Meganuclease Recombination System

The invention relates to a set of genetic constructs which allow the efficient and reproducible introduction of a specific nucleotide sequence at a fixed position in the genome. The present invention also relates to methods using these constructs and to these materials in the form of a kit.

Since the first gene targeting experiments in yeast more than 25 years ago (1, 2), homologous recombination has been used to insert, replace or delete genomic sequences in a variety of cells (3-5). However, targeted events occur at a very low frequency in mammalian cells. The frequency of homologous recombination can be significantly increased by a specific DNA double-strand break (DSB) in the targeted locus (6, 7). Such DSBs can be created using Meganucleases, which are sequence-specific endonucleases that recognize large DNA target sites (>12 bp). These proteins can cleave a unique chromosomal sequence without affecting overall genome integrity. Natural Meganucleases are essentially represented by homing endonucleases, a widespread class of proteins found in eukaryotes, bacteria and archae (8). Early studies of the I-SceI and HO homing endonucleases have illustrated how the cleavage activity of these proteins initiates homologous recombination (HR) events in living cells and demonstrated the recombinogenic properties of chromosomal DSBs (9, 10). Since then, Meganuclease-induced recombination has been successfully used for genome engineering purposes in bacteria (11), mammalian cells (6, 7, 12-14), mice (15) and plants (16, 17).

Gene insertion can be used, for example, to introduce genes of interest in specific loci, for heterologous protein production. Recombinant therapeutic proteins are today mostly produced in mammalian cells such as CHO, mouse SP2/0 and NS0 cells, or the human PerC.6 cell line, stably transfected with the gene of interest (18). In the process of selecting highly expressing clones, the level and stability of protein expression are two major criteria. Obtaining reproducible results from one clone to another would be an advantage in terms of improving screening efforts. These principles also apply to the generation of cells for screening of specific drug targets. The same principle can also be applied to the expression of various genes in the same genomic context to comparatively study and analyze the resulting cell

lines one to another. Such cell lines can furthermore be subjected to the effect of compounds libraries in screening programs.

At the present time however no means exist to induce a DSB at a locus wherein the insertion/deletion of heterologous sequences can be easily ascertained.

The Inventors have developed a new set of genetic constructs which allow the reproducible integration and expression of a gene of interest (GOI) or a series of genes in otherwise isogenic cell lines.

According to a first aspect of the present invention there is provided a set of genetic constructs comprising:

a) Construct (i) encoded by a nucleic acid molecule, which comprises at least the following components:

A1 – A2 – A3 – A4 – A5 (i)

wherein A1 is a first promoter; A2 is a first homologous portion; A3 is a meganuclease cleavage site; A4 is a first marker gene; A5 is a second homologous portion; and wherein construct (i) is configured to be stably integrated into the genome of at least one target cell;

b) Construct (ii) encoded by a nucleic acid molecule, which comprises at least the following components:

A2' – B1– B2 – B3 – B4 – A5' (ii)

wherein A2' comprises a portion of said first homologous portion A2; B1 is a second marker gene different to said first marker gene; B2 is a second promoter; B3 is a multiple cloning site; B4 is a third promoter; A5' comprises a portion of said second homologous portion A5;

c) At least one construct selected from the group comprising, constructs (iii) or (iv) encoded by nucleic acid molecules, which comprise at least the following components:

C1 – C2 (iii);

C3 (iv); or

Construct (v) which is an isolated or recombinant protein which comprises at least the following component:

C4 (v);

wherein C1 is a fourth promoter; C2 is the open reading frame (ORF) of a meganuclease; C3 are messenger RNA (mRNA) versions of said meganuclease; C4 is an isolated or recombinant protein of said meganuclease; wherein said meganuclease from constructs (iii), (iv) or (v) recognize and cleave A3; and wherein constructs (iii), (iv) or (v) are configured to be co-transfected with construct (ii) into said at least one target cell.

This system of genetic constructs allows the integration and expression of a GOI in an engineered cell at a specific genomic location. Construct (ii) containing a GOI which can be cloned into portion B3 is integrated into the genome via Meganuclease induced Recombination at a specific site corresponding to the genomic integration position of construct (i). The insertion event occurs at a very high frequency and is very specific.

Each of the genetic constructs consists of the above essential components, A1 to A5, A2' and A5', B1 to B4 and C1 to C2, but between these other nucleotide sequences may be present so long as they do not affect the properties of the claimed components as defined herein.

In the present invention, a promoter is a nucleotide sequence which when placed in combination with a second nucleotide sequence encoding an open reading frame causes the transcription of the open reading frame. In addition in the case of a RNA molecule, a promoter can also refer to a non-coding sequence which acts to increase the levels of translation of the RNA molecule.

In the present invention, a homologous portion refers to a nucleotide sequence which shares nucleotide residues in common with another nucleotide sequence so as to lead to a homologous recombination between these sequences, more particularly having at least 95 % identity, preferably 97 % identity and more preferably 99 % identity. The first and second homologous portions of construct (i) and the first and second homologous portions construct (ii) can be 100 % identical or less as indicated.

In particular the overlap between the portions A2 and A5 from construct (i) and portions A2' and A5' from construct (ii) is at least 200 bp and no more than 6000 bp. Preferably the overlap is between 1000 bp and 2000 bp.

In particular therefore components A2' and A5' from construct (ii), comprise at least 200 bp and no more than 6000 bp of components A2 and A5 from construct (i) respectively.

Most particularly components A2' and A5' from construct (ii), comprise at least 1000 bp and no more than 2000 bp of components A2 and A5 from construct (i) respectively.

The amounts of overlap necessary to allow efficient levels of homologous recombination are known in the art (49), starting from these known levels the inventors have identified the most efficient ranges of overlap for use with the set of constructs according to the present invention.

In the present invention, a meganuclease cleavage site is intended to mean a 22 to 24 bp double-stranded palindromic, partially palindromic (pseudo-palindromic) or non-palindromic polynucleotide sequence that is recognized and cleaved by a LAGLIDADG homing endonuclease. These terms refer to a distinct DNA location, preferably a genomic location, at which a double stranded break (cleavage) is to be induced by the meganuclease.

The meganuclease cleavage site can be the DNA sequence recognized and cleaved by a wild type meganuclease such as I-*Cre*I or I-*Dmo*I. Alternativley the meganuclease cleavage site can be the DNA sequence recognized and cleaved by altered meganucleases which recognize and cleave different DNA target sequences.

The inventors and others have shown that meganucleases can be engineered so as to recognize different DNA targets. The I-*Cre*I enzyme in particular has been studied extensively and different groups have used a semi-rational approach to locally alter the specificity of I-*Cre*I (26-28).

In addition, hundreds of I-*Cre*I derivatives with locally altered specificity were engineered by combining the semi-rational approach and High Throughput Screening:

- Residues Q44, R68 and R70 or Q44, R68, D75 and I77 of I-*Cre*I were mutagenized and a collection of variants with altered specificity at positions ± 3 to 5 of the DNA target (5NNN DNA target) were identified by screening (27, 28).

- Residues K28, N30 and Q38 or N30, Y33, and Q38 or K28, Y33, Q38 and S40 of I-*Cre*I were mutagenized and a collection of variants with altered specificity at positions ± 8 to 10 of the DNA target (10NNN DNA target) were identified by screening (29, 30).

All such variant meganucleases and the variant DNA targets which they recognize and cleave, are included in the present Patent Application and any combination of a particular meganuclease and its target can be used as the meganuclease target sequence represented by feature A3 from construct (i) and the meganuclease encoded variously by constructs (iii), (iv) and (v).

In the present invention a marker gene is a gene product which when expressed allows the differentiation of a cell or population of cells expressing the marker gene versus a cell or population of cells not expressing the marker gene.

In the present invention a multiple cloning site is a short segment of DNA which contains several restriction sites so as to allow the sub-cloning of a fragment of interest into the plasmid comprising the multiple cloning site.

In the present invention a meganuclease is intended to mean an endonuclease having a double-stranded DNA target sequence of 12 to 45 bp. This may be a wild type version of a meganuclease such as I-CreI or I-DmoI or an engineered version of one of these enzymes as described above or fusion proteins comprising portions of one or more meganuclease(s) (31-33).

The inventors have shown that this system can work with a number of diverse model mammalian cell lines for a number of GOIs.

Preferably component A5 comprises a marker gene or a portion thereof.

In accordance with this preferred embodiment of the present invention component A5 must encode a marker gene or a portion thereof such that following the homologous recombination event the detection of altered cells can be detected.

Alternatively a DNA sequence encoding a marker gene can be positioned after component A5, wherein this further portion encodes a marker gene and allows the detection of cells which have undergone homologous recombination.

Preferably component A5' comprises a 3' end deletion of said component A5.

Preferably the components of each of said constructs (i), (ii), (iii), (iv) and (v) are selected from the following groups:

Component	Group
A1	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
A2	EF1α intron 1 complete sequence, 5' homology (SEQ ID NO: 3)
A2'	EF1α intron 1 short sequence, 5' homology (SEQ ID NO: 29)
A3*	Meganuclease cleavage site (SEQ ID NO: 8)
A4	Hygromycin resistance gene (SEQ ID NO: 2); Neomycin resistance gene (SEQ ID NO: 7); Puromycin resistance gene (SEQ ID NO: 21)
A5	Neomycin resistance gene (SEQ ID NO: 7)
A5'	Inactive neomycin resistance gene deleted of its 3' end, 3' homology (SEQ ID NO: 13)
B1	Hygromycin resistance gene (SEQ ID NO: 2); Neomycin resistance gene (SEQ ID NO: 7); Puromycin resistance gene (SEQ ID NO: 21)
B2	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
B3	Multiple cloning site (SEQ ID NO: 23)
B4	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
C1	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
C2*	Meganuclease ORFs (SEQ ID NO: 14); (SEQ ID NO: 15)
C3*	Meganuclease ORFs (SEQ ID NO: 14); (SEQ ID NO: 15) and (SEQ ID NO 35)
C4*	Meganuclease peptide encoded by (SEQ ID NO: 14); (SEQ ID NO: 15) and (SEQ ID NO: 58)

* The meganuclease cleavage site used in the set of constructs according to the present invention must be recognized and cleaved by the meganuclease also included in the set of constructs. As pointed out above the meganuclease cleavage site can be a wild type meganuclease target site, such as SEQ ID NO: 8 the wild type cleavage site of

the wild type I-CreI meganuclease (provided herein in various forms as SEQ ID NO: 14, 15 and 58). If however component A3 is altered then the meganuclease of component C2, C3 or C4 will also be altered.

The above components are only examples and it is not intended that the present invention be limited to these specific sequences or combinations thereof. The characteristics of the claimed components are defined herein and the selection of other suitable components, such as resistance genes or promoter sequences is therefore encompassed by the present invention.

Preferred promoting sequences are pCMV promoter (SEQ ID NO: 25), pSV40 promoter (SEQ ID NO: 20), pEF1α (SEQ ID NO: 1) and Ubiquitin sub-unit c promoter (SEQ ID NO: 52).

Preferred marker genes are Neomycin resistance gene (SEQ ID NO: 7); Puromycin resistance gene (SEQ ID NO: 21), Hygromycin resistance gene (SEQ ID NO: 2); blasticidin resistance gene, zeocin resistance gene and phleomycin resistance gene. Many other selectable marker genes exist all these can be used in the present Patent Application

Most preferably, the construct (i) comprises SEQ ID NO: 6, which consists of the Inventors preferred construct which was used to create cGPS (cellular Genome Positioning System) cell lines. In this construct, a specific Meganuclease target site has been inserted into the host cell genome at a unique locus. This site is the precise insertion locus of the gene(s) of interest. This site has been inserted at a single copy into the host cell line as part of a larger construct. In the final cGPS cell line, the Meganuclease target site is located near the hygromycin resistance gene and downstream the EF1α promoter. The cGPS cell line is then resistant to hygromycin. Furthermore, the neomycin resistance gene is located just downstream the hygromycin gene but lacks a promoter making the cGPS cell line G418 sensitive (see Fig. 1).

The important features of cGPS locus are listed in table 1 below.

