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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 March 2011 (31.03.2011)

PCT

(10) International Publication Number
WO 2011/036510 A1

(51) International Patent Classification:
C12N 9/22 (2006.01)

(21) International Application Number:
PCT/IB2009/007171

(22) International Filing Date:
24 September 2009 (24.09.2009)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **CELLECTIS** [FR/FR]; 102, avenue Gaston Roussel, F-93230 Cedex Romainville (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHOULIKA, André** [FR/FR]; 164 rue de la Croix Nivert, F-75015 Paris (FR). **CEDRONE, Frédéric** [FR/FR]; 83 Boulevard Diderot, F-75012 Paris (FR).

(74) Agent: **CABINET ORES**; 36 rue de St Pétersbourg, F-75008 Paris (FR).

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

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

Published:

— *with international search report (Art. 21(3))*

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

WO 2011/036510 A1

(54) Title: MEGANUCLEASE VARIANTS CLEAVING THE GENOME OF THE HERPES SIMPLEX VIRUS AND USES THEREOF

(57) Abstract: An I-*Crel* variant, wherein at least one of the two I-Crel monomers has at least two substitutions, one in each of the two functional subdomains of the LAGLIDADG core domain situated from positions 26 to 40 and 44 to 77 of I-Crel, said variant being able to cleave a DNA target sequence from the herpes simplex virus (HSV) genome. Use of said variant and derived products for genome engineering and for in vivo and ex vivo (gene cell therapy) genome therapy as well as the treatment of a Herpesviridae infection.

MEGANUCLEASE VARIANTS CLEAVING THE GENOME OF THE HERPES SIMPLEX VIRUS AND USES THEREOF

The invention relates to a meganuclease variant cleaving the genome of a Herpesviridae virus and in particular the genome of a Herpes Simplex Virus. The present invention also relates to a vector encoding said variant, as well as to a cell, animal or plant modified by this vector and to the use of these meganuclease variants and derived products for genome engineering and for *in vivo* and *ex vivo* (gene cell therapy) genome therapy as well as the treatment of a Herpesviridae infection.

Of the more than 100 known Herpesviridae viruses, only 8 routinely infect humans: herpes simplex virus types 1 and 2, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human herpes virus 6 (variants A and B), human herpes virus 7, Kaposi's sarcoma virus and human herpes virus 8. A simian virus, called B virus, occasionally infects humans. All herpes viruses can establish latent infection within specific tissues, which are characteristic for each virus (Medical Microbiology, 4th Edition, Virology, Herpes viruses, Whitley RJ, 1996).

Herpes viruses infect members of all groups of vertebrates, as well as some invertebrates. Herpes viruses have been typically classified into three groups based upon details of tissue tropism, pathogenicity and viral behaviour under conditions of culture in the laboratory. The three types include: the alpha-herpes viruses which are neurotropic, have a rapid replication cycle and a broad host and cell range; and the beta- and gamma-herpes viruses which differ in genome size and structure but which both replicate more slowly and in a much more restricted range of cells of glandular and/or lymphatic origin. To date, eight discrete human herpes viruses have been described; each causing a characteristic disease (Norberg et al, J Clin Microbiol, 2006, 44, 4511-4514).

Herpes simplex virus types 1 and 2 (HSV-1 and -2) will be used to illustrate the problems presented by Herpesviridae viruses. In the present Patent Application references to Herpes Simplex Virus and/or HSV refer to both HSV-1 and HSV-2. HSV-1 and HSV-2 are the primary agents of recurrent facial and genital herpetic lesions. Infections although mild in terms of the severity of symptoms, can lead to significant psychological trauma. They are also a major cause of encephalitis. Herpes simplex virus -1/-2 are highly adapted human pathogens with a rapid lytic

replication cycle and also exhibit the ability to invade sensory neurons without showing any cytopathology. Latent infections are subject to reactivation whereby infectious virus can be recovered in peripheral tissue enervated by the latently infected neurons following a specific physiological stress. A major factor in these "switches" from lytic to latent infection and back involves changes in transcription patterns, mainly as a result of the interaction between viral promoters, the viral genome and cellular transcriptional machinery.

HSV is a nuclear replicating DNA virus. The HSV envelope contains at least 8 glycoproteins. The capsid itself is made up of 6 proteins. The major one is the capsid protein U_L19. The matrix which contacts both the envelope and the capsid contains at least 15-20 proteins.

The HSV-1 genome is a linear, double stranded DNA duplex 152,261 base pairs (bp) in length, and with a base composition of 68% G + C which circularizes upon infection. The virus encodes nearly 100 transcripts and more than 70 open translational reading frames (ORFs). Most ORFs are expressed by a single transcript. About 40 genes are considered as essential for virus replication in culture and these are listed in Table I below.

Name	kinetics	Required for replication in culture?	Function
"a"	-	Yes	cis genome cleavage, packaging signal
TR_L	-	Yes	Terminal Long Repeat
gL (U_L1)	early	Yes	viral entry, associates with gH- polyadenylation usage changes with time
Helicase/	early	Yes	DNA replication
U_L6	late	Yes	capsid protein, capsid maturation, DNA packaging
Helicase/Primase (UL8)	early	Yes	DNA replication
a0	IE	Yes	immediate-early transcription regulator (mRNA spliced)
Ori binding protein (U_L9)	early	Yes	DNA replication
U_L11	late	Yes	tegument protein, capsid egress and envelopement
Alkaline exonuclease (U_L12)	early	Yes	DNA packaging, capsid egress
U_L15	late	Yes	DNA packaging, cleavage of replicating DNA, spliced mRNA
U_L17	late	Yes	cleavage and packaging of DNA
Capsid (U_L18)	late	Yes	Triplex
Capsid (U_L19)	late	Yes	major capsid protein, hexon
U_L20	late	Yes	membrane associated, virion egress

gH (U_L22)	late	Yes	viral entry, functions with gL
U_L25	late	Yes	tegument protein, capsid maturation, DNA packaging
U_L26	early	Yes	Maturational protease
$U_L26.5$	early	Yes	Scaffolding protein
gB (U_L27)	early	Yes	Glycoprotein required for virus entry
U_L28	early	Yes	capsid maturation, DNA packaging
U_L29	early	Yes	single-stranded DNA binding protein, DNA replication
DNA pol (UL30)	early	Yes	DNA replication
U_L32	late	Yes	capsid maturation, DNA packaging
$_L33$	late	Yes	capsid maturation, DNA packaging
U_L35	late	Yes	capsid protein, capsomer tips
U_L38	late	Yes	Capsid protein, triplex
U_L39	early	Yes	Large subunit ribonucleotide reductase
U_L40	early	Yes	Small subunit ribonucleotide reductase
U_L42	late	Yes	Polymerase accessory protein, DNA replication
a-TIF (UL48)	-	Yes	virion-associated transcriptional activator, enhances immediate-early transcription through cellular Oct-1 and CTF binding at "TATGARAT" sites
Helicase/primase (UL52)	early	Yes	DNA replication
a27 (UL54)	immediate-	Yes	Immediate-early regulatory protein, inhibits splicing
IR_L	-	Yes	Internal Long Repeat
R_L/R_S Junction	-	Yes	Joint region, contains "a" sequences
IR_S	-	Yes	Internal Short Repeat
a4	Immediate-	Yes	immediate-early transcriptional activator
Ori_S	-	Yes	Origin of replication
gD (U_S6)	late	Yes	virus entry, binds HSV4EM
TR_S	-	Yes	Terminal Short Repeat
"a"	-	Yes	cis genome cleavage, packaging signal

The HSV-1 genome is divided into six important regions (Figure 1): 1) the ends of the linear molecules, the "a" sequences: these are important in both circularization of the viral DNA, and in packaging the DNA in the virion; 2) the 9,000 bp long repeat (R_L), which encode both an important immediate early regulatory protein (a0) and the promoter of most of the "gene" for the latency associated transcript (LAT); (3) the long unique region (U_L), which is 108,000 bp long, encodes at least 56 distinct proteins (actually more because some ORFs are spliced and

expressed in redundant ways); it contains genes for the DNA replication enzymes and the capsid proteins, as well as many other proteins; 4) the 6,600 bp short repeats (R_S) encode the very important "a" immediate early protein; this is a very powerful transcriptional activator which acts along with a0/ICP0 and a27 (ICP27/UL54) (in the U_L) to stimulate the infected cell for all viral gene expression that leads to viral DNA replication; 5) the origins of replication: the ori_L is in the middle of the U_L region; the ori_S is in the R_S and thus, is present in two copies. All sets of ori's operate during infection to give a very complicated replication complex, very similar to that seen in the replication of phage T4; 6) the 13,000 bp unique short region (U_S) encodes 12 ORFs, a number of which are glycoproteins important in viral host range and response to host defence.

Five HSV-1 genes (a4 or ICP4, a0 or ICP0, a27 or ICP27/U_L 54, a22 or ICP22/U_S1, and a47 or ICP47/U_S12) are expressed and function at the earliest stages of the productive infection cycle. The "immediate-early" or "a" phase of gene expression is mediated by the action of α-TIF through its interaction with cellular transcription factors at specific enhancer elements associated with the individual a-transcript promoters. Activation of the host cell transcriptional machinery by the action of "a" gene products, results in the expression of the "early" or "b" genes. Seven of these are necessary and sufficient for viral DNA replication under all conditions: DNA polymerase (U_L30), DNA binding proteins (U_L42 and U_L29 or ICP8), ORI binding protein (U_L9), and the helicase/primase complex (U_L5, 8, and 52). When sufficient levels of these proteins have accumulated within the infected cell, viral DNA replication ensues. Accessory or "non-essential" proteins for virus replication can be substituted for their function by one or another cell type.

HSV can adopt two different post-infection phenotypes: (i) productive infection or (ii) latent infection. The most recent models posit that when viral DNA migrates to nuclear pods, which are PML-associated subnuclear structures, it is either circularized by cellular DNA repair enzymes acting on the "a" sequences or remains linear through the action of the immediate-early ICP0 protein, which inhibits cellular DNA repair. In the former case, latent infection ensues while in the latter, productive replication takes place.

The vegetative replication of viral DNA which occurs during productive infection, represents a critical and central event in the viral replication cycle. High level of DNA replication irreversibly drives a cell to producing virus, which eventually results in its destruction. DNA replication also has a significant influence on viral gene expression. Early expression is significantly reduced or shut off following the start of DNA replication, while late genes begin to be expressed at high levels.

In a latent infection the viral genome is maintained intact in specific sensory neurons where it is genetically equivalent to that present in the viral particle, but the highly regulated productive cycle cascade of gene expression, so characteristic of herpes virus infections, does not occur. As a consequence, any transcription during latent infection with most herpes viruses is from a very restricted portion of the viral genome, and this transcription is important in some aspect of the process itself. During the latent phase, productive cycle genes are generally transcriptionally and functionally quiescent and only the latency associated transcript (LAT) is expressed. The promoter for the LAT contains neuron-specific cis-acting elements. The maintenance of the HSV genome in latently infected neurons requires no viral gene expression. HSV DNA is maintained as a nucleosomal, circular episome in latent infections and low levels of genome replication may occur or be necessary for the establishment or maintenance of a latent infection from which virus can be efficiently reactivated. The process of reactivation from latency is triggered by stress as well as other signals which are thought to transiently lead to increased transcriptional activity in the harboring neuron. The sensory nerve ganglia survive repeated reactivation without losing function. It appears to also occur without either extensive cytopathology associated with normal vegetative viral replication or with the death of only a very few cells. This process may be augmented by viral genes known to interfere with apoptosis, such as ICP34.5, which act to prevent neuronal death during reactivation where limited replication occurs (Maryam Ahmed et al., J Virol. . 2002 January; 76(2): 717–729. doi: 10.1128/JVI.76.2.717-729.2002.; Guey-Chuen Perng et al., J Virol. . 2002 February; 76(3): 1224–1235. doi: 10.1128/JVI.76.3.1224-1235.2002.; Ling Jin et al., J Virol. . 2005 October; 79(19): 12286–12295. doi: 10.1128/JVI.79.19.12286-12295.2005).

To date, HSV treatments have been limited to antiviral substances that can reduce the level of infection by reducing the level of virus proliferation during vegetative infection. However, such antiviral substances have no effect on quiescent virus during the latency phase.

Other possible treatments which have been investigated include improving the immune response so as to keep the number of viral particles below the proliferation limit and so make episodes of virus replication of less severity, duration or asymptomatic. Some clinical trials are currently running for vaccines but the question of their efficiency on quiescent HSV is still uncertain. (Clin Vaccine Immunol. 2008 Nov;15 (11):1638-43; Pediatric Research, 2001,49:4; Curr Pharm Des. 2007;13(19):1975-88)

The inventors of the present invention have developed a new approach and have created a new type of anti-Herpesviridae agent which can target and eliminate the virus whilst it is inside a target cell by targeting the viral genome with one or more highly specific DNA restriction enzyme. Such highly specific DNA restriction enzymes recognizing specific viral sequences could act on proliferating virus as well as on latent DNA. By this way the inventors of the present invention propose an issue to reminiscences and to herpetic infection.

These materials can be used to manipulate the virus genome so as to elucidate aspects of virus biology and/or as a medicament to directly target and eliminate virus genomic material from the nuclei of infected cells.

According to a first aspect of the present invention there is provided an I-*Cre*I variant, characterized in that at least one of the two I-*Cre*I monomers has at least two substitutions, one in each of the two functional subdomains of the LAGLIDADG core domain situated from positions 26 to 40 and 44 to 77 of I-*Cre*I, said variant being able to cleave a DNA target sequence from the genome of a Herpesviridae virus, and being obtainable by a method comprising at least the steps of:

(a) constructing a first series of I-*Cre*I variants having at least one substitution in a first functional subdomain of the LAGLIDADG core domain situated from positions 26 to 40 of I-*Cre*I,

(b) constructing a second series of I-*Cre*I variants having at least one substitution in a second functional subdomain of the LAGLIDADG core domain situated from positions 44 to 77 of I-*Cre*I,

(c) selecting and/or screening the variants from the first series of step (a) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions -10 to -8 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the Herpesviridae virus genome and (ii) the nucleotide triplet in positions +8 to +10 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the Herpesviridae virus genome,

(d) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions -5 to -3 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the Herpesviridae Virus genome and (ii) the nucleotide triplet in positions +3 to +5 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in position -5 to -3 of said DNA target sequence from the Herpesviridae Virus genome,

(e) selecting and/or screening the variants from the first series of step (a) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions +8 to +10 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the Herpesviridae Virus genome and (ii) the nucleotide triplet in positions -10 to -8 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the Herpesviridae Virus genome,

(f) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions +3 to +5 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the Herpesviridae Virus genome and (ii) the nucleotide triplet in positions -5 to -

3 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the Herpesviridae Virus genome,

(g) combining in a single variant, the mutation(s) in positions 26 to 40 and 44 to 77 of two variants from step (c) and step (d), to obtain a novel homodimeric I-*Cre*I variant which cleaves a sequence wherein (i) the nucleotide triplet in positions -10 to -8 is identical to the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the Herpesviridae Virus genome, (ii) the nucleotide triplet in positions +8 to +10 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the Herpesviridae Virus genome, (iii) the nucleotide triplet in positions -5 to -3 is identical to the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the Herpesviridae Virus genome and (iv) the nucleotide triplet in positions +3 to +5 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the Herpesviridae Virus genome, and/or

(h) combining in a single variant, the mutation(s) in positions 26 to 40 and 44 to 77 of two variants from step (e) and step (f), to obtain a novel homodimeric I-*Cre*I variant which cleaves a sequence wherein (i) the nucleotide triplet in positions +8 to +10 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the Herpesviridae Virus genome and (ii) the nucleotide triplet in positions -10 to -8 is identical to the reverse complementary sequence of the nucleotide triplet in positions +8 to +10 of said DNA target sequence from the Herpesviridae Virus genome, (iii) the nucleotide triplet in positions +3 to +5 is identical to the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the Herpesviridae Virus genome, (iv) the nucleotide triplet in positions -5 to -3 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the Herpesviridae Virus genome,

(i) combining the variants obtained in steps (g) and (h) to form heterodimers, and

(j) selecting and/or screening the heterodimers from step (i) which are able to cleave said DNA target sequence from the Herpesviridae Virus genome.

In the present Patent Application the terms meganuclease (s) and variant (s) and variant meganuclease (s) will be used interchangeably herein.

The inventors have therefore created a new class of meganuclease based reagents which are useful for studying a Herpesviridae Virus *in vitro* and *in vivo*; this class of reagents also represent a potential new class of anti- Herpesviridae Virus medicament, which instead of acting upon the virion or any component thereof, acts upon the intracellular genome of the virus.

To validate their invention, the Inventors have identified a series of DNA targets in the genome of the Herpesviridae Virus, Herpes Simplex Virus (HSV), that are cleaved by I-*Cre*I variants (Table II to VIII and Figures 3 and 24).

Target sequences can be chosen in any region of the Herpesviridae Virus genome, for instance in the coding sequence of a virus gene and in particular in a gene (s) which is essential for the virus. In the present Patent Application essential genes are those genes which must remain active in order for the virus to be able to direct the manufacture and assembly of further virus particles which are able to exit the host cell and infect further cells. In addition to essential genes, other types of essential genetic elements can exist such as the regulatory elements of essential genes and/or structural sequence elements of the virus genome that are necessary for its packaging. For instance if the structure of the virus genetic material can be disrupted for instance by linearization or a strand break, this could make the viral genome susceptible to degradation by the innate anti-viral *in vivo* systems such as nuclease digestion.

For most viruses the majority of genes encoded by the virus are essential and hence inactivation of one or more of these viral genes either directly for instance by a truncation event or indirectly by for instance interrupting a regulatory sequence prevents this virus genome from producing further infective virus particles.

A combinatorial approach was used to entirely redesign the DNA binding domain of the I-*Cre*I protein and thereby engineer novel meganucleases with fully engineered specificity.

In particular the heterodimer of step (i) may comprise monomers obtained in steps (g) and (h), with the same DNA target recognition and cleavage activity properties.

Alternatively the heterodimer of step (i) may comprise monomers obtained in steps (g) and (h), with different DNA target recognition and cleavage activity properties.

In particular the first series of I-*Cre*I variants of step (a) are derived from a first parent meganuclease.

In particular the second series of variants of step (b) are derived from a second parent meganuclease.

In particular the first and second parent meganucleases are identical.

Alternatively the first and second parent meganucleases are different.

In particular the variant may be obtained by a method comprising the additional steps of:

(k) selecting heterodimers from step (j) and constructing a third series of variants having at least one substitution in at least one of the monomers of said selected heterodimers,

(l) combining said third series variants of step (k) and screening the resulting heterodimers for enhanced cleavage activity against said DNA target from the Herpesviridae Virus genome.

The inventors have found that although specific meganucleases can be generated to a particular target in the Herpesviridae Virus genome using the above method, that such meganucleases can be improved further by additional rounds of substitution and selection against the intended target.

In particular in step (k) the substitutions in the third series of variants are introduced by site directed mutagenesis in a DNA molecule encoding said third series of variants, and/or by random mutagenesis in a DNA molecule encoding said third series of variants.

In the additional rounds of substitution and selection, the substitution of residues in the meganucleases can be performed randomly, that is wherein the chances of a substitution event occurring are of equal chance across all the residues of

the meganuclease. Or on a site directed basis wherein the chances of certain residues being subject to a substitution is higher than other residues.

In particular steps (k) and (l) are repeated at least two times and wherein the heterodimers selected in step (k) of each further iteration are selected from heterodimers screened in step (l) of the previous iteration which showed increased cleavage activity against said DNA target from the Herpesviridae Virus genome.

The inventors have found that the meganucleases can be further improved by using multiple iterations of the additional steps (k) and (l).

In particular said substitution(s) in the subdomain situated from positions 26 to 40 of I-*Cre*I are in positions 26, 28, 30, 32, 33, 38 and/or 40.

Through the inventors work they have identified the residues in the first subdomain which when altered have most effect upon altering the I-*Cre*I enzymes specificity.

In particular said substitution(s) in the subdomain situated from positions 44 to 77 of I-*Cre*I are in positions 44, 68, 70, 75 and/or 77.

Through the inventors work they have identified the residues in the second subdomain which when altered have most effect upon altering the I-*Cre*I enzymes specificity.

In particular the variant comprises one or more substitutions in positions 137 to 143 of I-*Cre*I that modify the specificity of the variant towards the nucleotide in positions ± 1 to 2, ± 6 to 7 and/or ± 11 to 12 of the target site in the Herpesviridae Virus genome.

In particular the variant comprises one or more substitutions on the entire I-*Cre*I sequence that improve the binding and/or the cleavage properties of the variant towards said DNA target sequence from the Herpesviridae Virus genome.

As well as specific mutations at the residue identified above, the present invention also encompasses the substitution of any of the residues present in the I-*Cre*I enzyme.

In particular wherein said substitutions are replacement of the initial amino acids with amino acids selected in the group consisting of A, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, Y, C, W, L and V.

In particular the variant is a heterodimer, resulting from the association of a first and a second monomer having different mutations in positions 26 to 40 and 44 to 77 of I-*Cre*I, said heterodimer being able to cleave a non-palindromic DNA target sequence from the Herpesviridae Virus genome.

In particular the variant may be characterized in that it recognizes and cleaves a target sequence which comprises a specific nucleotide or group(s) of nucleotide(s) at one or more of positions ± 1 to 12 which differs from the C1221 target (SEQ ID NO: 2) at least by one nucleotide.

