccaggt tctgcccaga gctctggtca atgatgtcta aaactcctct	5280
gattggtggt ctcggcctta tccattgcca ccaaaaccct ctttttacta agaaacagtg	5340
agccttggtc tggcagtcca gagaatgaca cgggaaaaaa gcagatgaag agaaggtggc	5400
aggagagggc acgtggccca gcctcagtct ctccaactga gttcctgcct gcctgccttt	5460
gctcagactg tttgcccctt actgctcttc taggcctcat tctaagcccc ttctccaagt	5520
tgctctcct tatttctccc tgtctgcaa aaaatctttc ccagctcact aagtcagtct	5580
cacgcagtca ctcatcacc caccaatcac tgattgtgcc ggcacatgaa tgcaccaggt	5640
gttgaagtgg aggaattaaa aagtcagatg aggggtgtgc ccagagggaag caccattcta	5700
gttgggggag cccatctgtc agctgggaaa agtccaaata acttcagatt ggaatgtgtt	5760
ttaactcagg gttgagaaaa cagccacctt caggacaaaa gtcagggaag ggctctctga	5820
agaaatgcta cttgaagata ccagccctac caagggcagg gagaggacca attgatggag	5880

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgcccgggc aaagcccggg	5940
cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgag agagggagtg	6000
gccaacggcg cgcctgcagg tctcaaaaat agctaccctc tccggcatga atttatcagc	6060
tagaacggtt gaatatcata ttgatggtga tttgactgtc tccggccttt ctacccggtt	6120
tgaatcttta cctacacatt actcaggcat tgcatttaaa atatatgagg gttctaaaaa	6180
tttttatcct tgcgttgaaa taaaggcttc tcccgcaaaa gtattacagg gtcataatgt	6240
ttttggtaca accgatttag ctttatgctc tgaggcttta ttgcttaatt ttgctaattc	6300
tttgccctgc ctgtatgatt tattggatgt tggaattcct gatgcggtat tttctcctta	6360
cgcactctgtg cggtatttca caccgcatat ggtgcactct cagtacaatc tgctctgatg	6420
ccgcatagtt aagccagccc cgacaccgc caacaccgc tgacgcgccc tgacgggctt	6480
gtctgctccc ggcatccgct tacagacaag ctgtgaccgt ctccgggagc tgcattgtgc	6540
agaggttttc accgtcatca ccgaaacgcg cgagacgaaa gggcctcgtg atacgcctat	6600
ttttataggt taatgtcatg ataataatgg tttcttagac gtcagggtggc acttttcggg	6660
gaaatgtgcg cggaaccctt atttgtttat ttttctaaat acattcaaatt atgtatccgc	6720
tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta	6780
ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgccctt cctgtttttg	6840
ctcaccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg	6900
gttacatcga actggatctc aacagcggtg agatccttga gagttttcgc cccgaagaac	6960
gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtattg	7020
acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt	7080
actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg	7140
ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac	7200
cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt	7260
gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag	7320
caatggcaac aacgttgcg aaactattaa ctggcgaact acttactcta gcttcccggc	7380
aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc	7440

```

ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta 7500
tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg 7560
ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga 7620
ttaagcattg gtaactgtca gaccaagttt actcatatat acttttagatt gatttaaaac 7680
ttcatTTTTa atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaaa 7740
tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat 7800
cttcttgaga tcctTTTTTT ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc 7860
taccagcggg ggtttgtttg ccgatcaag agctaccaac tctttttccg aaggtaactg 7920
gcttcagcag agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc 7980
acttcaagaa ctctgtagca ccgcctacat acctcgtctt gctaatacctg ttaccagtgg 8040
ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg 8100
ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa 8160
cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg 8220
aaggagaaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga 8280
gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccacctct 8340
gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca 8400
gcaacgcggc ctttttacgg ttcttggcct tttgctggcc ttttgctcac atgttctttc 8460
ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg 8520
ctcgccgag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc 8580
caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatg 8626

```

<210> 25

<211> 6827

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 25

```

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac agttgcgcag 60

