



- (51) **International Patent Classification:**
C12N 15/63 (2006.01) *C12N 15/86* (2006.01)
- (21) **International Application Number:**
PCT/EP2016/067883
- (22) **International Filing Date:**
27 July 2016 (27.07.2016)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
1513242.6 28 July 2015 (28.07.2015) GB
- (71) **Applicant:** GLAXOSMITHKLINE INTELLECTUAL
PROPERTY DEVELOPMENT LIMITED [GB/GB];
980 Great West Road, Brentford Middlesex TW8 9GS
(GB).
- (72) **Inventors:** BHINDER, Tejinder, Kaur; GlaxoSmithK-
line, Gunnels Wood Road, Stevenage Hertfordshire SG1
2NY (GB). LEWIS, Alan, Peter; GlaxoSmithKline, Gun-
nels Wood Road, Stevenage Hertfordshire SG1 2NY (GB).
PALLANT, Celeste; GlaxoSmithKline, Gunnels Wood
Road, Stevenage Hertfordshire SG1 2NY (GB). VINK,
Conrad; GlaxoSmithKline, Gunnels Wood Road, Steven-
age Hertfordshire SG1 2NY (GB).
- (74) **Agent:** O'FARRELL, Damien, John; GlaxoSmithKline,
Global Patents (CN925.1), 980 Great West Road, Brent-
ford Middlesex TW8 9GS (GB).

- (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, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
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).

Published:

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

(54) **Title:** CODON OPTIMISED TET REPRESSOR PROTEINS

(57) **Abstract:** The invention provides a codon optimised tetracycline repressor (TetR) sequence comprising the nucleic acid se-
quence of SEQ ID NO: 4 or SEQ ID NO: 5. The invention also provides a cell comprising the vector and a method of transfecting a
cell with the vector.



Codon Optimised Tet Repressor Proteins

TECHNICAL FIELD

5 The present invention relates to improved tet repressor proteins, nucleotide sequences encoding such repressor proteins, and vectors containing such sequences, for use in inducible gene expression systems in human cells.

BACKGROUND

10 The antibiotic effect of tetracycline antibiotics is due to their ability to inhibit protein synthesis in bacteria, mainly by binding to the mRNA translation complex. However, bacterial resistance to such antibiotics has evolved. Tetracycline resistance genes may be located on mobile genetic elements such as plasmids. The expression of certain tet resistance genes is controlled by transcriptional regulators known as tetracycline Repressors
15 (Tet repressors or TetR proteins). In the absence of a tetracycline antibiotic, the TetR protein forms a homodimer that binds with high affinity to DNA sequences (*tet* operators; *tetO*) in the promoter region of the resistance gene and represses transcription of the gene. However, if present, the antibiotic binds to TetR and reduces the affinity of TetR to the *tetO* binding site; the TetR:tetracycline complex dissociates from the *TetO* and allows
20 transcription of the bacteria's resistance gene.

 The tetracycline (*tet*) resistance system from *Escherischia coli* has been utilized to provide inducible gene expression constructs suitable for use in eukaryotic cells (see, e.g., Gossen and Bujard, PNAS USA 89(12):5547-51 (1992)). In one such system, a target gene of interest is operably linked to a tet-inducible promoter. Binding of tetracycline to tetR
25 homodimers results in derepression of the promoter (Yao *et al.*, Hum. Gen. Ther. 9, 1939-1950 (1998)).

 Improvements in tet-inducible systems that increase the level of expression of the tetR protein in the eukaryotic cell, or that increase the level of induced expression of a target gene of interest, or that decrease the basal expression of the target gene of interest,
30 are desirable.

SUMMARY OF THE INVENTION

One aspect of the present invention are codon optimised nucleotide sequences encoding a tetracycline repressor protein, and vectors comprising such sequences.

5 A further aspect of the present invention is host cells comprising an expression vector comprising a codon optimised nucleotide sequence encoding a tetR protein.

A further aspect of the present invention is a method of transfecting a cell with an expression vector comprising a codon optimised nucleotide sequence encoding a tetR protein.

0 A further aspect of the present invention is a method of producing a lentiviral vector packaging cell comprising transfecting a host cell with (a) an expression vector comprising a codon optimised nucleotide sequence encoding a tetR protein, and (b) expression vectors comprising a lentiviral GagPol sequence, a lentiviral Rev sequence, and a VSVg sequence, where one or more of said lentiviral sequences is tet-inducible.

5 A further aspect of the present invention is a method of producing a lentiviral vector, the method comprising (1) transfecting a host cell with (a) an expression vector comprising a codon optimised nucleotide sequence encoding a tetR protein, and (b) expression vectors comprising a lentiviral GagPol sequence, a lentiviral Rev sequence, and a VSVg sequence, where one or more of said lentiviral sequences is tet-inducible; (2) culturing said transfected
0 host cell, and (3) inducing expression of the tet-inducible gene(s) under conditions suitable for the production of lentiviral vectors.

BRIEF DESCRIPTION OF THE FIGURES

5 **Figure 1** provides a generalized schematic of lentiviral vector production and host cell transduction. A packaging host cell is transfected with plasmid expression vectors encoding lentiviral packaging genes and a transgene of interest (Green Fluorescent Protein or GFP in this schematic). Viral protein expression results in the assembly of proteins at the host membrane, and generation of a viral vector.

Figure 2A is a graph of TetR protein expression in cells transfected with different vector constructs containing *TetR*. Cells were transfected with 1.25ug of plasmid construct per 10^6 cells. Untransfected cells were used as controls.

Figure 2B is a graph of TetR protein expression in cells transfected with one of four different vector constructs containing *TetR*. Cells were transfected with 2.5ug of plasmid construct per 10^6 cells. Untransfected cells were used as controls.

Figure 3 illustrates viral titre produced in cells transfected with BACmod-WT lenti packaging constructs and *TetR* vectors containing either WT (SEQ ID NO:1), hCO (SEQ ID NO:3) or GSKCO (SEQ ID NO:4) *TetR*. Cells that did not contain the *TetR* vector (No *TetR*) and untransfected cells (UT) were used as controls. The Y-axis is viral titre produced (TU/mL).

Figure 4 graphs viral titer averages from Figure 3 on a linear scale.

Figure 5A – 5D provides schematics of the pcDNA6/TR plasmid (Figure 5A), pMA-BACmod-WT-TR-Ir-ZeoR plasmid (Figure 5B), pMA-BACmod-humTR-Ir-ZeoR plasmid (Figure 5C), and pMA-BACmod-GSKTR-Ir-ZeoR plasmid (Figure 5D).

DETAILED DESCRIPTION

Tet-inducible gene expression systems. The tetracycline (tet) resistance system from *Escherichia coli* has been utilized to provide inducible gene expression constructs suitable for use in eukaryotic cells (see, e.g., Gossen and Bujard, PNAS USA 89(12):5547-51 (1992)). The TetR protein is highly specific for its tetO DNA binding sites, and can recognize such binding sequences even in more complex eukaryotic genomes.

In one such system, a host cell is designed to stably express the Tet repressor, such as by containing a plasmid encoding the TetR protein. The host cell also contains plasmids or other genetic vectors that carry a gene of interest under the control of a promoter, where the promoter contains one or more tetracycline operator (tetO) sites. Due to the presence of TetR protein, expression of the gene of interest is repressed in the absence of tetracycline (though a basal level of transcription may occur) and induced in the presence of tetracycline. Additionally, the system may comprise a control expression plasmid, such as one that expresses B-galactosidase upon induction with tetracycline.

One such tetracycline-regulated expression system for mammalian cells is the T-Rex(TM) system (Invitrogen by Life Technologies), which utilizes a regulatory vector

(pcDNATM6/TR) which encodes a tet repressor under the control of the human CMV promoter.

As used herein, “induced” expression means the expression of an identified gene of interest in response to an inducer, where the induced level of expression is greater than the basal (un-induced) level of expression of that gene. Various inducers are known in the art; the appropriate inducer depends on the genetic expression system, as would be known by one of skill in the art. In tet-inducible systems, the inducer is a tetracycline antibiotic.

Tetracyclines are a group of antibiotics having a common molecular structure containing four hydrocarbon rings, and include naturally-occurring compounds as well as synthetic and semi-synthetic compounds. As used herein, “tetracyclines” or “tetracycline compound” comprises, but is not limited to: tetracycline, oxytetracycline, doxycycline, minocycline, chlortetracycline, tigecycline.

Codon usage and Codon Optimization: Codon bias refers to differences in the frequency of occurrence of synonymous nucleotide triplets in the coding DNA. The degeneracy found in the genetic code allows each amino acid to be encoded by between one and six synonymous codons allowing many alternative nucleic acid sequences to encode the same protein (Gustafsson, Govindarajan, and Minshull, Trends Biotechnol. 22(7): 346-53 (2004)). The frequencies with which different codons are used within genes is referred to as codon usage and can vary significantly between two genera or species of organisms, or among multiple genera or species of organisms. Different species often show particular preferences for one of the several codons that encode the same amino acid.