In particular in the positions ± 3 to 5, ± 8 to 10 or ± 1 to 2.

In particular wherein the sequence of nucleotides at the specified position is selected from the following groups:

± 3 to 5 – CAC, GCC, GTG, GGC, GGT, ACC, CTG;

± 8 to 10 – AAA, AGG, TTT, CCT, AAG, ACT, CTT, AGT;

± 1 to 2 – GTAC.

As explained above the I-*Cre*I enzyme acts as a dimer, by ensuring that the variant is a heterodimer this allows a specific combination of two different I-*Cre*I monomers which increases the possible targets cleaved by the variant.

In particular the heterodimeric variant is an obligate heterodimer variant having at least one pair of mutations in corresponding residues of the first and the second monomers which mediate an intermolecular interaction between the two I-*Cre*I monomers, wherein the first mutation of said pair(s) is in the first monomer and the second mutation of said pair(s) is in the second monomer and said pair(s) of mutations impairs the formation of functional homodimers from each monomer without preventing the formation of a functional heterodimer, able to cleave the genomic DNA target from the Herpesviridae Virus genome.

The inventors have previously established a number of residue changes which can ensure an I-*Cre*I monomer is an obligate heterodimer (WO2008/093249).

In particular the monomers have at least one of the following pairs of mutations, respectively for the first and the second monomer:

a) the substitution of the glutamic acid in position 8 with a basic amino acid, preferably an arginine (first monomer) and the substitution of the lysine in

position 7 with an acidic amino acid, preferably a glutamic acid (second monomer); the first monomer may further comprise the substitution of at least one of the lysine residues in positions 7 and 96, by an arginine.

b) the substitution of the glutamic acid in position 61 with a basic amino acid, preferably an arginine (first monomer) and the substitution of the lysine in position 96 with an acidic amino acid, preferably a glutamic acid (second monomer); the first monomer may further comprise the substitution of at least one of the lysine residues in positions 7 and 96, by an arginine

c) the substitution of the leucine in position 97 with an aromatic amino acid, preferably a phenylalanine (first monomer) and the substitution of the phenylalanine in position 54 with a small amino acid, preferably a glycine (second monomer); the first monomer may further comprise the substitution of the phenylalanine in position 54 by a tryptophane and the second monomer may further comprise the substitution of the leucine in position 58 or lysine in position 57, by a methionine, and

d) the substitution of the aspartic acid in position 137 with a basic amino acid, preferably an arginine (first monomer) and the substitution of the arginine in position 51 with an acidic amino acid, preferably a glutamic acid (second monomer).

In particular the variant, which is an obligate heterodimer, wherein the first and the second monomer, respectively, further comprises the D137R mutation and the R51D mutation.

In particular the variant, which is an obligate heterodimer, wherein the first monomer further comprises the K7R, E8R, E61R, K96R and L97F or K7R, E8R, F54W, E61R, K96R and L97F mutations and the second monomer further comprises the K7E, F54G, L58M and K96E or K7E, F54G, K57M and K96E mutations.

Alternatively there is provided a single-chain chimeric meganuclease which comprises two monomers or core domains of one or two variant(s) according to the first aspect of the present invention, or a combination of both.

The single chain meganuclease of the present invention further comprises obligate heterodimer mutations as described above so as to obtain single chain obligate heterodimer meganuclease variants.

An alternative approach to ensuring that the variant consists of a specific combination of monomers is to link the selected monomers for instance using a peptide linker.

In particular the single-chain meganuclease comprises a first and a second monomer according to the first aspect of the present invention, connected by a peptidic linker.

In particular the DNA target is within an essential gene or regulatory element or structural element of the Herpesviridae Virus genome.

Most particularly the Herpesviridae Virus is a virus which causes a disease in higher animals and in particular mammals.

In particular the Herpesviridae Virus is a virus selected from the group comprising: herpes simplex virus type 1, herpes simplex virus type 2, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human herpes virus 6 (variants A and B), human herpes virus 7, Kaposi's sarcoma virus and human herpes virus 8.

Multiple examples of genomic sequences for all these viruses are available from public databases such as the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) or the virus genomics and bioinformatics resources centre at University College London (http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html).

These publicly available resources together with the detailed materials and methods described in the present Patent Application mean that meganuclease variants cleaving appropriate targets in their genomes can be generated and that in turn these variants can be used to cleave the viral genomic material *in vivo* for therapeutic and/or research purposes in accordance with the various aspects of the present invention.

In particular the herpes simplex virus is Herpes Simplex Virus (HSV) Type 1 or Type 2.

In particular the DNA target sequence is from a Herpes Simplex Virus Type 1 or Type 2.

In particular the variants may be selected from the group consisting of SEQ ID NO: 25 to 36, 40 to 90, 93 to 151, 153 to 168, 171 to 246, 249 to 252, 267 to 273, 275 to 288, 290 to 433, 436 to 445, 458 to 463 and 470 to 471.

In particular the single chain variants may be selected from the group consisting of SEQ ID NO: 253 to 261 and 446 to 454.

In particular said DNA target is selected from the group consisting of the sequences SEQ ID NO: 8 to13 and 17 to 24.

In particular said DNA target is within a DNA sequence essential for HSV replication, viability, packaging or virulence.

In particular the DNA target is within an open reading frame of the HSV genome, selected from the group: RL2, RS1, US2 or UL19.

In the present Patent Application the inventors provide meganuclease variants which can cleave targets in the RL2/ICP0 gene (targets HSV 12 and 4, SEQ ID NO: 20 and 17 respectively); in the RS1 gene (targets HSV 13 and 14, SEQ ID NO: 21 and 22 respectively); in the US2 gene (target HSV 1, SEQ ID NO: 23) and in the UL19 gene (target HSV 2, SEQ ID NO: 24). The cleavage of these sites in the HSV genome *in vivo* would therefore disrupt the sequence encoding the corresponding gene and thereby following a disruption and/or alteration of these gene sequences inactivate the HSV genome.

The RL2 gene encodes an important immediate early transcription factor acting as a regulatory protein (a0). This gene is considered as non essential due to its possible replacement by cellular transcription factors. However, it has been considered of major interest due to its localization in TRL, which is essential for HSV-1. Moreover, the central role of a0 during acute infection, latency establishment and virus reactivation as lead us to consider ICP0 as an integrator of essential signals. ICP0 gene is located in the 9 kb RL region repeated twice in HSV genome. This RL region encodes most of the gene for the latency associated transcript. This region is the unique active region during latency phase. Thus, targeting ICP0 gene would allow targeting an "opened" genomic sequence of quiescent virus and an important immediate early protein during virus infection and vegetative production. Many meganucleases can be built to recognize sequences in ICP0 gene. HSV4 described latter is one of them.

HSV12 is an example of a target from within the RL2 gene for which meganuclease variants can be generated. The HSV12 target sequence (atgttccccgtctccatgtccagg, SEQ ID NO: 20) is located at positions 5168-5194bp and 121180-121203bp in exon 3 of the RL2 gene repeated from positions 2086 to 5698 and from positions 120673 to 124285. Shown in Table II are two heterodimeric I-*Cre*I variants which recognize and cleave the HSV12 target. Throughout the present Patent Application the sequence of the I-*Cre*I variants described herein may be made using the following notation 24V33C etc. In this notation the numeral refers to the amino acid number in the I-*Cre*I monomer and the letter refers to the amino acid present in this variant. If a residue is not explicitly listed this means this residue is identical to the residue in the wild type or parent I-*Cre*I monomer as appropriate.

Table II: Example of meganuclease variants cleaving the HSV12 (atgttccccgtctccatgtccagg) target
HSV12.3-M1 (SEQ ID NO: 25)
24V33C38S44I50R70S75N77R132V
+
HSV12.4-ME-132V (SEQ ID NO: 26)
19S8K30R33S44K66H68Y70S77T87I132V139R163S
HSV12.3-M1-80K (SEQ ID NO: 27)
24V33C38S44I50R70S75N77R80K132V
+
HSV12.4-ME-132V (SEQ ID NO: 26)
19S8K30R33S44K66H68Y70S77T87I132V139R163S

ICP4 (RS1) gene is located in the RS region (6.6 kb) repeated twice in HSV-1 genome. IRS and TRS are located from positions 125974 to 132604 and from 145585 to 152259. The ICP4 virus essential gene, functions at the earliest stages of the productive infection cycle. RS1 encodes the immediate early transcription activator (a4) which, upon infection, directs cellular machinery to viral gene expression. This protein functions in association with ICP0 (a0) and ICP27 (a27) to improve viral gene expression and viral mRNA translation. Thus targeting ICP4 gene is of major interest in a meganuclease mediated antiviral approach.

The ICP4 gene can be targeted by many meganucleases. For example, sequences aggggacggggaacagcgggtggt (SEQ ID NO: 21) and ctcttcttcgtcttcgggggtcgc (SEQ ID NO: 22) are recognized and efficiency cleaved by I-Cre I variants HSV13 and HSV14. HSV13 target sequence is located from positions 128569 to 128592 and from 149641 to 149664 (NC_001806). An example of I-*Cre*I variant targeting HSV13 is shown in Table III.

Table III: example of meganuclease variants cleaving the HSV13 (atgttccccgtctccatgtccagg) target

HSV13-3-M15-19S (SEQ ID NO: 28)

6S 19S 28E 33R 38R 40K 43L 44N 68H 70S 75Y 77N 79G 80K

+

HSV13-4-MD (SEQ ID NO: 29)

30R 44R 60E 68Y 70S 75N 77D 80G

HSV13-3-M16-19S (SEQ ID NO: 30)

6S 19S 28E 33R 38R 40K 44N 68H 70S 75Y 77N 79G 80K 105A

+

HSV13-4-MD (SEQ ID NO: 29)

30R 44R 60E 68Y 70S 75N 77D 80G

HSV14 target sequence is located from positions 128569 to 128592 and from 149641 to 149664 (NC_001806). An example of I-*Cre*I variant targeting HSV14 is shown in Table IV.

Table IV: example of meganuclease variants cleaving the HSV14 (ctcttcttcgtcttcgggggtcgc) sequence

HSV14.3-MA-19S (SEQ ID NO: 31)

19S33G38C44K66H68Y70S77T

+

HSV14.4 –MB (SEQ ID NO: 32)

33H40R43L44K54I68A70S115V129A

The US2 gene is located in the US region of the HSV-1 genome. The 12 open reading frames contained in this 13kb region are implicated in virus

defense against host response, most of gene products are glycoproteins. The US2 gene is located from positions 134053 to 134928, less than 2kb downstream the IRS region coding a4. This gene encodes a possible envelope-associated protein which interacts with cytokeratin 18. By targeting this gene the inventors of the present invention wanted to evaluate the accessibility of this locus as well as have an evaluation of the cleavage effect of this non essential viral gene toward HSV infection.

Among the multiple sequences recognized by I-*Cre*I variants, the HSV1 target sequence atgggacgtcgtaaggggcctgg, (SEQ ID NO: 23) (134215 – 134238) is targeted by meganuclease as detailed in Table V below.

Table V: example of meganuclease variants cleaving the HSV1 (atgggacgtcgtaaggggcctgg) target

HSV1.3-M5 (SEQ ID NO: 470)	
30R33G38T106P	
+	
HSV1.4-MF (SEQ ID NO: 471)	
30G38R44K57E70E75N108V	

HSV2 is a 24 bp (non-palindromic) target present in the UL19 gene encoding the HSV-1 major capsid protein. This 5.7kb gene in present in one copy in the locus 35023 to 40768 of the UL region. The HSV1-major capsid protein is expressed without maturation from an ORF located from 36404 to 40528. The target HSV2 is located from nucleotide 36966 to 36989 (accession number NC_001806. The HSV2 target is recognized and cleaved by the meganuclease shown in Table VI below.

Table VI: example of meganuclease variants cleaving the HSV2 (ataaactcacacacggcgtcctgg) target

HSV2.3-M1(SEQ ID NO: 33)	
44D68T70S75R77R80K	
+	
HSV2.4-MC (SEQ ID NO: 34)	
28E38R40K44K54I70S75N	

HSV4 is a 24 bp (non-palindromic) target present in the RL2 gene encoding the ICP0 or a0 protein. This 3,6kb gene repeated twice in TRL (2086 to 5698) and IRL (120673 to 124285) regions is formed of three exons : position 2261 to 2317, 3083 to 3749, 3886 to 5489 and 120882 to 122485,122622 to 123288, 124054 to 124110. The target sequence present in exon 2 corresponds to positions 3498 to 3521 and 122850 to 122873 in the two copies of the HSV-1 ICP_0 gene (accession number NC_001806). The HSV4 target is recognized and cleaved by the meganuclease shown in Table VII below.

Table VII: example of meganuclease variants cleaving the HSV4 (ccaagctggtgtacctgatagtgg) target

HSV4.3 optimised variant (SEQ ID NO: 35)44M70A80K132V146K156G

+

HSV4.4 optimised variant (SEQ ID NO: 36)

32E38Y44A68Y70S75Y77K105A

According to a second aspect of the present invention there is provided a polynucleotide fragment encoding the variant according to the first aspect of the present invention.

According to a third aspect of the present invention there is provided an expression vector comprising at least one polynucleotide fragment according to the second aspect of the present invention.

In particular the expression vector, includes a targeting construct comprising a sequence to be introduced flanked by sequences sharing homologies with the regions surrounding said DNA target sequence from the Herpesviridae Virus genome.

One important use of a variant according to the present invention is in increasing the incidence of homologous recombination events at or around the site where the variant cleaves its target. The present invention therefore also relates to a unified genetic construct which encodes the variant under the control of suitable regulatory sequences as well as sequences homologous to portions of the Herpesviridae Virus genome surrounding the variant DNA target site. Following cleavage of the target site by the variant these homologous portions can act as a

complementary sequences in a homologous recombination reactions with the Herpesviridae Virus genome replacing the existing Herpesviridae Virus genome sequence with a new sequence engineered between the two homologous portions in the unified genetic construct.

Preferably, homologous sequences of at least 50 bp, preferably more than 100 bp and more preferably more than 200 bp are used. Shared DNA homologies are located in regions flanking upstream and downstream the site of the break and the DNA sequence to be introduced should be located between the two arms.

Therefore, the targeting construct is preferably from 200 bp to 6000 bp, more preferably from 1000 bp to 2000 bp; it comprises: a sequence which has at least 200 bp of homologous sequence flanking the target site, for repairing the cleavage and a sequence for inactivating the Herpesviridae Virus genome and/or a sequence of an exogeneous gene of interest.

For the insertion of a sequence, DNA homologies are generally located in regions directly upstream and downstream to the site of the break (sequences immediately adjacent to the break; minimal repair matrix). However, when the insertion is associated with a deletion of ORF sequences flanking the cleavage site, shared DNA homologies are located in regions upstream and downstream the region of the deletion.

A vector which can be used in the present invention 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. adenoassociated 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, cytomegalovirus), 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 leukosissarcoma, 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).

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 (HRPT) for eukaryotic cell culture; TRP1 for *S. cerevisiae*; tetracycline, rifampicin or ampicillin resistance in *E. coli*.

In particular for the purposes of gene therapy and in accordance with a preferred embodiment of the present invention, the viral vector is selected from the group comprising lentiviruses, Adeno-associated viruses (AAV) and Adenoviruses.

A particular advantage of using virus vectors to deliver a variant which cleaves a virus target for a therapeutic purpose, is that the administration of the virus vector *per se* will illicit an immune response from the treated organism which in turn will impede the virus infection.

In accordance with another aspect of the present invention the variant and targeting construct may be on different nucleic acid constructs.

In accordance with another aspect of the present invention the variant in a peptide form and the targeting construct as a nucleic acid molecule may be used in combination.

In particular, wherein the sequence to be introduced is a sequence which inactivates the Herpesviridae Virus genome.

In particular, wherein the sequence which inactivates the Herpesviridae Virus genome comprises in the 5' to 3' orientation: a first transcription termination sequence and a marker cassette including a promoter, the marker open reading frame and a second transcription termination sequence, and said sequence interrupts the transcription of the coding sequence.

In particular, wherein said sequence sharing homologies with the regions surrounding DNA target sequence is from the Herpesviridae Virus genome is a fragment of the Herpesviridae Virus genome comprising sequences upstream and downstream of the cleavage site, so as to allow the deletion of coding sequences flanking the cleavage site.

According to a fourth aspect of the present invention there is provided a host cell which is modified by a polynucleotide according to a second aspect of the present invention or a vector according to a third aspect of the present invention.

A cell according to the present invention may be made according to a method, comprising at least the step of:

(a) introducing into a cell, a meganuclease, as defined above, so as to induce a double stranded cleavage at a site of interest of the Herpesviridae Virus genome comprising a DNA recognition and cleavage site of said meganuclease, and thereby generate a cell comprising at least one modified Herpesviridae Virus genome, in particular having repaired the double-strands break, by non-homologous end joining, and

(b) isolating the cell of step(a), by any appropriate mean.

The cell which is modified may be any cell of interest. For making transgenic/knock-out animals, the cells are pluripotent precursor cells such as embryo-derived stem (ES) cells, which are well-known in the art. For making recombinant cell lines, the cells may advantageously be human cells, for example HSV infecting cell lines such as human hepatoblastoma cell lines, hepatocellular carcinoma (Fellig et al., (2004) Biochemical and Biophysical Research Communications, Volume 321, Issue 2, Pages 269-274) or a more general cell line such as CHO or HEK293 (ATCC # CRL-1573) cells. The meganuclease can be provided directly to the cell or through an expression vector comprising the polynucleotide sequence encoding said meganuclease linked to regulatory sequences suitable for directing its expression in the cell used.

In addition to generating cells comprising modified Herpesviridae Virus genomes, the present invention also relates to modifying a copy(ies) of the Herpesviridae Virus genome which have been genomically integrated into the host

cell genome. Such modified cell lines are useful for elucidating aspects of virus biology amongst many other potential uses.

Such a modified cell line would have a number of potential uses including the elucidation of aspects of the biology of the modified Herpesviridae Virus genome as well as a model for screening compounds and other substances for therapeutic effects against cells comprising the modified Herpesviridae Virus genome.

The present invention therefore also relates to meganuclease variants which can recognise and cleave targets comprised in genomic insertions of viruses which do not normally insert into the host cell genome. The non-specific insertion of viral genetic material into the host cell genome as a disease causing mechanism is currently being investigated.

According to a fifth aspect of the present invention there is provided a non-human transgenic animal or plant which is modified by a polynucleotide according to a second aspect of the present invention or a vector according to a third aspect of the present invention. In particular these non-human transgenic animals or transgenic plants comprise a copy of the Herpesviridae Virus genome integrated into the genome of the host organism.

The subject-matter of the present invention is also a method for making a transgenic animal comprising an integrated Herpesviridae Virus genome, comprising at least the step of:

(a) introducing into a pluripotent precursor cell or an embryo of an animal, a meganuclease, as defined above, so as to induce a double stranded cleavage at a site of interest of the integrated Herpesviridae Virus genome comprising a DNA recognition and cleavage site of said meganuclease, and thereby generate a genomically modified precursor cell or embryo having repaired the double-strands break by non-homologous end joining,

(b) developing the genomically modified animal precursor cell or embryo of step (a) into a chimeric animal, and

(c) deriving a transgenic animal from a chimeric animal of step (b).

Alternatively, the Herpesviridae Virus genome may be inactivated by insertion of a sequence of interest by homologous recombination between the

genome of the animal and a targeting DNA construct according to the present invention.

Such transgenic animals/plants therefore can be used as model organisms to study the effects of genomically integrated virus genetic material which has been either introduced using a meganuclease based homologous recombination system or alternatively has been altered using a specific meganuclease variant.

In particular the targeting DNA is introduced into the cell under conditions appropriate for introduction of the targeting DNA into the site of interest.

In particular, step (b) comprises the introduction of the genomically modified precursor cell obtained in step (a), into blastocysts, so as to generate chimeric animals.

Such a transgenic animal could be used as a multicellular animal model to elucidate aspects of HSV biology by means of engineering the provirus present in the progenitor cell line. Such transgenic animals also could be used to screen and characterise the effects of novel anti-HSV medicaments.

In particular the targeting DNA construct is inserted in a vector.

For making transgenic animals/recombinant cell lines, including human cell lines expressing an heterologous protein of interest, the targeting DNA comprises the sequence of the exogenous gene encoding the protein of interest, and eventually a marker gene, flanked by sequences upstream and downstream of and essential gene in the Herpesviridae Virus genome, as defined above, so as to generate genomically modified cells (animal precursor cell or embryo/animal or human cell) having replaced the HSV gene by the exogenous gene of interest, by homologous recombination.

The exogenous gene and the marker gene are inserted in an appropriate expression cassette, as defined above, in order to allow expression of the heterologous protein/marker in the transgenic animal/recombinant cell line.

The meganuclease can be used either as a polypeptide or as a polynucleotide construct encoding said polypeptide. It is introduced into somatic cells of an individual, by any convenient means well-known to those in the art, which are appropriate for the particular cell type, alone or in association with either at least an appropriate vehicle or carrier and/or with the targeting DNA.

According to the present invention, the meganuclease (polypeptide) can be associated with:

- liposomes, polyethyleneimine (PEI); in such a case said association is administered and therefore introduced into somatic target cells.

- membrane translocating peptides (Bonetta, The Scientist, 2002, 16, 38; Ford et al., Gene Ther., 2001, 8, 1-4 ; Wadia and Dowdy, Curr. Opin. Biotechnol., 2002, 13, 52-56); in such a case, the sequence of the variant/single-chain meganuclease is fused with the sequence of a membrane translocating peptide (fusion protein).