```

cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cctttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgccc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat ttttaaaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtgggaaa tgagatcatg	1140
tcctaaccct gatcctcttg tcccacagat atccagaacc ctgaccctgc cgtgtaccag	1200
ctgagagact ctaaatccag tgacaagtct gtctgcctat tcaccgattt tgatttctcaa	1260
acaaatgtgt cacaaagtaa ggattctgat gtgtatatca cagacaaaac tgtgctagac	1320
atgaggtcta tggacttcaa gagcaacagt gctgtggcct ggagcaacta gtgatccaga	1380
catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaatg	1440
ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa	1500
acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga	1560
ggttttttta agcaagtaaa ctggtacctc atcgaggtgg taaggcttgc atgtgcagcg	1620

cgtcgtatgt gtccttggtc gcggtggaaa gcccctggta taagccgtcg tgtccctttc	1680
cgcgacgtct ctcgcccttc ataccaatct ccgagtagcg ctcagccatc ttgtccttct	1740
ggagttcggt gtacaggccc tcctgggggt tctttctccg aggcttgcct cccatctccg	1800
ggtccctacc tctacgcttg tctaacacgt cgtactcctc ccggcgcccc aggtttaact	1860
cgttgatatag ctggttctga ccctgctggt aagcgggtgc gtccgcggag cgagagaact	1920
tcactctcag ttgcagctg ctttcttctt cgctcggcgc gcggcagctg ctcgcatctt	1980
cttcctgcgc gctgctgcc gggctcgaa acgggctttt agccgccgcc gccgcttttt	2040
tgcggctcgc tttgcagtag agcgttatta ccaagctcaa caacaacact ccgcaagtcc	2100
cggctagtgg tgcccagatg tagatgtcgc aggcaagtc tagtcctcgc gtgtgaacgg	2160
cgccgcccgc cgcgggcctg caagcctcag ggcgaagaga taagggtga cttgctatcg	2220
ttggagctgg cgtgggggga cgtggggctg gggtagtagt gggcttcgct ggggatacat	2280
ttgtagaaac atttgaaaat gttccctggg taggtaactc gctggatacg gtcacagaag	2340
tgccctgacc ccagtagtcc attgcgtaac tcccgccgta gtagtagtgc ttgcgcagt	2400
agtaaatagc ggtgtcgtca gtctggagag agttcatctt taagaaaacc tgcgacttac	2460
tgttgctctt aatgatggtg aggcggctct tgagagctga gttgtagtat gtggtttcac	2520
ttccccagat aacccccagc cactctaagc cttttctagg cggctgtcgt atccatgata	2580
ccccgtagtc tgggagactg actccagaca cagtgcacgt gactgagagg ctctggctag	2640
gtgcgaccaa gcccgggtccg ctctcctgta acttgacctc cgagccgcca cccccgtgc	2700
cgccgccacc gctgccaccg cctccggtaa tctccagctt agttccccca ccgaatgtgt	2760
acgggagcgt gttgccctgc tgacagaagt aggtggcgat gtcctcctgc tcaaggttac	2820
tgattgtcag gctgtagtcc gtaccggaac cggaacccga aaatcggcta ggaacgcctg	2880
agtgtaatct cgacgtgtgg taaatgagca acttaactgt tccgtccggc ttctgctgat	2940
accagtttaa gtattttgag atgtcctggc tggcccggca agatattgtt actcggtcgc	3000
ctaaagaagc agatagggaa gaggtagtct gagtcatctg tatgtcagga cgagccgcgt	3060
gcaataataa cgccagaggt aaaagtaagg ctgtcactgg gagcgccatg gtggctttcc	3120
ggaacaaaca agaaccgcga cccaaatccc ggctgcgacg gaactagctg tgccacaccc	3180