Codon optimisation is a technique used in to modify genetic sequences with the intent of increasing the rate of expression of a gene in a heterologous expression system; typically the nucleotide sequence encoding a protein of interest is codon optimized such that the codon usage more closely resembles the codon bias of the host cell, while still coding for the same amino acid sequence. In the context of viral vector production in a eukaryotic host cell (packaging cell), it is therefore of interest to develop a highly efficient vector production process in order to achieve a suitable viral titre.

Codon optimisation tailored to an expression system increases the speed and efficiency of translation. However, it has been reported that increasing the rate of translation by codon optimisation can lead to misfolding of the resultant protein, negatively

impacting its stability or substrate specificity (Zhou et al, Nature 495(7439): 111–115 (2013); Kimchi-Sarfaty, Science 315(5811):525-8 (2007)).

In host cells carrying heterologous genes, codon optimization of the heterologous gene sequence to suit the host cell (e.g., to correspond to the codon usage bias of the host cell's species or genus) may increase the rate of gene expression. In the context of tet-inducible systems, vectors providing efficient expression of functional Tet repressor proteins in eukaryotic host cells (including human cells) that contain tet-inducible genes of interest, are desirable.

Viral Vectors/ packaging cell lines. Viral vectors designed to modify a subject's stem cells by the incorporation of a therapeutic gene are based on members of the retrovirus family due to the ability of retrovirus to integrate their genetic payload into the host's genome (provirus). Retroviral vectors are designed to keep the essential proteins required for packaging and delivery of the viral genome, but any non-essential accessory proteins including those responsible for the viral disease profile are removed. Retroviral vectors are thus Retroviruses modified to carry a therapeutic gene for delivery into target cells. Highly efficient vector production processes are desirable in order to achieve a suitable viral titre.

The use of lentiviral vectors in gene therapy of diseases like Diabetes mellitus or Murine haemophilia A has been described. The aim of gene therapy is to modify the genetic material of living cells for therapeutic purposes, and it involves the insertion of a functional gene into a cell to achieve a therapeutic effect. Lentiviral vectors produced according to the present invention can be used to transfect target cells and induce the expression of the gene of potential therapeutic interest.

Lentiviral vectors, such as those based upon Human Immunodeficiency Virus Type 1 (HIV-1) are widely used as they are able to integrate into non-proliferating cells. Both coding and cis-acting sequences are required for viral vector construction, however, these sequences are separated to prevent the generation of replication-competent retroviruses. Once the viral genome has been reverse transcribed from the transfer vector into double-stranded DNA (dsDNA) and has entered the nucleus of the cell, this nucleic acid is integrated into the host genome aided by the viral enzyme, integrase.

The lentiviral vector has been made replication defective by splitting the genome into separate parts (Dull et al. 8463-71), which can then be individually expressed by plasmid

vectors (transient transfection), or expressed in specifically engineered packaging cell lines. Producer cell lines can be generated by transfecting a cell line capable of packaging viral vectors, such as the 293T cell, with the lentiviral packaging genes on individual plasmids which also carry unique eukaryotic selection markers. These producer cells are then kept in constant selection and are described as being stably transfected. Alternatively, the packaging genes can be integrated into the packaging cell line's genome.

By isolating the lentiviral gag, pol, env and rev genes and by permitting their expression with accompanying (operationally linked) promoters, the viral proteins generated can package the separate transfer vector, including an LTR flanked therapeutic gene, but renders the resulting particle unable to generate further virions. This multipart genome, also known as the 3rd generation lentiviral vector packaging system (Dull et al. 8463-71) also avoids the risk of recombining with other wild type retroviruses that may be present in the target cell.

The various lentivirus generations are described in the following references. First generation: Naldini et al., Science, 272(5259):263-7 (1996). Second generation: Zufferey et al., Nat Biotechnol. 15(9):871-5 (1997). Third generation: Dull et al., J Virol. 72(11):8463-7 (1998). A review on the development of lentiviral vectors can be found in Sakuma et al., Biochem J. 443(3):603-18 (2012); Picanço-Castro et al., Lentiviral-mediated gene transfer—a patent review. Exp Opin Therap Patents 18(5):525–539 (2008).

Figure 1 provides a generalized schematic of lentiviral vector production and host cell transduction. A packaging host cell is transfected with plasmid expression vectors encoding lentiviral packaging genes and a transgene of interest (Green Fluorescent Protein or GFP in Figure 1). Viral protein expression results in the assembly of proteins at the host membrane, and generation of a viral vector

The skilled person will appreciate that lentiviral production is normally carried out in mammalian cells. Cells for use in the invention include, but are not limited to, HEK 293, 293T, and T_REX™ (Life Technologies) cell lines.

Tet system in Lentiviral packaging cells: Commercial-scale production of lentiviral vectors comprising therapeutic genes requires packaging cell lines capable of generating large quantities of viral vector. Inducible expression of viral packaging proteins is desirable

in order to minimize metabolic burden on the host cell and/or where the expressed viral protein is cytotoxic to the cell. According to an aspect of the present invention, this can be achieved by placing expression of one or more viral genes under the control of a tet-inducible promoter, and introducing a vector encoding the TetR protein into the packaging
5 host cell, where the *tetR* nucleotide sequence has been codon-optimized for that host cell.

The codon optimized tet-repressor sequence of the present invention can be carried by a vector and operationally linked to a promoter for expression in mammalian cells. One such promoter is the human Cytomegalovirus (CMV) promoter. The vector may further comprise a selection marker to aid in generating stable cell lines transfected with the vector.

0 The skilled person will appreciate that a vector is a DNA molecule used as a vehicle to artificially carry foreign genetic material into another cell, where it can be replicated and/or expressed. The four major types of vectors are plasmids, viral vectors, cosmids, and artificial chromosomes. The present invention includes all vectors as covered by this definition. In one embodiment of the present invention, the vector carrying the codon
5 optimized *TetR* sequence is a plasmid; in another embodiment, the vector is a Bacterial Artificial Chromosome (BAC).

One aspect of the present invention is plasmid expression vectors, whose specific function is to express TetR in a mammalian host cell. Expression vectors generally have a promoter sequence that drives expression of the transgene, and a marker gene that will
0 help select cells actively expressing the transgene. Accordingly, the vectors of the present invention may include a promoter such as the human cytomegalovirus (CMV) immediate early promoter, spleen focus-forming virus (SFFV) promoter, Rous sarcoma virus (RSV) promoter, or human elongation factor 1-alpha (EF1a) promoter. Furthermore, the vectors of the present invention may also include a selection marker such as enzymes encoding
5 resistance to an antibiotic, e.g., neomycin, puromycin, hygromycin blasticidin, or zeocin.

ACE and BAC system

Artificial chromosomes have the capacity to accommodate and express heterologous genes inserted therein. Furthermore, large segments of DNA, multiple copies of the
0 heterologous gene, and linked promoter element(s) can be retained in the artificial chromosomes, thereby providing a high-level of expression of the retroviral protein(s).

Artificial Chromosome Expression (ACE) refers to a fully functional mammalian chromosomal structure that can stably replicate and segregate alongside endogenous mammalian chromosomes so that it is transmittable to host cell progeny. Therefore, the ACE provides an extra genomic locus for targeted integration of heterologous DNA into a host cell.

A Bacterial Artificial Chromosome (BAC) is a functional bacterial chromosomal structure, capable of accommodating up to 350kb of heterologous DNA that can stably replicate within a bacterial cell alongside the bacterial genome. Therefore, the BAC can be used to amplify large segments of DNA in a bacterial system and capable to shuttle this DNA to mammalian cells for either transient transfection, stable transfection or targetted integration into the host cell chromosomes, including the ACE.

Codon Optimisation of T repressor

The present inventors constructed a codon-optimized (CO) Tet repressor suitable for use in viral packaging cell lines where at least one viral gene is under the control of a tet-inducible promoter.

Accordingly, in a first embodiment, the invention provides an isolated nucleotide comprising a codon optimised (CO) Tet repressor sequence of SEQ ID NO:3 or SEQ ID NO: 4.

In another aspect, the invention provides a vector comprising a codon optimised (CO) Tet repressor sequence of SEQ ID NO:3 or SEQ ID NO: 4. The vector comprises a promoter operably linked to the CO tet repressor sequence. A further aspect of the present invention is a host cell comprising such an expression vector.

In another aspect, the invention provides a eukaryotic host cell comprising at least one heterologous target gene of interest, where: (a) the host cell comprises a vector comprising a codon optimised (CO) Tet repressor sequence of the present invention; and (b) the target gene is under the control of a promoter that includes at least one Tet Operon, such that the expression of the target gene is tet-inducible. The host cell may be a human cell. In one embodiment the host cell is a Human Embryonic Kidney 293 cell (HEK-293) or a HEK 293T cell. In one aspect of the invention, the promoter that is operably linked to a target gene is a Cytomegalovirus promoter that additionally contains one or more Tet

Operon sequences; such a promoter may comprise a nucleic acid sequence selected from the group consisting of SEQ ID NO: 7-8.