Alternatively, the meganuclease (polynucleotide encoding said meganuclease) and/or the targeting DNA is inserted in a vector. Vectors comprising targeting DNA and/or nucleic acid encoding a meganuclease can be introduced into a cell by a variety of methods (e.g., injection, direct uptake, projectile bombardment, liposomes, electroporation). Meganucleases can be stably or transiently expressed into cells using expression vectors. Techniques of expression in eukaryotic cells are well known to those in the art. (See Current Protocols in Human Genetics: Chapter 12 "Vectors For Gene Therapy" & Chapter 13 "Delivery Systems for Gene Therapy"). Optionally, it may be preferable to incorporate a nuclear localization signal into the recombinant protein to be sure that it is expressed within the nucleus.

Once in a cell, the meganuclease and if present, the vector comprising targeting DNA and/or nucleic acid encoding a meganuclease are imported or translocated by the cell from the cytoplasm to the site of action in the nucleus or the cytoplasm.

According to a sixth aspect of the present invention there is provided a transgenic plant which is modified by a polynucleotide according to a second aspect of the present invention or a vector according to a third aspect of the present invention.

According to a further aspect of the present invention there is provided the use of at least one variant or at least one single-chain chimeric meganuclease according to the first aspect of the present invention, or at least one vector according to the third aspect of the present invention, for Herpesviridae Virus genome engineering, for non-therapeutic or therapeutic purposes.

In particular the variant or single-chain chimeric meganuclease, or vector is associated with a targeting DNA construct.

In particular the use of the variant is for inducing a double-strand break in a site of interest of the Herpesviridae Virus genome comprising a Herpesviridae Virus genomic DNA target sequence, thereby inducing a DNA recombination event, a DNA loss or DNA degradation.

According to the invention, said double-strand break is for: modifying a specific sequence in the Herpesviridae Virus genome, so as to induce cessation of a Herpesviridae Virus genome function such as replication, attenuating or activating the Herpesviridae Virus genome or a gene therein, introducing a mutation into a site of interest of a Herpesviridae Virus gene, introducing an exogenous gene or a part thereof, inactivating or deleting the Herpesviridae Virus genome or a part thereof or leaving the DNA unrepaired and degraded.

According to this aspect of the present invention the use of the meganuclease according to the present invention, comprises at least the following steps: 1) introducing a double-strand break at a site of interest of the Herpesviridae Virus genome comprising at least one recognition and cleavage site of said meganuclease, by contacting said cleavage site with said meganuclease ; 2) providing a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus. Said meganuclease can be provided directly to the cell or through an expression vector comprising the polynucleotide sequence encoding said meganuclease and suitable for its expression in the used cell. This strategy is used to introduce a DNA sequence at the target site, for example to generate knock-in or knock-out animal models or cell lines that can be used for drug testing.

According to a further aspect of the present invention the use of the meganuclease, comprises at least the following steps: 1) introducing a double-strand break at a site of interest of the Herpesviridae Virus genome comprising at least one recognition and cleavage site of said meganuclease, by contacting said cleavage site with said meganuclease; 2) maintaining said broken genomic locus under conditions appropriate for homologous recombination with chromosomal DNA sharing homologies to regions surrounding the cleavage site.

According to still further aspect of the present invention the use of the meganuclease, comprises at least the following steps: 1) introducing a double-strand break at a site of interest of the Herpesviridae Virus genome comprising at least one recognition and cleavage site of said meganuclease, by contacting said cleavage site with said meganuclease; 2) maintaining said broken genomic locus under conditions appropriate for repair of the double-strands break by non-homologous end joining.

According to a further aspect of the present invention the variant is used for genome therapy or the making of knock-out Herpesviridae Virus genomes, the sequence to be introduced is a sequence which inactivates the Herpesviridae Virus genome. All Herpesviridae Virus genomes present in the cell have to be targeted in order to totally inactivate the pathogenicity of the virus. In addition, the sequence may also delete the Herpesviridae Virus genome or part thereof, and introduce an exogenous gene or part thereof (knock-in/gene replacement). For making knock-in Herpesviridae Virus genomes the DNA which repairs the site of interest may comprise the sequence of an exogenous gene of interest, and a selection marker, such as the G418 resistance gene. Alternatively, the sequence to be introduced can be any other sequence used to alter the DNA in some specific way including a sequence used to modify a specific sequence, to attenuate or activate the endogenous gene of interest in the Herpesviridae Virus genome or to introduce a mutation into a site of interest in the Herpesviridae Virus genome.

Inactivation of the Herpesviridae Virus genome may occur by insertion of a transcription termination signal that will interrupt the transcription of an essential gene such as a viral DNA polymerase and result in a truncated protein. In this case, the sequence to be introduced comprises, in the 5' to 3' orientation: at least a transcription termination sequence (polyA1), preferably said sequence further comprises a marker cassette including a promoter and the marker open reading frame (ORF) and a second transcription termination sequence for the marker gene ORF (polyA2). This strategy can be used with any variant cleaving a target downstream of the relevant gene promoter and upstream of the stop codon.

Inactivation of the Herpesviridae Virus genome may also occur by insertion of a marker gene within an essential gene of Herpesviridae Virus, which

would disrupt the coding sequence. The insertion can in addition be associated with deletions of ORF sequences flanking the cleavage site and eventually, the insertion of an exogenous gene of interest (gene replacement).

In addition, inactivation of Herpesviridae Virus may also occur by insertion of a sequence that would destabilize the mRNA transcript of an essential gene.

The present invention also provides a composition characterized in that it comprises at least one variant as defined above (variant or single-chain derived chimeric meganuclease) and/or at least one expression vector encoding the variant, as defined above.

In particular the composition comprises a targeting DNA construct comprising a sequence which inactivates the Herpesviridae Virus genome, flanked by sequences sharing homologies with the Herpesviridae Virus genomic DNA cleavage site of said variant, as defined above.

Preferably, said targeting DNA construct is either included in a recombinant vector or it is included in an expression vector comprising the polynucleotide(s) encoding the variant according to the invention.

The subject-matter of the present invention is also the use of at least one meganuclease and/or one expression vector, as defined above, for the preparation of a medicament for preventing, improving or curing a Herpesviridae Virus and in particular a HSV infection in an individual in need thereof.

The subject-matter of the present invention is also the use of at least one variant and/or one expression vector, as defined above, for the preparation of a medicament for preventing, improving or curing a pathological condition associated with a Herpesviridae Virus infection in an individual in need thereof.

In particular compositions according to the present invention may comprise more than one variant. The genome of a virus is subject to more changes than the genome of a higher organism such as a prokaryotic or eukaryotic cell. Therefore in a population of viruses in an infected individual it is possible that the DNA target recognized by the variant will be altered and hence the variant will not cut this target. To lessen the potential effects of such mutants, compositions according to the present invention may comprise variants which recognize and cleave different

targets in the Herpesviridae Virus genome. The chances of a particular virus having mutations in all the various targets cleaved by the variants contained in the composition are very low and hence the virus will be recognized and acted upon by at least one of the variants present in the composition.

The use of the meganuclease may comprise at least the step of (a) inducing in at least one Herpesviridae Virus genome contained in an at least one cell of infected individual a double stranded cleavage at a site of interest of the Herpesviridae Virus genome comprising at least one recognition and cleavage site of said meganuclease by contacting said cleavage site with said meganuclease, and (b) introducing into said at least one cell a targeting DNA, wherein said targeting DNA comprises (1) DNA sharing homologies to the region surrounding the cleavage site and (2) DNA which inactivates the Herpesviridae Virus genome upon recombination between the targeting DNA and the Herpesviridae Virus genome, as defined above. The targeting DNA is introduced into the Herpesviridae Virus genome under conditions appropriate for introduction of the targeting DNA into the site of interest. The targeting construct may comprise sequences for deleting the Herpesviridae Virus genome or a portion thereof and introducing the sequence of an exogenous gene of interest (gene replacement).

Alternatively, the Herpesviridae Virus genome may be inactivated by the mutagenesis of an open reading frame therein, by the repair of the double-strands break by non-homologous end joining. In the absence of a repair matrix, the DNA double-strand break in an exon will be repaired essentially by the error-prone Non Homologous End Joining pathway NHEJ, resulting in small deletions (a few nucleotides), that will inactivate the cleavage site, and result in frame shift mutation.

In this case the use of the meganuclease comprises at least the step of: inducing in virus infected tissue(s) of the an individual a double stranded cleavage at a site of interest of in the Herpesviridae Virus genome comprising at least one recognition and cleavage site of the meganuclease by contacting the cleavage site with the meganuclease, and thereby inducing mutagenesis of an open reading frame in the Herpesviridae Virus genome by repair of the double-strands break by non-homologous end joining.

According to the present invention, said double-stranded cleavage may be induced, *ex vivo* by introduction of said meganuclease into infected cells isolated for instance from the circulatory system of the donor/individual and then transplantation of the modified cells back into the diseased individual.

The subject-matter of the present invention is also a method for preventing, improving or curing Herpes Simplex Virus Type 1 or Type 2 infection, in an individual in need thereof, said method comprising at least the step of administering to said individual a composition as defined above, by any means.

For purposes of therapy, the meganucleases and a pharmaceutically acceptable excipient are administered in a therapeutically effective amount. Such a combination is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of the recipient. In the present context, an agent is physiologically significant if its presence results in a decrease in the severity of one or more symptoms of the targeted Herpesviridae Virus and in particular Herpes Simplex Virus Type 1 or 2 infection.

In particular as far as possible the meganuclease comprising compositions should be non-immunogenic, i.e., engender little or no adverse immunological response. A variety of methods for ameliorating or eliminating deleterious immunological reactions of this sort can be used in accordance with the invention. One means of achieving this is to ensure that the meganuclease is substantially free of N-formyl methionine. Another way to avoid unwanted immunological reactions is to conjugate meganucleases to polyethylene glycol ("PEG") or polypropylene glycol ("PPG") (preferably of 500 to 20,000 Daltons average molecular weight (MW)). Conjugation with PEG or PPG, as described by Davis et al. (US 4,179,337) for example, can provide non-immunogenic, physiologically active, water soluble endonuclease conjugates with anti-viral activity. Similar methods also using a polyethylene--polypropylene glycol copolymer are described in Saifer et al. (US 5,006,333).

Definitions

Throughout the present Patent Application a number of terms and features are used to present and describe the present invention, to clarify the meaning

of these terms a number of definitions are set out below and wherein a feature or term is not otherwise specifically defined or obvious from its context the following definitions apply.

- 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.

- Amino acid substitution means the replacement of one amino acid residue with another, for instance the replacement of an Arginine residue with a Glutamine residue in a peptide sequence is an amino acid substitution.

- Altered/enhanced/increased cleavage activity, refers to an increase in the detected level of meganuclease cleavage activity, see below, against a target DNA sequence by a second meganuclease in comparison to the activity of a first meganuclease against the target DNA sequence. Normally the second meganuclease is a variant of the first and comprise one or more substituted amino acid residues in comparison to the first meganuclease.

- by "beta-hairpin" it is intended two consecutive beta-strands of the antiparallel beta-sheet of a LAGLIDADG homing endonuclease core domain ($\beta_1\beta_2$ or $\beta_3\beta_4$) which are connected by a loop or a turn,

- by "chimeric DNA target" or "hybrid DNA target" it 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 at least two separate subdomains (combined DNA target).

- Cleavage activity: the cleavage activity of the variant according to the invention may be measured by any well-known, *in vitro* or *in vivo* cleavage assay, such as those described in the International PCT Application WO 2004/067736; Epinat *et al.*, Nucleic Acids Res., 2003, 31, 2952-2962; Chames *et al.*, Nucleic Acids Res., 2005, 33, e178; Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458, and Arnould *et al.*, J. Mol. Biol., 2007, 371, 49-65. For example, 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. The reporter vector comprises two truncated, non-functional copies of a reporter gene (direct repeats) and the genomic

(non-palindromic) DNA target sequence within the intervening sequence, cloned in a yeast or a mammalian expression vector. Usually, the genomic DNA target sequence comprises one different half of each (palindromic or pseudo-palindromic) parent homodimeric meganuclease target sequence. Expression of the heterodimeric variant results in a functional endonuclease which is able to cleave the genomic DNA target sequence. This cleavage induces homologous recombination between the direct repeats, resulting in a functional reporter gene (LacZ, for example), whose expression can be monitored by an appropriate assay. The specificity of the cleavage by the variant may be assessed by comparing the cleavage of the (non-palindromic) DNA target sequence with that of the two palindromic sequences cleaved by the parent homodimeric meganucleases or compared with wild type meganuclease.

- by "selection or selecting" it is intended to mean the isolation of one or more meganuclease variants based upon an observed specified phenotype, for instance altered cleavage activity. This selection can be of the variant in a peptide form upon which the observation is made or alternatively the selection can be of a nucleotide coding for selected meganuclease variant.

- by "screening" it is intended to mean the sequential or simultaneous selection of one or more meganuclease variant (s) which exhibits a specified phenotype such as altered cleavage activity.

- by "derived from" it is intended to mean a meganuclease variant which is created from a parent meganuclease and hence the peptide sequence of the meganuclease variant is related to (primary sequence level) but derived from (mutations) the sequence peptide sequence of the parent meganuclease.

- by "domain" or "core domain" it is intended the "LAGLIDADG homing endonuclease core domain" which is the characteristic $\alpha_1\beta_1\beta_2\alpha_2\beta_3\beta_4\alpha_3$ 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 ($\beta_1\beta_2\beta_3\beta_4$) 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-*Cre*I (163 amino acids), the LAGLIDADG homing endonuclease core domain corresponds to the residues 6 to 94.

- by "DNA target", "DNA target sequence", "target sequence", "target-site", "target", "site"; "site of interest"; "recognition site", "recognition sequence", "homing recognition site", "homing site", "cleavage site" it is intended a 20 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 such as I-*Cre*I, or a variant, or a single-chain chimeric meganuclease derived from I-*Cre*I. 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 DNA target is defined by the 5' to 3' sequence of one strand of the double-stranded polynucleotide, as indicated for C1221 (see figure 3, SEQ ID NO: 2). Cleavage of the DNA target occurs at the nucleotides at 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 an I-*Cre*I 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" it is intended the portion of the DNA target which is bound by each LAGLIDADG homing endonuclease core domain.

- by "DNA target sequence from the HSV genome" it is intended a 20 to 24 bp sequence of the HSV genome which is recognized and cleaved by a meganuclease variant. In particular the DNA target sequence from then HSV genome is in an essential gene sequence and/or within an essential regulatory sequence and/or within an essential structural sequence of the HSV genome.

- by "first/second/third/n^{th} series of variants" it is intended a collection of variant meganucleases, each of which comprises one or more amino acid substitution in comparison to a parent meganuclease from which all the variants in the series are derived.

- by "functional variant" it is intended a variant which is able to cleave a DNA target sequence, preferably said target is a new target which is not cleaved by the parent meganuclease. 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 "heterodimer" it is intended to mean a meganuclease comprising two non-identical monomers. In particular the monomers may differ from each other in their peptide sequence and/or in the DNA target half-site which they recognise and cleave.

- 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 %.

- by "*I-CreI*" it is intended the wild-type I-*Cre*I having the sequence of pdb accession code 1g9y, corresponding to the sequence SEQ ID NO: 1 in the sequence listing.

- by "*I-CreI* variant with novel specificity" it is intended a variant having a pattern of cleaved targets different from that of the parent meganuclease. 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. In the present Patent Application the *I-CreI* variants described comprise an additional Alanine after the first Methionine of the wild type *I-CreI* sequence and three additional amino acid residues (SEQ ID NO: 3). In the present Application, *I-CreI* variants may be homodimers (meganuclease comprising two identical monomers) or heterodimers (meganuclease comprising two non-identical monomers).

These variants also comprise two additional Alanine residues and an Aspartic Acid residue after the final Proline of the wild type *I-CreI* sequence. These additional residues do not affect the properties of the enzyme and to avoid confusion these additional residues do not affect the numeration of the residues in *I-CreI* or a variant referred in the present Patent Application, as these references exclusively refer to residues of the wild type *I-CreI* enzyme (SEQ ID NO: 1) as present in the variant, so for instance residue 2 of I-CreI is in fact residue 3 of a variant which comprises an additional Alanine after the first Methionine.

- by "*I-CreI* site" it is intended a 22 to 24 bp double-stranded DNA sequence which is cleaved by *I-CreI*. *I-CreI* sites include the wild-type (natural) non-palindromic *I-CreI* homing site and the derived palindromic sequences such as 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: 2), also called C1221.

- "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.

- by "meganuclease", it is intended an endonuclease having a double-stranded DNA target sequence of 12 to 45 bp. The meganuclease is either a dimeric enzyme, wherein each domain is on a monomer or a monomeric enzyme comprising the two domains on a single polypeptide.

- by "meganuclease domain", it is intended the region which interacts with one half of the DNA target of a meganuclease and is able to associate with the other domain of the same meganuclease which interacts with the other half of the DNA target to form a functional meganuclease able to cleave said DNA target.

- by "meganuclease variant" or "variant" it is intended a meganuclease obtained by replacement of at least one residue in the amino acid sequence of the parent meganuclease (natural or variant meganuclease) with a different amino acid.

- by "monomer" it is intended to mean a peptide encoded by the open reading frame of the I-CreI gene or a variant thereof, which when allowed to dimerise forms a functional I-CreI enzyme. In particular the monomers dimerise via interactions mediated by the LAGLIDADG motif.

- by "mutation" is intended the substitution, deletion, insertion of one or more nucleotides/amino acids in a polynucleotide (cDNA, gene) or a polypeptide sequence. Said mutation can affect the coding sequence of a gene or its regulatory sequence. It may also affect the structure of the genomic sequence or the structure/stability of the encoded mRNA.

- 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 "parent meganuclease" it is intended to mean a wild type meganuclease or a variant of such a wild type meganuclease with identical properties or alternatively a meganuclease with some altered characteristic in comparison to a wild type version of the same meganuclease. In the present invention the parent meganuclease can refer to the initial meganuclease from which the first series of variants are derived in step a. or the meganuclease from which the second series of variants are derived in step b., or the meganuclease from which the third series of variants are derived in step k.

- by "peptide linker" it is intended to mean a peptide sequence of at least 10 and preferably at least 17 amino acids which links the C-terminal amino acid residue of the first monomer to the N-terminal residue of the second monomer and which allows the two variant monomers to adopt the correct conformation for activity and which does not alter the specificity of either of the monomers for their targets.

- by "subdomain" it 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.

- by "single-chain meganuclease", "single-chain chimeric meganuclease", "single-chain meganuclease derivative", "single-chain chimeric meganuclease derivative" or "single-chain derivative" it is intended a meganuclease comprising two LAGLIDADG homing endonuclease domains or core domains linked by a peptidic

spacer. The single-chain meganuclease is able to cleave a chimeric DNA target sequence comprising one different half of each parent meganuclease target sequence.

- by "single-chain obligate heterodimer", it is intended a single-chain derived from an obligate heterodimer, as defined above.

- by "targeting DNA construct/minimal repair matrix/repair matrix" it is intended to mean a DNA construct comprising a first and second portions which are homologous to regions 5' and 3' of the DNA target *in situ*. The DNA construct also comprises a third portion positioned between the first and second portion which comprise some homology with the corresponding DNA sequence *in situ* or alternatively comprise no homology with the regions 5' and 3' of the DNA target *in situ*. Following cleavage of the DNA target, a homologous recombination event is stimulated between the genome containing the Herpesviridae Virus genome and the repair matrix, wherein the genomic sequence containing the DNA target is replaced by the third portion of the repair matrix and a variable part of the first and second portions of the repair matrix.

- by "vector" is intended a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked into a host cell *in vitro*, *in vivo* or *ex vivo*.

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: HSV-1 genome schematic representation. Gene considered as accessory (upper) and essential (down) are represented from both parts of linear form of virus DNA.

Figure 2: HSV-1 genome schematic representation with HSV2 and UL19 localization

Figure 3: The HSV2 and C1221 I-*Cre*I target sequences and their derivatives. 10AAA_P, 5CAC_P, 10AGG_P, 5GCC_P are close derivatives found to be cleaved by previously obtained I-*Cre*I mutants. They differ from C1221 by the boxed motives. C1221, 10AAA_P, 5CAC_P, 10AGG_P, 5GCC_P were first described as 24 bp sequences, but structural data suggest that only the 22 bp are

relevant for protein/DNA interaction. However, positions ±12 are indicated in parenthesis. In the HSV2.2 target, the ACAC sequence in the middle of the target is replaced with GTAC, the bases found in C1221. HSV2.3 is the palindromic sequence derived from the left part of HSV2.2, and HSV2.4 is the palindromic sequence derived from the right part of HSV2.2. As shown in the Figure, the boxed motives from 10AAA_P, 5CAC_P, 10AGG_P, 5GCC_P are found in the HSV2 series of targets

Figure 4: pCLS1055

Figure 5: pCLS0542

Figure 6: pCLS1107

Figure 7: Cleavage of HSV2.2 and HSV2 by heterodimeric mutants from database. **A.** Secondary screening of combinations of I-CreI mutants with the HSV2.2. target. **B**. Secondary screening of the same combinations of I-CreI mutants with the HSV2 target.

Figure 8: Improvement of HSV2.5 cleavage : A series of I-CreI N75 mutants cutting HSV2.3 and HSV2.5 were optimized by random mutagenesis. Cleavage is tested with the HSV2.5 target. Mutants displaying high specific cleavage activity of HSV2.5 (and HSV2.3) are circled. H10 is a negative control. H11 and H12 are positive controls.