ggcgcgtcct tatataatca tcggcggttca ccgcccattc tccgcccagc cataaaaggc	3240
aactttcggga acggcgcacg ctgattggct ccgccctaac tccgcccact agtgcggccg	3300
cactgcaaac ccagggtgc cttggaaaag gcgcaacccg ggccccctcg agccggcgcc	3360
aaagtggatc tctgctgtcc ctgtaataaa ccgaaaatt ttgaattttt gtaatttggt	3420
tttgtaattc tttagtttgt atgtctgttg ctattatgtc tactattctt tcccctgcac	3480
tgtaccccc aatccccct tttcttttaa aattgtggat gaatactgcc atttgtctca	3540
agatctagaa ttcaaaaaa ggtttgaaga tgccgcattt ctcgagaaat gcggcatctt	3600
caaacctccg gtacctcgtc ctttccacaa gatataataa gccaagaaat cgaaatactt	3660
tcaagttacg gtaagcatat gatagtccat tttaaaacat aattttaaaa ctgcaaacta	3720
ccaagaaat tattactttc tacgtcacgt attttgtact aatatctttg tgtttacagt	3780
caaattaatt ctaattatct ctctaacagc cttgtatcgt atatgcaa atgaaggaat	3840
catgggaaat aggccctctc tgggtcccct ggatccgcaa caaatctgac tttgcatgtg	3900
caaacgcctt caacaacagc attattccag aagacacctt cttccccagc ccaggtaagg	3960
gcagctttgg tgccttcgca ggctgtttcc ttgcttcagg aatggccagg ttctgcccag	4020
agctctggtc aatgatgtct aaaactcctc tgattgggtg tctcggcccc aattgatgga	4080
gttggccact ccctctctgc gcgctcgctc gctcactgag gccgcccggg caaagcccgg	4140
gcgtcgggag acctttggtc gccgggcctc agtgagcgag cgagcgcgca gagagggagt	4200
ggccaacggc gcgcctgcag gtctcaaaaa tagctaccct ctccggcatg aatttatcag	4260
ctagaacggg tgaatatcat attgatggtg atttgactgt ctccggcctt tctcaccgt	4320
ttgaatcttt acctacacat tactcaggca ttgcatttaa aatatatgag ggttctaaaa	4380
atttttatcc ttgcgttgaa ataaaggctt ctcccga aa agtattacag ggtcataatg	4440
tttttggtac aaccgattta gctttatgct ctgaggcttt attgcttaat tttgctaatt	4500
ctttgccttg cctgtatgat ttattggatg ttggaattcc tgatgcggta ttttctcctt	4560
acgcatctgt gcggtatttc acaccgcata tggtgcactc tcagtacaat ctgctctgat	4620
gccgcatagt taagccagcc ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct	4680
tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt	4740

cagaggtttt caccgtcatc accgaaacgc gcgagacgaa agggcctcgt gatacgcccta	4800
tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcagggtgg cacttttcgg	4860
ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg	4920
ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt	4980
attcaacatt tccgtgtcgc cttattccc tttttgcgg cattttgcct tcctgttttt	5040
gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagt	5100
ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg cccgaagaa	5160
cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt	5220
gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag	5280
tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt	5340
gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac gatcggagga	5400
ccgaaggagc taaccgcttt ttgacacaac atgggggatc atgtaactcg ccttgatcgt	5460
tggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta	5520
gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg	5580
caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc	5640
cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt	5700
atcattgcag cactggggcc agatggtaag ccctcccgtg tcgtagttat ctacacgacg	5760
gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg	5820
attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat tgatttaaaa	5880
cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct catgaccaa	5940
atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga	6000
tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg	6060
ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc gaaggtaact	6120
ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta gttaggccac	6180
cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatcct gttaccagt	6240
gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg	6300

gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga 6360
 acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc 6420
 gaaggagaaa aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg 6480
 agggagcttc cagggggaaa cgcctgggtat ctttatagtc ctgtcgggtt tcgccacctc 6540
 tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc 6600
 agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt 6660
 cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg agctgatacc 6720
 gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc 6780
 ccaatacgca aaccgcctct ccccgcgcggt tggccgattc attaattg 6827

<210> 26

<211> 6762

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 26

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ctttcccaac agttgcgag 60
 cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc 120
 gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat 180
 agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca 240
 acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaac 300
 acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt 360
 agtccccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata 420
 gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 480
 cgctacactt gccagcgcgc tagcgccgcg tcctttcgct ttcttccctt cttttctcgc 540
 cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 600
 tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg 660
 gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag 720

tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cgggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgcgcc cgggctttgc ccgggcggcc tcagtgcgcg agcgcgcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtgggaaa tgagatcatg	1140
tcctaaccct gatcctcttg tcccacagat atccagaacc ctgaccctgc cgtgtaccag	1200
ctgagagact ctaaattccag tgacaagtct gtctgcctat tcaccgattt tgattctcaa	1260
acaaatgtgt cacaaagtaa ggattctgat gtgtatatca cagacaaaac tgtgctagac	1320
atgaggctta tggacttcaa gagcaacagt gctgtggcct ggagcaacta gtgatccaga	1380
catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaatg	1440
ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa	1500
acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga	1560
ggttttttta agcaagtaaa ctggtacctc atcgagggtg taaggcttgc atgtgcagcg	1620
cgctcgtatg gtccttggtc gcggtggaaa gcccctggta taagccgtcg tgtccctttc	1680
cgcgacgtct ctgcacctc ataccaatct ccgagtacgc ctgagccatc ttgtccttct	1740
ggagtctgtt gtacaggccc tcctgggggt tctttctccg aggcttgcct cccatctccg	1800
ggcccctacc tctacgcttg tctaacacgt cgtactcctc ccggcgcccc aggtttaact	1860
cgttgatatg ctggttctga ccctgctggt aagcgggtgc gtccgcggag cgagagaact	1920
tcactctcag ttgcagctg ctttcttctt cgctcggcgc gcggcagctg ctcgcatctt	1980
cttcctgcgc gctgctgcc gggctcgcaa acgggctttt agccgccgcc gccgcttttt	2040
tgcggctcgc tttgcagtag agcgttatta ccaagctcaa caacaacact ccgcaagtcc	2100
cggctagtgg tgcccagatg tagatgtcgc aggcgaaagtc tagtcctcgc gtgtgaacgg	2160
cgccgccgc gcggggcctg caagcctcag ggcgaagaga taagggctga cttgctatcg	2220
ttggagctgg cgtgggggga cgtggggctg gggtagtagt gggcttcgct ggggatacat	2280

ttgtagaaac atttgaaaat gttccctggg taggtaactc gctggatacg gtcacagaag	2340
tgccctgacc ccagtagtcc attgcgtaac tcccgccgta gtagtagtgc ttcgcgcagt	2400
agtaaatagc ggtgtcgtca gtctggagag agttcatctt taagaaaacc tgcgacttac	2460
tgttgtcctt aatgatggtg aggcggctct tgagagctga gttgtagtat gtggtttcac	2520
ttccccagat aacccccagc cactctaagc cttttctagg cggctgtcgt atccatgata	2580
ccccgtagtc tgggagactg actccagaca cagtgcacgt gactgagagg ctctggctag	2640
gtgcgaccaa gcccgggtccg ctctcctgta acttgacctc cgagccgcca cccccgctgc	2700
cgccgccacc gctgccaccg cctccggtaa tctccagctt agttccccca ccgaatgtgt	2760
acgggagcgt gttgccctgc tgacagaagt aggtggcgat gtcctcctgc tcaaggttac	2820
tgattgtcag gctgtagtcc gtaccggaac cggaaccgga aaatcggcta ggaacgcctg	2880
agtgtaatct cgacgtgtgg taaatgagca acttaactgt tccgtccggc ttctgctgat	2940
accagtttaa gtattttgag atgtcctggc tggcccgga agatattgtt actcggtcgc	3000
ctaaagaagc agataggga gaggtagtct gagtcatctg tatgtcagga cgagccgcgt	3060
gcaataataa cgccagaggt aaaagtaagg ctgtcactgg gagcgccatg gtggctttcc	3120
ggaacaaaca agaaccgcga cccaaatccc ggctgcgacg gaactagctg tgccacaccc	3180
ggcgcgtcct tatataatca tcggcggttca ccgcccattc tccgcccagc cataaaaggc	3240
aactttcgga acggcgcacg ctgattggct ccgccctaac tccgcccact agtgcggccg	3300
cgggcccgtt taaacgctag cgagagggcc tatttcccat gattccttca tatttgcata	3360
tacgatacaa ggctgttaga gagataatta gaattaattt gactgtaaac acaaagatat	3420
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaaat	3480
tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg	3540
ctttatatat cttgtggaaa ggacgaggta ccggagggtt gaagatgccg catttctcga	3600
gaaatgcggc atcttcaaac ctttttttga attctagatc ttgagacaaa tggcagtatt	3660
catccacaat tttaaaagaa aaggggggat tggggggtac agtgcagggg aaagaatagt	3720
agacataata gcaacagaca tacaactaa agaattacaa aaacaaatta caaaaattca	3780
aaattttcct aggtacgtat ctagtggatc cgcaacaaat ctgactttgc atgtgcaaac	3840