Table 1

Feature	Benefit
pEF1α (referred as A1)	Promoter from the human Elongation Factor I alpha gene driving the transcription of the puromycin resistance gene after HR in cGPS cells (SEQ ID NO: 1)
EF1α exon 1	Exon 1 of the human Elongation Factor I alpha gene (SEQ ID NO: 11)
EF1α exon 2	Exon 2 of the human Elongation Factor I alpha gene (SEQ ID NO: 12)
EF1α intron 1 (referred as A2, 5' homology)	Intron 1 of the human Elongation Factor I alpha gene, composed of a 1kb fragment (SEQ ID NO: 10)
Meganuclease cleavage site (referred as A3)	Meganuclease cleavage site for targeted insertion of the GOI
HygroR (referred as A4)	Hygromycin resistance gene
NeoR (referred as A5, 3' homology)	Neomycin resistance gene (inactive because lacking a promoting sequence to drive its transcription)
SV40 pA	Polyadenylation signal from SV40 virus (SEQ ID NO: 4, SEQ ID NO: 5), allowing efficient transcription termination and polyadenylation of hygromycin and neomycin resistance genes

Most preferably, the construct (ii) comprises SEQ ID NO: 22, which consists of pTV-DS-MCS2 which is a 6932 bp vector that expresses a GOI under the control of the CMV promoter. It also contains two homology arms for efficient HR and insertion of the GOI at the cGPS locus (see Fig. 2).

The important features of pTV-DS-MCS2 are described in table 2 below. All features have been functionally tested.

Table 2

Feature	Benefit
EF1α intron 1 (referred as A2', part of the 5' homology)	0,8 kb fragment (SEQ ID NO 28) for efficient homologous recombination at the cGPS site composed of the intron 1 of the human Elongation Factor I alpha gene; once reconstituted after HR at the cGPS site, it allows puromycin selection of stable cGPS expressing cell clones
NeoR Del3' (referred as A5', part of the 3'homology)	0.6 kb fragment (SEQ ID NO 26) for efficient homologous recombination at the cGPS site composed of an inactive neomycin resistance gene deleted of its 3' end; once reconstituted after HR at the cGPS site, it allows neomycin selection of stable expressing cell clones
PuroR	Puromycin resistance gene (inactive because lacking a promoting sequence to drive its transcription); once reconstituted after HR at the cGPS site, it allows puromycin selection of stable cell clones
pCMV	Human cytomegalovirus (CMV, SEQ ID NO: 25) immediate early promoter, driving high-level expression of the GOI
MCS2	Multiple cloning site containing NheI, BmtI, Bsu36I, AscI, BglII, BsrGI, BstBI, EcoRV, PacI, NotI restriction sites for the molecular cloning of GOIs
SV40 pA	Polyadenylation signal from SV40 virus (SEQ ID NO: 19), allowing efficient transcription termination and polyadenylation of the puromycin resistance gene
BGH pA	Polyadenylation signal from bovine growth hormone gene (SEQ ID NO: 27), allowing efficient transcription termination and polyadenylation of the mRNA of interest
pSV40	SV40 promoter (SEQ ID NO: 20) driving high-level expression of the neomycin resistance gene only after HR
pMB1 ORI	Permits high-copy number replication and growth in *E. coli bla* promoter
AmpR	Ampicillin (*bla*) resistance gene (β-lactamase), for selection of transformants in *E. coli*

Most preferably constructs (iii) comprise SEQ ID NO: 38 and SEQ ID NO: 39, which consist of pCLS1088 (Figure 33) or pCLS2147 (Figure 34),

respectively. These 5647 bp vectors contain two different ORFs of the Meganuclease under the control of the CMV promoter.

The important features of pCLS1088 and pCLS2147 are described in table 3 below. All features have been functionally tested.

Table 3

Feature	Benefit
pCMV	Human cytomegalovirus immediate early promoter; allowing high-level expression of your GOI (Andersson *et al.*, 1989; Boshart *et al.*, 1985; Nelson *et al.*, 1987)
Meganuclease(s)	Meganuclease ORFs (SEQ ID NO: 14 or SEQ ID NO: 15), improving HR events at the cGPS locus
TK pA	Polyadenylation signal from herpes simplex virus thymidine kinase gene, allowing efficient transcription termination and polyadenylation of the meganuclease mRNA
pUC & f1 origins	Permits high-copy number replication and growth in *E. coli*
AmpR	Ampicillin resistance gene (β-lactamase) for selection of transformants in *E. coli*

Most preferably constructs (iv) comprise ORFs of the Meganuclease (SEQ ID NO: 14 and SEQ ID NO: 15).

Wherein constructs (iv) consist of Meganuclease polyadenylated mRNAs (SEQ ID NO 34, SEQ ID NO 35), from which the ribosomal scanning is mediated either by 7-methyl-guanine capped sequence or by internal ribosome entry site (IRES). (see fig. 3).

Wherein constructs (v) consist of a cell penetrating peptide fused to the N-terminal part of Meganuclease. An example of a meganuclease according to this aspect of the present invention is provided as SEQ ID NO: 58, this sequence encodes an I-CreI monomer with the cell-penetrating peptide DPV15b (SEQ ID NO: 56) fused to the N-terminal of the meganuclease and a 6x hisitidine tag fused at the C-terminal of the meganuclease. The Inventors have also evaluated another cell-penetrating peptide DPV1047 (SEQ ID NO: 57).

Cell penetrating peptides were initially developed following the observation that certain proteins, including the HIV-1 protein Tat, could cross the cell

membrane (34). The HIV-1 transcriptional activator Tat is a multifunctional protein that, in addition to acting as a powerful inducer of viral gene expression, is transported in and out of the cells (35). This cell penetration property relies on the integrity of a highly basic arginine-rich sequence (amino acids 49–58).

Peptides containing this arginine-rich sequence have been developed, named Tat peptides, that after conjugation to a range of macromolecules can facilitate cellular entry of the conjugate. This method of intracellular delivery has been used successfully *in vitro* for a range of macromolecules including fluorochromes, enzymes, antibodies and liposomes (41, 42, 43, 45, 47, 48). The Tat peptide has also been shown to facilitate cellular entry of functional proteins such as β-galactosidase *in vivo* (46).

A number of other proteins and their peptide derivatives have been found to possess similar cell internalization properties including the herpes virus tegument protein VP22 (37), the homoeotic protein of *Drosophila melanogaster* antennapedia (Antp), (the internalizing peptide derived from full length Antp is called penetratin) (36), the protegrin 1 (PG-1) antimicrobial peptide SynB (40) and the basic fibroblast growth factor (39). The carrier peptides derived from these proteins show little sequence homology with each other, but are all highly cationic and arginine- or lysine-rich. Following on from this observation, synthetic polyarginine peptides have been shown to be internalized with a high level of efficiency (38, 44).

All such cell-penetrating peptides which can cause an increase in the rate of internalisation of a meganuclease linked thereto are incorporated in the present Patent Application.

According to a second aspect of the present invention there is provided a kit to introduce a sequence encoding a GOI into at least one cell, comprising the set of genetic constructs according to the first aspect of the present invention; and instructions for the generation of a transformed cell using said set of genetic constructs.

Preferably the kit, further comprising construct (vi) consisting of SEQ ID NO: 17 (Lac-Z) which consists of pTV-DS-LacZ.

pTV-DS-LacZ is a 9981 bp vector that expresses LacZ (as a positive control) in place of the GOI under the control of the CMV promoter as previously described (23). It also contains two homology arms for efficient homologous recombination and insertion of the GOI at the cGPS locus. Figure 4 summarizes the features of the vector. Features of pTV-DS-LacZ are composed of the very same features as pTV-DS-MCS2 but this plasmid contains the LacZ gene encoding the β-galactosidase protein under the control of the CMV promoter. It can be used as a positive control for HR at the cGPS site.

Preferably, the kit further comprises at least one cell stably transformed with said construct (i).

Most preferably the at least one cell is selected from the group comprising: CHO-K1 cells (Sigma-Aldrich); HEK-293-derived cells (Invitrogen); Caco2 cells (Invitrotech); U2-OS cells (Invitrogen); NIH 3T3 cells (Invitrogen); NSO cells (Sigma-Aldrich); SP2 cells (Sigma-Aldrich); CHO-S cells (Invitrogen); DG44 cells (Invitrogen).

According to a third aspect of the present invention there is provided a method for transforming by HR at least one cell comprising the steps of:

A method for transforming by homologous recombination at least one cell comprising the steps of:

a) stably transforming at least one cell by inserting construct (i) as defined above into the genome of said at least one cell;

b) cloning a sequence coding for a gene of interest into position B3 of construct (ii) as defined above;

c) co-transfecting said cell of step a), with said construct (ii) of step b) and constructs (iii), (iv) or (v) as defined above;

d) following homologous recombination between said construct (ii) and said stably inserted construct (i), selecting at least one cell from step c) based upon: the absence of a first marker gene encoded by component A4 of said construct (i) and the activity of a second marker gene encoded by component B1 and the activity of a third marker gene encoded by component A5.

Most preferably, selection in step d) is carried out sequentially for each of said first marker, said second marker and said third marker.

For a better understanding of the invention and to show how the same may be carried into effect, there will now be shown by way of example only, specific embodiments, methods and processes according to the present invention with reference to the accompanying drawings in which:

Figure 1 shows a schematic representation of the cGPS locus.

Figure 2 shows a schematic representation of the pTV-DS-MCS2 vector.

Figure 3 shows a schematic representation of a Meganuclease capped polyadenylated mRNA.

Figure 4 shows a schematic representation of the pTV-DS-LacZ vector.

Figure 5 shows a schematic representation of a transfection protocol for the cGPS CHO-K1 cell line according to the present invention.

Figure 6 shows a schematic representation of the clonal selection protocol (left column) and batch selection protocol (right column) for the cGPS CHO-K1 cell line.

Figure 7 shows a schematic representation of a transfection protocol for the cGPS NIH 3T3 cell line according to the present invention.

Figure 8 shows a schematic representation of the clonal selection protocol (left column) and batch selection protocol (right column) for the cGPS NIH 3T3 cell line.

Figure 9 shows the production of double resistant lacZ targeted clones in the cGPS NIH 3T3 system and their molecular characterization.

Figure 10 shows a schematic representation of a transfection protocol for the cGPS HEK 293 cell line according to the present invention.

Figure 11 shows a schematic representation of the clonal selection protocol (left column) and batch selection protocol (right column) for the cGPS HEK 293 cell line.

Figure 12 shows the production of double resistant lacz targeted clones in the cGPS NIH 3T3 system and their molecular characterization.

Figure 13 shows a Southern blot analysis of 25 selected clones targeted with the pTV-DS-LacZ vector.

Figure 14 shows the results of experiments to determine by FACS analysis the stability of human CD4 expression over time.

Figure 15 shows the results of a functional assay for cAMP production inhibition performed upon clones targeted with the somatostatin receptor (GPCR-SSTR2) as GOI.

Figure 16 shows the results of hATX expression from cGPS CHO-K1 hATX.

Figure 17 shows the results of radioligand saturation experiments performed on cGPS CHO-K1 hMT1-targeted clones.

Figure 18 shows the results of radioligand saturation experiments performed on cGPS CHO-K1 hMT2-targeted clones.

Figure 19 shows the results of experiments to determine the homogeneity of expression levels of several clones expressing a monoclonal antibody that is controlled by the Ubiquitin sub unit c promoter (pUbc).

Figure 20 shows the stability of expression of β-galactosidase (mean value for 4 cGPS CHO-K1 lacz targeted clones) and luciferase (mean value for 4 cGPS CHO-K1 luciferase targeted clones) over a period of 23 weeks in the presence of the selecting agents and over a period of 15 weeks in the absence of selecting agents.

Figure 21 shows the efficacy of gene targeting through the use of a I-CreI recombinant protein fused to a penetrating peptide (DPV15b/I-CreI N75/6xHis).

Figure 22 shows a schematic representation of the cGPS custom CHO-K1 lacz integration matrix vector.