Figure 9: Improvement of HSV2.6 cleavage : A series of I-CreI N75 mutants cutting HSV2.4 and HSV2.6 were optimized by random mutagenesis. Cleavage is tested with the HSV2.6 target (panel A) and HSV2.4 (panel B). Mutants displaying specific cleavage activity of HSV2.6 (and HSV2.4) are circled. D10 is a negative control. D11 and D12 are positive controls.

Figure 10: Cleavage of HSV2 by optimized heterodimeric mutants from random mutagenesis. Combinations displaying high cleavage activity of HSV2 are circled.

Figure 11: pCLS1058

Figure 12: pCLS2437

Figure 13: pCLS2733 and pCLS2735

Figure 14: pCLS1853

Figure 15: pCLS0001

Figure 16: pCLS2222 positive control expressing SCOH-RAG1.10 meganuclease.

Figure 17: pCLS1069 (empty vector) and pCLS1090 (positive control expressing I-*Sce*I)

Figure 18: Example of activity cleavage in CHO cells of designed single chain SCOH-HSV2 variants compared to initial heterodimer, I-*Sce*I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 19: Example of activity cleavage in CHO cells of single chain SCOH-HSV2 variants compared to initial heterodimer, I-Sce I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 20: Example of activity cleavage in CHO cells of single chain SCOH-HSV2-M1-105A132V-MC132V compared to initial heterodimer, I-*Sce*I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 21: Example of activity cleavage in CHO cells of single chain SCOH-HSV2-M1-MC-80K105A132V (pCLS2459) compared to initial heterodimer, I-*Sce*I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 22: Example of activity cleavage in CHO cells of single chain SCOH-HSV2-M1-MC-132V (pCLS2457) compared to initial heterodimer, I-*Sce*I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 23: HSV-1 genome schematic representation with HSV4 and ICP0 (or RL2) genes localization

Figure 24: The HSV4 and C1221 I-Cre I target sequences and their derivatives. 10AAG_P, 5GGT_P, 5CAG_P, 10ACT_P are close derivatives found to be cleaved by previously obtained I-*Cre*I mutants. They differ from C1221 by the boxed motives. C1221, 10AAG_P, 5GGT_P, 5CAG_P, 10ACT_P were first described as 24 bp sequences, but structural data suggest that only the 22 bp are relevant for protein/DNA interaction. However, positions ±12 are indicated in parenthesis. In the HSV4 target, the GTAC sequence in the middle of the target is found in C1221. HSV4.3 is the palindromic sequence derived from the left part of HSV4, and HSV4.4 is the palindromic sequence derived from the right part of HSV4. As shown in the

Figure, the boxed motives from 10AAG_P, 5GGT_P, 5CAG_P, 10ACT_P are found in the HSV4 series of targets

Figure 25: Cleavage of HSV4 by heterodimeric combinations of mutants obtained after combinatorial process.

Figure 26: Improvement of HSV4.3 cleavage : A series of I-*Cre*I N75 mutants cutting HSV4.3 were optimized by random mutagenesis. Cleavage is tested with the HSV4.3 target. Mutants displaying high specific cleavage activity of HSV4.3 are circled. H10 is a negative control. H11 and H12 are positive controls.

Figure 27: Improvement of HSV4.4 cleavage : A series of I-*Cre*I N75 mutants cutting HSV4.4 were optimized by random mutagenesis. Cleavage is tested with the HSV4.4 target. 14 mutants displaying higher specific cleavage activity of HSV4.4 than best starting one are circled. H10 is a negative control. H11 and H12 are positive controls.

Figure 28: Cleavage of HSV4 by optimized heterodimeric mutants from random mutagenesis. All combinations are displaying high cleavage activity of HSV4.

Figure 29: pCLS1768

Figure 30: pCLS2266 and pCLS2267

Figure 31: pCLS0491

Figure 32: pCLS2222, positive control expressing SCOH-RAG-CLS meganuclease under pCMV promoter, and pCLS2294, positive control expressing SCOH-RAG-CLS meganuclease under pEF1alpha promoter.

Figure 33: Example of activity cleavage in CHO cells of designed single chain SCOH-HSV4 variants compared to initial heterodimer, I-Sce I and SCOH-RAG-CLS meganucleases as positive controls.

Figure 34: Example of activity cleavage in CHO cells of single chain SCOH-HSV4 variants compared to initial heterodimer, I-Sce I and SCOH-RAG1.10 meganucleases as positive controls.

Figure 35: Example of activity cleavage in CHO cells of single chain SCOH-HSV4-M2-54L-MF (pCLS2474) compared to initial heterodimer, I-*Sce*I and SCOH-RAG-CLS meganucleases as positive controls.

Figure 36: Example of activity cleavage in CHO cells of single chain SCOH-HSV4- M2-105A-MF-80K132V (pCLS2481) compared to initial heterodimer, I-*Sce*I and SCOH-RAG-CLS meganucleases as positive controls.

Figure 37: Example of activity cleavage in CHO cells of single chain SCOH-HSV4- M2-MF-132V (pCLS2472) compared to initial heterodimer, I-*Sce*I and SCOH-RAG-CLS meganucleases as positive controls.

Figure 38: Example of activity cleavage in CHO cells of single chain SCOH-HSV4- M2-MF (pCLS2470) compared to initial heterodimer, I-*Sce*I and SCOH-RAG-CLS meganucleases as positive controls.

Figure 39: Genomic structure of recombinant virus. The overall structure of the HSV-1 genome is shown with unique long (U_L) and unique short (U_S) regions flanked by inverted terminal repeats. The LAT region located in the terminal repeats has been expanded and the location of the LAT transcript are shown. An expression cassette containing the CMV promoter and the LacZ coding sequence was inserted in the major LAT gene. I-SceI target site was cloned between the CMV promoter and the LacZ gene.

Figure 40: pCLS0126

Figure 41: Example of inhibition of viral replication by I-CreI single chain obligate heterodimer variants cleaving HSV2, HSV4 or HSV12 target sequences. COS-7 cells were transfected with empty vector, plasmid expressing I-SceI or plasmid expressing I-CreI variants cleaving HSV2, HSV4 or HSV12 target sequences. Twenty-four hours later the transfected cells were infected with rHSV-1 which expresses the LacZ gene. Beta-galactosidase activity levels, indicative of LacZ gene expression, was assayed twenty-four hours after infection. The detected activity levels are depicted in the histogram with the percent activity compared to empty vector indicated below the histogram.

Example 1. Strategy for engineering novel meganucleases cleaving target from the UL_{19} gene in HSV-1 genome.

HSV2 is a 24 bp (non-palindromic) target (SEQ ID NO: 24) present in the UL19 gene encoding the HSV-1 major capsid protein. This 5.7kb gene is present in one copy at position 35023 to 40768 of the UL region. The HSV1-major capsid protein is expressed without maturation from an ORF located from 36404 to

40528. The target HSV2 is located from nucleotide 36966 to 36989 (accession number NC_001806; Figure 2).

The 10AAA_P, 5CAC_P, 10AGG_P, 5GCC_P targets sequences are 24 bp derivatives of C1221, a palindromic sequence cleaved by I-*Cre*I (Arnould et al., precited). However, the structure of I-*Cre*I bound to its DNA target suggests that the two external base pairs of these targets (positions -12 and 12) have no impact on binding and cleavage (Chevalier et al., Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier et al., J. Mol. Biol., 2003, 329, 253-269), and in this study, only positions -11 to 11 were considered. Consequently, the HSV2 series of targets were defined as 22 bp sequences instead of 24 bp. HSV2 differs from C1221 in the 4 bp central region. According to the structure of the I-*Cre*I protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the I-CreI protein (Chevalier et al., Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier et al., J. Mol. Biol., 2003, 329, 253-269). Thus, the bases at these positions should not impact the binding efficiency. However, they could affect cleavage, which results from two nicks at the edge of this region. Thus, the ACAC sequence in -2 to 2 was first substituted with the GTAC sequence from C1221, resulting in target HSV2.2 (Figure 3). Then, two palindromic targets, HSV2.3 and HSV2.4, were derived from HSV2.2 (Figure 3). Since HSV2.3 and HSV2.4 are palindromic, they should be cleaved by homodimeric proteins. Thus, proteins able to cleave the HSV2.3 and HSV2.4 sequences as homodimers were first designed (examples 1.1 and 1.2) and then co-expressed to obtain heterodimers cleaving HSV2 (example 1.3). Heterodimers cleaving the HSV2.2 and HSV2 targets could be identified. In order to improve cleavage activity for the HSV2 target, a series of variants cleaving HSV2.3 and HSV2.4 was chosen, and then refined. The chosen variants were subjected to random mutagenesis, and used to form novel homodimers (examples 1.4 and 1.5). Several improved mutants were then chosen to form heterodimers that were screened against the HSV2 target (example 1.6). Heterodimers could be identified with an improved cleavage activity for the HSV2 target. Chosen heterodimers were then cloned into mammalian expression vectors for HSV2 cleavage in CHO cells (example 1.7). These results were then utilized to design single chain molecules directed against the HSV2

target that were cloned into mammalian expression vectors and tested for HSV2 cleavage in CHO cells (example 1.8). Strong cleavage activity of the HSV2 target could be observed for these single chain molecules in mammalian cells.

Example 1.1: Identification of meganucleases cleaving HSV2.3 and HSV2.5 targets

This example shows that I-*Cre*I variants can cut the HSV2.3 and HSV2.5 DNA target sequences derived from the left part of the HSV2 target in a palindromic form. Target sequences described in this example are 22 bp palindromic sequences. Therefore, they will be described only by the first 11 nucleotides, followed by the suffix _P (For example, target HSV2.3 will be noted HSV2.3 TAAACTCACGT_P SEQ ID NO: 10).

HSV2.3 and HSV2.5 are similar to 10AAA_P at positions ±10, ±9, ±8 and to 5CAC_P at positions ±5, ±4, ±3. It was hypothesized that positions ±7 and ±11 would have little effect on the binding and cleavage activity. Variants able to cleave 10AAA-5CAC_P target were previously obtained by mutagenesis on I-CreI N75 at positions 24, 44, 68, 70, 75 and 77 as described in Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458; Smith *et al.* Nucleic Acids Res., 2006, 34, e149; International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156. 192 of these variants were stored in our database and ready to be assayed for HSV2.3 and HSV2.5 cleavage.

A) Material and Methods

a) Construction of target vector

The target was cloned as follows: an oligonucleotide corresponding to the HSV2.3 and HSV2.5 targets sequences flanked by gateway cloning sequences was ordered from PROLIGO: HSV2.3

5'TGGCATACAAGTTTATAAACTCACGTACGTGAGTTTATCAATCGTCTGTC A3' (SEQ ID NO: 38);

HSV2.5

5'TGGCATACAAGTTTATAAACTCACACACGTGAGTTTATCAATCGTCTGTC A3' (SEQ ID NO: 39). Double-stranded target DNA, generated by PCR amplification of the single stranded oligonucleotide, was cloned using the Gateway protocol (INVITROGEN) into the yeast reporter vector (pCLS1055, Figure 4). Yeast reporter

vector was transformed into Saccharomyces cerevisiae strain FYBL2-7B (MAT a, ura3△851, trp1△63, leu2△1, lys2△202), resulting in a reporter strain. (MilleGen)

b) Mating of meganuclease expressing clones and screening in yeast

Screening of variants from our data bank was performed as described previously (Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony gridder (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain. Membranes were placed on solid agar YPD rich medium, and incubated at 30 °C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking leucine and tryptophan, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6% dimethyl formamide (DMF), 7 mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software.

B) Results

Examples of variants able to cleave 10AAA-5CAC_P target are displayed in Table VIII. Among 192 unique variants, 156 clones were found on HSV2.3 which correspond to 156 different endonucleases (Table IX), 55 of them where able to cut HSV2.5 as well. Examples of positives are shown in Table IX.

Table VIII: Panel of variants extracted from our database

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
44N70S75R77Y/	40
44V68E75N77R/24V80K	41
44N68E70S75R77K/	42
44K68Q/	43
44R68T/	44
44N68Y70S75R77V/	45
44A70S75R77Y/24V	46
44N70S75R77N/	47
44A70S75R77Y/	48
44N68E70S75R77R/	49
44R68N/	50
44K68A/	51
444N70S75R77N/24V	52
44T68E70S75R77R/	53
44A68Y70S75Y77K/	54
44N68Y70S75R77Y/	55
44A70S75R77L/	56
44T68K70S75R77R/	57
44N68Y70S75R/	58
44N68E70S75R77R/24V	59
44A68Y70S75R/24V	60
44N68Y70S75R77V/24V	61
44K68H70S75N/	62
44R68A/	63
44N68A70S75R77Y/	64
44D68A70S75K77R/	65
44R68Y70S75N77T/24V	66
44R68Y70S75N/24V	67
44A68Y70S75R/	68
44I68E75N77R/24V	69
44N70S75R/	70
44T68E70S75R77R/24V	455
44R68Y70S75Y77T/24V	456
44N68Y70S75R77Q/	457
44N68Y70S75Y77K/24V	458
44N68K70S75R77N/	459
44R68Y70S75Y77N/	460
44R68H/	461
44A68Y70S75R77V/24V	462
44S68E70S75R77K/24V	463

Table IX: I-*Cre*I variants capable of cleaving the HSV2.3 as well as HSV2.5 DNA targets.

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
44N68Y70S75R77Y/	71
44A68Y70S75R/	72
44N70S75R77Y/	73
44N68Y70S75R/	74
44T68T70S75K77E/	75
24V44T68T70S75K77E/	76
44N70S75R77N/	77
44A70S75R77Y/	78
44R68A/	79
44R68T/	80
24V44N68Y70S75Y77R/	81
44N68A70S75R77Y/	82
44N68Y70S75R77V/	83
44N68T70S75R77Y/	84
44T68Y70S75Y77R/	85
44A68Y70S75R77R/	86
44N68Y70S75Y77R/	87
44K68A/	88
44R68H/	89
44A68T70S75Q77R/	90

Example 1.2: Identification of meganucleases cleaving HSV2.4 and HSV2.6

This example shows that I-*Cre*I variants can cleave the HSV2.4 and HSV2.6 DNA target sequences derived from the right part of the HSV2 target in a palindromic form (Figure 3). All target sequences described in this example are 22 bp palindromic sequences. Therefore, they will be described only by the first 11 nucleotides, followed by the suffix _P (for example, HSV2.4 will be called *CAGGACGCCGT*_P).

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 1.1, with the exception that an oligonucleotide corresponding to the HSV2.4 and HSV2.6 target sequences were used: 5' TGGCATACAAGTTTCCAGGACGCCGTACGGCGTCCTGGCAATCGTCTGTCA 3' (SEQ ID NO: 91).

and

5'TGGCATACAAGTTTCCAGGACGCCACACGGCGTCCTGGCAATCGTCTGT CA3' (SEQ ID NO: 92) (resp. HSV2.4 and HSV2.6)

b) Mating of meganuclease expressing clones and screening in yeast

Screening was performed as described previously (Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony gridder (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain. Membranes were placed on solid agar YPD rich medium, and incubated at 30 °C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking tryptophan, adding G418, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6 % dimethyl formamide (DMF), 7 mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software. Positives resulting clones were verified by sequencing (MILLEGEN) as described in example 1.1.

B) Results

Examples of variants able to cleave 10AGG-5GCC_P target are displayed in Table X. Among 57 clones, 33 clones were positives on HSV2.4, 3 of them where able to cut HSV2.6 too. Examples of positives are shown in Table XI.

Table X: Panel of variants extracted from our data bank

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
28E33R38R40K44K70D75N/	93
28E33R38R40K44K70E75N/	94
28E33R38R40K44K68T70G75N/	95
28E33R38R40K44K68T70S75N/	96
28E38R40K44K68S70S75N/	97
30D33A38H44K70D75N/	98
30D33A38H44K70E75N/	99
D33H38K44K70D75N/	100
30D33H38K44K70D75N/87L	101
30D33H38K44K70E75N/	102
30D33H38K44K70T75N/	103
30D33R38G44K70D75N/	104
30D33R38G44K70D75N/17A	105
30D33R38G44K70E75N/	106
30G38G44K70E75N/	107
30G38G44K70S75N/54I	108
30G38H44K68A70G75N/	109
30G38H44K68A70Q75N/	110
30G38H44K68A70S75N/	111
28k30G38H44K68G70A75N/	112
28k30G38H44K68N70A75N/	113
30G38H44K68Q70S75N/	114
30G38H44K70D75N/	115
30G38H44K70E75N/	116
30G38H44K70N75N/	117
30G38H44K70S75N/	118
30G38H44K68S70D75N/	119
30G38H44K68T70S75N/	120
30G38R44K68A70G75N/	121
30G38R44K68A70N75N/	122
30G38R44K68A70Q75N/	123
30G38R44K68G70S75N/	124
30G38R44K68G70S75N/3A	125
30G38R44K68H70S75N/	126
30G38R44K68Q70D75N/	127
30G38R44K68Q70S75N/	128
30G38R44K70D75N/	129
30G38R44K70E75N/	130
30G38R44K70N75N/	131

Table XI: I-*Cre*I variants capable of cleaving the HSV2.4 and/or HSV2.6 DNA targets.

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
28E38R40K44K70E75N/54L81V96R153V160R	132
28E38R40K44K68S70S75N/94S132V150T	133
30G38G44K70S75N/54I	134
28E38R40K44K70E75N/163T	135
28E38R40K44K70E75N/	136
28E38R40K44K68S70S75N/	137
30G38H44K70S75N/	138
28E38R40K44K70E75N77T/	139
30G38R44K70S75N/162F	140
30G38R44K70D75N/	141
30G38R44K68S70S75N/	142
30G38R44K70S75N/	143
30G38R44K68T70S75N/	144
30G38R44K70S75N/62V	145
30G38R44K68Q70S75N/	146
30G38H44K70D75N/	147
30G38R44K70E75N/	148
30G38R44K68T70G75N/	149
30G338R44K68G70S75N/3A	150
30G38R44K70N75N/	151

Example 1.3: Identification of meganucleases cleaving HSV2

I-*Cre*I variants able to cleave each of the palindromic HSV2 derived targets (HSV2.3/2.5 and HSV2.4/2.6) were identified in example 1.1 and 1.2. Pairs of such variants (one cutting HSV2.3 and one cutting HSV2.4) were co-expressed in yeast. Upon co-expression, there should be three active molecular species, two homodimers, and one heterodimer. It was assayed whether the heterodimers that should be formed, cut the non palindromic HSV2 target.

A) Materials and Methods

a) Construction of target vector

The experimental procedure is as described in example 1.2, with the exception that an oligonucleotide corresponding to the HSV2 target sequence: 5' TGGCATACAAGTTTATAAACTCACACACGGCGTCCTGGCAATCGTCTGTCA 3' (SEQ ID NO: 152) was used.

b) Co-expression of variants

Yeast DNA was extracted from variants cleaving the HSV2.4 target in the pCLS1107 (Figure 6) expression vector using standard protocols and was used to transform *E. coli*. The resulting plasmid DNA was then used to transform yeast strains expressing a variant cutting the HSV2.3 target in the pCLS542 expression vector. Transformants were selected on synthetic medium lacking leucine and containing G418.

c) Mating of meganucleases coexpressing clones and screening in yeast

Mating was performed using a colony gridder (QpixII, Genetix). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of different reporter-harboring yeast strains for each target. Membranes were placed on solid agar YPD rich medium, and incubated at 30°C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking leucine and tryptophan, adding G418, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6% dimethyl formamide (DMF), 7mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software.

B) Results

Co-expression of variants cleaving the HSV2.4 target (6 variants chosen among those described in Table XI) and 10 variants cleaving the HSV2.3 target (described in Table IX) resulted in cleavage of the HSV2 target in 14 cases (Figure 7). Functional combinations are summarized in Table XII.

Table XII: Cleavage of the HSV2 target by the heterodimeric variants

		Amino acids positions and residues of the I-CreI variants cleaving the HSV2.3 target									
Amino acids positions and residues of I-CreI variants cleaving the HSV2.4 target		44A68Y70S75R (SEQ ID NO: 159)	44N70S75R77N (SEQ ID NO: 160)	44N70S75R77Y (SEQ ID NO: 161)	44A70S75R77Y (SEQ ID NO: 162)	44T68T70S75K77E (SEQ ID NO: 163)	44N68Y70S75R77Y (SEQ ID NO: 164)	44N68Y70S75R (SEQ ID NO: 165)	24V44T68T70S75K77E (SEQ ID NO: 166)	44R68T (SEQ ID NO: 167)	44R68A (SEQ ID NO: 168)
	28E38R40K44K68S70S75N/94S132V150T (SEQ ID NO: 153)			+	+	+	+				
	28E38R40K44K70E75N/(SEQ ID NO: 154)	+							+		
	28E38R40K44K70E75N/77T (SEQ ID NO: 155)					+					+
	28E38R40K44K70E75N/163T (SEQ ID NO: 156)										
	30G38G44K70S75N/54I (SEQ ID NO: 157)	+									
	28E38R40K44K68S70S75N7 (SEQ ID NO: 158)		+	+	+	+	+				

+ indicates a functional combination

Example 1.4: Improvement of meganucleases cleaving HSV2.5 by random mutagenesis

I-*Cre*I variants able to cleave the palindromic HSV2.5 target have been previously identified in example 1.1. Some of them can cleave the HSV2 target when associated with variants able to cut HSV2.6 (examples 1.2 and 1.3).