gccttcaaca acagcattat tccagaagac accttcttcc ccagcccagg taagggcagc	3900
tttgggtgcct tcgcaggctg tttccttgct tcaggaatgg ccaggttctg cccagagctc	3960
tggatcaatga tgtctaaaac tcctctgatt ggtgggtctcg gcccgaattg atggagttgg	4020
ccactccctc tctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	4080
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	4140
acggcgcgcc tgcaggtctc aaaaatagct accctctccg gcatgaattt atcagctaga	4200
acggttgaat atcatattga tggatgattg actgtctccg gcctttctca cccgtttgaa	4260
tctttaccta cacattactc aggcattgca tttaaaatat atgagggttc taaaaatttt	4320
tatccttgcg ttgaaataaa ggcttctccc gcaaaagtat tacagggtca taatgttttt	4380
ggtacaaccg atttagcttt atgctctgag gctttattgc ttaattttgc taattctttg	4440
ccttgctgt atgatttatt ggatgttgga attcctgatg cggtattttc tccttacgca	4500
tctgtgcggt atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc	4560
atagttaagc cagccccgac acccgccaac acccgctgac gcgcccctgac gggcttgtct	4620
gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag	4680
gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac gcctattttt	4740
ataggttaat gtcataataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa	4800
tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaataatgt atccgctcat	4860
gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca	4920
acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg tttttgctca	4980
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta	5040
catcgaactg gatctcaaca gcggtgaagat ccttgagagt tttcgccccg aagaacgttt	5100
tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc	5160
cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc	5220
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtctgc	5280
cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa	5340
ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga	5400

P109070023W000-SEQ-MJT.txt

accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat	5460
ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca	5520
attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc	5580
ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat	5640
tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag	5700
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa	5760
gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca	5820
ttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaaatccc	5880
ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc	5940
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc	6000
agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt	6060
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt	6120
caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc	6180
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa	6240
ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg agcgaacgac	6300
ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg	6360
gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgagggga	6420
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact	6480
tgagcgtcga tttttgtgat gctcgtcagg ggggaggagc ctatggaaaa acgccagcaa	6540
cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc	6600
gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgaactg ataccgctcg	6660
ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcccaat	6720
acgcaaaccg cctctccccg cgcgttggcc gattcattaa tg	6762

<210> 27

<211> 7227

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic polynucleotide

<400> 27

cagcagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgag	60
cctgaatggc gaatggaatt ccagacgatt gagcgtcaaa atgtaggtat ttccatgagc	120
gtttttcctg ttgcaatggc tggcggtaat attgttctgg atattaccag caaggccgat	180
agtttgagtt cttctactca ggcaagtgat gttattacta atcaaagaag tattgcgaca	240
acggttaatt tgcgtgatgg acagactctt ttactcgggtg gcctcactga ttataaaaaac	300
acttctcagg attctggcgt accgttcctg tctaaaatcc ctttaatcgg cctcctgttt	360
agctcccgct ctgattctaa cgaggaaagc acgttatacg tgctcgtcaa agcaaccata	420
gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac	480
cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cttttctcgc	540
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt	600
tagtgcttta cggcacctcg acccaaaaaa acttgattag ggtgatgggt cacgtagtgg	660
gccatcgcgc tgatagacgg tttttcgccc ttgacgttg gagtccacgt tctttaatag	720
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt	780
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt	840
taacgcgaat tttaacaaaa tattaacgtt tacaatttaa atatttgctt atacaatctt	900
cctgtttttg gggcttttct gattatcaac cggggtacat atgattgaca tgctagtttt	960
acggcgcgcc gggttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	1020
accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	1080
cagagaggga gtggccaact ccatcactag gggttcctac gcgtgggaaa tgagatcatg	1140
tcctaaccct gatcctcttg tcccacagat atccagaacc ctgaccctgc cgtgtaccag	1200
ctgagagact ctaaattccag tgacaagtct gtctgcctat tcaccgattt tgattctcaa	1260
acaaatgtgt cacaaagtaa ggattctgat gtgtatatca cagacaaaac tgtgctagac	1320
atgaggctta tggacttcaa gagcaacagt gctgtggcct ggagcaacta gtgatccaga	1380
catgataaga tacattgatg agtttgaca aaccacaact agaatgcagt gaaaaaaatg	1440

ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa	1500
acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga	1560
ggttttttta agcaagtaaa ctggtacctc atcgagggtg taaggcttgc atgtgcagcg	1620
cgtcgtatgt gtccttggtc gcggtggaaa gcccctggta taagccgtcg tgtccctttc	1680
cgcgacgtct ctcgcccttc ataccaatct ccgagtagcg ctcagccatc ttgtccttct	1740
ggagttcgtt gtacaggccc tcctgggggt tctttctccg aggcttgcct cccatctccg	1800
ggtccttacc tctacgttg tctaacacgt cgtactcctc ccggcgcccc aggtttaact	1860
cgttgatatag ctggttctga ccctgctggt aagcgggtgc gtccgcggag cgagagaact	1920
tcactctcag ttgcagctg ctttcttctt cgctcggcgc gcggcagctg ctcgcatctt	1980
cttcctgcgc gctgctcgcc gggctcgcaa acgggctttt agccgccgcc gccgcttttt	2040
tgcggtcgc tttgcagtag agcgttatta ccaagctcaa caacaacact ccgcaagtcc	2100
cggctagtgg tgcccagatg tagatgtcgc aggcgaagtc tagtcctcgc gtgtgaacgg	2160
cgccgcccgc cgcgggcctg caagcctcag ggcgaagaga taagggtga cttgctatcg	2220
ttggagctgg cgtgggggga cgtggggctg gggtagtagt gggcttcgct ggggatacat	2280
tttagaaaac atttgaaaat gttccctggg taggtaactc gctggatacg gtcacagaag	2340
tgccctgacc ccagtagtcc attgcgtaac tcccgccgta gtagtagtgc ttcgcgcagt	2400
agtaaatagc ggtgtcgtca gtctggagag agttcatctt taagaaaacc tgcgacttac	2460
tgttgtcctt aatgatggtg aggcggctct tgagagctga gttgtagtat gtggtttcac	2520
ttccccagat aacccccagc cactctaagc cttttctagg cggctgtcgt atccatgata	2580
ccccgtagtc tgggagactg actccagaca cagtgcacgt gactgagagg ctctggctag	2640
gtgcgaccaa gcccgggtccg ctctcctgta acttgacctc cgagccgcca cccccgctgc	2700
cgccgccacc gctgccaccg cctccggtaa tctccagctt agttccccca ccgaatgtgt	2760
acgggagcgt gttgccctgc tgacagaagt aggtggcgat gtcctcctgc tcaaggttac	2820
tgattgtcag gctgtagtcc gtaccggaac cggaaccgga aaatcggcta ggaacgcctg	2880
agtgtaatct cgacgtgtgg taaatgagca acttaactgt tccgtccggc ttctgctgat	2940
accagtttaa gtattttgag atgtcctggc tggcccgga agatattgtt actcggtcgc	3000