SEQ ID NO:6 (CMV promoter sequence)

```

5  TGGCCATTGC ATACGTTGTA TCCATATCAT AATATGTACA TTTATATTGG CTCATGTCCA
   ACATTACCGC CATGTTGACA TTGATTATTG ACTAGTTATT AATAGTAATC AATTACGGGG
   TCATTAGTTC ATAGCCCATATA TATGGAGTTC CGCGTTACAT AACTTACGGT AAATGGCCCCG
   CCTGGCTGAC CGCCCAACGA CCCCCGCCCA TTGACGTCAA TAATGACGTA TGTTCCCATA
   GTAACGCCAA TAGGGACTTT CCATTGACGT CAATGGGTGG AGTATTTACG GTAAACTGCC
0  CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA CGTCAATGAC
   GGTAAATGGC CCGCCTGGCA TTATGCCCAG TACATGACCT TATGGGACTT TCCTACTTGG
   CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA TCGGTTTTTG GCAGTACATC
   AATGGGCGTG GATAGCGGTT TGA CTCACGG GGATTTCCAA GTCTCCACCC CATTGACGTC
   AATGGGAGTT TGT TTTGGCA CAAAATCAA CGGGACTTTC CAAAATGTGC TAACA ACTCC
5  GCCCCATTGA CGCAAATGGG CGGTAGGCGT GTACGGTGGG AGGTCTATAT AAGCAGAGCT
   CGTTTAGTGA ACCG

```

SEQ ID NO:7 (CMV-TO2(IH); CMV promoter sequence modified to contain tet operons)

```

   GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA
0  GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC
   CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG
   GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC TTGGCAGTAC
   ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCCG
   CCTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG
5  TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT
   AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT
   TTTGGAACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC
   AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCTC CCTATCAGTG
   ATAGAGATCT CCCTATCAGT GATAGAGATC GTCGACGAGC TCGTTTAGTG AACCGTCAGA
0  TCGCCTGGAG ACGCCATCCA CGCTGTTTTG ACCTCCATAG AAGACACCGG GACCGATCCA
   GCCTCCG

```

SEQ ID NO:8 (CMV-TO2(LT); CMV promoter sequence modified to contain tet operons)

```

   GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA
5  GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC
   CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG

```

GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC TTGGCAGTAC
 ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCCG
 CCTGGCATTG TGGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG
 TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT
 5 AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT
 TTTGGAACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC
 AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCTC CCTATCAGTG
 ATAGAGATCT CCCTATCAGT GATAGAGATC GTCGACGAGC TCGTTTAGTG AACCGTCAGA
 TCGCCTGGAG ACGCCATCCA CGCTGTTTTG ACCTCCATAG AAGACACCGG GACCGATCCA
 0 GCCTCCG

In one embodiment of the invention, the codon optimised sequence(s) or vector(s) are present episomally. In another embodiment of the invention, the codon optimised sequence(s) or vector(s) are integrated into the cell's genome.

5 A further embodiment of the invention provides a method of transfecting a cell with one or more of the expression vectors of the invention, wherein the transfection method comprises: (i) contacting the cell with a plasmid expression vector solution containing one or more plasmid expression vectors according to the invention; and (ii) growing the cells in a growth media.

0 A further aspect of the present invention is a method of producing a lentiviral vector packaging cell comprising transfecting a host cell with (a) an expression vector according to the present invention comprising a codon optimised nucleotide sequence encoding a tetR protein, and (b) expression vectors comprising a lentiviral GagPol sequence, a lentiviral Rev sequence, and a VSVg sequence, where one or more of said lentiviral sequences is tet-
 5 inducible.

A further aspect of the present invention is a method of producing a lentiviral vector, the method comprising (1) transfecting a host cell with (a) an expression vector according to the present invention comprising a codon optimised nucleotide sequence encoding a tetR protein, and (b) expression vectors comprising a lentiviral GagPol sequence, a lentiviral Rev
 0 sequence, and a VSVg sequence, where one or more of said lentiviral sequences is tet-inducible; (2) culturing said transfected host cell, and (3) inducing expression of the tet-inducible gene(s) under conditions suitable for the production of lentiviral vectors.

Insertion of a vector into the target cell is usually called transformation for bacterial cells and transfection for eukaryotic cells, although insertion of a viral vector may also be called transduction. The skilled person will be aware of the different viral transfection methods commonly used, which include, but are not limited to, the use of physical methods, chemical reagents or cationic lipids. Many transfection methods require the contact of solutions of plasmid DNA to the cells, which are then grown and selected for a marker gene expression. The viral supernatants may then be harvested, purified and titred for further use.

As used herein, the term “comprising” encompasses “including” as well as “consisting” e.g. a composition “comprising” X may consist exclusively of X or may include something additional e.g. X + Y.

References cited herein are incorporated by reference herein in their entirety.

The term “about” in relation to a numerical value x means, for example, X plus or minus 10%.

The invention will be further described by reference to the following, non-limiting, figures and examples.

EXAMPLES

Example 1: Codon optimisation of *TetR*

A wild-type (WT) Tet Repressor nucleotide sequence was selected for codon optimization (SEQ ID NO:1, NCBI Gene ID: 4924774).

ATGTCTAGAT	TAGATAAAAG	TAAAGTGATT	AACAGCGCAT	TAGAGCTGCT
TAATGAGGTC	GGAATCGAAG	GTTTAACAAC	CCGTAAACTC	GCCCAGAAGC
TAGGTGTAGA	GCAGCCTACA	TTGTATTGGC	ATGTAAAAAA	TAAGCGGGCT
TTGCTCGACG	CCTTAGCCAT	TGAGATGTTA	GATAGGCACC	ATACTCACTT
TTGCCCTTTA	GAAGGGGAAA	GCTGGCAAGA	TTTTTTACGT	AATAACGCTA
AAAGTTTTAG	ATGTGCTTTA	CTAAGTCATC	GCGATGGAGC	AAAAGTACAT
TTAGGTACAC	GGCCTACAGA	AAAACAGTAT	GAAACTCTCG	AAAATCAATT
AGCCTTTTTA	TGCCAACAAG	GTTTTTCACT	AGAGAATGCA	TTATATGCAC
TCAGCGCTGT	GGGGCATTTT	ACTTTAGGTT	GCGTATTGGA	AGATCAAGAG
CATCAAGTCG	CTAAAGAAGA	AAGGGAAACA	CCTACTACTG	ATAGTATGCC
GCCATTATTA	CGACAAGCTA	TCGAATTATT	TGATCACCAA	GGTGCAGAGC
CAGCCTTCTT	ATTCGGCCTT	GAATTGATCA	TATGCGGATT	AGAAAAACAA

CTTAAATGTG AAAGTGGGTC TTAA (624)

The WT TetR gene encodes the following amino acid sequence (SEQ ID NO:2):

5 MSRLDKSKVI NSALELLNEV GIEGLTTRKL AQKLGVEQPT LYWHVKNKRA
LLDALAIEML DRHHTHFCPL EGESWQDFLR NNAKSFRCL LSHRDGAKVH
LGTRPTEKQY ETLENQLAFL CQQGFSLENA LYALSAVGHF TLGCVLEDQE
HQVAKEERET PTTDSMPPLL RQAIELFDHQ GAEP AFLFGL ELIICGLEKQ
LKCESGS (207)

0 The WT sequence was altered to provide two codon optimised sequences based on
codon usage. The sequences were further optimised for ~3750 iterations using all of the
following parameters: codon tandem repeats (repetition of identical codons); secondary
structure (hairpin loop formation); GC distribution; long range repeats; DNA motifs (ARE
motifs); Open Reading Frame (ORF) optimisation; restriction sites - AsiSI, BamHI, BsiWI,
5 EcoRI, MluI, NheI, NheI, PaeI, PmeI, PvuI, Sall, XbaI, XhoI; cryptic splice sites – human.

A human codon optimized (humCO) sequence was generated (SEQ ID NO:3) where a
70% Leto frequency threshold for rare codons was set, discarding any codons below 70% of
the theoretical ratio (all codons in equal amounts).