Therefore 6 selected variants cleaving HSV2.5 were mutagenized, and variants were screened for activity improvement on HSV2.5. According to the structure of the I-*Cre*I protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the I-*Cre*I protein (Chevalier *et al.*, Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier *et al.*, J. Mol. Biol., 2003, 329, 253-269). Thus, it is difficult to rationally choose a set of positions to mutagenize, and mutagenesis was performed on the whole protein. Random mutagenesis results in high complexity libraries.

Therefore, to limit the complexity of the variant libraries to be tested, the two components of the heterodimers cleaving HSV2 were mutagenized and screened in parallel.

A) Material and Methods

a) Construction of libraries by random mutagenesis

Random mutagenesis was performed on a pool of chosen variants, by PCR using Mn^{2+}. PCR reactions were carried out that amplify the I-*Cre*I coding sequence using the primers preATGCreFor (5'-gcataaattactatacttctatagacacgcaaacacaaatacacagcggccttgccacc-3'; SEQ ID NO: 169) and ICreIpostRev (5'-ggctcgaggagctcgtctagaggatcgctcgagttatcagtcggccgc-3'; SEQ ID NO: 170), which are common to the pCLS0542 (Figure 5) and pCLS1107 (Figure 6) vectors. Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLS0542) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MATα, trp1Δ63, leu2Δ1, his3Δ200*) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Expression plasmids containing an intact coding sequence for the I-*Cre*I variant were generated by *in vivo* homologous recombination in yeast.

b) Target vector yeast strains

The yeast strain FYBL2-7B (*MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*) containing the HSV2.5 target in the yeast reporter vector (pCLS1055 Figure 4) was constructed as described in example 1.1.

c) mating of meganuclease expressing clones, screening in yeast and sequencing

Mating HSV2.3 target strain and mutagenized variant clones and screening were performed as described in example 1.1. One variant from first generation was added as control on filter during screening steps for activity improvement evaluation.

B) Results

Six variants cleaving HSV2.5, (Table XIII), were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then mated with a yeast strain that contains the HSV2.5 target in a reporter plasmid. After mating with this yeast strain, 761 clones were found to cleave the HSV2.5 target. 93 of them were characterized. 72 of them shown high activity and retain HSV2.5/2.3 specificity.

An example of positives is shown in Figure 8. Sequencing of these 46 positive clones indicates that 32 distinct variants listed in Table XIV were identified.

Table XIII: pool of variants cleaving HSV2.3/2.5 and sequences used as template for random mutagenesis

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
44A68Y70S75R/	171
44N70S75R77N/	172
44N70S75R77Y/	173
44A70S75R77Y/	174
44T68T70S75K77E/	175
44N68Y70S75R77Y/	176

Table XIV: Improved variants displaying strong cleavage activity for HSV2.5

Amino acids positions and residues of the I-*Cre*I variants	SEQ ID NO:
44A70S75R77Y/	177
44D68T70S75R77H/	178
44D68T70S75R77P/37Y	179
44D68T70S75R77R/	180
44D68T70S75R77R/100R	181
44D68T70S75R77R/12H	182
44D68T70S75R77R/156G	183
44D68T70S75R77R/2D37Y129A	184
44D68T70S75R77R/37C105A	185
44D68T70S75R77R/37Y	186
44D68T70S75R77R/37Y114T	187
44D68T70S75R77R/37Y121R	188
44D68T70S75R77R/37Y129A	189
44D68T70S75R77R/37Y66H	190
44D68T70S75R77R/37Y82R	191
44D68T70S75R77R/4R151A	192
44D68T70S75R77R/4R37Y151A	193
44D68T70S75R77R/57E159R	194
44D68T70S75R77R/64A	195
44D68T70S75R77R/6S37Y	196
44D68T70S75R77R/80K	197
44H68T70P75R77R/	198
44N70S75R77Y/	199

44N70S75R77Y/157G	200
44N68Y70S75R/	201
44N68Y70S75R/132V	202
44N68Y70S75R77Y/	203
44R68A75d/54S	204
44R68T/111R	205
44R68T/54L121E	206
44T68T70S75R77Q/	207
24V44D68T70S75R77R/	208

Example 1.5: Improvement of meganucleases cleaving HSV2.6 by random mutagenesis

I-*Cre*I variants able to cleave the palindromic HSV2.4 target has been previously identified in example 1.2. Some of them can cleave HSV4 target when associated with variants able to cut HSV2.3 (examples 1.1 and 1.3).

Six of the selected variants cleaving HSV2.6 and 2.4 were mutagenized, and variants were screened for activity improvement on HSV2.6. As described in example 1.4, mutagenesis was performed on the whole protein and HSV2.4 variants were screened in parallel of HSV2.3 (example 1.4).

A) Material and Methods

a) Construction of libraries by random mutagenesis

Random mutagenesis was performed as described in example 1.4, on a pool of chosen variants, by PCR using the same primers and Mn^{2+} conditions (preATGCreFor SEQ ID NO: 169 and ICreIpostRev SEQ ID NO: 170). Approximately 25 ng of the PCR product and 75 ng of vector DNA pCLS1107) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MATα, trp1Δ63, leu2Δ1, his3Δ200*) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Expression plasmids containing an intact coding sequence for the I-*Cre*I variant were generated by *in vivo* homologous recombination in yeast.

b) Target vector yeast strains

The yeast strain FYBL2-7B (*MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*) containing the HSV2.6 target in the yeast reporter vector (pCLS1055 Figure 4) was constructed as described in example 1.2.

c) mating of meganuclease expressing clones, screening in yeast and sequencing

Mating HSV2.6 target strain and mutagenized variant clones and screening were performed as described in example 1.2. One variant from first generation was added as control on filter during screening steps for activity improvement evaluation.

B) Results

Six chosen variants cleaving HSV2.6 and 2.4, (Table XV), were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then mated with a yeast strain that contains the HSV4.4 target in a reporter plasmid. After mating with this yeast strain, 32 clones were found to cleave the HSV2.6 target. An example of positives is shown in Figure 9. Sequencing 32 positive clones indicates that 19 distinct variants listed in Table XVI were identified.

Table XV: pool of variants cleaving HSV2.6 and 2.4 and sequences used as template for random mutagenesis

SEQ ID NO:	Amino acids positions and residues of the I-*Cre*I variants
209	28E38R40K44K68S70S75N/94S132V150T
210	28E38R40K44K70E75N/
211	28E38R40K44K70E75N/77T
212	28E38R40K44K70E75N/163T
213	30G38G44K70S75N/54I
214	28E38R40K44K68S70S75N/

Table XVI: Improved variants displaying cleavage activity for HSV2.6

SEQ ID NO:	Amino acids positions and residues of the I-CreI variants
215	28E38R40K44K70E75N
216	28E38R40K44K70E75N77T/50R
217	28E38R40K44K70E75N77T/80K
218	28E38R40K44K70E75N77T132V150T/17A
219	28E38R40K44K70G75N132V150T/
220	28E38R40K44K68S70S75N/50R
221	28E38R40K44K68S70S75N132V
222	28E38R40K44K68S70S75N132V163T/2D
223	28E38R40K44K68S70S75N132V150T
224	28E38R40K44K68S70S75N94S132V
225	28E38R40K44K68S70S75N94S132V150T
226	28E38R40K44K68S70S75N94S132V150T/50R
227	28E38R40K44K68S70S75N94S132V150T/66H
228	28E38R40K44K68S70S75N77T132V
229	28E38R40K44K70E75N77T94S
230	28E38R40K44K70S75N
231	28E38R40K44K70S75N150T
232	28E38R40K44K54L70E75N94S
233	28E38R40K44K54L70E75N77T

Example 1.6: Identification of improved meganucleases cleaving HSV2

Improved I-*Cre*I variants able to cleave each of the palindromic HSV2 derived targets (HSV2.3/2.5 and HSV2.4/2.6) were identified in example 1.4 and example 1.5. As described in example 1.3, pairs of such variants (one cutting HSV2.3/2.5 and one cutting HSV2.4/2.6) were co-expressed in yeast. The heterodimers that should be formed were assayed for cutting the non palindromic HSV2 target.

A) Materials and Methods

a) Construction of target vector

The HSV2 target vector was constructed as described in example 1.3.

b) Co-expression of variants

Yeast DNA was extracted from variants cleaving the HSV2.6target in the pCLS1107 expression vector using standard protocols and was used to

transform *E. coli*. The resulting plasmid DNA was then used to transform yeast strains expressing a variant cutting the HSV2.5target in the pCLS542 expression vector. Transformants were selected on synthetic medium lacking leucine and containing G418.

c) Mating of meganucleases coexpressing clones and screening in yeast

Mating and screening of meganucleases coexpressing clones were performed as described in example 1.3. Results were analyzed by scanning and quantification was performed using appropriate software.

B) Results

Co-expression of improved variants cleaving the HSV2.6 target (6 variants chosen among those described in Table XIV) and 7 improved variants cleaving the HSV2.5 target (described in Table XVI) resulted in cleavage of the HSV2 target in all except one case (Figure 10). All assayed combinations are summarized in Table XVII.

Table XVII: Cleavage of the HSV2 target by the heterodimeric improved variants

		Amino acids positions and residues of the I-CreI variants cleaving the HSV2.3 target						
Amino acids positions and residues of I-CreI variants cleaving the HSV2.4 target		44D68T70S75R77R/80K (SEQ ID NO: 240)	44D68T70S75R77R/12H (SEQ ID NO: 241)	44R68T/54L121E (SEQ ID NO: 242)	44D68T70S75R77R (SEQ ID NO: 243)	44D68T70S75R77R/156G (SEQ ID NO: 244)	44D68T70S75R77R/100R (SEQ ID NO: 245)	44D68T70S75R77R/64A (SEQ ID NO: 246)
	28E38R40K44K68S70S75N132V150T (SEQ ID NO: 234)	+	+	+		+	+	+
	28E38R40K44K54L70E75N77T (SEQ ID NO: 235)	+	+	+	+	+	+	+
	28E30n38R40K44K54l68r70S75N77i94f132i150a163p (SEQ ID NO: 236)	+	+	+	+	+	+	+
	28E38R40K44K68S70S75N/50R (SEQ ID NO: 237)	+	+	+	+	+	+	+
	28E38R40K44K70E75N77T132V150T/17A (SEQ ID NO: 238)	+	+	+	+	+	+	+
	28E38R40K44K70E75N77T94S (SEQ ID NO: 239)	+	+	+	+	+	+	+

+ indicates a functional combination

Example 1.7: Validation of HSV2 target cleavage in an extrachromosomal model in CHO cells

I-*Cre*I variants able to efficiently cleave the HSV2 target in yeast when forming heterodimers were described in examples 1.3 and 1.7. In order to identify heterodimers displaying maximal cleavage activity for the HSV2 target in CHO cells, the efficiency of chosen combinations of variants to cut the HSV2 target was compared, using an extrachromosomal assay in CHO cells. The screen in CHO cells is a single-strand annealing (SSA) based assay where cleavage of the target by the meganucleases induces homologous recombination and expression of a LagoZ reporter gene (a derivative of the bacterial lacZ gene).

1) Materials and methods

a) Cloning of HSV2 target in a vector for CHO screen

The target was cloned as follows: oligonucleotide corresponding to the HSV2 target sequence flanked by gateway cloning sequence was ordered from PROLIGO 5'TGGCATACAAGTTTATAAACTCACACACGGCGTCCTGGCAATCGTCTGT CA3' (SEQ ID NO: 152). Double-stranded target DNA, generated by PCR amplification of the single stranded oligonucleotide, was cloned using the Gateway protocol (INVITROGEN) into CHO reporter vector (pCLS1058, Figure 11). Cloned target was verified by sequencing (MILLEGEN).

b) Re-cloning of meganucleases

The ORF of I-CreI variants cleaving the HSV2.3 and HSV2.4 targets identified in examples 1.4 and 1.5 were sub-cloned in pCLS2437 (Figure 12). ORFs were amplified by PCR on yeast DNA using the AT1CA1F (5'-AAAAAGCAGGCTGGCGCGCCTACACAGCGGCCTTGCCACCATG-3' SEQ ID NO: 247) and AT2CA2R (5'-AGAAAGCTGGGTGCTAGCGCTCGAGTTATCAGTCGG-3' SEQ ID NO: 248) primers. PCR products were cloned in the CHO expression vector pCLS2437 (Figure 12) using the Asc I and Xho I for internal fragment replacement. Selected clones resulting from ligation and *E. coli* transformation steps were verified by sequencing (MILLEGEN).

c) Extrachromosomal assay in mammalian cells

CHO K1 cells were transfected with Polyfect® transfection reagent according to the supplier's protocol (QIAGEN). 72 hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β -galactosidase liquid assay was added (typically 1 liter of buffer contained: 100 ml of lysis buffer (Tris-HCl 10 mM pH7.5, NaCl 150 mM, Triton X100 0.1 %, BSA 0.1 mg/ml, protease inhibitors), 10 ml of Mg 100X buffer ($MgCl_2$ 100 mM, β-mercaptoethanol 35 %), 110 ml ONPG 8 mg/ml and 780 ml of sodium phosphate 0.1M pH7.5). After incubation at 37°C, OD was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform.

Per assay, 150 ng of target vector was cotransfected with 12.5 ng of each one of both mutants (12.5 ng of mutant cleaving palindromic HSV2.3 target and 12.5 ng of mutant cleaving palindromic HSV2.4 target).

2) Results

2 variants cleaving HSV2.5 and 2 variants cleaving HSV2.6 described in example 1.4, 1.5 and 1.6 were re-cloned in pCLS2437 (Figure 12). Then, I-*Cre*I variants cleaving the HSV2.5 or HSV2.6 targets were assayed together as heterodimers against the HSV2 target in the CHO extrachromosomal assay.

Table XVIII shows the functional combinations obtained for 24 heterodimers. Analysis of the efficiencies of cleavage and recombination of the HSV4 sequence demonstrates that 9 combinations of I-CreI variants are able to transpose their cleavage activity from yeast to CHO cells without additional mutation.

Table XVIII: Functional heterodimeric combinations cutting the HSV2 target in CHO cells.

		Optimized variants cleaving HSV2.5	
		44D68T70S75 R77R/80K (SEQ ID NO: 251)	44D68T70S75R77 R/ (SEQ ID NO: 252)
Optimized variants cleaving HSV2.6	28E38R40K44K70S7 5N (SEQ ID NO: 249)	+	+
	28E38R40K44K70E7 5N77T132V150T/17 A (SEQ ID NO: 250)	+	+

+ indicates a functional combination

Example 1.8: Covalent assembly as single chain and improvement of meganucleases cleaving HSV2 by site-directed mutagenesis

Co-expression of the cutters described in example 1.5, 1.6, 1.7 leads to a high cleavage activity of the HSV2 target in yeast. Some of them have been validated for HSV2 cleavage in a mammalian expression system. One of them is shown in Table XIX.

Table XIX: Example of functional heterodimer cutting the HSV2 target in CHO cells.

HSV2.3-M1 (SEQ ID NO: 33)	44D 68T 70S 75R 77R **80K**
HSV2.4-MC (SEQ ID NO: 34)	28E 38R 40K 44K 54I 70S 75N

The M1 x MC HSV2 heterodimer gives high cleavage activity in yeast. M1 is a HSV2.5 cutter that bears the following mutations in comparison with the I-*Cre*I wild type sequence: 44D 68T 70S 75R 77R **80K**. MC is a HSV2.6 cutter that bears the following mutations in comparison with the I-*Cre*I wild type sequence: 28E 38R 40K 44K 54I 70S 75N.

Single chain constructs were engineered using the linker RM2 (AAGGSDKYNQALSKYNQALSKYNQALSGGGGS) (SEQ ID NO: 464) resulting in the production of the single chain molecule: M1-RM2-MC. During this design step, the G19S mutation was introduced in the C-terminal MC mutant. In addition, mutations K7E, K96E were introduced into the M1 mutant and mutations E8K, E61R into the MC mutant to create the single chain molecule: M1(K7E K96E)-RM2-MC(E8K E61R) that is called further SCOH-HSV2-M1-MC.

Four additional amino-acid substitutions have been found in previous studies to enhance the activity of I-CreI derivatives: these mutations correspond to the replacement of Phenylalanine 54 with Leucine (F54L), Glutamic acid 80 with Lysine (E80K), Valine 105 with Alanine (V105A) and Isoleucine 132 with Valine (I132V). Only E80K is already present in HSV2.5 variant (i.e. HSV2.5-M1). Some additional combinations were introduced into the coding sequence of N-terminal and C-terminal protein fragment (an example is shown in Table XX), and the resulting proteins were tested for their ability to induce cleavage of the HSV2 target. The twelve single chain constructs were then tested in CHO for cleavage of the HSV2 target.

Table XX: Single Chain I-Cre I variants for HSV2 cleavage in CHO cells.

Construct	Single chain	Mutations on N-terminal segment	Mutations on C-terminal segment	SEQ ID NO
pCLS2456	SCOH-HSV2-M1-MC	7E 44D 68T 70S 75R 77R **80K 96E**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N	**253**
pCLS2457	SCOH-HSV2-M1-MC-132V	7E 44D 68T 70S 75R 77R **80K 96E**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **132V**	**254**
pCLS2458	SCOH-HSV2-M1-MC-80K132V	7E 44D 68T 70S 75R 77R **80K 96E**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **80K 132V**	**255**
pCLS2459	SCOH-HSV2-M1-MC-80K105A132V	7E 44D 68T 70S 75R 77R **80K 96E**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **80K 105A 132V**	**256**
pCLS2460	SCOH-HSV2-M1-132V-MC	7E 44D 68T 70S 75R 77R **80K 96E 132V**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N	**257**
pCLS2462	SCOH-HSV2-M1-132V-MC-80K105A	7E 44D 68T 70S 75R 77R **80K 96E 132V**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **80K 105A**	**258**
pCLS2463	SCOH-HSV2-M1-132V-MC-80K132V	7E 44D 68T 70S 75R 77R **80K 96E 132V**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **80K 132V**	**259**
pCLS2464	SCOH-HSV2-M1-132V-MC-80K105A132V	7E 44D 68T 70S 75R 77R **80K 96E 132V**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **80K 105A 132V**	**260**
pCLS2465	SCOH-HSV2-M1-105A132V-MC-132V	7E 44D 68T 70S 75R 77R **80K 96E 105A 132V**	**8K 19S** 28E 38R 40K 44K 54I 61R 70S 75N **132V**	**261**

1) Material and Methods

a) Cloning of the SC_OH single chain molecule

A series of synthetic gene assembly was ordered to MWG-EUROFINS. Synthetic genes coding for the different single chain variants targeting HSV2 were cloned in pCLS1853 (figure 14) using AscI and XhoI restriction sites.

b) Extrachromosomal assay in mammalian cells

CHO K1 cells were transfected as described in example 1.8. 72 hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β -galactosidase liquid assay was added. After incubation at 37°C, OD was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform.

Per assay, 150 ng of target vector was cotransfected with an increasing quantity of variant DNA from 0.75 to 25 ng (25 ng of single chain DNA corresponding to 12,5ng + 12,5ng of heterodimer DNA). Finally, the transfected DNA variant DNA quantity was 0.78ng, 1.56ng, 3.12ng, 6.25ng, 12.5ng and 25ng. The total amount of transfected DNA was completed to 175ng (target DNA, variant DNA, carrier DNA) using empty vector (pCLS0001).

2) Results

The activity of the SCOH-HSV2 single chain molecules (Table XX) against the HSV2 target was monitored using the previously described CHO assay by comparison to the HSV2.3-M1 x HSV2.4-MC heterodimer (pCLS2733 x pCLS2735) and our internal control SCOH-RAG and I-Sce I meganucleases. All comparisons were done at 0.78ng, 1.56ng, 3.12ng, 6.25ng, 12.5ng, and 25ng transfected variant DNA (Figures 18 and 19).

Variants shared specific behaviour upon assayed dose depending on the mutation profile they bear (Figures 18 and 19). For example, SCOH-HSV2-M1-105A132V-MC-132V (pCLS2465) has a similar profile to our internal standard SCOH-RAG (SEQ ID NO: 468): its activity increases from low quantity to high quantity (Figure 20). SCOH-HSV2-M1-MC-80K105A132V (pCLS2459) has an activity maximum at low quantity of transfected DNA (1.56ng) and its activity quickly decreases with dose (Figure 21). SCOH-HSV2-M1-MC-132V (pCLS2457) shares an intermediate profile between the two previous ones, it has an activity

maxima at a low DNA dose (3.12ng) which slowly decreases with dose (Figure 22). All of these variants could be used for HSV-1 genome targeting depending on the tissue infected.

Example 2. Strategy for engineering novel meganucleases cleaving targets from the ICP_0 gene in HSV-1 genome.

HSV4 is a 24 bp (non-palindromic) target present in the RL2 gene encoding the ICP0 or a0 protein. This 3,6kb gene repeated twice in TRL (2086 to 5698) and IRL (120673 to 124285) regions is formed of three exons : position 2261 to 2317, 3083 to 3749, 3886 to 5489 and 120882 to 122485,122622 to 123288, 124054 to 124110. The target sequence present in exon 2 corresponds to positions 3498 to 3521 and 122850 to 122873 in the two copies of the HSV-1 ICP_0 gene (accession number NC_001806; Figure 23).

The HSV4 sequence is partly a patchwork of the 10AAG_P, 5GGT_P, 5CAG_P, 10ACT_P targets (Figure 24).