Figure 23 shows a schematic representation of the Sc MA17-RM2-G19H33 meganuclease expression vector.

Figure 24 shows a molecular characterization by southern blot of double targeted clones.

Figure 25 shows the stability of expression of β-galactosidase and luciferase (mean value for 4 double targeted clones) over a period of 11 weeks.

Figure 26 shows a schematic representation of the pTV-DS-CD4 vector.

Figure 27 shows a schematic representation of the pTV-DS-SSTR2 vector.

Figure 28 shows a schematic representation of the pTV-DS-hATX vector.

Figure 29 shows a schematic representation of the pTV-DS-hMT1 vector.

Figure 30 shows a schematic representation of the pTV-DS-hMT2 vector.

Figure 31 shows a schematic representation of the pTV-DS-5F11 vector.

Figure 32 shows a schematic representation of the pTV-DS-luciferase vector.

Figure 33 shows a schematic representation of the I-CreI N75 meganuclease expression vector.

Figure 34 shows a schematic representation of the I-CreI N75 105A 132V meganuclease expression vector.

There will now be described by way of example a specific mode contemplated by the Inventors. In the following description numerous specific details are set forth in order to provide a thorough understanding. It will be apparent however, to one skilled in the art, that the present invention may be practiced without limitation to these specific details. In other instances, well known methods and structures have not been described so as not to unnecessarily obscure the description.

Definitions

- Amino acid residues in a polypeptide sequence are designated herein according to the one-letter code, in which, for example, Q means Gln or Glutamine residue, R means Arg or Arginine residue and D means Asp or Aspartic acid residue.

- hydrophobic amino acid refers to leucine (L), valine (V), isoleucine (I), alanine (A), methionine (M), phenylalanine (F), tryptophane (W) and tyrosine (Y).

- Nucleotides are designated as follows: one-letter code is used for designating the base of a nucleoside: a is adenine, t is thymine, c is cytosine, and g is guanine. For the degenerated nucleotides, r represents g or a (purine nucleotides), k represents g or t, s represents g or c, w represents a or t, m represents a or c, y represents t or c (pyrimidine nucleotides), d represents g, a or t, v represents g, a or c, b represents g, t or c, h represents a, t or c, and n represents g, a, t or c.

- by "meganuclease" is intended an endonuclease having a double-stranded DNA target sequence of 12 to 45 bp. Examples include *I-Sce I, I-Chu I, I-Cre I, I-Csm I, PI-Sce I, PI-Tli I, PI-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, PI-Civ I, PI-Ctr I, PI-Aae I, PI-Bsu I, PI-Dha I, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu I, PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, PI-Tag I, PI-Thy I, PI-Tko I, PI-Tsp I, I-MsoI.*

- by "parent LAGLIDADG homing endonuclease" is intended a wild-type LAGLIDADG homing endonuclease or a functional variant thereof. Said parent LAGLIDADG homing endonuclease may be a monomer, a dimer (homodimer or heterodimer) comprising two LAGLIDADG homing endonuclease core domains which are associated in a functional endonuclease able to cleave a double-stranded DNA target of 22 to 24 bp.

- by "homodimeric LAGLIDADG homing endonuclease" is intended a wild-type homodimeric LAGLIDADG homing endonuclease having a single LAGLIDADG motif and cleaving palindromic DNA target sequences, such as *I-CreI* or I-*Mso*I or a functional variant thereof.

- by "LAGLIDADG homing endonuclease variant" or "variant" is intended a protein obtained by replacing at least one amino acid of a LAGLIDADG homing endonuclease sequence, with a different amino acid.

- by "functional variant" is intended a LAGLIDADG homing endonuclease variant which is able to cleave a DNA target, preferably a new DNA target which is not cleaved by a wild type LAGLIDADG homing endonuclease. For example, such variants have amino acid variation at positions contacting the DNA target sequence or interacting directly or indirectly with said DNA target.

- by "homing endonuclease variant with novel specificity" is intended a variant having a pattern of cleaved targets (cleavage profile) different from that of the parent homing endonuclease. The variants may cleave less targets (restricted profile) or more targets than the parent homing endonuclease. Preferably, the variant is able to cleave at least one target that is not cleaved by the parent homing endonuclease.

The terms "novel specificity", "modified specificity", "novel cleavage specificity", "novel substrate specificity" which are equivalent and used indifferently, refer to the specificity of the variant towards the nucleotides of the DNA target sequence.

- by "*I-CreI*" is intended the wild-type *I-CreI* having the sequence SWISSPROT P05725 or pdb accession code 1g9y (SEQ ID NO: 36) .

-by "*I-DmoI*" is intended the wild-type *I-DmoI* having the sequence SWISSPROT number P21505 (SEQ ID NO: 37) or the structure PDB code 1b24

- by "domain" or "core domain" is intended the "LAGLIDADG homing endonuclease core domain" which is the characteristic αββαββα fold of the homing endonucleases of the LAGLIDADG family, corresponding to a sequence of about one hundred amino acid residues. Said domain comprises four beta-strands folded in an antiparallel beta-sheet which interacts with one half of the DNA target. This domain is able to associate with another LAGLIDADG homing endonuclease core domain which interacts with the other half of the DNA target to form a functional endonuclease able to cleave said DNA target. For example, in the case of the dimeric homing endonuclease *I-CreI* (163 amino acids), the LAGLIDADG homing

endonuclease core domain corresponds to the residues 6 to 94. In the case of monomeric homing endonucleases, two such domains are found in the sequence of the endonuclease; for example in *I-DmoI* (194 amino acids), the first domain (residues 7 to 99) and the second domain (residues 104 to 194) are separated by a short linker (residues 100 to 103).

- by "subdomain" is intended the region of a LAGLIDADG homing endonuclease core domain which interacts with a distinct part of a homing endonuclease DNA target half-site. Two different subdomains behave independently or partly independently, and the mutation in one subdomain does not alter the binding and cleavage properties of the other subdomain, or does not alter it in a number of cases. Therefore, two subdomains bind distinct part of a homing endonuclease DNA target half-site.

- by "beta-hairpin" is intended two consecutive beta-strands of the antiparallel beta-sheet of a LAGLIDADG homing endonuclease core domain which are connected by a loop or a turn,

- by "C1221" it is intended to refer to the first half of the *I-CreI* target site '12' repeated backwards so as to form a palindrome '21'.

- by "cleavage activity" the cleavage activity of the variant of the invention may be measured by a direct repeat recombination assay, in yeast or mammalian cells, using a reporter vector, as described in the PCT Application WO 2004/067736; Epinat *et al.*, Nucleic Acids Res., 2003, 31, 2952-2962; Chames et al., Nucleic Acids Res., 2005, 33, e178, and Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458. The reporter vector comprises two truncated, non-functional copies of a reporter gene (direct repeats) and a chimeric DNA target sequence within the intervening sequence, cloned in a yeast or a mammalian expression vector. The DNA target sequence is derived from the parent homing endonuclease cleavage site by replacement of at least one nucleotide by a different nucleotide. Preferably a panel of palindromic or non-palindromic DNA targets representing the different combinations of the 4 bases (g, a, c, t) at one or more positions of the DNA cleavage site is tested (4^n palindromic targets for n mutated positions). Expression of the variant results in a functional endonuclease which is able to cleave the DNA target sequence. This

cleavage induces homologous recombination between the direct repeats, resulting in a functional reporter gene, whose expression can be monitored by appropriate assay.

- by "DNA target", "DNA target sequence", "target sequence", "target-site", "target", "site"; "recognition site", "recognition sequence", "homing recognition site", "homing site", "cleavage site" is intended a 22 to 24 bp double-stranded palindromic, partially palindromic (pseudo-palindromic) or non-palindromic polynucleotide sequence that is recognized and cleaved by a LAGLIDADG homing endonuclease. These terms refer to a distinct DNA location, preferably a genomic location, at which a double stranded break (cleavage) is to be induced by the endonuclease. The DNA target is defined by the 5' to 3' sequence of one strand of the double-stranded polynucleotide. For example, the palindromic DNA target sequence cleaved by wild type *I-CreI* is defined by the sequence 5'- $t_{-12}c_{-11}a_{-10}a_{-9}a_{-8}a_{-7}c_{-6}g_{-5}t_{-4}c_{-3}g_{-2}t_{-1}a_{+1}c_{+2}g_{+3}a_{+4}c_{+5}g_{+6}t_{+7}t_{+8}t_{+9}t_{+10}g_{+11}a_{+12}$ (SEQ ID NO:8). Cleavage of the DNA target occurs at the nucleotides in positions +2 and -2, respectively for the sense and the antisense strand. Unless otherwise indicated, the position at which cleavage of the DNA target by a meganuclease variant occurs, corresponds to the cleavage site on the sense strand of the DNA target.

- by " DNA target half-site", "half cleavage site" or half-site" is intended the portion of the DNA target which is bound by each LAGLIDADG homing endonuclease core domain.

- by "chimeric DNA target"or "hybrid DNA target" is intended the fusion of a different half of two parent meganuclease target sequences. In addition at least one half of said target may comprise the combination of nucleotides which are bound by separate subdomains (combined DNA target).

- by "mutation" is intended the substitution, the deletion, and/or the addition of one or more nucleotides/amino acids in a nucleic acid/amino acid sequence.

- by "homologous" is intended a sequence with enough identity to another one to lead to a homologous recombination between sequences, more particularly having at least 95 % identity, preferably 97 % identity and more preferably 99 %.

- "Identity" refers to sequence identity between two nucleic acid molecules or polypeptides. Identity can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base, then the molecules are identical at that position. A degree of similarity or identity between nucleic acid or amino acid sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. Various alignment algorithms and/or programs may be used to calculate the identity between two sequences, including FASTA, or BLAST which are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with, e.g., default settings.

- "individual" includes mammals, as well as other vertebrates (e.g., birds, fish and reptiles). The terms "mammal" and "mammalian", as used herein, refer to any vertebrate animal, including monotremes, marsupials and placental, that suckle their young and either give birth to living young (eutharian or placental mammals) or are egg-laying (metatharian or nonplacental mammals). Examples of mammalian species include humans and other primates (e.g., monkeys, chimpanzees), rodents (e.g., rats, mice, guinea pigs) and ruminants (e.g., cows, pigs, horses).

- “gene of interest” or “GOI” refers to any nucleotide sequence encoding a known or putative gene product.

- “genetic disease” refers to any disease, partially or completely, directly or indirectly, due to an abnormality in one or several genes. Said abnormality can be a mutation, an insertion or a deletion. Said mutation can be a punctual mutation. Said abnormality can affect the coding sequence of the gene or its regulatory sequence. Said abnormality can affect the structure of the genomic sequence or the structure or stability of the encoded mRNA. This genetic disease can be recessive or dominant. Such genetic disease could be, but are not limited to, cystic fibrosis, Huntington’s chorea, familial hyperchoiesterolemia (LDL receptor defect), hepatoblastoma, Wilson’s disease, congenital hepatic porphyrias, inherited disorders of hepatic metabolism, Lesch Nyhan syndrome, sickle cell anemia, thalassaemias, xeroderma pigmentosum, Fanconi’s anemia, retinitis pigmentosa, ataxia

telangiectasia, Bloom's syndrome, retinoblastoma, Duchenne's muscular dystrophy, and Tay-Sachs disease.

- "cGPS site" or "cGPS locus" refers to the genomic location at which the essential components of construct (i) have been introduced stably.

- "cGPS cell line" refers to at least one cell in which the "cGPS site" or "cGPS locus" is present.

- "cell-penetrating peptide" or "CPP" refers to peptides that facilitate cellular uptake of various molecular cargo in particular proteins and large macromolecules which would not normally be able to pass through the cell membrane at a rate sufficient for the cargo to have any effect upon the target cell.

- "EF1α" refers to the human gene which encodes an isoform of the alpha subunit of the elongation factor-1 complex, which is responsible for the enzymatic delivery of aminoacyl tRNAs to the ribosome. The sequence of the human EF1α promoter, first and second exon and first intron is provided as SEQ ID NO: 9; the sequence of the human EF1α first intron is provided as SEQ ID NO: 10; the sequence of the human EF1α first exon is provided as SEQ ID NO: 11 and the sequence of the human EF1α second exon is provided as SEQ ID NO: 12.