0 ATGTCTAGGC TGGACAAGTC CAAGGTGATC AACTCCGCCC TGGAGCTCCT
TAACGAGGTG GGGATTGAAG GGTGACTAC CCGGAAACTG GCTCAGAAAC
TGGGCGTCGA ACAGCCAACC CTGTACTGGC ACGTGAAGAA CAAGCGCGCT
TTGCTGGATG CACTTGCCAT CGAGATGCTG GACAGGCACC ACACACACTT
TTGCCCCTTG GAGGGCGAAT CCTGGCAGGA CTTCTGAGG AACAACGCCA
AGAGTTTCCG CTGTGCCCTG TTGAGTCACA GGGACGGAGC AAAGGTTTAC
5 CTTGGAACCA GACCCACCGA GAAGCAGTAC GAGACTCTGG AGAACCAGCT
GGCCTTCCTG TGTCAGCAGG GCTTTAGCCT GGAGAACGCT CTGTACGCAC
TGCTGCCGT TGGCCACTTT AACTCGGAT GCGTGCTGGA AGACCAGGAG
CATCAGGTTG CCAAGGAAGA GAGGGAGACC CCTACCACTG ACAGCATGCC
TCCACTTCTG AGACAGGCCA TCGAGCTCTT TGACCACCAG GGAGCAGAAC
0 CCGCTTTCTT CTTCCGGGCTG GAACTGATCA TCTGCGGCCT CGAAAAGCAG
CTCAAGTGCG AAAGCGGCAG CTGA (624)

A GSK Codon Optimized (GSKCO) sequence was generated (SEQ ID NO:4) where a 90% Leto
frequency threshold for rare codons was set, discarding any codons below 70% of the
5 theoretical ratio (all codons in equal amounts). Sequence differences in SEQ ID NO:4
(compared to WT *TetR* SEQ ID NO:1) are indicated by underlining:

ATGAG CAGGC TGGAC AAGAG CAAGG TGATC AACAG CGCCC TGGAG CTGCT
GAACG AGGTC GGCAT CGAGG GCCTG ACAAC CAGGA AGCTG GCCCA GAAAC
TCCGA GTGGA GCAGC CCACC CTGTA CTGGC ACGTG AAGAA CAAGC GCGCC

CTCCT GGACG CCCTG GCTAT TGAGA TGCTG GACAG GCACC ACACC CACTT
 CTGCC CACTG GAGGG CGAGA GCTGG CAGGA CTTCC TGAGG AACAA CGCCA
 AGAGC TTCAG GTGTG CCCTG CTGAG CCATA GGGAC GGCGC CAAGG TGCAC
 CTGGG CACCA GGCCC ACCGA GAAGC AGTAC GAGAC CCTGG AGAAC CAGCT
 5 GGCCCT TCCTC TGCCA GCAGG GCTTC AGCCT GGAGA ATGCA CTGTA CGCCC
 TGAGC GCCGT GGGCC ACTTC ACCCT GGGGT GCGTG CTGGA AGACC AGGAA
 CACCA GGTGG CCAAG GAGGA GAGGG AGACC CCCAC CACCG ACTCT ATGCC
 CCCCC TGCTG AGGCA GGCCA TCGAA CTGTT CGATC ACCAG GGCGC CGAGC
 CCGCC TTTCT GTTCG GCCTG GAGCT GATCA TCTGC GGCCT GGAGA AGCAG
 0 CTGAA GTGCG AGAGC GGCTC CTGA (624)

The codon optimisation of both SEQ ID NO:3 and SEQ ID NO:4 preserves the amino acid sequence for the Tet R protein (SEQ ID NO:2)

5

The pcDNATM6/TR vector is commercially available from Invitrogen. The pcDNATM6/TR vector contains the *TetR* gene under the control of the Human CMV immediate early promoter, and the blasticidin (bsd) resistance gene under the control of the SV40 early promoter (for expression of BSD in mammalian cells to allow for selection of stably transfected mammalian cells).

0

Examination of the *TetR* sequence contained in the pcDNATM6/TR vector revealed that it includes a sequence comprising multiple cloning sites immediately upstream of the *TetR* stop codon (SEQ ID NO:5):

ATGTCTAGAT TAGATAAAAG TAAAGTGATT AACAGCGCAT TAGAGCTGCT
 5 TAATGAGGTC GGAATCGAAG GTTTAACAAC CCGTAAACTC GCCCAGAAGC
 TAGGTGTAGA GCAGCCTACA TTGTATTGGC ATGTAAAAAA TAAGCGGGCT
 TTGCTCGACG CCTTAGCCAT TGAGATGTTA GATAGGCACC ATACTCACTT
 TTGCCCTTTA GAAGGGGAAA GCTGGCAAGA TTTTTTACGT AATAACGCTA
 AAAGTTTTAG ATGTGCTTTA CTAAGTCATC GCGATGGAGC AAAAGTACAT
 0 TTAGGTACAC GGCCTACAGA AAAACAGTAT GAAACTCTCG AAAATCAATT
 AGCCTTTTTA TGCCAACAAG GTTTTTCACT AGAGAATGCA TTATATGCAC
 TCAGCGCTGT GGGGCATTTT ACTTTAGGTT GCGTATTGGA AGATCAAGAG
 CATCAAGTCG CTAAAGAAGA AAGGGAAACA CCTACTACTG ATAGTATGCC
 GCCATTATTA CGACAAGCTA TCGAATTATT TGATCACCAA GGTGCAGAGC
 5 CAGCCTTCTT ATTCGGCCTT GAATTGATCA TATGCGGATT AGAAAAACAA
 CTTAAATGTG AAAGTGCGTC CGCGTACAGC GGATCCCGGG AATTCAGATC
TTATTAA

The multiple cloning site (MCS) sequence is at residues 621-653, and the stop codon at residues 654-657, of SEQ ID NO:5 (referred to below as the WT-Invitrogen *TetR* sequence).

0

Example 2: Comparing TetR protein expression with WT-*TetR* and CO-*TetR*

Human 293T cells were transfected with four different plasmid constructs (see Figures 5A-%D), in two doses, as shown in Table 2. Each transfection (plasmid + dose) was run in triplicate. Un-transfected 293T cells were used as controls.

5

Table 2

Plasmid construct	DNA(ug)/N ^o Cells	<i>TetR</i> sequence
<i>none</i>	-	-
pcDNA6/TR	1.25ug/10 ⁶ cells	WT-Invitrogen (SEQ ID NO:5)
pcDNA6/TR	2.5ug/10 ⁶ cells	WT-Invitrogen (SEQ ID NO:5)
pMA-BACmod-WT-TR-lr-ZeoR	1.25ug/10 ⁶ cells	WT (SEQ ID NO:1)
pMA-BACmod-WT-TR-lr-ZeoR	2.5ug/10 ⁶ cells	WT (SEQ ID NO:1)
pMA-BACmod-humTR-lr-ZeoR	1.25ug/10 ⁶ cells	humCO (SEQ ID NO:3)
pMA-BACmod-humTR-lr-ZeoR	2.5ug/10 ⁶ cells	humCO (SEQ ID NO:3)
pMA-BACmod-GSKTR-lr-ZeoR	1.25ug/10 ⁶ cells	GSKCO (SEQ ID NO:4)
pMA-BACmod-GSKTR-lr-ZeoR	2.5ug/10 ⁶ cells	GSKCO (SEQ ID NO:4)

Hek 293T cells were plated at 10⁶ cells per well in a six-well plate, cultured overnight at 37°C and 5% CO₂, and transfected with 1.2ug or 2.5ug of DNA per construct.

0 Transfections were set up in triplicate and performed using the transfection reagent lipofectamine 2000 according to manufacturer's instruction. The transfected and control cells were cultured for a further 48 hours and subsequently harvested for western blot analysis. Protein was isolated from the transfected and control cells, separated by SDS-PAGE and the TetR protein band detected with a TetR specific antibody and visualised with a
5 secondary antibody conjugated to HRP and HRP substrate. Protein loading was verified by a GAPDH control.

Total TetR protein band intensity was quantified by the programme ImageJ and the values plotted in figure 2.

0 Where the individual DNA constructs were transfected at a quantity of 1.25ug/transfection, little distinction could be made between the individual constructs and there was no difference in the background as compared to the UT control. However, when 2.5ug DNA was used per transfection, there appeared to be a greater amount of TetR

protein in the humCO-TetR and GSKCO-TetR samples compared to the WT-TetR and pcDNA6/TR. Furthermore, the codon optimisation of GSKTetR appeared to be capable of increasing protein expression 2 fold.

TetR protein expression was compared using Western blot band density. Band density was quantified and compared using ImageJ, a Java-based image analysis software package available for download from the National Institutes of Health at <http://rsbweb.nih.gov/ij/download.html>. Results are shown in Figure 2A and 2B, where UT = untransfected control; pcDNATM6/TR is the regulatory vector pcDNATM6/TR (Invitrogen); WT is *TetR* of SEQ ID NO:1; hCO is *TetR* of SEQ ID NO:3; and GSKCO is *TetR* of SEQ ID NO:4. Each experimental condition was run in either duplicate or triplicate; each bar represents one experiment.

These results indicate that cells transfected with vectors containing the GSKCO optimized *TetR* sequence reliably provided the greatest TetR protein expression, compared to WT and hCO optimized *TetR* sequences.

Example 3: Control of tet-inducible lentiviral packaging genes

Viral vector was made using lenti packaging constructs. 293T cells were transiently transfected with plasmid vectors carrying lentiviral packaging genes including GagPol, VSVg and Rev, each individually driven by a TetO2 controlled CMV promoter. Also included in the transfection were a GFP expressing transfer vector and *TetR* vectors containing either WT (SEQ ID NO:1), hCO (SEQ ID NO:3) or GSKCO (SEQ ID NO:4) *TetR* all utilising a non-Tet operon controlled CMV promoter. Cells that did not contain the *TetR* vector (No TetR) and untransfected cells (UT) were used as positive and negative viral vector controls respectively. In this example, where cells are not exposed to Doxycycline, the presence of TetR protein is expected to decrease viral production levels. 48 hours post transfection, the viral vector supernatants were harvested. Viral titre of the supernatants was measured by the ability of the viral vector to deliver the GFP expression cassette into 293T cell chromatin. 293T cells were transduced with viral vector supernatant, cultured for 72hours and the percentage GFP positive cells analysed by FACS.