The 10AAG_P, 5GGT_P, 5CAG_P, 10ACT_P targets sequences are 24 bp derivatives of C1221, a palindromic sequence cleaved by I-CreI (Arnould et al., precited). However, the structure of I-CreI bound to its DNA target suggests that the two external base pairs of these targets (positions -12 and 12) have no impact on binding and cleavage (Chevalier et al., Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier et al., J. Mol. Biol., 2003, 329, 253-269), and in this study, only positions -11 to 11 were considered. Consequently, the HSV4 series of targets were defined as 22 bp sequences instead of 24 bp. HSV4 do not differs from C1221 in the 4 bp central region. According to the structure of the I-CreI protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the I-CreI protein (Chevalier et al., Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier et al., J. Mol. Biol., 2003, 329, 253-269). Thus, the bases at these positions should not impact the binding efficiency. However if different, they could have affected cleavage, which results from two nicks at the edge of this region. Thus, the sequence gtac in -2 to 2 was not modified during process.

Two palindromic targets, HSV4.3 and HSV4.4, were derived from HSV4 (Figure 24). Since HSV4.3 and HSV4.4 are palindromic, they should be

cleaved by homodimeric proteins. Thus, proteins able to cleave the HSV4.3 and HSV4.4 sequences as homodimers were first designed (examples 2.1 and 2.2) and then co-expressed to obtain heterodimers cleaving HSV4 (example 2.3). Heterodimers cleaving the HSV4 target could be identified. In order to improve cleavage activity for the HSV4 target, a series of variants cleaving HSV4.3 and HSV4.4 was chosen, and then refined. The chosen variants were subjected to random or site-directed mutagenesis, and used to form final heterodimers that were assayed against the HSV4 target (examples 2.4, 2.5 and 2.6). Heterodimers could be identified with an improved cleavage activity for the HSV4 target. Chosen heterodimers were subsequently cloned into mammalian expression vectors and screened against the HSV4 target in CHO cells (example 2.7). From positive heterodimer combinations in CHO cells, single chain variants with additional mutations were designed as final constructs for HSV4 targeting in mammalian cells. Strong cleavage activity of the HSV4 target could be observed for these heterodimers and single chain variants (example 2.8).

Example 2.1: Identification of meganucleases cleaving HSV4.3

This example shows that I-*Cre*I variants can cut the HSV4.3 DNA target sequence derived from the left part of the HSV4 target in a palindromic form. Target sequences described in this example are 22 bp palindromic sequences. Therefore, they will be described only by the first 11 nucleotides, followed by the suffix _P (For example, target HSV4.3 will be noted HSV4.3 CAAGCTGGTGT_P SEQ ID NO: 18).

A) Material and Methods

a) Construction of target vector

The target was cloned as follows: an oligonucleotide corresponding to the HSV4.3 target sequence flanked by gateway cloning sequences was ordered from (PROLIGO): 5'TGGCATACAAGTTTCCAAGCTGGTGTACACCAGCTT GGCAATCGTCTGTCA3' (SEQ ID NO: 262). Double-stranded target DNA, generated by PCR amplification of the single stranded oligonucleotide, was cloned using the Gateway protocol (INVITROGEN) into the yeast reporter vector (pCLS1055, Figure 4). Yeast reporter vector was transformed into *Saccharomyces cerevisiae* strain FYBL2-7B (*MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*), resulting in a reporter strain. (MilleGen)

b) Construction of combinatorial mutants

I-*Cre*I variants cleaving 10AAG_P or 5GGT_P were previously identified, as described in Smith *et al.* Nucleic Acids Res., 2006, 34, e149; International PCT Applications WO 2007/060495 and WO 2007/049156, and Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458; International PCT Applications WO 2006/097784 and WO 2006/097853, respectively for the 10AAG_P and 5GGT_P targets. In order to generate I-*Cre*I derived coding sequences containing mutations from both series, separate overlapping PCR reactions were carried out that amplify the 5' end (aa positions 1-43) or the 3' end (positions 39-167) of the I-*Cre*I coding sequence. For both the 5' and 3' end, PCR amplification is carried out using primers For both the 5' and 3' end, PCR amplification is carried out using primers (Gal10F 5'-gcaactttagtgctgacacatacagg-3' (SEQ ID NO: 263) or Gal10R 5'-acaaccttgattggagacttgacc-3'(SEQ ID NO: 264)) specific to the vector (pCLS0542, Figure 5) and primers (assF 5'-ctannnttgaccttt-3' (SEQ ID NO: 265) or assR 5'-aaaggtcaannntag-3'(SEQ ID NO: 266)), where nnn codes for residue 40, specific to the I-*Cre*I coding sequence for amino acids 39-43. The PCR fragments resulting from the amplification reaction realized with the same primers and with the same coding sequence for residue 40 were pooled. Then, each pool of PCR fragments resulting from the reaction with primers Gal10F and assR or assF and Gal10R was mixed in an equimolar ratio. Finally, approximately 25 ng of each final pool of the two overlapping PCR fragments and 75 ng of vector DNA (pCLS0542, Figure 5) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATα, trp1△63, leu2△1, his3△200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). An intact coding sequence containing both groups of mutations is generated by *in vivo* homologous recombination in yeast.

c) Mating of meganuclease expressing clones and screening in yeast

Screening was performed as described previously (Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony gridder (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-

harboring yeast strain. Membranes were placed on solid agar YPD rich medium, and incubated at 30 °C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking leucine and tryptophan, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6% dimethyl formamide (DMF), 7 mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software.

d) Sequencing of variants

To recover the variant expression plasmids, yeast DNA was extracted using standard protocols and used to transform *E. coli*. Sequencing of variant ORFs was then performed on the plasmids by MILLEGEN SA. Alternatively, ORFs were amplified from yeast DNA by PCR (Akada et al., Biotechniques, 2000, 28, 668-670), and sequencing was performed directly on the PCR product by MILLEGEN SA.

B) Results

I-*Cre*I combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5GGT_P with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10AAG_P on the I-CreI scaffold, resulting in a library of complexity 1680. Examples of combinatorial variants are displayed in Table XXI. This library was transformed into yeast and 3348 clones (2 times the diversity) were screened for cleavage against the HSV4.3 DNA target (CCAAGCTGGTGTACACCAGCTTGG). 9 positive clones were found which after sequencing turned out to correspond to 7 different novel endonuclease variants (Table XXII). Examples of positives are shown in Table XXII. The sequences of three variants identified display non parental combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77. These variants may be I-CreI combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast. Moreover, two of the selected variants display additional mutations to parental combinations (see examples Table XXII). Such mutations likely result from PCR artifacts during the combinatorial process.

Table XXI: Panel of variants theoretically present in the combinatorial library

Amino acids at positions 44, 68, 70, 75 and 77 (ex : ARNNI stands for A44, R68, N70, N75 and I77)	Amino acids at positions 28, 30,32, 33, 38 and 40 (ex: KHSSQS stands for K28, H30, S32, S33, Q38 and S40)										
	NTSYDS	RNSAYQ	SNSYQK	KASYQS	KASHQS	KASTQS	KDSRQS	KGSYQS	KGAYQS	KGSYGS	KGSYHS
ASSDR											
ARHDI											
DQSYR											
DRHDI											
DRRNI											
HKKDI											
IRKNV											
KYSNV											
KSTDI											
LRKNV											
LRNNI											
MRANI											
MRCNI											
NKSHF											
TRHDI											
TRKDI											
YRSDI											
YRSAT											
YRSEI											
YRSNI											
YRSNV											
YRSYQ											
YRSYT											
QRSNL											

*Only 264 out of the 1680 combinations are displayed. None of them were identified in the positive clones.

Table XXII: I-*Cre*I variants with and without additional mutations capable of cleaving the HSV4.3 DNA target.

Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-*Cre*I variants (ex: KRSRES/TYSNI stands for K28, R30, S32 , R33 , E38, S40/ T44, Y68, S70, N75 and I77)	SEQ ID NO:
KNSYQS/MRANV/132V	267
KNSYQS/MRCNI/12H	268
SNSYQK/MRNNI/80K94L	269
KNSYQS/LRNNI/80K	270
KGSYQS/IRKNV/132V	271
KNSYQS/YRKDI	272
SNSYQK/LRNNI/80K	273

Example 2.2: Identification of meganucleases cleaving HSV4.4

This example shows that I-*Cre*I variants can cleave the HSV4.4 DNA target sequence derived from the right part of the HSV4 target in a palindromic

form (Figure 24). All target sequences described in this example are 22 bp palindromic sequences. Therefore, they will be described only by the first 11 nucleotides, followed by the suffix _P (for example, HSV4.4 will be called *CACTATCAGGT*_P).

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 2.1, with the exception that an oligonucleotide corresponding to the HSV4.4 target sequence was used: 5'TGGCATACAAGTTTCCACTATCAGGTACCTGATAGTGGCAATCGTCTGTC A3' (SEQ ID NO: 274).

b) Construction of combinatorial variants

I-*Cre*I variants cleaving 10ACT_P or 5CAG_P were previously identified, as described in Smith *et al.* Nucleic Acids Res., 2006, 34, e149; International PCT Applications WO 2007/060495 and WO 2007/049156, and Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458; International PCT Applications WO 2006/097784 and WO 2006/097853, respectively for the 10ACT_P and 5CAG_P targets. In order to generate I-*Cre*I derived coding sequences containing mutations from both series, separate overlapping PCR reactions were carried out that amplify the 5' end (aa positions 1-43) or the 3' end (positions 39-167) of the I-*Cre*I coding sequence. For both the 5' and 3' end, PCR amplification is carried out using primers (Gal10F 5'-gcaactttagtgctgacacatacagg-3' (SEQ ID NO: 263) or Gal10R 5'-acaaccttgattggagacttgacc-3' (SEQ ID NO: 264)) specific to the vector (pCLS1107, Figure 6) and primers (assF 5'-ctannnttgaccttt-3' (SEQ ID NO: 265) or assR 5'-aaaggtcaannntag-3'(SEQ ID NO: 266)), where nnn codes for residue 40, specific to the I-*Cre*I coding sequence for amino acids 39-43. The PCR fragments resulting from the amplification reaction realized with the same primers and with the same coding sequence for residue 40 were pooled. Then, each pool of PCR fragments resulting from the reaction with primers Gal10F and assR or assF and Gal10R was mixed in an equimolar ratio. Finally, approximately 25 ng of each final pool of the two overlapping PCR fragments and 75 ng of vector DNA (pCLS1107, Figure 6) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATα, trp1Δ63, leu2Δ1, his3Δ200)

using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). An intact coding sequence containing both groups of mutations is generated by *in vivo* homologous recombination in yeast.

c) Mating of meganuclease expressing clones and screening in yeast

Screening was performed as described previously (Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony gridder (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain. Membranes were placed on solid agar YPD rich medium, and incubated at 30 °C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking tryptophan, adding G418, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6 % dimethyl formamide (DMF), 7 mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software. Positives resulting clones were verified by sequencing (MILLEGEN) as described in example 2.2.

B) Results

I-*Cre*I combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5CAG_P with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10ACT_P on the I-CreI scaffold, resulting in a library of complexity 1600. Examples of combinatorial variants are displayed in Table XXIII. This library was transformed into yeast and 3348 clones (2.1 times the diversity) were screened for cleavage against the HSV4.4 DNA target (*CACTATCAGGT*_P). A total of 20 positive clones were found to cleave HSV4.4. Sequencing and validation by secondary screening of these I-*Cre*I variants resulted in the identification of 14 different novel endonucleases. The sequence of 4 of the variants identified display non parental combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77 as well as additional mutations (see examples in Table XXIV). Such variants likely result from PCR artifacts during the combinatorial process. Alter-

natively, the variants may be I-CreI combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.

Table XXIII: Panel of variants theoretically present in the combinatorial library

Amino acids at positions 44, 68, 70, 75 and 77 (ex : HNRDI stands for H44, N68, R70, D75 and I77)	Amino acids at positions 28, 30,32, 33, 38 and 40 (ex: KRGYQS stands for K28, R30, G32, Y33, Q38 and S40)									
	KASTQS	KCSCQS	KGSYGS	KHSSQS	KKSAQS	KKSHQS	KKSRQS	KKSSQS	KKSTQS	KNDYYS
ARGNI										
ARSNI										
ATNNI										
ANNNI										
NRNNI										
ARNNI										
AQRNI										
ARHNI										
QGGNI										
AASYK										
ARSYT										
RYSEV										
DHSYI										
RYSDT										
ASSYK										
SYSYV										
NHSYN										
ATSDR										
AYSYI										
RYSYV										
ARSDR										
RYSYN										

* Only 220 out of the 1600 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HSV4.4 target was found among the identified positives.

Table XXIV: I-*Cre*I variants with and without additional mutations capable of cleaving the HSV4.4 DNA target.

Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-*Cre*I variants (ex: KRGYQS/KYSNI stands for K28, R30, G32 , Y33 , Q38, S40/ K44, Y68, S70, N75 and I77)	SEQ ID NO:
KNSRES / NHSYN / 80K	275
KNTYSS / TYSYV / 80K	276
KNTYWS / ARSYY	277
KNTCQS / AYSYK	278
KNTYSS / AYSYK	279
KRDYQS / ACSYI / 115V	280
KRSYES / AYSYK	281
KNEYYS / NYSYK	282
KNAYYS / AYSYK	283
KRDYQS / AYSYK	284
KNEYYS / AYSYK	285
KNSRES / AYSYK	286
KRDYQS / ARNNI	287
KNDYYS / AYSYK	288

Example 2.3: Identification of meganucleases cleaving HSV4

I-*Cre*I variants able to cleave each of the palindromic HSV4 derived targets (HSV4.3 and HSV4.4) were identified in example 2.2. Pairs of such variants (one cutting HSV4.3 and one cutting HSV4.4) were co-expressed in yeast. Upon co-expression, there should be three active molecular species, two homodimers, and one heterodimer. It was assayed whether the heterodimers that should be formed, cut the non palindromic HSV4 target.

A) Materials and Methods

a) Construction of target vector

The experimental procedure is as described in example 2.2, with the exception that an oligonucleotide corresponding to the HSV4 target sequence: 5'TGGCATACAAGTTTCCAAGCTGGTGTACCTGATAGTGGCAATCGTCTGTC A3' (SEQ ID NO: 289) was used.

b) Co-expression of variants

Yeast DNA was extracted from variants cleaving the HSV4.4 target in the pCLS1107 expression vector using standard protocols and was used to transform *E. coli*. The resulting plasmid DNA was then used to transform yeast strains expressing a variant cutting the HSV4.3 target in the pCLS542 expression vector.

Transformants were selected on synthetic medium lacking leucine and containing G418.

c) Mating of meganucleases coexpressing clones and screening in yeast

Mating was performed using a colony gridder (QpixII, Genetix). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm^2). A second gridding process was performed on the same filters to spot a second layer consisting of different reporter-harboring yeast strains for each target. Membranes were placed on solid agar YPD rich medium, and incubated at 30°C for one night, to allow mating. Next, filters were transferred to synthetic medium, lacking leucine and tryptophan, adding G418, with galactose (2 %) as a carbon source, and incubated for five days at 37 °C, to select for diploids carrying the expression and target vectors. After 5 days, filters were placed on solid agarose medium with 0.02 % X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1 % SDS, 6% dimethyl formamide (DMF), 7mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Results were analyzed by scanning and quantification was performed using appropriate software.

B) Results

Co-expression of variants cleaving the HSV4.4 target (9 variants chosen among those described in Table XXIV) and 6 variants cleaving the HSV4.3 target (described in Table XXII) resulted in cleavage of the HSV4 target in some cases (Figure 25). Functional combinations are summarized in Table XXV.

Table XXV: Cleavage of the HSV4 target by the heterodimeric variants

		Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-CreI variants cleaving the HSV4.3 target (ex: KRSRES/TYSNI stands for K28, R30, S32 , R33 , E38, S40/ T44, Y68, S70, N75 and I77)					
Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 Of I-CreI variants cleaving the HSV4.4 target (ex: KRGYQS/RHRDI stands for K28, R30, G32 , Y33 , Q38, S40/ R44, H68, R70, D75 and I77)		SNSYQK/ MRNNI/ 80K94L (SEQ ID NO: 299)	SNSYQK/ LRNNI/ 80K (SEQ ID NO: 300)	KNSYQS/ MRANV/ 132V (SEQ ID NO: 301)	KNSYQS/ LRNNI/ 80K (SEQ ID NO: 302)	KNSYQS/ YRKDI (SEQ ID NO: 303)	KGSYQS / IRKNV / 132V (SEQ ID NO: 304)
	KNDYYS / AYSYK (SEQ ID NO: 290)	+	+	+	+	+	+
	KNAYYS / AYSYK (SEQ ID NO: 291)						
	KRDYQS / ARNNI (SEQ ID NO: 292)	+	+	+	+	+	+
	KNTYSS / AYSYK (SEQ ID NO: 293)	+	+	+	+	+	+
	KNEYYS / AYSYK (SEQ ID NO: 294)	+	+	+	+	+	+
	KNSRES / AYSYK (SEQ ID NO: 295)						
	KNSRES / NHSYN / 80K (SEQ ID NO: 296)	+	+	+	+	+	+
	KNTYWS / ARSYY (SEQ ID NO: 297)	+	+	+	+	+	+
	KRDYQS / ACSYI / 115V (SEQ ID NO: 298)	+	+	+	+	+	+

+ indicates a functional combination

Example 2.4: Improvement of meganucleases cleaving HSV4.3 by random mutagenesis

I-*Cre*I variants able to cleave the palindromic HSV4.3 target has been previously identified in example 2.1. Some of them can cleave HSV4 target when associated with variants able to cut HSV4.4 (examples 2.2 and 2.3).

Therefore the 6 combinatorial variants cleaving HSV4.3 were mutagenized, and variants were screened for activity improvement on HSV4.3. According to the structure of the I-*Cre*I protein bound to its target, there is no contact

between the 4 central base pairs (positions -2 to 2) and the I-*Cre*I protein (Chevalier *et al.*, Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3757-3774; Chevalier *et al.*, J. Mol. Biol., 2003, 329, 253-269). Thus, it is difficult to rationally choose a set of positions to mutagenize, and mutagenesis was performed on the whole protein. Random mutagenesis results in high complexity libraries. Therefore, to limit the complexity of the variant libraries to be tested, the two components of the heterodimers cleaving HSV4 were mutagenized and screened in parallel.

A) Material and Methods

a) Construction of libraries by random mutagenesis

Random mutagenesis was performed on a pool of chosen variants, by PCR using Mn^{2+}. PCR reactions were carried out that amplify the I-*Cre*I coding sequence using the primers preATGCreFor (5'-gcataaattactatacttctatagacacgcaaacacaaatacacagcggccttgccacc-3'; SEQ ID NO: 169) and ICreIpostRev (5'-ggctcgaggagctcgtctagaggatcgctcgagttatcagtcggccgc-3'; SEQ ID NO: 170), which are common to the pCLS0542 (Figure 5) and pCLS1107 (Figure 6) vectors. Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLS0542) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MATα, trp1Δ63, leu2Δ1, his3Δ200*) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Expression plasmids containing an intact coding sequence for the I-*Cre*I variant were generated by *in vivo* homologous recombination in yeast.

b) Target vector yeast strains

The yeast strain FYBL2-7B (*MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*) containing the HSV4.3 target in the yeast reporter vector (pCLS1055 Figure 4) was constructed as described in example 2.1.

c) mating of meganuclease expressing clones, screening in yeast and sequencing

Mating HSV4.3 target strain and mutagenized variant clones and screening were performed as described in example 2.1. One variant from first generation was added as control on filter during screening steps for activity improvement evaluation.

B) Results

The 6 variants cleaving HSV4.3, (Table XXVI), were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then mated with a yeast strain that contains the HSV4.3 target in a reporter plasmid. After mating with this yeast strain, 86 clones were found to cleave the HSV4.3 target and 46 of them shown higher activity than the best original variant. An example of positives is shown in Figure 26. Sequencing of these 46 positive clones indicates that 38 distinct variants listed in Table XXVII were identified.