- "vectors": a vector which can be used in the present invention for instance as construct (ii) or (iii) as defined above includes, but is not limited to, a viral vector, a plasmid, a RNA vector or a linear or circular DNA or RNA molecule which may consists of a chromosomal, non chromosomal, semi-synthetic or synthetic nucleic acids. Preferred vectors are those capable of autonomous replication (episomal vector) and/or expression of nucleic acids to which they are linked (expression vectors). Large numbers of suitable vectors are known to those of skill in the art and commercially available.

Viral vectors include retrovirus, adenovirus, parvovirus (e. g. adeno-associated viruses), coronavirus, negative strand RNA viruses such as orthomyxovirus (e. g., influenza virus), rhabdovirus (e. g., rabies and vesicular stomatitis virus), paramyxovirus (e. g. measles and Sendai), positive strand RNA viruses such as picornavirus and alphavirus, and double-stranded DNA viruses including adenovirus, herpesvirus (e. g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomega-

lovirus), and poxvirus (e. g., vaccinia, fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, and hepatitis virus, for example. Examples of retroviruses include: avian leukosis-sarcoma, mammalian C-type, B-type viruses, D type viruses, HTLV-BLV group, lentivirus, spumavirus (Coffin, J. M., Retroviridae: The viruses and their replication, In Fundamental Virology, Third Edition, B. N. Fields, et al., Eds., Lippincott-Raven Publishers, Philadelphia, 1996). The term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors. A vector according to the present invention comprises, but is not limited to, a YAC (yeast artificial chromosome), a BAC (bacterial artificial), a baculovirus vector, a phage, a phagemid, a cosmid, a viral vector, a plasmid, a RNA vector or a linear or circular DNA or RNA molecule which may consist of chromosomal, non chromosomal, semi-synthetic or synthetic DNA. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. Large numbers of suitable vectors are known to those of skill in the art.

Vectors can comprise selectable markers, for example: neomycin phosphotransferase, histidinol dehydrogenase, dihydrofolate reductase, hygromycin phosphotransferase, herpes simplex virus thymidine kinase, adenosine deaminase, glutamine synthetase, and hypoxanthine-guanine phosphoribosyl transferase for eukaryotic cell culture; TRP1 for *S. cerevisiae*; tetracycline, rifampicin or ampicillin resistance in *E. coli*. These selectable markers can also be used as a part of the constructs (i) and (ii) according to the present invention.

Preferably said vectors are expression vectors, wherein a sequence encoding a polypeptide of the invention is placed under control of appropriate transcriptional and translational control elements to permit production or synthesis of

said protein. Therefore, said polynucleotide is comprised in an expression cassette. More particularly, the vector comprises a replication origin, a promoter operatively linked to said encoding polynucleotide, a ribosome site, an RNA-splicing site (when genomic DNA is used), a polyadenylation site and a transcription termination site. It also can comprise enhancer or silencer elements. Selection of the promoter will depend upon the cell in which the polypeptide is expressed.

EXAMPLE 1: GENERATING STABLE EUKARYOTIC CELL LINES TRANSFECTED WITH CONSTRUCT (i)

Construct (i) can be stably transfected into cells using known techniques. There are various methods of introducing foreign DNA into a eukaryotic cell and many materials have been used as carriers for transfection, which can be divided into three kinds: (cationic) polymers, liposomes and nanoparticles. Other methods of transfection include nucleofection, electroporation, heat shock, magnetofection and proprietary transfection reagents such as Lipofectamine®, Dojindo Hilymax®, Fugene®, JetPEI®, Effectene®, DreamFect®, PolyFect®, Nucleofector®, Lyovec®, Attractene®, Transfast®, Optifect®.

1.1 CHO-K1

In the present example construct (i), consists of SEQ ID NO: 6, which comprises a specific Meganuclease target site which has been inserted into the CHO-K1 genome at a unique locus. This site is the precise insertion locus of GOIs. This site has been inserted as a single copy into the CHO-K1 Cell line as part of a larger construct. In the final Cell Line, termed cGPS-CHO-K1, the Meganuclease target site is located near the hygromycin resistance gene and downstream the EF1 alpha promoter. The cGPS-CHO-K1 cell line is then resistant to hygromycin. Furthermore, the neomycin resistance gene is located just downstream the hygromycin gene but lacks a promoter making the cGPS-CHO-K1 cell line G418 sensitive (see Fig 1).

1.2 NIH 3T3

In the present example construct (i) consists of SEQ ID NO: 6, which comprises a specific Meganuclease target site which has been inserted into the NIH 3T3 genome at a unique locus. This site is the precise insertion locus of GOIs. This

site has been inserted as a single copy into the NIH 3T3 Cell line as part of a larger construct. In the final Cell Line, termed cGPSNIH 3T3, the Meganuclease target site is located near the hygromycin resistance gene and downstream the EF1 alpha promoter. The cGPSNIH 3T3 cell line is then resistant to hygromycin. Furthermore, the neomycin resistance gene is located just downstream the hygromycin gene but lacks a promoter making the cGPSNIH 3T3 cell line G418 sensitive (see Fig 1).

1.3 HEK 293

In the present example construct (i), consists of SEQ ID NO: 6, which comprises a specific Meganuclease target site which has been inserted into the human HEK 293 genome at a unique locus. This site is the precise insertion locus of GOIs. This site has been inserted as a single copy into the HEK 293 Cell line as part of a larger construct. In the final Cell Line, termed cGPSHEK 293, the Meganuclease target site is located near the hygromycin resistance gene and downstream the EF1 alpha promoter. The cGPSHEK 293 cell line is then resistant to hygromycin. Furthermore, the neomycin resistance gene is located just downstream the hygromycin gene but lacks a promoter making the cGPSHEK 293 cell line G418 sensitive (see Fig 1).

<u>EXAMPLE 2:</u> HIGHLY EFFICIENT TARGETED INSERTION OF GENE OF INTEREST (GOI) IN cGPS CELL LINE.

2.1 CLONING OF GOI INTO CONSTRUCT (ii)

The first step for generating a cell line expressing a GOI is to sub-clone the GOI into the pTV-DS-MCS2 vector (SEQ ID NO: 22; see Fig. 2). For this purpose a multiple cloning site has been introduced therein.

The expression of the gene of interest will be controlled by a CMV promoter and the bovine growth hormone (BGH) polyadenylation signal. The pTV-DS-MCS2 plasmid contains all the characteristics to favor a highly efficient HR event at the cGPS locus (SEQ ID NO: 6). A left homology arm (corresponding to portion A2 of construct (i) and A2' of construct (ii)) is composed of a 0,8kb fragment homologous to the genomic 1kb upstream the Meganuclease target site in the cGPS Cell Line. A right homology arm (corresponding to A5 of construct (i) and to A5' of

construct (ii)) is composed of a 0.6kb fragment homologous to the genomic 0.8kb downstream the Meganuclease target site in the cGPS Cell Line.

Both homology arms are separated by (i) the puromycin resistance gene (which lacks a promoter on the plasmid), (ii) a CMV promoter for the expression of the GOI, (iii) a multiple cloning site for the insertion of the GOI, (iv) a polyadenylation signal controlling the stability of the mRNA for the GOI, and (v) a modified neomycin resistance gene. By itself, the pTV-DS-MCS2 plasmid cannot induce a puromycin and neomycin resistance phenotype.

For example, the pTV-DS-LacZ plasmid (SEQ ID NO: 17; see Fig.4) has been obtained by cloning the lacZ gene into the multiple cloning site of the pTV-DS-MCS2. The pTV-DS-LacZ plasmid (SEQ ID NO: 17) can be used as a positive control vector for mammalian cell transfection. It may be used to assay for expression levels in the cGPS cell lines. The sequence of the cGPS site following insertion of the lacZ gene is provided as SEQ ID NO: 18, the sequence of the cGPS site prior to insertion of the lacZ gene is provided as SEQ ID NO: 24.

Insertion of LacZ at the cGPS locus can be monitored as if it was the GOI.

2.2 cGPS CHO-K1 CELL LINE

2.2.1 cGPSCHO-K1 CULTURE CONDITIONS AND TRANSFECTION

cGPSCHO-K1 cells are sub-cultured in F-12K complete medium supplemented with 0.6 mg/ml of hygromycin. cGPSCHO-K1 cells are passed twice a week at 1:10 – 1:40 ratio.

Media and supplements

- Complete medium: F-12K medium (Invitrogen-Life Science) is supplemented with 2 mM L-glutamine, penicilline (100 UI/ml), streptomycine (100 µg/ml), amphotericine B (Fongizone) (0.25 µg/ml) and 10% FBS.
- PBS
- Hygromycin B solution (Sigma).
- Puromycin dichloride (Sigma).
- G418 sulfate (Invitrogen-Life Science).
- Trypsin-EDTA solution (Invitrogen-Life Science).

- Freezing medium: F12K complete medium supplemented with 10% DMSO.

Transfection.

One day prior to transfection, the cGPSCHO-K1 cells are seeded in 10 cm tissue culture dishes ($2x10^5$ cells per dish) in complete F-12K medium.

On D day, 2µg of pTV plasmid versions (pTV-DS-MCS2 containing any GOI or pTV-DS-LacZ) and 1µg of meganuclease constructs (pCLS1088 or pCLS2147 plasmid DNAs, or meganuclease-encoding mRNAs) are diluted in EC-R buffer. 6 µl of Enhancer Reagent is added (ratio nucleic acid (µg):enhancer (µl) should be 1:2).

Total volume DNA: enhancer EC-R buffer should be 100µl. Vortex 10'' and incubate 5' at room temperature.

Add 24 µl of TransMessenger™ (Qiagen) reagent (ratio nucleic acid (µg):TransMessenger™ (µl) should be 1:8) to the mix. Vortex 10'' and incubate 10' at room temperature.

Meantime, replace culture medium with 9 ml of fresh medium.

Add 900µl of serum- and antibiotic-free medium of the transfection mix and dispense over plated cells.

Incubate dish in a 37°C, 5% CO_2 humidified incubator.

A schematic representation of the transfection protocol is shown in Fig. 5.

2.2.2 cGPS CHO-K1 TARGETED CLONES SELECTION

2.2.2.1 cGPS CHO-K1 CLONAL SELECTION

Clonal selection is a longer but better protocol to select the proper cell line expressing the GOI. Fig. 6 shows a schematic representation of the clonal selection protocol (left column).

cGPS CHO-K1 cells are transfected with the protocol described above (2.2.1). 24 hours after transfection, the cells are washed and fresh medium supplemented with 0.6mg/ml of G418 is added.

After 10 days of G418 selection, the culture medium is replaced with complete medium supplemented with G418 at 0.6 mg/ml and puromycin at 10µg/ml.

3 to 4 days later, double resistant clones are picked up and seeded in a 96 well plate. Double resistant clones are amplified to reach confluence of a 10cm culture dish

7 to 10 days later, double resistant clones can be characterized by analytical PCR and Southern blotting experiments. Positive control clones can be assayed for β-galactosidase activity, if pTV-DS-LacZ has been used as a positive control.

2.2.2.2 cGPS CHO-K1 POPULATION SELECTION

In addition to the clonal selection described above, a population selection procedure can be used to retrieve the insertion clones. This procedure is faster and much easier to handle; however, the inventors believe that the clonal procedure is better to obtain pure single insertion clones.

For example, cGPS CHO-K1 cells are transfected with the protocol described above (2.2.1).

24 hours after transfection, wash the cells and add fresh medium supplemented with 0.6mg/ml of G418.

10 days after G418 selection, wash the cells and add fresh medium supplemented with G418 at 0.6 mg/ml and puromycin at 10μg/ml.

4 to 10 days later, double resistant population can be amplified in complete medium supplemented with the two selective agents.

Fig. 6 shows a schematic representation of the population selection protocol (right column).

Targeted insertion for different GOIs in the cGPS CHO-K1 are presented in examples 3, 4, 5, 6 and 7.