Results are shown in Figure 3, where the Y-axis is viral titre produced (TU/mL). Background fluorescence is established by the Untransfected (UT) cells. It is seen that both

the hCO and GSKCO *TetR* vectors are capable of reducing gene expression to below that obtained with WT *TetR* and, in some cases, below that of UT cells.

Figure 4 graphs viral titer averages from Figure 3 on a linear scale.

5 Conclusion

Codon optimised *TetR* sequences useful in tet-inducible gene expression systems in mammalian, specifically human, cells are provided herein.

We claim:

- 5 1. A codon optimised tetracycline repressor (*TetR*) sequence comprising the nucleic acid sequence of SEQ ID NO:4 or SEQ ID NO: 5.
2. A vector comprising a codon optimised *TetR* sequence according to claim 1.
- 0 3. A vector according to claim 2, wherein the vector comprises a human cytomegalovirus (CMV) promoter.
4. A vector according to claim 4, wherein the CMV promoter sequence is selected from SEQ ID NO:7 and SEQ ID NO:8.
- 5 5. A vector according to any one of claims 2-4, further comprising a selection marker gene.
6. A vector according to any one of claims 2-5, where said vector is a plasmid expression vector.
- 0 7. A vector according to claim 6, further comprising a selection marker gene.
8. A cell comprising a vector according to any one of claims 2-7.
- 5 9. A cell according to claim 8, wherein the cell is selected from the group consisting of HEK 293 and 293T cells.
10. The cell according to claim 8 or claim 9, wherein the vector is present episomally.
- 0 11. A method of transfecting a cell with a plasmid expression vector according to claim 7, wherein the transfection method comprises:

- (i) contacting a cell with a plasmid expression vector solution containing one or more plasmid expression vectors according to claim 6; and
- (ii) growing the cells in a growth media under conditions which allow for selection of cells which express the selection marker gene.

5

12. A method of transfection according to claim 11 wherein the plasmid expression vector solution used for the transfection further comprises (i) a plasmid containing a lentiviral gene.