Table XXVI: pool of variants cleaving HSV4.3 and sequences used as template for random mutagenesis

SEQ ID NO:	Amino acids positions and residues of the I-*Cre*I variants
305	28S40K44M70N75N/80K94L
306	28S40K44L70N75N/80K
307	44M70A75N77V/132V
308	44L70N75N/80K
309	44Y70K/
310	30G44I70K75N77V/132V

Table XXVII: Improved variants displaying strong cleavage activity for HSV4.3

SEQ ID NO:	Amino acids positions and residues of the I-*Cre*I variants
311	28S30N32S33Y38Q40K44M70N75N77I/**80K94L**
312	28S30N40K44M70N75N77I80K94F132I/**56A**
313	28K30N40S44Y70K75N77I80K94F**132V**/
314	28S30N40K44M70N75N77I**80R94L**132I/
315	28S30N40K44M70N75N77I**80K94L132V**/
316	28K30G40S44M70A75N77V80E94F132V/**4N155Q**
317	28K30N40S44M70N75N77I80K94F132I/**82R163S**
318	28K30N40S44M70A75N77I80K94F132I/**100R**
319	28S30N40K44L70N75N77I80R94F132V/
320	28K30N40S44Y70K75D77I80E94F132V/
321	28S30N40K44M70N75N77I80K94L132I/
322	28S30N40K44M70N75N77I80K94F132I/**31R79C128R**
323	28K30N40S44Y70K75D77I80K94F132I/**157D**
324	28K30N40S44M70N75N77I80K94F132I/**3P68H69E**
325	28S30N40K44L70N75N77I80K94F132V/**86D**
326	28S30N40K44M70N75N77I80K94F132I/
327	28K30N40S44Y70K75D77I80K94F132I/
328	28K30N40S44M70A75N77V80E94F132V/**54L**
329	28S30N40K44M70N75D77I80K94F132V/**108T154G**
330	28S30N40K44L70N75N77I80K94F132V/**105D**

331	28S30N40K44L70N75N77I80K94F132V/**43L**
332	28K30N40S44M70N75N77I80K94F132I/**56V151A**
333	28K30G40S44M70A75N77V80E94F132V/**156G**
334	28S30N40K44L70N75N77T80K94F132I/
335	28K30G40S44M70A75N77V80E94F132V/
336	28K30N40S44M70A75N77I80K94F132V/**146K156G**
337	28S30N40K44M70K75N77V80E94F132V/
338	28K30N40S44M70A75N77I80K94F132I/**117V**
339	28K30G40S44I70K75D77I80E94F132I/**160E**
340	28S30N40K44L70N75N77I80K94F132V/
341	28S30N40K44L70N75N77I80K94F132I/**7R79T102V**
342	28S30N40K44L70N75N77I80K94F132I/**54V**
343	28K30N40S44M70N75N77I80K94F132I/
344	28S30N40K44M70N75N77M80R94F132I/**71E105A**
345	28K30N40S44M70K75N77V80E94F132V/**7E**
346	28S30N40K44M70N75N77I80K94F132I/**116R**
347	28K30N40S44M70K75D77I80E94F132V/**2Y**
348	28K30N40S44Y70K75D77I80E94F132V/**79C120A152V**
349	28S30N40K44L70N75N77I80K94F132I/**50R**

* Mutations resulting from random mutagenesis are in bold.

Example 2.5: Improvement of meganucleases cleaving HSV4.4 by random mutagenesis

I-*Cre*I variants able to cleave the palindromic HSV4.4 target has been previously identified in example 2.2. Some of them can cleave HSV4 target when associated with variants able to cut HSV4.4 (examples 2.1 and 2.3).

Therefore 9 of the 14 combinatorial variants cleaving HSV4.4 were mutagenized, and variants were screened for activity improvement on HSV4.4. As described in example 2.5, mutagenesis was performed on the whole protein and HSV4.4 variants were screened in parallel of HSV4.3 (example 2.5).

A) Material and Methods

a) Construction of libraries by random mutagenesis

Random mutagenesis was performed as described in example 2.5, on a pool of chosen variants, by PCR using the same primers and Mn^{2+} conditions (preATGCreFor SEQ ID NO: 169 and ICreIpostRev SEQ ID NO: 170). Approximately 25 ng of the PCR product and 75 ng of vector DNA pCLS1107) linearized by digestion with *Nco*I and *Eag*I were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MATα, trp1Δ63, leu2Δ1, his3Δ200*) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002,

350, 87-96). Expression plasmids containing an intact coding sequence for the I-*Cre*I variant were generated by *in vivo* homologous recombination in yeast.

b) Target vector yeast strains

The yeast strain FYBL2-7B (*MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*) containing the HSV4.4 target in the yeast reporter vector (pCLS1055 Figure 4) was constructed as described in example 2.2.

c) mating of meganuclease expressing clones, screening in yeast and sequencing

Mating HSV4.4 target strain and mutagenized variant clones and screening were performed as described in example 2.2. One variant from first generation was added as control on filter during screening steps for activity improvement evaluation.

B) Results

Nine chosen variants cleaving HSV4.4, (Table XXVIII), were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then mated with a yeast strain that contains the HSV4.4 target in a reporter plasmid. After mating with this yeast strain, 262 clones were found to cleave the HSV4.4 target and 17 of them shown higher activity than the best original variant. An example of positives is shown in Figure 27. Sequencing 93 of these 262 positive clones indicates that 63 distinct variants listed in Table XXIX were identified. 14 of them shown higher activity than the best original variant.

Table XXVIII: pool of variants cleaving HSV4.4 and sequences used as template for random mutagenesis

SEQ ID NO:	Amino acids positions and residues of the I-*Cre*I variants
350	28K30N32D33Y38Y40S44A68Y70S75Y77K/
351	28K30N32S33R38E40S44N68Y70S75Y77K/
352	28K30R32D33Y38Q40S44A68R70N75N77I/
353	28K30N32T33Y38S40S44A68Y70S75Y77K/
354	28K30N32E33Y38Y40S44A68Y70S75Y77K/
355	28K30N32S33R38E40S44N68H70S75Y77N/80K
356	28K30N32T33Y38W40S44A68R70S75Y77Y/
357	28K30R32D33Y38Q40S44A68C70S75Y77I/115V
358	28K30N32S33R38E40S44A68Y70S75Y77K/

Table XXIX: Improved variants displaying strong cleavage activity for HSV4.4

SEQ ID NO:	Amino acids positions and residues of the I-CreI variants
359	32E38Y44A68Y70S75Y77K**105A**
360	**24F**32E38Y44A68Y70S75Y77K**107R**
361	32E38Y44A68Y70S75Y77K**132V**
362	32D38Y44A68Y70S75Y77K
363	32E38Y44A68Y70S75Y77K**114Y**
364	**1V**32G38Y44A68Y70S75Y77K**80K103S**
365	32D38Y44A68Y**69G**70S75Y77K**80A105A162P**
366	32E38Y44A**50R**68Y70S75Y77K**160E**
367	32E38Y44A68Y70S75Y77K
368	33R38E44A**54L**68Y70S75Y77K**103I**
369	32T38S44A68Y70S75Y77K**124Q132V**
370	33R38E44A68Y70S75Y77K**129A**
371	**6S**32E38Y44A68Y70S75Y77K**111R**
372	32D38Y44A68Y70S75Y77N**80K132V**
373	32D38Y44A68Y70S75Y77K**164T**
374	32E38Y44A68Y70S75Y77K**163R**
375	32D38Y44A68Y70S75Y77K**82M110D117G132V163Q**
376	32E38Y44A68Y70S75Y77K**162P**
377	32T**36N**38Y44A**60G**68S70S75Y77K**107E110K132V**
378	32E38Y44A68Y70S75Y77K**160N**
379	32E38Y44A68Y70S75Y77K**163Q**
380	32E38Y**43L**44A68Y70S75Y77K
381	32E38Y44A68Y70S75Y77K**107R**
382	32E38Y44A68Y70S75Y77K**151A**
383	**2S**32D38Y44A68Y70S75Y77K
384	32D38Y44A**66H**68Y70S75Y77K**152Q158E**
385	**6S**32D38Y44A68Y70S75Y77K
386	32D38Y44A68Y70S75Y77K**78I**
387	**33R**38Y44A68Y70S75Y77K**100R161P**
388	32E38Y44A**64A**68Y70S**72C**75Y77K**129A135Q156R**
389	32D38Y44A**60N**68Y70S75Y77K
390	32E38Y44A68Y70S**73I**75Y77K**87L**
391	32D38Y**40C**44A68Y70S75Y77K
392	32E38Y44A**49A**68Y70S75Y77K
393	33R38E44A68Y70S75Y77K**85R**
394	32E38Y44A68Y70S75Y77K**159R**
395	32K38Y44A68Y70S75Y77N**134T**

396	32T**33R**38E44A68Y70S75Y77K
397	30K32E38Y44A68Y70S75Y77K
398	32E38Y44A68Y70S75Y77K**96Q**
399	**3S**32E38Y44A68Y70S75Y77K
400	32E38Y44A68Y70S75Y77K**105G**
401	33R38E44A68Y70S75Y77K
402	**7E**32D36N38Y44A68Y70S75Y77K
403	30R**31H**32D44A**49A**68Y70S75Y77K
404	32T38S44A68Y70S75Y77K**114T**
405	33R38E44A68Y70S75Y77K**163S**
406	**33H**38E44A**66H**68Y70S75Y77K
407	32D38Y44A68Y70S72Y75Y77K
408	33R38E44A68Y70S75Y77R97V**161A**
409	**13M**32E38Y44A68Y70S75Y77K
410	32E38Y44A68Y70S75Y77K**107E**
411	30R32D44A68Y70S75Y77K**100N**
412	30R32D44A68Y70S75Y77K**114T120A**
413	32E38Y44A64A68Y70S75Y77K**82M**
414	30R32D44A68Y70S75Y77K
415	32D38Y44A68Y70S**71E**75Y77K
416	33R38E44A57R68Y70S75Y77K
417	32T38W44A70S75Y77Y**80K110D**
418	30R32D44A68Y70S75Y77K**107R**
419	32T38S44A68Y70S75Y77K**120G**
420	33R38E44N68H70S75Y77N**80K111H131R132V**
421	32D38Y44A54L68Y70S75Y77K**79G129A156G**

* Mutations resulting from random mutagenesisare in bold.

Example 2.6: Identification of improved meganucleases cleaving HSV4

Improved I-*Cre*I variants able to cleave each of the palindromic HSV4 derived targets (HSV4.3 and HSV4.4) were identified in example 2.5 and example 2.6. As described in example 2.3, pairs of such variants (one cutting HSV4.3 and one cutting HSV4.4) were co-expressed in yeast. The heterodimers that should be formed were assayed for cutting the non palindromic HSV4 target.

A) Materials and Methods

a) Construction of target vector

The HSV4 target vector was constructed as described in example 2.3.

b) Co-expression of variants

Yeast DNA was extracted from variants cleaving the HSV4.4 target in the pCLS1107 expression vector using standard protocols and was used to transform *E. coli*. The resulting plasmid DNA was then used to transform yeast strains expressing a variant cutting the HSV4.3 target in the pCLS542 expression vector. Transformants were selected on synthetic medium lacking leucine and containing G418.

c) Mating of meganucleases coexpressing clones and screening in yeast

Mating and screening of meganucleases coexpressing clones were performed as described in example 2.3. Results were analyzed by scanning and quantification was performed using appropriate software.

B) Results

Co-expression of improved variants cleaving the HSV4.4 target (6 variants chosen among those described in Table XXIX) and 6 improved variants cleaving the HSV4.3 target (described in Table XXVII) resulted in cleavage of the HSV4 target in all of cases (Figure 28). All assayed combinations are summarized in Table XXX.

Table XXX: Cleavage of the HSV4 target by the heterodimeric improved variants

		Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-CreI variants cleaving the HSV4.3 target (ex: KRSRES/TYSNI stands for K28, R30, S32 , R33 , E38, S40/ T44, Y68, S70, N75 and I77)					
Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 Of I-CreI variants cleaving the HSV4.4 target (ex: KRGYQS/RHRDI stands for K28, R30, G32 , Y33 , Q38, S40/ R44, H68, R70, D75 and I77)		SNSYQK / LRNNI / 80R132V (SEQ ID NO: 428)	KNSYQS / MRANI / 80K132V146 K156G (SEQ ID NO: 429)	SNSYQK / LRNNI / 80K132V/43 L (SEQ ID NO: 430)	KNSYQS / MRANV / 54L132V (SEQ ID NO: 431)	SNSYQK / MRNNI / 80K (SEQ ID NO: 432)	KNSYQS / MRKDI / 2Y132V (SEQ ID NO: 433)
	KNEYYS / AYSYK /50R160E (SEQ ID NO: 422)	+	+	+	+	+	+
	KNDYYS / AYSYK / 69G80A105A16 2P (SEQ ID NO: 423)	+	+	+	+	+	+
	KNEYYS / AYSYK / 132V (SEQ ID NO: 424)	+	+	+	+	+	+
	KNEYYS / AYSYK / 24F107R (SEQ ID NO: 425)	+	+	+	+	+	+
	KNGYYS / AYSYK /1V80K103S (SEQ ID NO: 426)	+	+	+	+	+	+
	KNEYYS / AYSYK /105A (SEQ ID NO: 427	+	+	+	+	+	+

+ indicates a functional combination

Example 2.7: Validation of HSV4 target cleavage in an extrachromosomal model in CHO cells

I-*Cre*I variants able to efficiently cleave the HSV4 target in yeast when forming heterodimers were described in examples 2.3 and 2.7. In order to identify heterodimers displaying maximal cleavage activity for the HSV4 target in CHO cells, the efficiency of chosen combinations of variants to cut the HSV4 target was compared, using an extrachromosomal assay in CHO cells. The screen in CHO cells is a single-strand annealing (SSA) based assay where cleavage of the target by the meganucleases induces homologous recombination and expression of a LagoZ reporter gene (a derivative of the bacterial lacZ gene).

1) Materials and methods

a) Cloning of HSV4 target in a vector for CHO screen

The target was cloned as follows: oligonucleotide corresponding to the HSV4 target sequence flanked by gateway cloning sequence was ordered from

PROLIGO 5'TGGCATACAAGTTTCCAAGCTGGTGTACCTGATAGTGGCAATCGTCTGTC A3' (SEQ ID NO: 289). Double-stranded target DNA, generated by PCR amplification of the single stranded oligonucleotide, was cloned using the Gateway protocol (INVITROGEN) into CHO reporter vector (pCLS1058, Figure 11). Cloned target was verified by sequencing (MILLEGEN).

b) Re-cloning of meganucleases

The ORF of I-CreI variants cleaving the HSV4.3 and HSV.4 targets identified in examples 2.5 and 2.6 were re-cloned in pCLS1768 (Figure 29). ORFs were amplified by PCR on yeast DNA using the attB1-ICreIFor (5'-ggggacaagtttgtacaaaaaagcaggcttcgaaggagatagaaccatggccaataccaaatataacaaagagttcc-3'; SEQ ID NO: 434) and attB2-ICreIRev (5'-ggggaccactttgtacaagaaagctgggtttagtcggccgccggggaggatttcttcttctcgc-3'; SEQ ID NO: 435) primers. PCR products were cloned in the CHO expression vector pCLS1768 (Figure 29) using the Gateway protocol (INVITROGEN). Resulting clones were verified by sequencing (MILLEGEN).

c) Extrachromosomal assay in mammalian cells

CHO K1 cells were transfected with Polyfect® transfection reagent according to the supplier's protocol (QIAGEN). 72 hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β -galactosidase liquid assay was added (typically 1 liter of buffer contained: 100 ml of lysis buffer (Tris-HCl 10 mM pH7.5, NaCl 150 mM, Triton X100 0.1 %, BSA 0.1 mg/ml, protease inhibitors), 10 ml of Mg 100X buffer ($MgCl_2$ 100 mM, β-mercaptoethanol 35 %), 110 ml ONPG 8 mg/ml and 780 ml of sodium phosphate 0.1M pH7.5). After incubation at 37°C, OD was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform.

Per assay, 150 ng of target vector was cotransfected with 12.5 ng of each one of both mutants (12.5 ng of mutant cleaving palindromic HSV4.3 target and 12.5 ng of mutant cleaving palindromic HSV4.4 target).

2) Results

6 variants cleaving HSV4.3 and 4 variants cleaving HSV4.4 described in example 2.5, 2.6 and 2.7 were re-cloned in pCLS1768 (Figure 29). Then,

I-*Cre*I variants cleaving the HSV4.3 or HSV4.4 targets were assayed together as heterodimers against the HSV4 target in the CHO extrachromosomal assay.

Table XXXIII shows the functional combinations obtained for 24 heterodimers. Analysis of the efficiencies of cleavage and recombination of the HSV4 sequence demonstrates that 9 combinations of I-CreI variants are able to transpose their cleavage activity from yeast to CHO cells without additional mutation.

Table XXXI: Functional heterodimeric combinations cutting the HSV4 target in CHO cells.

		Optimized variants cleaving HSV4.3					
		28S40K44L 70N80R132V (SEQ ID NO: 440)	44M70A80K 132V146K156G (SEQ ID NO: 441)	28S40K43L44L 70N80K132V (SEQ ID NO: 442)	44M54L70A 77V132V (SEQ ID NO: 443)	28S40K44M 70N80K (SEQ ID NO: 444)	2Y44M70K 132V (SEQ ID NO: 445)
Optimized variants cleaving HSV4.4	32D38Y44A68Y 69G70S75Y77 K80A105A162P (SEQ ID NO: 436)		+	+	+	+	
	24F32E38Y44A 68Y70S75Y77K 107R (SEQ ID NO: 437)						
	70S75Y77K 80K103S (SEQ ID NO: 438)						
	32E38Y44A68Y 70S75Y77K105A (SEQ ID NO: 439)		+	+	+	+	+

+ indicates a functional combination

Example 2.8: Covalent assembly as single chain and improvement of meganucleases cleaving HSV4 by site-directed mutagenesis

Coexpression of the cutters described in example 2.6, 2.7, 2.8 leads to a high cleavage activity of the HSV4 target in yeast. Some of them are able to cleave HSV4 in a mammalian expression system. One of them is shown as example in Table XXXII.

Table XXXII: Functional heterodimer cutting the HSV4 target in CHO cells.

HSV4.3-M2 (SEQ ID NO: 35)	44M 70A 80K 132V 146K 156G
HSV4.4-MF (SEQ ID NO: 36)	32E 38Y 44A 68Y 70S 75Y 77K 105A

The M2 / MF HSV4 heterodimer gives high cleavage activity in yeast and CHO cells. M2 is a HSV4.3 cutter that bears the following mutations in

comparison with the I-*Cre*I wild type sequence: 44M, 70A, 80K, 132V, 146K, 156G. MF is a HSV4.4 cutter that bears the following mutations in comparison with the I-*Cre*I wild type sequence: 32E, 38Y, 44A, 68Y, 70S, 75Y, 77K, 105A.

Single chain constructs were engineered using the linker RM2 resulting in the production of the single chain molecule: M2-RM2-MF. During this design step, the G19S mutation was introduced in the C-terminal MF mutant. In addition, mutations K7E, K96E were introduced into the M2 mutant and mutations E8K, E61R into the MF mutant to create the single chain molecule: M2(K7E K96E)-RM2-MF(E8K E61R) that is called further SCOH-HSV4-M2-MF.

Four additional amino-acid substitutions have been found in previous studies to enhance the activity of I-CreI derivatives: these mutations correspond to the replacement of Phenylalanine 54 with Leucine (F54L), Glutamic acid 80 with Lysine (E80K), Valine 105 with Alanine (V105A) and Isoleucine 132 with Valine (I132V). Some of these are already present in HSV4.3 and HSV4.4 variants (i.e. HSV4.3-M2 and HSV4.4-MF). Some additional combinations were introduced into the coding sequence of N-terminal and C-terminal protein fragment (example in Table XXXIII), and the resulting proteins were tested for their ability to induce cleavage of the HSV4 target. Twelve single chain constructs were then tested in CHO for cleavage of the HSV4 target.

Table XXXIII: Example of single chain I-*Cre*I variants assayed for HSV4 cleavage in CHO cells.

Construct	Single Chain	Mutations on N-terminal segment	Mutations on C-terminal segment	SEQ ID NO
pCLS2470	SCOH-HSV4-M2-MF	7E 44M 70A **80K 96E 132V** 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **105A**	446
pCLS2471	SCOH-HSV4-M2-MF-54L	7E 44M 70A **80K 96E 132V** 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **105A 132V**	447
pCLS2472	SCOH-HSV4-M2-MF-132V	7E 44M 70A **80K 96E 132V** 146K 156G	**8K 19S** 32E 38Y 44A **54L** 61R 68Y 70S 75Y 77K **105A**	448
pCLS2474	SCOH-HSV4-M2-54L-MF	7E 44M **54L** 70A **80K 96E 132V** 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **105A**	449
pCLS2476	SCOH-HSV4-M2-54L-MF-132V	7E 44M **54L** 70A **80K 96E 132V** 146K 156G	**8K 19S** 32E 38Y 44A **54L** 61R 68Y 70S 75Y 77K **105A**	450
pCLS2477	SCOH-HSV4-M2-54L-MF-80K132V	7E 44M **54L** 70A **80K 96E** 132V 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **80K 105A 132V**	451
pCLS2478	SCOH-HSV4-M2-105A-MF	7E 44M 70A **80K 96E 105A 132V** 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **105A**	452
pCLS2479	SCOH-HSV4-M2-105A-MF-132V	7E 44M 70A **80K 96E 105A 132V** 146K 156G	**8K 19S** 32E 38Y 44A 61R 68Y 70S 75Y 77K **105A 132V**	453
pCLS2481	SCOH-HSV4-M2-105A-MF-80K132V	7E 44M 70A **80K 96E 105A 132V** 146K 156G	8K 19S 32E 38Y 44A 61R 68Y 70S 75Y 77K **80K 105A 132V**	454

1) Material and Methods

a) Cloning of the SC_OH single chain molecule

A series of synthetic gene assembly was ordered to TOPGENE TECHNOLOGY. Synthetic genes coding for the different single chain variants targeting HSV4 were cloned in pCLS0491 (figure 31) using Eco RI and Bam HI restriction sites.

b) Extrachromosomal assay in mammalian cells

CHO K1 cells were transfected as described in example 2.8. 72 hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β -galactosidase liquid assay was added. After incubation at 37°C, OD was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform.

Per assay, 150 ng of target vector was cotransfected with an increasing quantity of variant DNA from 0.75 to 50 ng (50 ng of single chain DNA corresponding to 25ng + 25ng of heterodimer DNA). Finally, the transfected DNA variant DNA quantity was 0.78ng, 1.56ng, 3.12ng, 6.25ng, 12.5ng, 25ng and 50ng. The total amount of transfected DNA was completed to 200ng (target DNA, variant DNA, carrier DNA) using empty vector (pCLS0001).