2.3 cGPS NIH 3T3 CELL LINE

2.3.1 cGPS NIH 3T3 CULTURE CONDITIONS AND TRANSFECTION

cGPS NIH 3T3 cells are sub-cultured in DMEM complete medium supplemented with 0.6 mg/ml of hygromycin. cGPS NIH 3T3 cells are passed twice a week at 1:3 – 1:10 ratio.

Media and supplements

- Complete medium: DMEM medium (Invitrogen-Life Science) is supplemented with 2mM L-glutamine, penicilline (100 UI/ml), streptomycine (100 µg/ml), amphotericine B (Fongizone) (0.25 µg/ml) and 10% FBS.
- PBS
- Hygromycin B solution (Sigma).
- Puromycin dichloride (Sigma).
- G418 sulfate (Invitrogen-Life Science).
- Trypsin-EDTA solution (Invitrogen-Life Science).
- Freezing medium: DMEM complete medium supplemented with 10% DMSO.

Transfection.

One day prior to transfection, the cGPS NIH 3T3 cells are seeded in 10 cm tissue culture dishes ($2.5x10^5$ cells per dish) in complete medium.

On D day, 1 µg of pTV plasmid versions (pTV-DS-MCS2 containing any GOI or pTV-DS-LacZ) and 1 µg of meganuclease constructs (pCLS1088 or pCLS2147 plasmid DNAs, or meganuclease-encoding mRNAs) are diluted in 300 µl of EC buffer. 16 µl of Enhancer Reagent is added (ratio nucleic acid (µg):enhancer (µl) should be 1:8).

Total volume DNA: EC buffer should be 300µl. Vortex lightly and incubate 5' at room temperature.

Add 40 µl of Effectene™ (Qiagen) reagent (ratio nucleic acid (µg):Effectene™ (µl) should be 1:20) to the mix. Vortex 10'' and incubate 10' at room temperature.

Meantime, replace culture medium with 9 ml of fresh medium.

Add 1 ml of complete medium of the transfection mix and dispense over plated cells.

Incubate dish in a 37°C, 5% CO_2 humidified incubator.

A schematic representation of the transfection protocol is shown in Fig. 7.

2.3.2 cGPS NIH 3T3 TARGETED CLONES SELECTION

Clonal selection is a longer but better protocol to select the proper cell line expressing the GOI. Fig. 8 shows a schematic representation of the clonal selection protocol (left column).

cGPS NIH 3T3 cells are transfected with the protocol described above (2.3.1). 24 hours after transfection, the cells are washed and fresh medium supplemented with 0.4mg/ml of G418 is added.

10 days after G418 selection, single colony clones are picked up and seeded in 96 well plates in complete medium supplemented with G418 at 0.4 mg/ml and puromycin at 3 µg/ml.

6 to 7 days later, double resistant clones can be amplified in complete medium supplemented with the two selective agents.

7 to 10 days later, double resistant clones can be characterized by analytical PCR and Southern blotting experiments. Positive control clones can be assayed for β-galactosidase activity, if pTV-DS-LacZ has been used as a positive control.

As shown on figure 9 panel A, double resistant clones are only present upon co-transfection with meganuclease expression vector and pTV-DS-Lacz. No double resistant clone is obtained in the controls (i.e. meganuclease expression vector transfected cells (negative control 1) or pTV-DS-lacZ (negative control 2) transfected cells). Furthermore, near all double resistant clones expressed the lacz gene as revealed by X-gal staining (see example 6 for materials and methods). Several double resistant clones are picked at random and amplified for molecular characterization by Southern blotting (see section 2.5). As shown on figure 9 panel B, genomic DNA from 14 double resistant clones is digested with the RsrII restriction enzyme, immobilized on nitrocellulose membrane and hybridized with a ^{32}P-neomycine probe. A correct targeted insertion is characterized by the identification of a band at 6.3 kb. Such band is identified for all double resistant clones analyzed, while a band at 1.6 kb is shown for the parental cGPS NIH 3T3. In addition to the 6.3 kb band, others bands are present for 4/14 clones and is probably due to an additional random insertion. Hence, these results demonstrate that cGPS NIH 3T3 double

resistant clones, obtained with the method described above, expressed the reported gene, present on the integration matrix, that is correctly targeted in the cGPS locus.

2.4 cGPS HEK 293 CELL LINE

2.4.1 cGPS HEK 293 CULTURE CONDITIONS AND TRANSFECTION

cGPS HEK 293 cells are sub-cultured in DMEM complete medium supplemented with 0.1 mg/ml of hygromycin. cGPS HEK 293 cells are passed twice a week at 1:3 – 1:10 ratio.

Media and supplements

- Complete medium: DMEM medium Glutamax (Invitrogen-Life Science) is supplemented with penicilline (100 UI/ml), streptomycine (100 µg/ml), amphotericine B (Fongizone) (0.25 µg/ml) and 10% FBS.
- PBS
- Hygromycin B solution (Sigma).
- Puromycin dichloride (Sigma).
- G418 sulfate (Invitrogen-Life Science).
- Trypsin-EDTA solution (Invitrogen-Life Science).
- Freezing medium: DMEM complete medium supplemented with 10% DMSO.

Transfection.

One day prior to transfection, the stable cGPSHEK 293 cells are seeded in 10 cm tissue culture dishes (10^6 cells per dish) in complete medium.

On D day, 3µg of pTV plasmid versions (pTV-DS-MCS2 containing any GOI or pTV-DS-LacZ) and 2µg of meganuclease constructs (pCLS1088 or pCLS2147 plasmid DNAs, or meganuclease-encoding mRNAs) are diluted in 300 µl of DMEM without serum. On the other hand, 10µl of Lipofectamine 2000 (Invitrogen) are mixed with 290µl of DMEM without serum.

The two mixes are incubated 5 min at room temperature. Then the DNA mix is added to the lipofectamine mix and incubated for 20 min. at room temperature.

Meantime, replace culture medium with 9 ml of fresh medium.

After the incubation period, add the total transfection mix (600 µl) over plated cells.

Incubate dish in a 37°C, 5% CO_2 humidified incubator.

Change medium 6 hours after transfection (optional)

A schematic representation of the transfection protocol is shown in Fig. 10.

2.4.2 cGPS HEK 293 TARGETED CLONES SELECTION

Clonal selection is a longer but better protocol to select the proper cell line expressing the GOI. Fig. 11 shows a schematic representation of the clonal selection protocol (left column).

cGPS HEK 293 cells are transfected with the protocol described above (2.4.1). 24 hours after transfection, culture medium is replaced with fresh medium supplemented with 0.4mg/ml of G418.

After 12 days of G418 selection, the second selective agent (puromycin) is added at the concentration of 0.4 µg/ml.

After 7 - 9 days of double selection, single colony clones are picked up and seeded in 96 well plates in complete medium supplemented with G418 at 0.4 mg/ml and puromycin at 0.4 µg/ml.

10 days later, double resistant clones can be characterized by analytical PCR and Southern blotting experiments. Positive control clones can be assayed for β-galactosidase activity, if pTV-DS-LacZ has been used as a positive control.

As shown on figure 12 panel A, double resistant clones are obtained upon cotransfection with meganuclease expression vector and pTV-DS-lacZ. These double resistant clones expressed the lacz gene as revealed by X-gal staining (see example 6 for materials and methods). Several double resistant clones are picked at random and amplified for molecular characterization by Southern blotting (see section 2.5). As shown on figure 12 panel B, gDNA from 13 double resistant clones is digested with the RsrII restriction enzyme, immobilized on nitrocellulose membrane and hybridized with a ^{32}P-neo probe. A correct targeted insertion is characterized by the identification of a band at 4.3 kb. Such band is identified for 11 double resistant clones out of 13 clones analyzed, while a band at 1.6 kb is shown for the parental cGPS HEK 293. In addition to the 4.3 kb band, a second band is present for 3 out of

11 clones and is probably due to an additional random insertion. Hence, these results demonstrate that cGPS HEK 293 double resistant clones, obtained with the method described above, expressed the reported gene, present on the integration matrix, that is correctly targeted in the cGPS locus.

2.5 MOLECULAR CHARACTERISATION OF INSERTION CLONES

A correct targeted insertion in double resistant clones can be easily identified at the molecular level by Southern blot analysis (Fig.13). Alternatively PCR primers can also be designed for a quicker characterization of targeted clones.

Materials and methods

Genomic DNA (gDNA) from targeted clones was purified from 10^7 cells (about a nearly confluent 10 cm dish) using the Blood and Cell culture DNA midi kit (Qiagen, 5 to 10 µg of gDNA are digested with a 10-fold excess of restriction enzyme by overnight incubation).

Digested gDNA was separated on a 0.8% agarose gel and transfer on nylon membrane.

Nylon membranes were then probed with a ^{32}P DNA probe specific for the EF1α intron.

After appropriate washes, the specific hybridization of the probe is revealed by autoradiography.

To check the left region of the targeted insertion:

Forward oligo (in the cGPS locus) F1_Prom:

CCCCGACCGGAGCTGAGAGTAATT (SEQ ID NO: 30)

Reverse oligo (in the pTV-DS-MCS2 vector) B1_Pur:

CAGGAGGCCTTCCATCTGTTG (SEQ ID NO: 31)

The amplification product is 1794 base pairs (bp) long.

For checking the right region of the targeted insertion :

Forward oligo (in the pTV-DS-MCS2 vector) SV40s :

CTGTGGAATGTGTGTCAGT (SEQ ID NO: 32)

Reverse oligo (in the cGPS locus) NEOr :

CAACGCTATGTCCTGATAGCGGTC (SEQ ID NO: 33)

The amplification product is 1073 bp long.

Results

For example (Fig.13), the targeted insertion of LacZ is checked in the 5' side with a double digest BglII (1 site upstream of pEF1α promoter) and EcoRV (a unique site in the LacZ gene). The probe is located within the EF1α intron. Thus, the native locus when digested within BglII/EcoRV, gives a band higher than 10kb. On the contrary, a targeted insertion will bring the EcoRV site from LacZ in the vicinity of the BglII site. Upon double digest, a 5 kb DNA fragment is generated that is identified with the intron probe. The same approach can be used for the 3' side of the insertion.

EXAMPLE 3 EXPRESSING GENE OF INTEREST IN cGPS CHO-K1 CELL LINE

3.1 CD4 EXPRESSION

The human CD4 ORF (SEQ ID NO: 40) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-CD4, Figure 26, SEQ ID NO: 41) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the CD4 gene is provided as SEQ ID NO: 62. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the transmembrane CD4 protein is evaluated.

Materials and methods

Cells from targeted clones are washed twice in PBS and incubated with 2 ml of Versene solution. After 5 min. incubation at 37°C, cells are collected in a 15 ml conical tube. The cells are counted.

10^6 cells are transferred in 5 ml tube (Falcon, 2058) and centrifuge at 300 g for 5min. at 4°C. Cells are washed once with FACS buffer. Cell pellets are re-suspended in 20µl of Biotin conjugated anti-CD4 or Biotin-conjugated isotype control antibody. After 30 min. of incubation on ice, cells are washed once in FACS buffer. Cell pellets are then incubated with 20 µl of Streptavidin-conjugated PE for 30 min. on ice and protected from light. The cells are washed once in FACS buffer and finally re-suspended in 0.5 ml of FACS buffer.

Results

The cells sample are analyzed on a FACS vantage II (BD Bioscience) using a 488 nm Ion-Argon laser. The emitted fluorescence (emission wave length at approximately 580 nm) is collected in the fluorescence 2 channel (Fig 14).

These experiments showed that the CD4 gene product could be reliably inserted into the cGPS locus and then be stably expressed over a prolonged period of time.

3.2 SOMATOSTATIN RECEPTOR (GPCR SSTR2) EXPRESSION

The human GPCR SSTR2 ORF (SEQ ID NO: 42) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-SSTR2, Figure 27, SEQ ID NO: 43) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the SSTR2 gene is provided as SEQ ID NO: 63. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the SSTR2 protein is evaluated.