0

13. A method of transfection according to claim 11 wherein the cell comprises a plasmid containing a lentiviral gene.

Figure 1

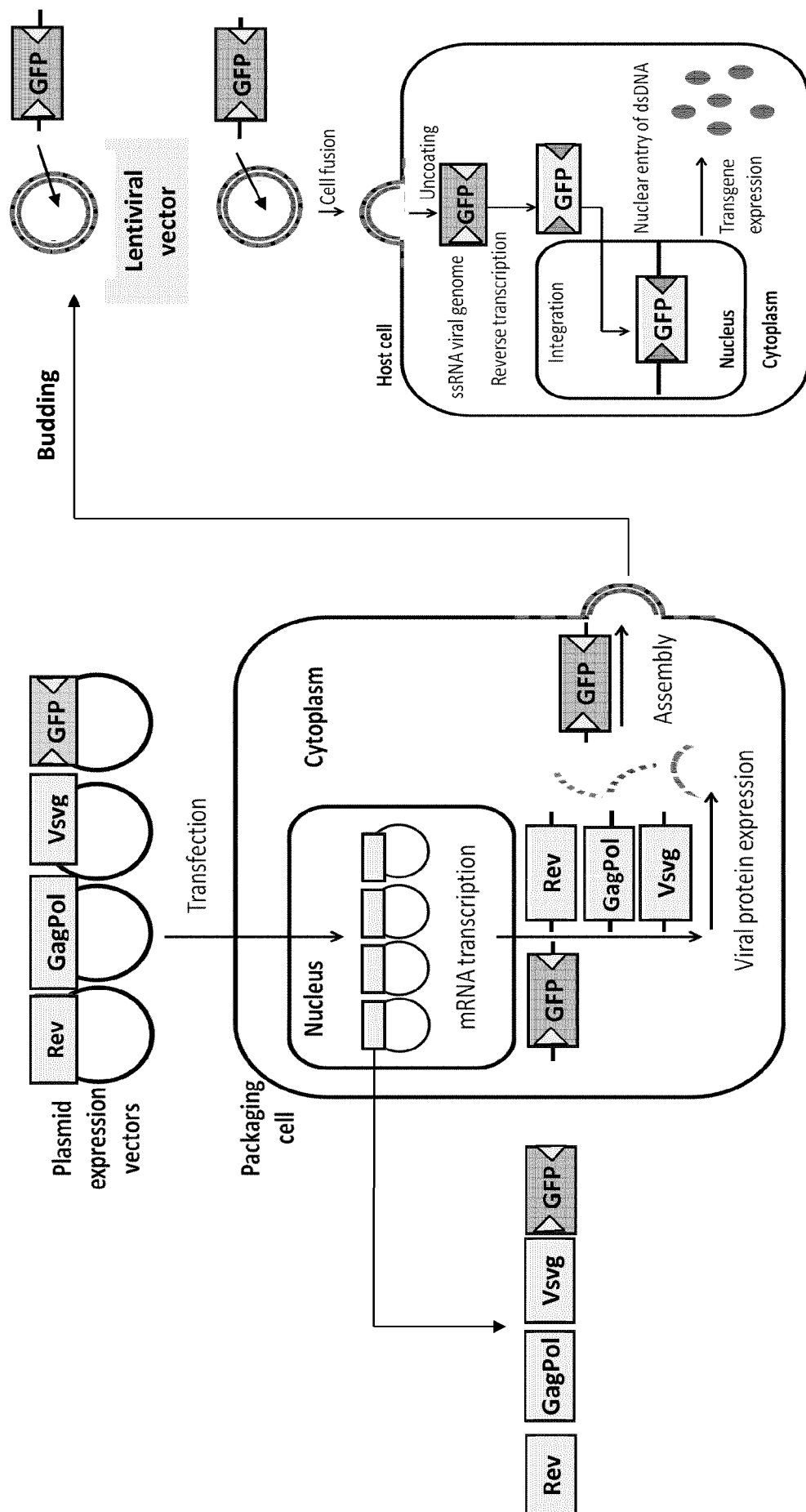


Figure 2A

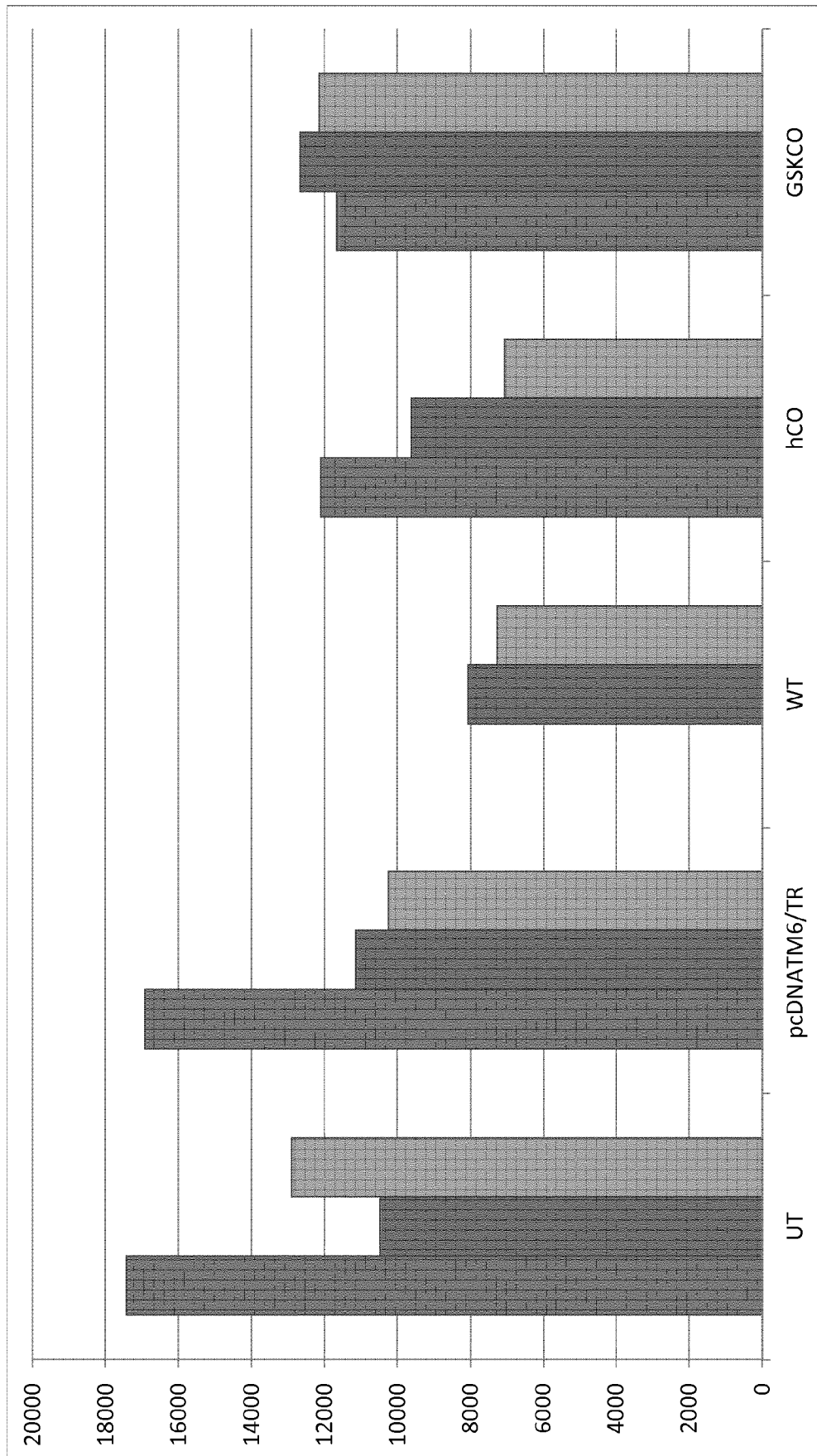


Figure 2B

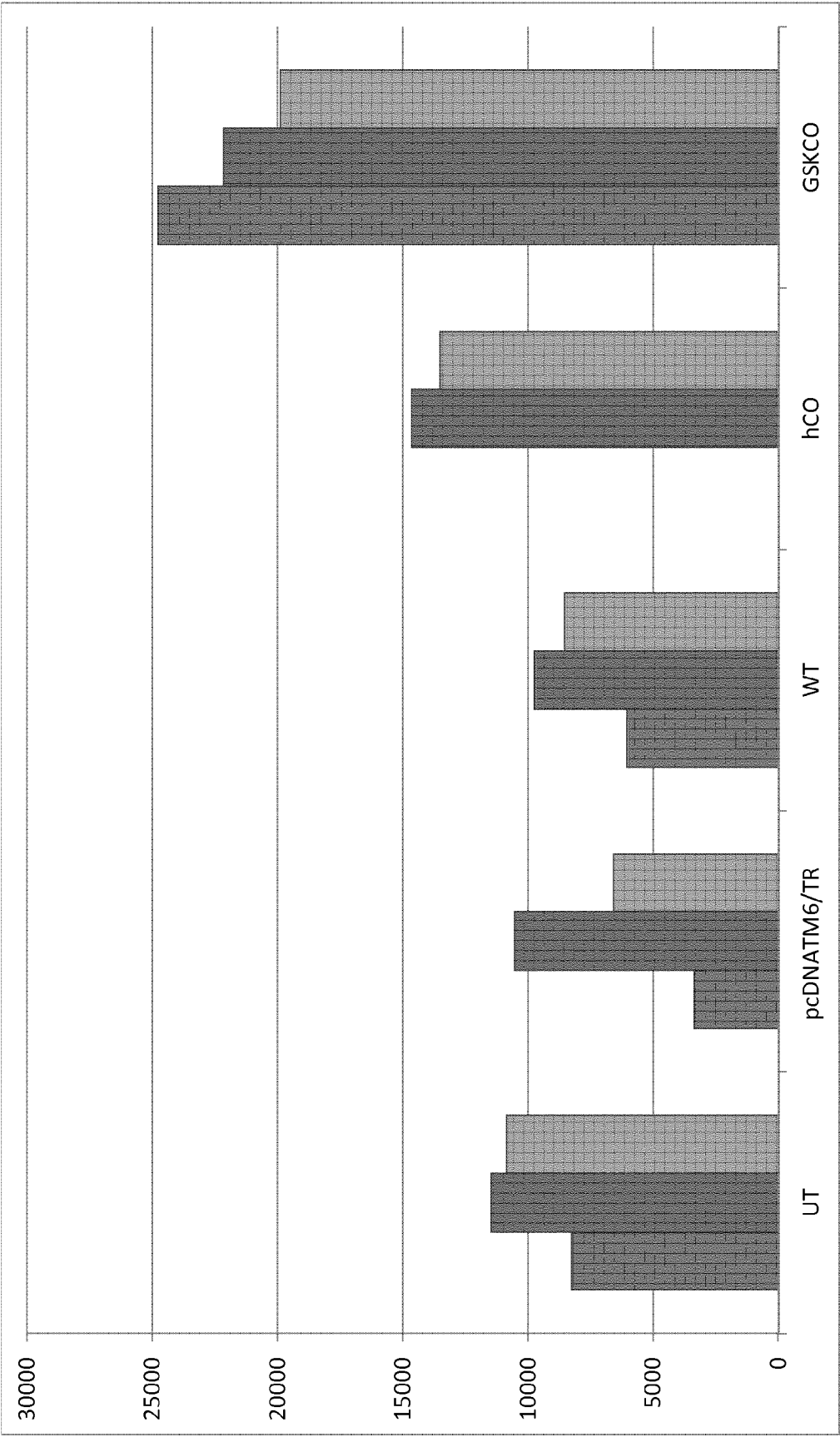


Figure 3

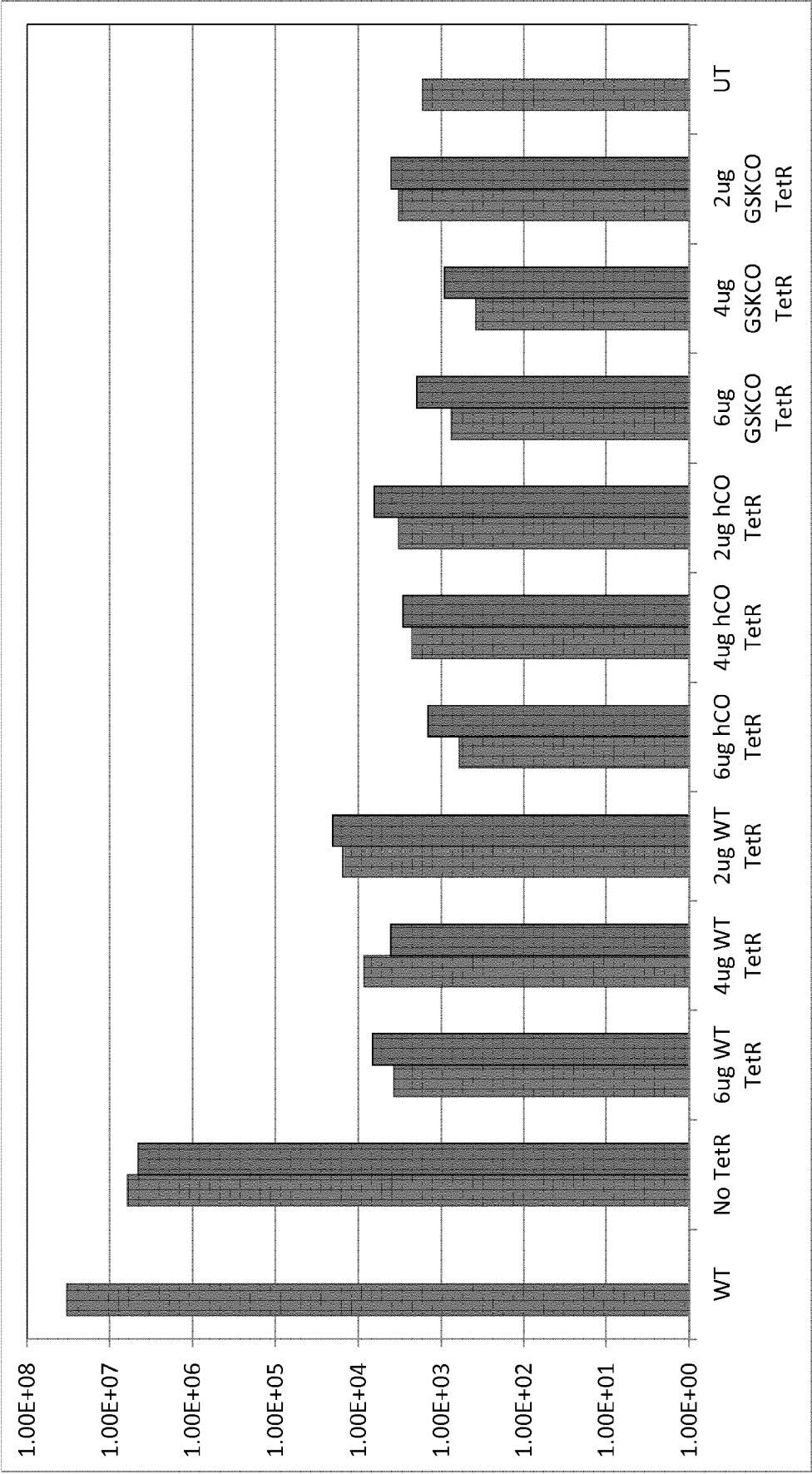


Figure 4

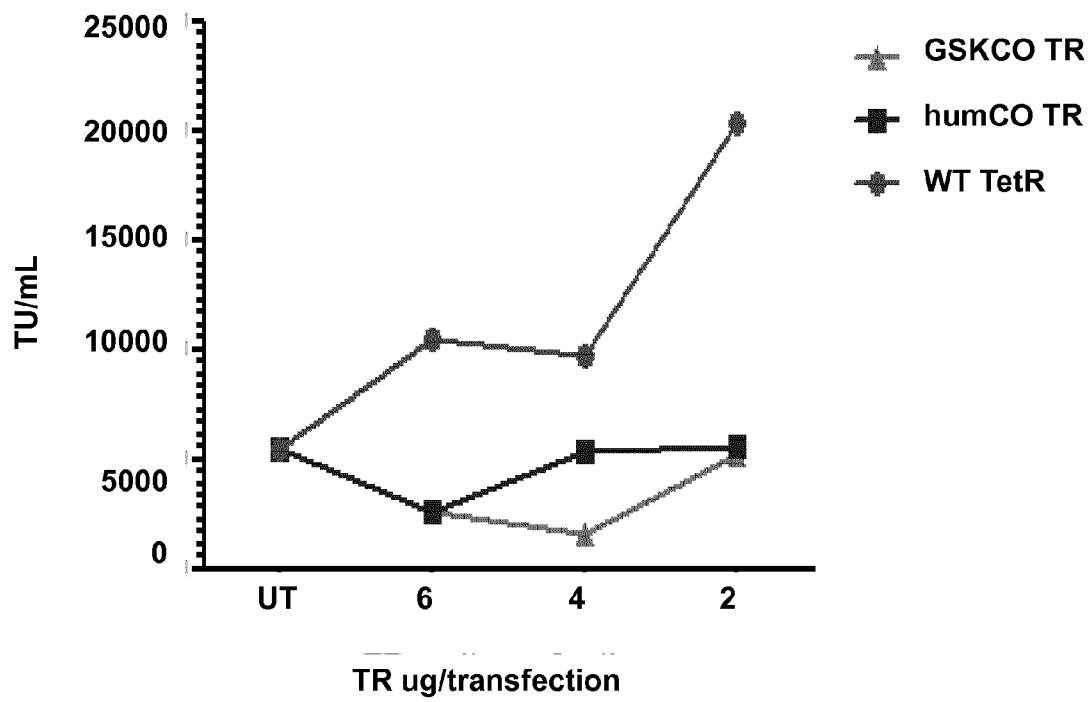


Figure 5A

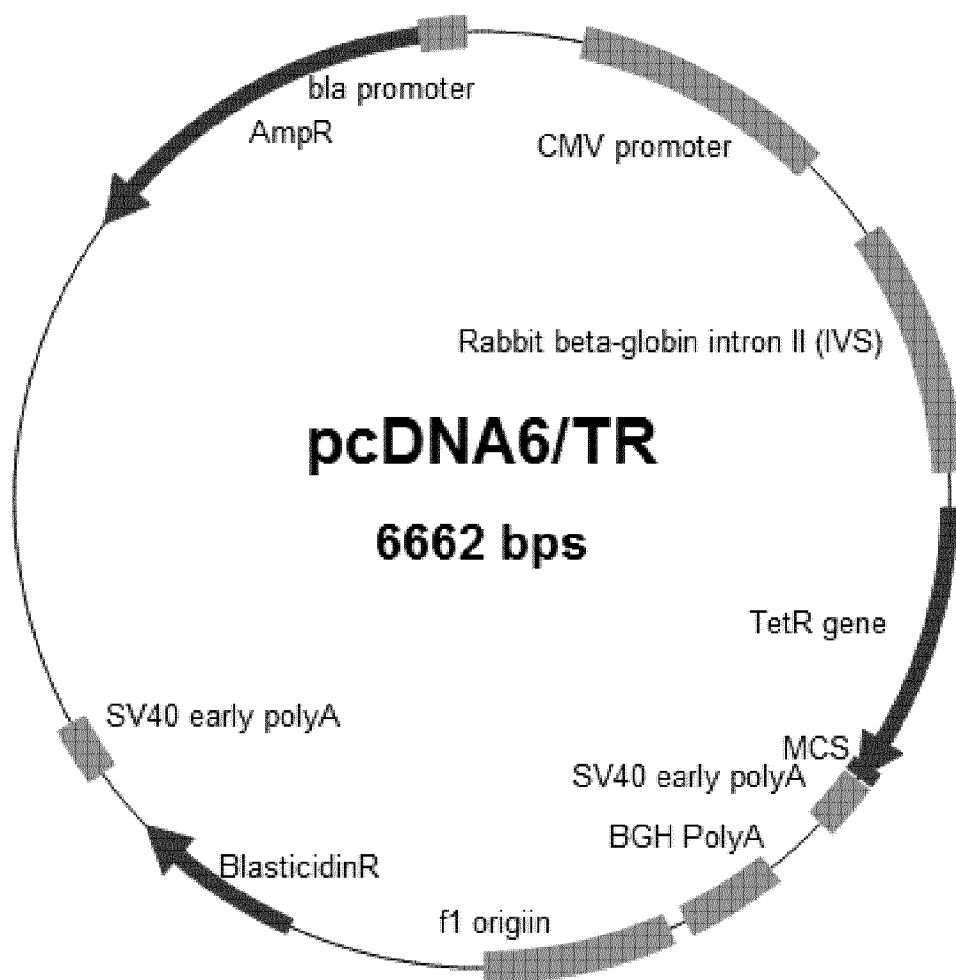


Figure 5B

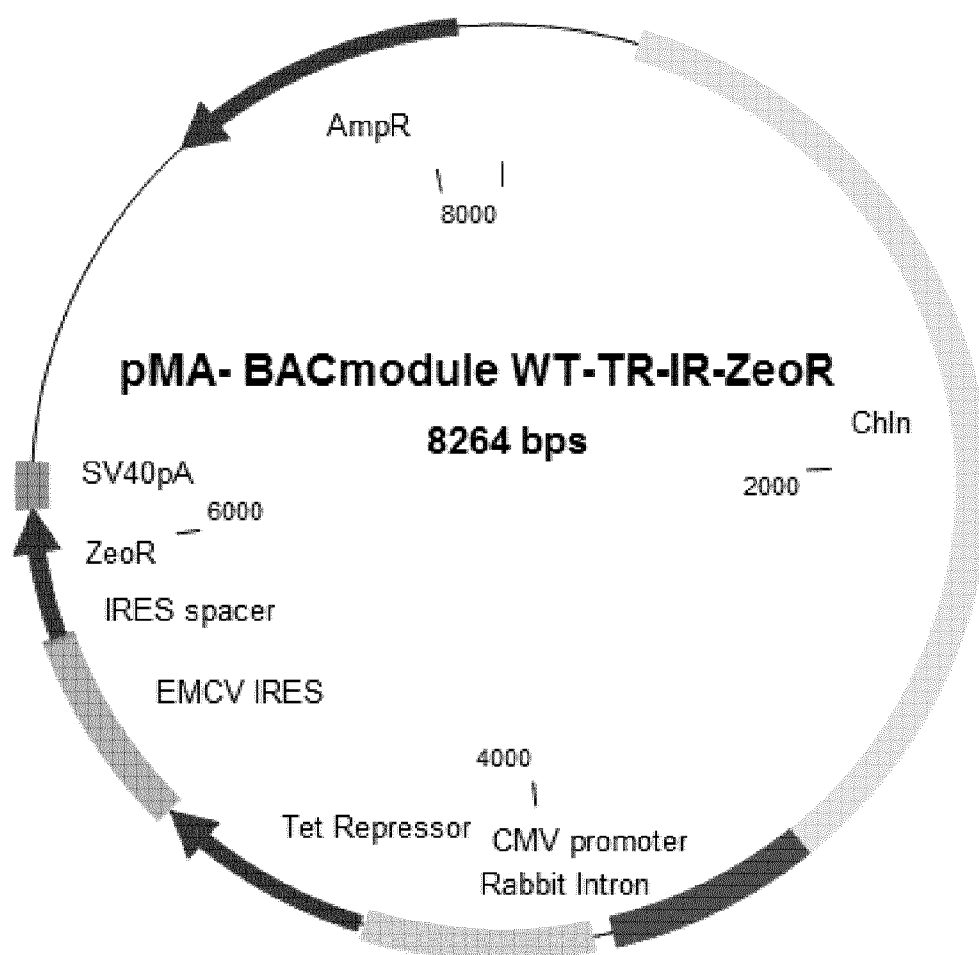


Figure 5C

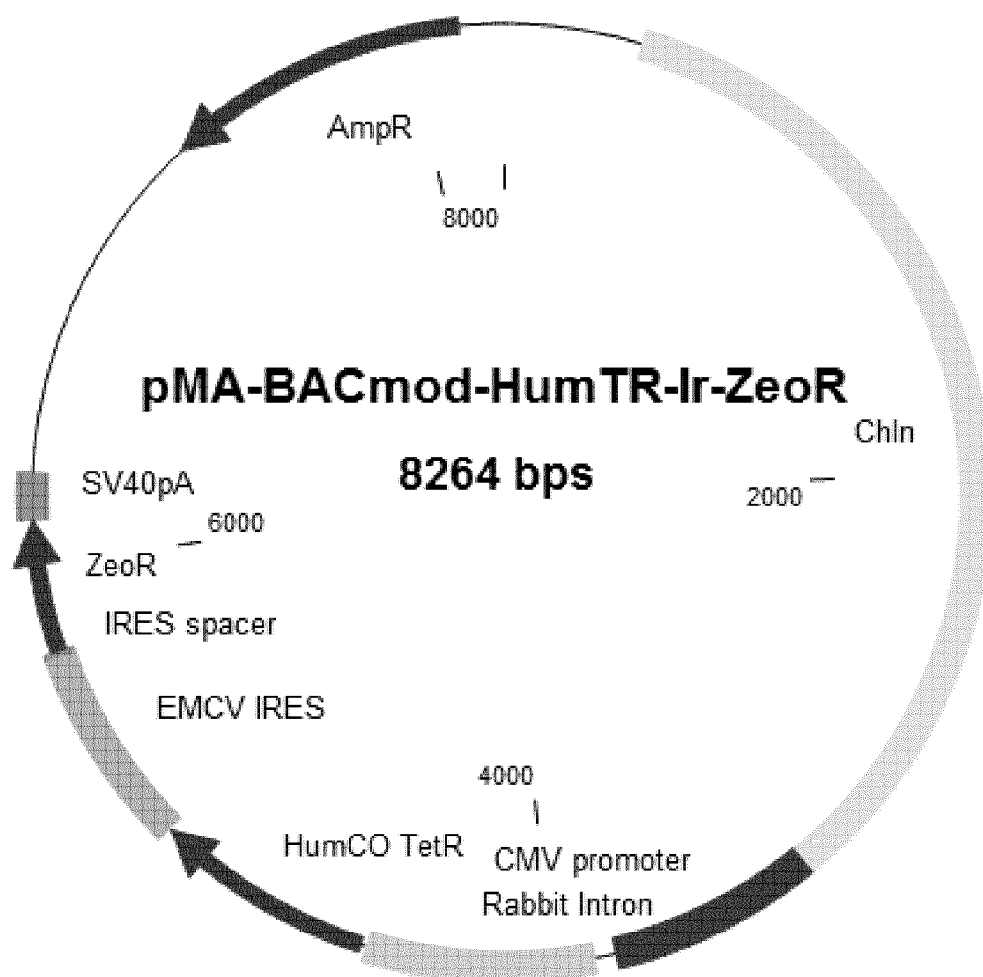
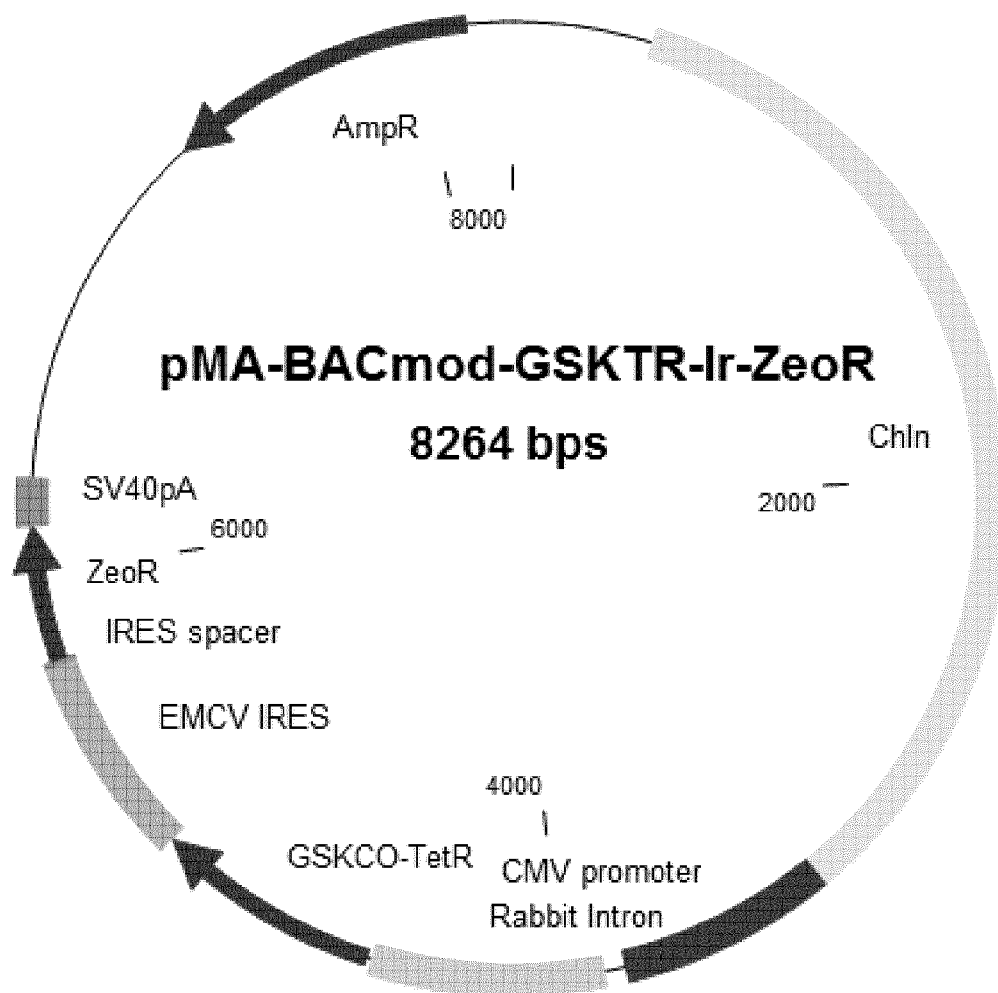


Figure 5D



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/067883

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N15/63 C12N15/86
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 2 348 119 A2 (OXFORD BIOMEDICA LTD [GB]) 27 July 2011 (2011-07-27) paragraphs [0016], [0096] - [0123], [0156], [0164], [0166], [0175], [0195] - [0212], [0344], [0415], [0597]; claims 3,9-11,16; figures 22,23; examples 13,19	1-13
Y	DE 10 2006 011918 A1 (UNIV ZU KOELN [DE]) 9 August 2007 (2007-08-09) sequence 8	3,4
Y	US 2007/196867 A1 (MANCEBO HELENA S [US] ET AL) 23 August 2007 (2007-08-23) sequence 20	3,4
	----- -/--	