ctaaagaagc agatagggaa gaggtagtct gagtcatctg tatgtcagga cgagccgcgt	3060
gcaataataa cgccagaggt aaaagtaagg ctgtcactgg gagcgccatg gtggctttcc	3120
ggaacaaaca agaaccgcga cccaaatccc ggctgcgacg gaactagctg tgccacaccc	3180
ggcgcgtcct tatataatca tcggcgttca ccgcccattc tccgcccagc cataaaaaggc	3240
aactttcggga acggcgcacg ctgattggct ccgccctaac tccgcccact agtgcggccc	3300
gtttaaacgc tagcgagagg gcctatttcc catgattcct tcatatttgc atatacgata	3360
caaggctgtt agagagataa ttagaattaa tttgactgta aacacaaaga tattagtaca	3420
aaatacgtga cgtagaaagt aataatttct tgggtagttt gcagttttaa aattatgttt	3480
taaaatggac tatcatatgc ttaccgtaac ttgaaagtat ttcgatttct tggctttata	3540
tatcttgtgg aaaggacgag gtaccggagg tttgaagatg ccgcatttct cgagaaatgc	3600
ggcatcttca aacctttttt tgaattctag atcttgagac aaatggcagt attcatccac	3660
aattttaaaa gaaaaggggg gattgggggg tacagtgcag gggaaagaat agtagacata	3720
atagcaacag acatacaaac taaagaatta caaaaacaaa ttacaaaaat tcaaaatttt	3780
cctagcgaga gggcctattt cccatgattc cttcatattt gcatatacga tacaaggctg	3840
ttagagagat aattagaatt aatttgactg taaacacaaa gatattagta caaaatacgt	3900
gacgtagaaa gtaataattt cttgggtagt ttgcagtttt aaaattatgt tttaaaatgg	3960
actatcatat gcttaccgta acttgaaagt atttcgattt cttggcttta tataatctgt	4020
ggaaaggacg aggtaccgga ggtttgaaga tgccgcattt ctcgagaaat gcggcatctt	4080
caaacctttt tttgaattct agatcttgag acaaatggca gtattcatcc acaattttaa	4140
aagaaaaggg gggattgggg ggtacagtgc aggggaaaga atagtagaca taatagcaac	4200
agacatacaa actaaagaat tacaaaaaca aattacaaaa attcaaaatt ttcctaggta	4260
cgtatctagt ggatccgcaa caaatctgac tttgcatgtg caaacgcctt caacaacagc	4320
attattccag aagacacctt cttccccagc ccaggtaagg gcagctttgg tgccttcgca	4380
ggctgtttcc ttgcttcagg aatggccagg ttctgcccag agctctggtc aatgatgtct	4440
aaaactcctc tgattggtgg tctcgcccc aattgatgga gttggccact ccctctctgc	4500
gcgctcgctc gtcactgag gccgcccggg caaagcccgg gcgtcgggcg acctttggtc	4560

gccccggcctc agtgagcgag cgagcgcgca gagagggagt ggccaacggc gcgcctgcag	4620
gtctcaaaaa tagctaccct ctccggcatg aatttatcag ctagaacggg tgaatatcat	4680
attgatgggtg atttgactgt ctccggcctt tctcaccgt ttgaatcttt acctacacat	4740
tactcaggca ttgcatttaa aatatatgag ggttctaaaa atttttatcc ttgcgttgaa	4800
ataaaggctt ctcccgaaa agtattacag ggtcataatg tttttggtac aaccgattta	4860
gctttatgct ctgaggcttt attgcttaat ttgctaatt ctttgccttg cctgtatgat	4920
ttattggatg ttggaattcc tgatgcggtt ttttctcctt acgcatctgt gcggtatttc	4980
acaccgcata tgggtgactc tcagtacaat ctgctctgat gccgcatagt taagccagcc	5040
ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc	5100
ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc	5160
accgaaacgc gcgagacgaa agggcctcgt gatacgcta tttttatagg ttaatgtcat	5220
gataataatg gtttcttaga cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc	5280
tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg	5340
ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc	5400
ccttattccc ttttttgcgg cattttgcct tcctgttttt gctcacccag aaacgctggg	5460
gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg gggtacatcg aactggatct	5520
caacagcggg aagatccttg agagttttcg cccgaagaa cgttttccaa tgatgagcac	5580
ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc aagagcaact	5640
cggtcgccgc atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa	5700
gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga	5760
taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt	5820
tttgacaac atgggggatc atgtaactcg cttgatcgt tgggaaccgg agctgaatga	5880
agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg	5940
caactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat	6000
ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat	6060
tgctgataaa tctggagccg gtgagcgtgg gtctcgcggg atcattgcag cactggggcc	6120

agatggtaag ccctcccgtg tcgtagttat ctacacgacg gggagtcagg caactatgga	6180
tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc	6240
agaccaagtt tactcatata tacttttagat tgatttataaa cttcattttt aatttataaag	6300
gatctaggtg aagatccttt ttgataatct catgacaaaa atcccttaac gtgagttttc	6360
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt	6420
tctgcgcgta atctgctgct tgcaaaaaaaa aaaaccaccg ctaccagcgg tggtttgttt	6480
gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat	6540
accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc	6600
accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa	6660
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg	6720
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag	6780
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa aggcggacag	6840
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa	6900
cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt	6960
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg	7020
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc	7080
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac	7140
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct	7200
ccccgcgcgt tggccgattc attaatag	7227