Materials and methods

GPCR SSTR2 activity is accomplished by measuring the inhibition of cAMP production after proper agonist stimulation of the SSTR2 receptor. The inventors used the protocol and reagents provided by the HitHunter ^{TM} cAMP XS+ assay kit (DiscoverX). Briefly, cells from targeted clones are seeded in white 96 well plates at the density of 10^4 cells per well. After co stimulation of cells with Forskolin (100 µM) and increasing concentrations of Somatostatin (from 10^{-12} M to 10^{-4} M), cells are lysed and cAMP level is measured using a microplate luminometer (Victor, Perkin Elmer) (Fig.15).

Results

In these experiments individual clones were seen to show essentially the same cAMP production inhibition profile in response to different levels of somatostatin.

3.3 HUMAN AUTOTAXIN (hATX).

The human AUTOTAXIN ORF (SEQ ID NO: 44) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-hATX, Figure 28 SEQ ID NO: 45) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the hATX gene is

provided as SEQ ID NO: 64. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the hATX protein is evaluated by western blotting.

Materials and methods

Preparation of ATX conditioned media

cGPS CHO-K1 hATX targeted clones were washed twice with PBS first, 3 times with serum-free FK12 medium supplemented with 1% glutamin in order to remove serum (2 ml per well and per wash for a 6 wells plate), and then incubated with the same medium (1 ml per well) 6 hours at 37°C in a humidified atmosphere containing 7% CO2. After incubation, conditioned-medium (CM) was separated from the cells, centrifuged to eliminate cell debris and then dialyzed overnight against 10 liters of 20 mM HEPES, pH 7.4, 6 mM D(+)-glucose, 1 mM $CaCl_2$, and 1.2 mM $MgSO_4$ using Spectra-Por 1.7 ml/cm tubing (Pierce Chemicals, Interchim, Montluçon, France). After dialysis CM are concentrated (about 15 fold) using an Amicon Ultra 10,000 (Millipore). Concentrated conditioned media (CCM) were aliquoted and stored at –20°C before use.

SDS-PAGE separation and Western Blotting

SDS-PAGE 4-12% was performed according to Laemmli (25) followed by Sypro Ruby staining and Western Blotting detection. After addition of sample buffer (Novex, Invitrogen) concentrated fractions of CM were boiled at 100°C for 5min. Electrophoretic separation of proteins was carried out on a 1mm-thick 18 x 10-cm gel 4-12% acrylamide. An equivalent amount of total protein in sample buffer was loaded into a 4-mm well of the gel and separated at 40mA. A total of 30µg of standards (Mark12, or Magic Mark, Invitrogen) migrated in a neighboring lane. One of the gel was stained with Sypro Ruby and the other was transferred to nitrocellulose membranes and stained with chicken anti-autotaxin antibody followed by an HRP-conjugated anti-chicken antibody (Sigma Aldrich) before chemiluminescence detection of the immuno-complexes.

Results

The detection of hATX in the supernatants of cGPS CHO-K1 hATX targeted clones is performed by western blot. Figure 16 shows the identification of a band over 100 kD using a specific antibody for the human ATX, in conditioned media

from 10 cGPS CHO-K1 hATX targeted clones supernatants or from a cGPS CHO-K1 hATX targeted cell population (poly). No band is detected in the supernatant from the negative control. These results indicate that this rather large secreted protein is expressed by all cGPS CHO-K1 hATX targeted clones.

3.4 HUMAN MELATONIN 1 RECEPTOR (hMT1) and HUMAN MELATONIN 2 RECEPTOR (hMT2)

The human GPCR MT1 ORF (SEQ ID NO: 46) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-hMT1, Figure 29, SEQ ID NO: 47) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the hMT1 gene is provided as SEQ ID NO: 65. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the MT1 protein is evaluated.

The human GPCR MT2 ORF (SEQ ID NO: 48) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-hMT2, Figure 30, SEQ ID NO: 49) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the hMT2 gene is provided as SEQ ID NO: 66. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the hMT2 protein is evaluated.

Materials and methods

Radio-ligand saturations with intact cells

cGPS CHO-K1 hMT1 targeted clones and cGPS CHO-K1 hMT2 targeted clones were resuspended in Tris/HCl 50mM pH 7.4, EDTA 1mM and MgCl2 5mM and dispensed in 96-well polypropylene plates at 13,000 cells/well. [^{125}I]-2-Iodomelatonine 5pM to 1.5nM was added to determine the total binding signal, while control wells contained an additional 1µM melatonin to determine non specific binding. The incubation was performed at 37°C for 2 hrs in a total volume of 250µL. Cells were then transferred to unifilter GF/B plates (Perkin Elmer) with a FilterMate cell harvester (Perkin Elmer) and washed 3 times with 1ml of ice-cold Tris 50mM. Microscint 20 (40µl/well, Perkin Elmer) was added before sealing plates. The radioligand associated with filter plates was evaluated by scintillation counting using a TopCount (Perkin Elmer). Experiments were conducted in triplicates, and data are

expressed as fmol radioligand specific binding sites (total minus non specific) per mg of total protein. Graphic representations and data analysis were generated using PRISM 4.03 (GraphPad).

Results

Ten cGPS CHO-K1 hMT1 targeted clones and 10 cGPS CHO-K1 hMT2 targeted clones were picked at random and functionally tested for radioligand saturations experiments using [^{125}I]-iodomelatonin. Results obtained for the hMT1 are presented on figure 17 and those for hMT2, on figure 18. From saturations curves, pKd values are obtained by Scatchard analysis (fig. 17, panel A and fig.18, panel A). From the saturations curves, the quantity (fmol) of specific binding sites per mg of total protein (Bmax, fig.17 panel B and fig.18, panel A) is measured. Similar results are obtained with both receptors. The data indicate that pKd values from each clones are very closed to each other and to polyclonal cGPS CHO-K1 hMT1- or hMT2-targeted cell population. However, some variations of hMT1 and hMT2 receptors expression is observed from clones to clones. pKd and Bmax values for both receptors are consistent with previous published observations.

EXAMPLE 4 EXPRESSING GOI UNDER THE CONTROL OF DIFFERENT PROMOTERS.

In this example, the heavy chain (SEQ ID NO: 50) and the light chain (SEQ ID NO: 51) of the 5F11 monoclonal antibody (Medarex Inc.) have been cloned in the pTV-DS-MCS2. Both chains are under the control of the Ubiquitin sub-unit c promoter (pUbc SEQ ID NO: 52). The resulting vector (pTV-DS-5F11, Figure 31, SEQ ID NO: 53) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the 5F11 gene is provided as SEQ ID NO: 67. Targeted clones surviving the selection process (2.2.2) are isolated and the expression of the 5F11 monoclonal antibody protein is evaluated.

Materials and methods

Cells from targeted clones are seeded in 96 well plates (Costar) at the density of 10^4 cells per well. After 48h of culture in complete medium, supernatants are collected and assayed for monoclonal antibody expression by ELISA. Briefly, 96

well plates are coated with a Goat-anti-Human kappa light chain (Southern Biotechnology Associates) in PBS overnight at 4°C. All washing steps are done in PBS, 0.1% Tween 20. After washing, plates are blocked in PBS, 1% BSA (PBA) for 90 min at 37°C under shaking. After washing, 50µl of diluted supernatant from samples are added and incubated for 90 min at 37°C under shaking. After washing, a goat-anti-human IgG Fc coupled to Alkaline Phosphatase (Jackson ImmunoResearch) in PBA is added. After washing, the Developing buffer (Pierce) containing 1 mg/ml PNPP (Pierce) is added. Optical density (OD) is read at 405 nm using a microplate Reader (Model 550, BioRad) (Fig 19).

Results

This study showed that the measured level of antibody expression was greater than 0.04 µg/ml in 48 hours for each of the generated clones and that expression levels were generally homogenous, although clone 29 showed higher expression levels in comparison to the other studied clones.

EXAMPLE 5 LONG-LASTING EXPRESSION OF GOI IN THE PRESENCE OR ABSENCE OF SELECTING DRUGS.

In this example, the inventors monitored the level of expression of four cGPS CHO-K1 targeted clones expressing the lacZ gene and of four cGPS CHO-K1 targeted clones expressing the luciferase gene. The lacZ ORF (SEQ ID NO: 16) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-lacz, Figure 4, SEQ ID NO: 17) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. Targeted clones surviving the selection process (2.2.2) are isolated and characterized according to section 2.5. The luciferase ORF (SEQ ID NO: 54) has been cloned in the pTV-DS-MCS2. The resulting vector (pTV-DS-luciferase, Figure 32, SEQ ID NO: 55) has been transfected in cGPS CHO-K1 cells according to the protocol described in 2.2.1. The sequence of the cGPS site following insertion of the luciferase gene is provided as SEQ ID NO: 68. Targeted clones surviving the selection process (2.2.2) are isolated and characterized according to section 2.5.

The 4 cGPS CHO-K1 lacz targeted clones and the 4 cGPS CHO-K1 luciferase targeted clones were maintained in culture over a period of 45 passages (two passages per week). Each clone was cultured in the presence of selecting drugs

(Puro; 10 µg/ml and G418: 0.6 mg/ml). Furthermore, the inventors evaluated the expression of the two reporter gene for the same clones but without selecting drugs (i.e. in complete F12K medium) over a period of time corresponding to 30 passages.

Materials and methods

Lacz expression: Cells from targeted clones are washed twice in PBS then incubated with 5 ml of trypsin-EDTA solution. After 5 min. incubation at 37°C, cells are collected in a 15 ml conical tube and counted.

Cells are then resuspended in complete F-12K medium at the density of 50000 cells/ml. 100 µl (5000 cells) are aliquoted in triplicate in a white 96 well plate (Perkin-Elmer). 100 µl of beta-Glo reagent (Promega) is added per well and after a 30 min. incubation period, the plate can be read on a luminometer (Viktor, Perkin-Elmer).

Luciferase expression: Cells from targeted clones are washed twice in PBS then incubated with 5 ml of trypsin-EDTA solution. After 5 min. incubation at 37°C, cells are collected in a 15 ml conical tube and counted.

Cells are then resuspended in complete F-12K medium at the density of 50000 cells/ml. 100 µl (5000 cells) are aliquoted in triplicate in a white 96 well plate (Perkin-Elmer). 100 µl of One-Glo reagent (Promega) is added per well and after a short incubation the plate can be read on a microplate luminometer (Viktor, Perkin-Elmer).

Results

The data are presented on fig.20. On panels A and B, the mean level of lacz expression for 4 cGPS CHO-K1 lacz targeted clones is measured as a function of time in the presence or absence of selecting agents, respectively. On panels C and D, the mean level of luciferase expression for 4 cGPS CHO-K1 luciferase targeted clones is shown as a function of time in the presence or absence of selecting agents, respectively. These data indicates that the expression of both reporters is remarkably stable even after a long period of culture. Furthermore the presence of the selecting agents is not necessary to ensure a long lasting expression of transgene since the stability of reporter expression is equivalent when the targeted clones are cultivated without selecting agents.

EXAMPLE 6 GENERATING TARGETED CLONES USING I-CreI RECOMBINANT PROTEIN.

In this example, the inventors addressed the possibility to use I-CreI as a recombinant protein instead of I-CreI expression vector plasmid or mRNA to achieve efficient gene targeting in cGPS CHO-K1 system. Two different cell-penetrating peptides, also termed DPVs have been demonstrated to transport reporter proteins to the nucleus in a variety of mammalian cell lines including epithelial (HeLa, HCT116), myeloid (HL-60), erythroid (K562), lymphoid (Molt4), fibroblast (NIH-3T3) cells, as well as primary hepatocyte cultures (24). They have been shown to efficiently mediate the internalization of molecules as little as a few Daltons, and up to 200kDa. These peptidic sequences, DPV15b (SEQ ID NO: 56) and DPV1047 (SEQ ID NO: 57) have been fused to the N-terminal part of the I-CreI N75 meganuclease protein (SEQ ID NO: 14) and resulting recombinant protein has been produced in E. coli and purified. In this example, the inventors used the purified recombinant protein termed DPV15b/I-CreI N75/6xHis (SEQ ID NO: 58).