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

24 October 2016

Date of mailing of the international search report

15/11/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Seroz, Thierry

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/067883

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 731 607 A1 (ARTEMIS PHARMACEUTICALS GMBH [DE]) 13 December 2006 (2006-12-13) the whole document -----	1-13
Y	WELLS ET AL.: "Codon optimization, genetic insulation, and an rtTA reporter improve performance of the tetracycline switch.", TRANSGENIC RESEARCH, vol. 8, no. 5, October 1999 (1999-10), pages 371-381, XP002763275, the whole document -----	1-13
Y	WO 2006/131543 A1 (ARTEMIS PHARMACEUTICALS GMBH [DE]; SEIBLER JOST [DE]; SCHWENK FRIEDER) 14 December 2006 (2006-12-14) the whole document -----	1-13
Y	ANASTASSIADIS K ET AL: "A predictable ligand regulated expression strategy for stably integrated transgenes in mammalian cells in culture", GENE, ELSEVIER, vol. 298, no. 2, 2 October 2002 (2002-10-02), pages 159-172, XP004391541, AMSTERDAM, NL ISSN: 0378-1119, DOI: 10.1016/S0378-1119(02)00979-4 the whole document -----	1-13
Y	US 2008/220471 A1 (DAVIS DAVID P [US] ET AL) 11 September 2008 (2008-09-11) the whole document -----	1-13
Y	EP 2 166 107 A1 (TACONICARTEMIS GMBH [DE]) 24 March 2010 (2010-03-24) the whole document -----	1-13
Y	GRAY DANIEL C ET AL: "pHUSH: a single vector system for conditional gene expression.", BMC BIOTECHNOLOGY, vol. 7, 61, 26 September 2007 (2007-09-26) , page 14PP, XP002763276, ISSN: 1472-6750, DOI: 10.1186/1472-6750-7-61 the whole document -----	1-13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/067883

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2348119	A2	27-07-2011	AU 2003238456 A1 02-09-2003
			DK 1504108 T3 10-06-2013
			EP 1504108 A2 09-02-2005
			EP 2348119 A2 27-07-2011
			ES 2411981 T3 09-07-2013
			JP 4598398 B2 15-12-2010
			JP 5601898 B2 08-10-2014
			JP 2005515783 A 02-06-2005
			JP 2010263904 A 25-11-2010
			PT 1504108 E 12-06-2013
			SI 1504108 T1 31-07-2013
			WO 03064665 A2 07-08-2003

DE 102006011918 A1	09-08-2007	NONE	

US 2007196867	A1	23-08-2007	US 2007196867 A1 23-08-2007
			US 2010330591 A1 30-12-2010

EP 1731607	A1	13-12-2006	NONE

WO 2006131543	A1	14-12-2006	EP 1888750 A1 20-02-2008
			US 2009210955 A1 20-08-2009
			WO 2006131543 A1 14-12-2006

US 2008220471	A1	11-09-2008	US 2008220471 A1 11-09-2008
			WO 2007014363 A2 01-02-2007

EP 2166107	A1	24-03-2010	NONE