Materials and methods

Transfection: On D day, 0.5µg of the pTV-DS-lacz (SEQ ID NO: 17) is transfected in cGPS CHO-K1 cells by using the PolyFect® reagent from Qiagen. One day after the transfection of the integration matrix (D+1), 1000µg of the purified DPV15b/I-CreI N75/6xHis (SEQ ID NO: 58) meganuclease protein is directly added into transfected cGPS CHO-K1 cells.

Selection: cGPS CHO-K1 cells are transfected with the protocol described above. 24 hours after transfection, the cells are washed and fresh medium supplemented with 0.6mg/ml of G418 is added. After 10 days of G418 selection, the culture medium is replaced with complete medium supplemented with G418 at 0.6 mg/ml and puromycin at 10µg/ml. 3 to 4 days later, double resistant clones are visualized through an inverted microscope. At this step, double resistant clones are either stained for lacz expression monitoring or picked up for amplification and molecular characterization. (see section 2.5)

X-gal staining: Culture medium is removed and LacZ-targeted double resistant cGPS CHO-K1 adherent cells are washed once with PBS. 5 ml of fixing

buffer (100 mM phosphate buffer, 1mM $MgCl_2$, 0.5 % (v/v) glutaraldehyde (Prolabo, 25% solution)) is added. After 10 min. of incubation on ice, fixing is replaced by 5 ml of washing buffer (100mM phosphate buffer, 1mM $MgCl_2$, 0.02% (v/v) NP40). Then, 5 ml of staining buffer (10mM phosphate buffer, 1mM $MgCl_2$, 33mM KFerri [Potassium hexacyanoferrate (III)], 33mM KFerro [Potassium hexacyanoferrate (II)], 0,1 % (v/v) X-Gal) is added for incubation at 37°C. Blue cells should appear within 24 hours.

Results

As shown in Figure 21 panel A, double resistant ($Neo^R/Puro^R$) cGPS CHO-K1 cell colonies are generated after the KI assay, from the moment that the I-CreI N75 meganuclease is expressed (pCLS1088) or added by fusion with a DPV cell penetrating peptide (DPV15b/I-CreI N75/6xHis) in association with the transfection of a LacZ-encoding integration matrix (pCLS1625). By contrast, the sole transfection of the integration matrix does not generate any cell colony, highlighting the importance of I-CreI N75 meganuclease to mediate homologous recombination in the cGPS CHO-K1 KI model.

Although a fewer number of cell colonies have been generated by the delivery of DPV15b/I-CreI N75/6xHis (SEQ ID NO: 58) recombinant proteins (around 30, leading to a selection frequency of $3x10^{-4}$) over those depicted from the transfection of I-CreI N75-encoding DNA plasmid (around 50, with a selection frequency of $5x10^{-4}$), these results clearly indicates that the DPV15b/I-CreI N75/6xHis (SEQ ID NO: 58) recombinant protein triggers targeted integration of the lacz gene. Moreover, all these cGPS CHO-K1 lacZ targeted clones are blue after X-Gal staining for both DNA and protein conditions, probably suggesting that all cell clones have been positively targeted.

However, since the expression of the reporter LacZ gene could arise from a random chromosomal integration, a Southern blot analysis has been settled in order to check the correct cGPS genetic pattern of integration for different double resistant cGPS CHO-K1 cell clones. Genomic DNAs extracted from nineteen individual cGPS CHO-K1 lacz targeted clones, as well as from CHO-K1 and

untargeted cGPS CHO-K1 cells have been submitted to the hybridization with a probe is specific to the cGPS chromosomal locus (i.e. the EF1α intron sequence).

As shown in Figure 21 panel B, a vast majority of LacZ-targeted cGPS CHO-K1 cell clones (18 out of 19) contain the expected cGPS modified locus, with a band around 5kb which demonstrates the targeted insertion of the LacZ gene. By contrast, the hybridization of genomic DNA extracts originating from the original untargeted cGPS CHO-K1 cell line reveals the presence of a higher band (10kb), and no band is detected in the negative control, i.e. the CHO-K1 cell lineage that does not contain the EF1α intron sequence. The targeting at the cGPS locus is highly specific since no additional band is detected from LacZ-targeted cGPS CHO-K1 cell clones, therefore revealing the absence of any random integration elsewhere in the genome.

EXAMPLE 7 GENERATING DOUBLE TARGETED CELL LINE BY COMBINING THE cGPS CHO-K1 SYSTEM WITH A CUSTOM MEGANUCLEASE GENE TARGETING SYSTEM.

In this example, the inventors addressed the possibility to use the cGPS CHO-K1 system in combination with a custom meganuclease cGPS system, leading to the targeting of two genes of interest into two distinct loci of the CHO-K1 genome. The insertion of the two GOI is sequential. The first GOI is inserted in the cGPS locus as described in section 2.2. Once a cGPS CHO-K1 targeted clone is identified as described in section 2.5, a second GOI is inserted in the cGPS CHO-K1 targeted clone through a custom meganuclease, named Sc MA17-RM2-G19H33 (SEQ ID NO: 60), that has been engineered to cleave in the 3rd exon of the CHO-K1 HPRT gene (WO2008/059382). The cGPS Custom CHO-K1 Integration Matrix containing the lacZ gene (Fig.22; SEQ ID NO: 59), and Meganuclease Expression Vector (Fig.23; SEQ ID NO: 61) are co-transfected into the cGPS CHO-K1 targeted clone. Upon co-transfection, the engineered meganuclease is expressed, recognizes its HPRT recognition site and induces a DNA double-strand break at this precise site. Homologous recombination occurs at the meganuclease recognition site. The gene of interest, cloned in the Integration Matrix in between the homology regions, is integrated at the meganuclease recognition site during this recombination event. Following meganuclease-induced homologous recombination, the hygromycin

resistant gene is transcribed via the endogenous HPRT promoter and expressed as a fusion protein with the first exons of HPRT (exons 1, 2 and part of exon 3). In addition to the newly acquired hygromycin-resistance phenotype, targeting the mono-allelic HPRT gene locus leads to its inactivation, therefore allowing resistance to 6-thioguanine (6-TG) nucleotides. Thus, stable cGPS CHO-K1 targeted clone can be selected for the double hygromycin / 6-TG resistance and expression of the recombinant protein of interest.

In the following example, the luciferase gene is chosen as the first GOI, to be inserted in the cGPS CHO-K1 locus, while the lacz gene is the second GOI, to be inserted in the HPRT locus. After selection of double targeted clones, the expression of the two reporter genes is monitored over 20 passages (40 weeks) in order to evaluate their stability.

Materials and methods

Transfection

One day prior to transfection, cGPS CHO-K1 targeted clone cells are seeded in a 10cm tissue culture dish ($2x10^5$ cells per dish). On transfection day, 1µg of the Meganuclease Expression Vector and 2µg of the Integration Matrix, containing the lacZ gene, are diluted in 275µl of medium without serum. 25µl of the PolyFect™ reagent is added to the diluted DNA and the transfection mix is vortexed for 10'' and incubated 10' at room temperature.

In the meantime culture medium is replaced with 9 ml of fresh medium. Then 700µl of complete medium is added to the transfection mix and the total volume is dispensed over plated cells.

Transfected cells are incubated in a 37°C, 5% CO_2 humidified incubator.

Selection

3 days after transfection, cells are washed and fresh medium, supplemented with 0.6mg/ml of hygromycin B, is added.

After 7 days of hygromycin selection (Day+10), fresh medium, supplemented with 0.6mg/ml of hygromycin B and 5µg/ml of 6-thioguanine (Hybrimax, Sigma), is added.

After 5 or 8 days of double selection (Day+15 to Day+18), single colony clones are picked and seeded in 96 well plates in complete medium supplemented with 0.6mg/ml of hygromycin B at and 5µg/ml of 6-thioguanine.

Double resistant clones are amplified in complete medium supplemented with the two selective agents. For downstream experiments (i.e. molecular characterization, lacZ expression, etc...) the inventors strongly recommend to maintain both selective agents to maintain homogeneous expression.

Southern Blot

Genomic DNA (gDNA) is purified from 10^7 cells (about a nearly confluent 10 cm dish). 5 to 10 µg of gDNA are digested with a 10-fold excess of EcoRV restriction enzyme by overnight incubation. Digested DNA is transferred on a nitrocellulose membrane and hybridization is performed with a ^{32}P-labeled-lacz probe (see section 2.5 for details).

Results

The inventors have previously produced cGPS CHO-K1 targeted clones expressing the luciferase reporter gene (see example 4 and figure 20 panel C and D). One of these cGPS CHO-K1 luciferase targeted clone has been used to perform a second targeted insertion of the lacz gene into the Hprt locus. cGPS CHO-K1 luciferase targeted cells are co-transfected with the integration matrix containing the lacz gene and the meganuclease expression vector specific for the hamster Hprt gene. Upon selection, as described in the materials and methods section, hygromycin and 6-TG resistant clones are analyzed for correct insertion of the lacz gene in the Hprt gene. As shown on fig.24 panel A, 18 clones out of 18 are correctly targeted. 5 clones out of 18 present additional bands corresponding probably to the random insertion of the integration matrix. These data are in accordance with the data obtained when the experiment is done in CHO-K1. Furthermore, the inventors verified by southern blot that the cGPS CHO-K1 locus, site of the first targeted insertion is still modified. As shown on fig.24 panel B, the 18 analyzed clones present a hybridization pattern compatible with a modified cGPS CHO-K1 locus. All together, these data demonstrate that the first targeted insertion in the cGPS CHO-K1 locus has no impact

on the efficiency of the second insertion, and the second targeted insertion is not prejudicial to the first, at least at the genomic level.

To verify whether the double insertion has an impact on the expression of the two reporter genes, four doubled targeted clones were maintained in culture over a period of 11 weeks (21 passages) and regularly checked for lacz and luciferase expression. As shown on fig. 25 panel A, the expression of the lacz gene is stable allover the study period. Similarly, the luciferase expression is stable (Fig.25 panel B) and comparable to those observed for single cGPS CHO-K1 targeted clones (Fig.20 panel C and D).

REFERENCES

1. Hinnen, A., Hicks, J. B., and Fink, G. R. (1978) Transformation of yeast. *Proc Natl Acad Sci U S A* 75, 1929-33.

2. Rothstein, R. J. (1983) One-step gene disruption in yeast. *Methods Enzymol* 101, 202-11.

3. Thomas, K. R., and Capecchi, M. R. (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51, 503-12.

4. Capecchi, M. R. (2001) Generating mice with targeted mutations. *Nat Med* 7, 1086-90.

5. Smithies, O. (2001) Forty years with homologous recombination. *Nat Med* 7, 1083-6.

6. Rouet, P., Smih, F., and Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* 14, 8096-106.

7. Choulika, A., Perrin, A., Dujon, B., and Nicolas, J. F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol Cell Biol* 15, 1968-73.

8. Chevalier, B. S., and Stoddard, B. L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* 29, 3757-74.

9. Dujon, B., Colleaux, L., Jacquier, A., Michel, F., and Monteilhet, C. (1986) Mitochondrial introns as mobile genetic elements: the role of intron-encoded proteins. *Basic Life Sci* 40, 5-27.

10. Haber, J. E. (1995) In vivo biochemistry: physical monitoring of recombination induced by site-specific endonucleases. *Bioessays* 17, 609-20.

11. Posfai, G., Kolisnychenko, V., Bereczki, Z., and Blattner, F. R. (1999) Markerless gene replacement in Escherichia coli stimulated by a double-strand break in the chromosome. *Nucleic Acids Res* 27, 4409-15.

12. Sargent, R. G., Brenneman, M. A., and Wilson, J. H. (1997) Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Mol Cell Biol* 17, 267-77.

13. Donoho, G., Jasin, M., and Berg, P. (1998) Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* 18, 4070-8.

14. Cohen-Tannoudji, M., Robine, S., Choulika, A., Pnto, D., El Marjou, F., Babinet, C., Louvard, D., and Jaisser, F. (1998) I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol Cell Biol* 18, 1444-8.

15. Gouble, A., Smith, J., Bruneau, S., Perez, C., Guyot, V., Cabaniols, J. P., Leduc, S., Fiette, L., Ave, P., Micheau, B., Duchateau, P., and Paques, F. (2006) Efficient in toto targeted recombination in mouse liver by meganuclease-induced double-strand break. *J Gene Med* 8, 616-22.

16. Siebert, R., and Puchta, H. (2002) Efficient Repair of Genomic Double-Strand Breaks by Homologous Recombination between Directly Repeated Sequences in the Plant Genome. *Plant Cell* 14, 1121-31.

17. Puchta, H., Dujon, B., and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci U S A* 93, 5055-60.

18. Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* 22, 1393-8.

19. Craig, NL. (1988) The mechanism of conservative site-specific recombination. *Annu Rev Genet* 22, 77-105.

20. Sauer, B. (1994) Site-specific recombination: developments and applications. *Curr Opin Biotechnol* 5, 521-7.

21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Plainview, New York: Cold Spring Harbor Laboratory Press).

22. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994). Current Protocols in Molecular Biology (New York: Greene Publishing Associates and Wiley-Interscience).

23. O'Gorman, S., Fox, D. T., and Wahl, G. M. (1991). Recombinase-Mediated Gene Activation and Site-Specific Integration in Mammalian Cells. *Science* 251, 1351-5.

24. de Coupade, C., Fittipaldi, A., Chagnas, V., Michel, M., Carlier, S., Tasciotti, E., Darmon, A., Ravel, D., Kearsley, J., Giacca, J. and Cailler, F. (2005) Novel human-derived cell-penetrating peptides for specific subcellular delivery of therapeutic biomolecules. *Biochem. J.* 390, 407-418.

25 Laemmli, U. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

26. Seligman *et al.*, Genetics, 1997, 147, 1653-1664; Sussman *et al.*, J. Mol. Biol., 2004, 342, 31-41.

27. International PCT Applications WO 2006/097784 and WO 2006/097853.

28. Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458; Rosen *et al.*, Nucleic Acids Res., 2006, 34, 4791-4800 ; Smith *et al.*, Nucleic Acids Res., 2006, 34, e149.

29. Smith *et al.*, Nucleic Acids Res., 2006, 34, e149;

30. International PCT Applications WO 2007/060495 and WO 2007/049156

31. Chevalier *et al.*, Mol. Cell., 2002, 10, 895-905.

32. Epinat *et al.*, Nucleic Acids Res, 2003, 31, 2952-62.

33. International PCT Applications WO 03/078619 and WO 2004/031346.

34. Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W. A. and Rosen, C. A. (1989) Structural and functional characterization of human immunodeficiency virus tat protein. J. Virol. **63**, 1–8

35. Tyagi, M., Rusnati, M., Presta, M. and Giacca, M. (2001) Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. J. Biol. Chem. **276**, 3254–3261

36. Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G. and Prochiantz, A. (1996) Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. J. Biol. Chem. **271**, 18188–18193
Res, 566, 131-67.

37. Elliott, G. and O'Hare, P. (1997) Intercellular trafficking and protein delivery by a herpesvirus structural protein. Cell (Cambridge, Mass.) **88**, 223–233

38. Futaki, S., Goto, S. and Sugiura, Y. (2003) Membrane permeability commonly shared among arginine-rich peptides. J. Mol. Recognit. **16**, 260–264

39. Jans, D. A. (1994) Nuclear signaling pathways for polypeptide ligands and their membrane receptors FASEB J. **8**, 841–847

40. Kokryakov, V. N., Harwig, S. S., Panyutich, E. A., Shevchenko, A. A., Aleshina, G. M., Shamova, O. V., Korneva, H. A. and Lehrer, R. I. (1993) Protegrins: leukocyte antimicrobial peptides that combine features of corticostatic defensins and tachyplesins. FEBS Lett. **327**, 231–236

41. Mie, M., Takahashi, F., Funabashi, H., Yanagida, Y., Aizawa, M. and Kobatake, E. (2003) Intracellular delivery of antibodies using TAT fusion protein A. Biochem. Biophys. Res. Commun. **310**, 730–734

42. Silhol, M., Tyagi, M., Giacca, M., Lebleu, B. and Vives, E. (2002) Different mechanisms for cellular internalization of the HIV-1 Tat-derived cell penetrating peptide and recombinant proteins fused to Tat. Eur. J. Biochem. **269**, 494–501

43. Stein, S., Weiss, A., Adermann, K., Lazarovici, P., Hochman, J. and Wellhoner, H. (1999) A disulfide conjugate between anti-tetanus antibodies and HIV (37-72)Tat neutralizes tetanus toxin inside chromaffin cells. FEBS Lett. **458**, 383–386

44. Suzuki, T., Futaki, S., Niwa, M., Tanaka, S., Ueda, K. and Sugiura, Y. (2001) Possible existence of common internalization mechanisms among arginine-rich peptides. J. Biol. Chem. **277**, 2437–2443

45. Torchilin, V. P., Rammohan, R., Weissig, V. and Levchenko, T. S. (2001) TAT peptide on the surface of liposomes affords their efficient intracellular delivery

even at low temperature and in the presence of metabolic inhibitors. Proc. Natl. Acad. Sci. U.S.A. **98**, 8786–8791

46. Schwarze, S. R., Ho, A., Vocero-Akbani, A. and Dowdy, S. F. (1999) In vivo protein transduction: delivery of a biologically active protein into the mouse. Science **285**, 1569–1572

47. Tasciotti, E., Zoppe, M. and Giacca, M. (2003) Transcellular transfer of active HSV-1 thymidine kinase mediated by an 11-amino-acid peptide from HIV-1 Tat. Cancer Gene Ther. **10**, 64–74

48. Vives, E., Brodin, P. and Lebleu, B. (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. J. Biol. Chem. **272**, 16010–16017

49. PEREZ C, GUYOT V, CABANIOLS J, GOUBLE A, MICHEAUX B, SMITH J, LEDUC S, PAQUES F, DUCHATEAU P, (2005) BioTechniques vol. 39, n°1, pp. 109-115

CLAIMS

1. A set of genetic constructs comprising:

a) Construct (i) encoded by a nucleic acid molecule, which comprises at least the following components:

A1 – A2 – A3 – A4 – A5 (i)

wherein A1 is a first promoter; A2 is a first homologous portion; A3 is a meganuclease cleavage site; A4 is a first marker gene; A5 is a second homologous portion; and wherein construct (i) is configured to be stably integrated into the genome of at least one target cell;

b) Construct (ii) encoded by a nucleic acid molecule, which comprises at least the following components:

A2' – B1– B2 – B3 – B4 – A5' (ii)

wherein A2' comprises a portion of said first homologous portion A2; B1 is a second marker gene different to said first marker gene; B2 is a second promoter; B3 is a multiple cloning site; B4 is a third promoter; A5' comprises a portion of said second homologous portion A5;

c) At least one construct selected from the group comprising, constructs (iii) or (iv) encoded by nucleic acid molecules, which comprise at least the following components:

C1 – C2 (iii);

C3 (iv); or

Construct (v) which is an isolated or recombinant protein which comprises at least the following component:

C4 (v);

wherein C1 is a fourth promoter; C2 is the open reading frame (ORF) of a meganuclease; C3 are messenger RNA (mRNA) versions of said meganuclease; C4 is an isolated or recombinant protein of said meganuclease; wherein said meganuclease from constructs (iii), (iv) or (v) recognize and cleave A3; and wherein constructs (iii), (iv) or (v) are configured to be co-transfected with construct (ii) into said at least one target cell.

2. The set of constructs of claim 1, wherein said components A2' and A5' from construct (ii), comprise at least 200 bp and no more than 6000 bp of said components A2 and A5 from construct (i) respectively.

3. The set of constructs of claim 1 or 2, wherein said components A2' and A5' from construct (ii), comprise at least 1000 bp and no more than 2000 bp of said components A2 and A5 from construct (i) respectively.

4. The set of constructs according to any one of claims 1 to 3, wherein said component A5 encodes a marker gene or a portion thereof.

5. The set of genetic constructs of any one of claims 1 to 4, wherein said component A5' comprises a 3' end deletion of said A5.

6. The set of genetic constructs of any one of claims 1 to 5, wherein said components of each of said constructs (i), (ii), (iii), (iv) and (v) are selected from the following groups:

Component	**Group**
A1	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
A2	EF1α intron 1 complete sequence, 5' homology (SEQ ID NO: 3)
A2'	EF1α intron 1 short sequence, 5' homology (SEQ ID NO: 29)
A3	Meganuclease cleavage site (SEQ ID NO: 8)
A4	Hygromycin resistance gene (SEQ ID NO: 2); Neomycin resistance gene (SEQ ID NO: 7); Puromycin resistance gene (SEQ ID NO: 21);
A5	Neomycin resistance gene (SEQ ID NO: 7)
A5'	Inactive neomycin resistance gene deleted of its 3' end, 3' homology (SEQ ID NO: 13)
B1	Hygromycin resistance gene (SEQ ID NO: 2); Neomycin resistance gene (SEQ ID NO: 7); Puromycin resistance gene (SEQ ID NO: 21);
B2	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ

	ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
B3	Multiple cloning site (SEQ ID NO: 23)
B4	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
C1	pEF1α promoter (SEQ ID NO: 1); pSV40 (SEQ ID NO: 20); pCMV (SEQ ID NO: 25); Ubiquitin sub-unit c promoter (SEQ ID NO: 52)
C2	Meganuclease ORFs (SEQ ID NO 14); (SEQ ID NO 15)
C3	(SEQ ID NO: 14); (SEQ ID NO: 15); (SEQ ID NO: 35)
C4	Meganuclease peptide encoded by (SEQ ID NO: 14); (SEQ ID NO: 15) and (SEQ ID NO: 58).

7. The set of genetic constructs of any one of claims 1 to 6, wherein construct (i) comprises SEQ ID NO: 6.

8. The set of genetic constructs of any one of claims 1 to 7, wherein construct (ii) comprises SEQ ID NO: 22.

9. The set of genetic constructs of any one of claims 1 to 8, wherein constructs (iii) comprise SEQ ID NO: 38 or SEQ ID NO: 39.

10. The set of genetic constructs of any one of claims 1 to 9, wherein constructs (iv) comprise SEQ ID NO: 14 or SEQ ID NO: 15 or SEQ ID NO: 34 or SEQ ID NO: 35.

11. The set of genetic constructs of any one of claims 1 to 10, wherein construct (v) comprises a cell penetrating peptide, in particular wherein said cell-penetrating peptide is selected from the group comprising: SEQ ID NO: 56, SEQ ID NO: 57.

12. A kit to introduce a sequence encoding a GOI into at least one cell, comprising the set of genetic constructs according to any one of claims 1 to 11; and instructions for the generation of a transformed cell using said set of genetic constructs.

13. The kit of claim 12, further comprising construct (vi) consisting of SEQ ID NO: 17 or SEQ ID NO: 54.

14. The kit of claim 12 or 13, further comprising at least one cell stably transformed with said construct (i).

15. The kit of claim 14, wherein said at least one cell is selected from the group comprising: CHO-K1 cells; HEK293 cells; Caco2 cells; U2-OS cells; NIH 3T3 cells; NSO cells; SP2 cells; CHO-S cells; DG44 cells.

16. A method for transforming by homologous recombination at least one cell comprising the steps of:

a) stably transforming at least one cell by inserting construct (i) of any one of claims 1 to 7 into the genome of said at least one cell;

b) cloning a sequence coding for a gene of interest into position B3 of construct (ii) of any one of claims 1 to 8;

c) co-transfecting said cell of step a), with said construct (ii) of step b) and constructs (iii), (iv) or (v) of any one of claims 1 to 11;

d) following homologous recombination between said construct (ii) and said stably inserted construct (i), selecting at least one cell from step c) based upon: the absence of a first marker gene encoded by component A4 of said construct (i) and the activity of a second marker gene encoded by component B1 and the activity of a third marker gene encoded by component A5.

17. The method of claim 16, wherein selection in step d) is carried out sequentially for each of said first marker, said second marker and said third marker.