2) Results

The activity of the SCOH-HSV4 single chain molecules (Table XXXIII) against the HSV4 target was monitored using the previously described CHO assay by comparison to the HSV4.3-M2 x HSV4.4-MF heterodimer and our internal control SCOH-RAG (SEQ ID NO: 468) and I-*Sce*I (SEQ ID NO: 469) meganucleases . All comparisons were done at 0.78ng, 1.56ng, 3.12ng, 6.25ng, 12.5ng, 25ng and 50ng transfected variant DNA.

All assayed single chain variants are more active than M2 x MF heterodimer and the internal control I-*Sce*I at standard dose (25ng). Variants shared specific behaviour upon assayed dose depending on the mutation profile they bear (Figures 33 and 34). For example, scOH-HSV4-M2-54L-MF (pCLS2474, SEQ ID NO: 449) has a similar profile to our internal standard SCOH-RAG: its activity increase from low quantity to high quantity (Figure 35). scOH-HSV4-M2-105A-MF-80K132V (pCLS2481, SEQ ID NO: 454) is highly active at low quantities of

transfected DNA (3.12ng) and its apparent activity decreases with dose (figure 36). scOH-HSV4-M2-MF-132V (pCLS2472, SEQ ID NO: 448) shares an intermediate profile between the two previous ones (figure 37). The profile of scOH-HSV4-M2-MF (pCLS2470, SEQ ID NO: 446), which is a common scaffold to all assayed single chain variants, is an average of individual behaviors at low DNA quantity (max at 6.25ng) and decreases quickly with DNA dose (the lowest at 50ng) (figure 38). All of these variants could be used for HSV-1 genome targeting depending on the tissue infected.

Example 3: Inhibition of viral replication by I-CreI variants cleaving HSV2, HSV4 or HSV12 target sequences

Single-chain obligate heterodimer I-*Cre*I variants able to efficiently cleave the HSV2 or HSV4 target sequences in yeast and CHO cells were described in examples 1 and 2. Single chain obligate heterodimer constructs were also generated for the I-CreI variants able to cleave the HSV12 target sequences described in table II. These single chain constructs were engineered using the linker RM2 (AAGGSDKYNQALSKYNQALSKYNQALSGGGGS) (SEQ ID NO: 464). During this design step, mutations K7E, K96E were introduced into the M1 or the M1-80K mutant and mutations E8K, E61R into the ME-132V mutant to create the single chain molecules: M1(K7E K96E)-RM2-ME-132V(E8K E61R) that is called SCOH-HSV12-M1-ME-132V and M1-80K(K7E K96E)-RM2-ME-132V(E8K E61R) that is called SCOH-HSV12-M1-80K-ME-132V (Table XXXIV).

Table XXXIV: Example of single chain I-*Cre*I variants for HSV12

Construct	Single Chain	Mutations on N-terminal segment	Mutations on C-terminal segment	SEQ ID NO
pCLS2633	SCOH-HSV12-M1-ME-132V	7E 24V 33C 38S 44I 50R 70S 75N 77R 96E 132V	8K 19S 30R 33S 44K 61R 66H 68Y 70S 77T 87I 132V 139R 163S	465
pCLS2635	SCOH-HSV12-M1-80K-ME-132V	7E 24V 33C 38S 44I 50R 70S 75N 77R 80K 96E 132V	8K 19S 30R 33S 44K 61R 66H 68Y 70S 77T 87I 132V 139R 163S	466

In order to further validate the cleavage activity of these single chain molecules, the ability of I-CreI variants cleaving HSV2, HSV4 or HSV12 target sequences to inhibit viral replication was examined using a recombinant Herpes Simplex Virus (rHSV-1). rHSV was constructed with a cassette containing a CMV

promoter driving the LacZ gene (Figure 39). An I-SceI target site was inserted between the CMV promoter and the LacZ gene and served as a positive control for inactivation of the virus. This expression cassette was introduced into the major LAT locus of HSV by homologous recombination resulting in LacZ expression during lytic infection of COS-7 cells. Thus to evaluate the inhibition of viral replication, the ability of I-SceI or the I-CreI variants cleaving HSV2, HSV4 or HSV12 target sequences to diminish LacZ expression after infection with rHSV-1 was evaluated.

1) Material and Methods

a) Single chain obligate heterodimer (SC OH) molecules

Single chain obligate heterodimer molecules were generated for the I-CreI variants able to cleave the HSV12 target sequences described in table II by custom gene synthesis (MWG-EUROFINS). Synthetic genes coding for the different single chain variants targeting HSV12 were cloned in pCLS1853 (figure 14) using AscI and XhoI restriction sites.

b) Cells and viruses

Viruses were grown and assayed on COS-7 cells. COS-7 cells were cultured in DMEM supplemented with 2mM L-glutamine, penicillin (100 IU/ml), streptomycin (100 mg/ml), amphotericin B (Fongizone: 0.25 mg/ml, Invitrogen-Life Science) and 10% FBS. HSV-1 was purchased from the American Type Culture Collection (ATCC). Viruses were propagated at a multiplicity of infection of 0.003 PFU/cell and virus titers were determined by plaque assays.

c) Construction of recombinant HSV-1

Recombinant virus was generated in a manner similar to that previously described (Lachmann, R.H., Efstathioun S., 1997, Journal of Virology, 3197-3207). An approximately 4,6 kb PstI-BamHI viral genomic fragment was cloned into pUC19. Based on HSV-1 sequence from the database (GenBank NC_001806) this represents nucleotides 118869-123461 and 7502-2910 in the inverted terminal repeats of the HSV-1 genome. A cassette containing the CMV promoter driving LacZ expression was introduced into a 19bp SmaI/HpaI deletion. This region is located within the major LAT locus of HSV-1. The I-SceI cleavage site (tagggataacagggtaat SEQ ID NO: 467) was inserted after the CMV promoter and before the ATG of the LacZ gene. This construct (pCLS0126, figure 40) was used to generate recombinant

viruses. Plasmid was linearized by XmnI digestion and 2μg of this plasmid was co-transfected with 10μg of HSV-1 genomic DNA into COS-7 cells using Lipofectamine 2000 (Invitrogen). After 3 days, infected cells were harvested and sonicated. An aliquot of the lysed cells was used to infect a COS monolayer and cells were overlayed with 1% agarose medium. After 3 days, 300μg/ml of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to the overlay. β-galactosidase positive 'blue' plaques were picked and subjected to three rounds of plaque purification.

d) Viral inhibition

6-well plates were seeded with 2 x 10^5 cells per well. The next day COS-7 cells were transfected using lipofectamine 2000 (Invitrogen) with either 1μg or 5μg of plasmid expressing I-SceI or the I-CreI variants cleaving HSV2, HSV4 or HSV12 target sequences, the total volume of DNA was completed to 5μg with empty vector pCLS0001 (Figure 15). The transfection efficiency was between 50-70% using this method. Twenty-four hours later, subconfluent transfected cells were infected with rHSV-1. For infection, rHSV-1 was diluted in DMEM without serum and adsorbed onto cells for 2 hour at 37°C in a humidified incubateur with 5% CO_2. 6 well plates were infected with 300 plaque forming units (PFU) per well. Cells were harvested twenty-four hours after infection and β-galactosidase activity was assayed on a total of 2.5 x 10^4 cells using a luminescent β-galactosidase assay (Beta-Glo assay, Promega).

2) Results

Three single chain variants cleaving the HSV4 target sequence (pCLS2472, SCOH-HSV4-M2-MF-132V, SEQ ID NO: 448; pCLS2474, SCOH-HSV4-M2-54L-MF, SEQ ID NO: 449 and pCLS2481, SCOH-HSV4-M2-105A-MF-80K132V, SEQ ID NO: 454) described in example 2.8, three single-chain variants cleaving the HSV2 target sequence described in example 1.8 (pCLS2457, SCOH-HSV2-M1-MC-132V, SEQ ID NO: 254; pCLS2459, SCOH-HSV2-M1-MC-80K105A132V, SEQ ID NO: 256 and pCLS2465, SCOH-HSV2-M1-105A132V-MC-132V, SEQ ID NO: 261) and two single chain variants cleaving the HSV2 target sequence (pCLS2633, SCOH-HSV12-M1-ME-132V (SEQ ID NO: 465) and

pCLS2635, SCOH-HSV12-M1-80K-ME-132V, SEQ ID NO: 466)) described in Table XXXIV were tested for their ability to inhibit viral replication of rHSV-1.

Figure 41 shows the results obtained for the eight single-chain variants as well as I-*Sce*I compared to cells treated with empty vector only. Transfection of 5μg I-SceI expression vector before viral infection results in a significant reduction in LacZ activity (greater than 3-fold), the levels of LacZ activity observed are only 31% of those observed following transfection of an empty vector. The single-chain obligate heterodimer variants cleaving the HSV4, HSV2 or HSV12 target sequences display reductions in LacZ activity similar to that of I-SceI (2- to 4-fold). The level of LacZ activity observed was 25-51% of that observed with an empty vector.

CLAIMS

1. An I-*Cre*I variant, characterized in that at least one of the two I-*Cre*I monomers has at least two substitutions, one in each of the two functional subdomains of the LAGLIDADG core domain situated from positions 26 to 40 and 44 to 77 of I-*Cre*I, said variant being able to cleave a DNA target sequence from the herpes simplex virus (HSV) genome, and being obtainable by a method comprising at least the steps of:

(a) constructing a first series of I-*Cre*I variants having at least one substitution in a first functional subdomain of the LAGLIDADG core domain situated from positions 26 to 40 of I-*Cre*I,

(b) constructing a second series of I-*Cre*I variants having at least one substitution in a second functional subdomain of the LAGLIDADG core domain situated from positions 44 to 77 of I-*Cre*I,

(c) selecting and/or screening the variants from the first series of step (a) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions -10 to -8 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the HSV genome and (ii) the nucleotide triplet in positions +8 to +10 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in position -10 to -8 of said DNA target sequence from the HSV genome,

(d) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions -5 to -3 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the HSV genome and (ii) the nucleotide triplet in positions +3 to +5 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in position -5 to -3 of said DNA target sequence from the HSV genome,

(e) selecting and/or screening the variants from the first series of step (a) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions +8 to +10 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the HSV genome and (ii) the nucleotide triplet in positions -10 to -8 has been

replaced with the reverse complementary sequence of the nucleotide triplet which is present in position +8 to +10 of said DNA target sequence from the HSV genome,

(f) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant I-*Cre*I site wherein at least one of (i) the nucleotide triplet in positions +3 to +5 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the HSV genome and (ii) the nucleotide triplet in positions -5 to -3 has been replaced with the reverse complementary sequence of the nucleotide triplet which is present in position +3 to +5 of said DNA target sequence from the HSV genome,

(g) combining in a single variant, the mutation(s) in positions 26 to 40 and 44 to 77 of two variants from step (c) and step (d), to obtain a novel homodimeric I-*Cre*I variant which cleaves a sequence wherein (i) the nucleotide triplet in positions -10 to -8 is identical to the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the HSV genome, (ii) the nucleotide triplet in positions +8 to +10 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the HSV genome, (iii) the nucleotide triplet in positions -5 to -3 is identical to the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the HSV genome and (iv) the nucleotide triplet in positions +3 to +5 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions -5 to -3 of said DNA target sequence from the HSV genome, and/or

(h) combining in a single variant, the mutation(s) in positions 26 to 40 and 44 to 77 of two variants from step (e) and step (f), to obtain a novel homodimeric I-*Cre*I variant which cleaves a sequence wherein (i) the nucleotide triplet in positions +8 to +10 of the I-*Cre*I site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the HSV genome and (ii) the nucleotide triplet in positions -10 to -8 is identical to the reverse complementary sequence of the nucleotide triplet in positions +8 to +10 of said DNA target sequence from the HSV genome, (iii) the nucleotide triplet in positions +3 to +5 is identical to the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the HSV genome, (iv) the nucleotide triplet

in positions -5 to -3 is identical to the reverse complementary sequence of the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the HSV genome,

(i) combining the variants obtained in steps (g) and (h) to form heterodimers, and

(j) selecting and/or screening the heterodimers from step (i) which are able to cleave said DNA target sequence from the HSV genome.

2. The variant of claim 1, wherein said variant may be obtained by a method comprising the additional steps of:

(k) selecting heterodimers from step (j) and constructing a third series of variants having at least one substitution in at least one of the monomers in said selected heterodimers,

(l) combining said third series variants of step (k) and screening the resulting heterodimers for altered cleavage activity against said DNA target from the HSV genome.

3. The variant of claim 2, wherein in said step (k) said at least one substitution are introduced by site directed mutagenesis in a DNA molecule encoding said third series of variants, and/or by random mutagenesis in a DNA molecule encoding said third series of variants.

4. The variants of claim 2 or 3, wherein steps (k) and (l) are repeated at least two times and wherein the heterodimers selected in step (k) of each further iteration are selected from heterodimers screened in step (l) of the previous iteration which showed increased cleavage activity against said DNA target from the HSV genome.

5. The variant of any one of claims 1 to 4, wherein said substitution(s) in the subdomain situated from positions 44 to 77 of I-*Cre*I are in positions 44, 68, 70, 75 and/or 77.

6. The variant of any one of claims 1 to 5, wherein said substitution(s) in the subdomain situated from positions 26 to 40 of I-*Cre*I are in positions 26, 28, 30, 32, 33, 38 and/or 40.

7. The variant of any one of claims 1 to 6, which comprises one or more substitutions in positions 137 to 143 of I-*Cre*I that modify the specificity of the

variant towards the nucleotide in positions ± 1 to 2, ± 6 to 7 and/or ± 11 to 12 of the target site in the HSV genome.

8. The variant of any one of claims 1 to 7, which comprises one or more substitutions on the entire I-*Cre*I sequence that improve the binding and/or the cleavage properties of the variant towards said DNA target sequence from the HSV genome.

9. The variant according to any one of claims 1 to 8, wherein said DNA target is located in an element of the HSV genome, selected from the group: an essential gene, a regulatory sequence for an essential gene, an essential structural sequence.

10. The variant of any one of claims 1 to 9, wherein said substitutions are replacement of the initial amino acids with amino acids selected in the group consisting of A, D, E, F, G, H, I, K, M, N, P, Q, R, S, T , Y, C, W, L and V.

11. The variant of any one of claims 1 to 10, which is a heterodimer, resulting from the association of a first and a second monomer having different mutations in positions 26 to 40 and 44 to 77 of I-*Cre*I, said heterodimer being able to cleave a non-palindromic DNA target sequence from the HSV genome.

12. The variant of claim 11, which is an obligate heterodimer, wherein the first and the second monomer, respectively, further comprises the D137R mutation and the R51D mutation.

13. The variant of claim 11, which is an obligate heterodimer, wherein the first monomer further comprises the K7R, E8R, E61R, K96R and L97F or K7R, E8R, F54W, E61R, K96R and L97F mutations and the second monomer further comprises the K7E, F54G, L58M and K96E or K7E, F54G, K57M and K96E mutations.

14. The variant according to any one of claim 1 to 13, wherein said variant consists of a single polypeptide chain comprising two monomers or core domains of one or two variant(s) of anyone of claims 1 to 13 or a combination of both.

15. The variant of claim 14 which comprises the first and the second monomer as defined in anyone of claims 1 to 13, connected by a peptide linker.

16. The variant of anyone of claims 1 to 15, wherein said HSV is Herpes Simplex Virus Type 1.

17. The variant of anyone of claims 1 to 16 wherein said DNA target is selected from the group consisting of the SEQ ID NO: 8 to 13 and 17 to 24.

18. The variant according to claim 17, wherein said variant is selected from the group consisting of SEQ ID NO: 25 to 36, 40 to 90, 93 to 151, 153 to 168, 171 to 246, 249 to 261, 267 to 273, 275 to 288, 290 to 433, 436 to 463 and 470 to 471.

19. A polynucleotide fragment encoding the variant of anyone of claims 1 to 18.

20. An expression vector comprising at least one polynucleotide fragment of claim 19.

21. The vector of claim 20, which includes a targeting construct comprising a sequence to be introduced flanked by sequences sharing homologies with the regions surrounding said DNA target sequence from the HSV genome.

22. The vector of claim 21, wherein said sequence to be introduced is a sequence which inactivates the HSV genome.

23. The vector of claim 22, wherein the sequence which inactivates the HSV genome comprises in the 5' to 3' orientation: a first transcription termination sequence and a marker cassette including a promoter, the marker open reading frame and a second transcription termination sequence, and said sequence interrupts the transcription of the coding sequence.

24. The vector of anyone of claims 20 to 23, wherein said sequence sharing homologies with the regions surrounding DNA target sequence from the HSV genome is a fragment of the HSV genome comprising sequences upstream and downstream of the cleavage site, so as to allow the deletion of coding sequences flanking the cleavage site.

25. A host cell which is modified by a polynucleotide of claim 19 or a vector of anyone of claims 20 to 24.

26. A non-human transgenic animal which is modified by a polynucleotide of claim 19 or a vector of anyone of claims 20 to 24.

27. A transgenic plant which is modified by a polynucleotide of claim 19 or a vector of anyone of claims 20 to 24.

28. Use of at least one variant of anyone of claims 1 to 18, or at least one vector according to anyone of claims 20 to 24, for genome engineering, for non-therapeutic purposes.

29. Use of a variant according to any one of claims 1 to 18, a nucleic acid molecule according to claim 19 or a vector according to any one of claims 20 to 24 to prepare a medicament to treat an HSV infection.

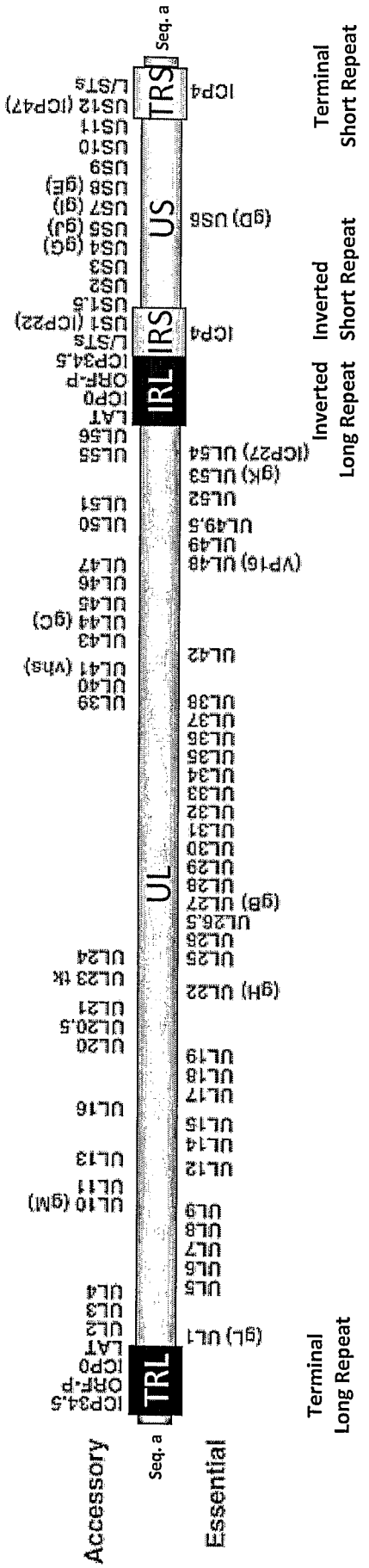

Seq. a
TRS
IRS
IRL
TRL
US
UL
Seq. a
Accessory
Essential
Terminal
Long Repeat
Inverted
Long Repeat
Inverted
Short Repeat
Terminal
Short Repeat
ICP4
US12 (ICP47)
US11
US10
US9
US8 (gE)
US7 (gI)
US6 (gJ)
US4 (gG)
US3
US2
US1.5
US1 (ICP22)
L/STs
ICP34.5
ORFs.p
ICP0
LAT
UL56
UL55
(ICP27) UL54
(gK) UL53
UL52
UL51
UL50
UL49.5
UL49
(VP16) UL48
UL47
UL46
UL45
UL44 (gC)
UL43
UL42
UL41 (vhs)
UL40
UL39
UL38
UL37
UL36
UL35
UL34
UL33
UL32
UL31
UL30
UL29
UL28
(gB) UL27
UL26.5
UL26
UL25
UL24
UL23 tk
(gH) UL22
UL21
UL20.5
UL20
UL19
UL18
UL17
UL16
UL15
UL14
UL13
UL12
UL10 (gM)
UL11
UL9
UL8
UL7
UL6
UL5
UL4
UL3
UL2
(gL) UL1
ICP34.5
ORFs.p
ICP0
LAT
UL3
UL4
UL5
UL6
UL7
UL8
UL9
UL10 (gM)
UL11
UL12
UL13
UL14
UL15
UL16
UL17
UL18
UL19
UL20
UL20.5
UL21
UL23 tk
UL24
UL25
UL26
UL26.5
(gB) UL27
UL28
UL29
UL30
UL31
UL32
UL33
UL34
UL35
UL36
UL37
UL38
UL39
UL40
UL41 (vhs)
UL42
UL43
UL44 (gC)
UL45
UL46
UL47
(VP16) UL48
UL49
UL49.5
UL50
UL51
UL52
(gK) UL53
(ICP27) UL54
UL55
UL56
LAT
ICP0
ORFs.p
ICP34.5
L/STs
US1 (ICP22)
US1.5
US2
US3
US4 (gG)
US6 (gJ)
US7 (gI)
US8 (gE)
US9
US10
US11
US12 (ICP47)
L/STs
ICP4

Figure 1

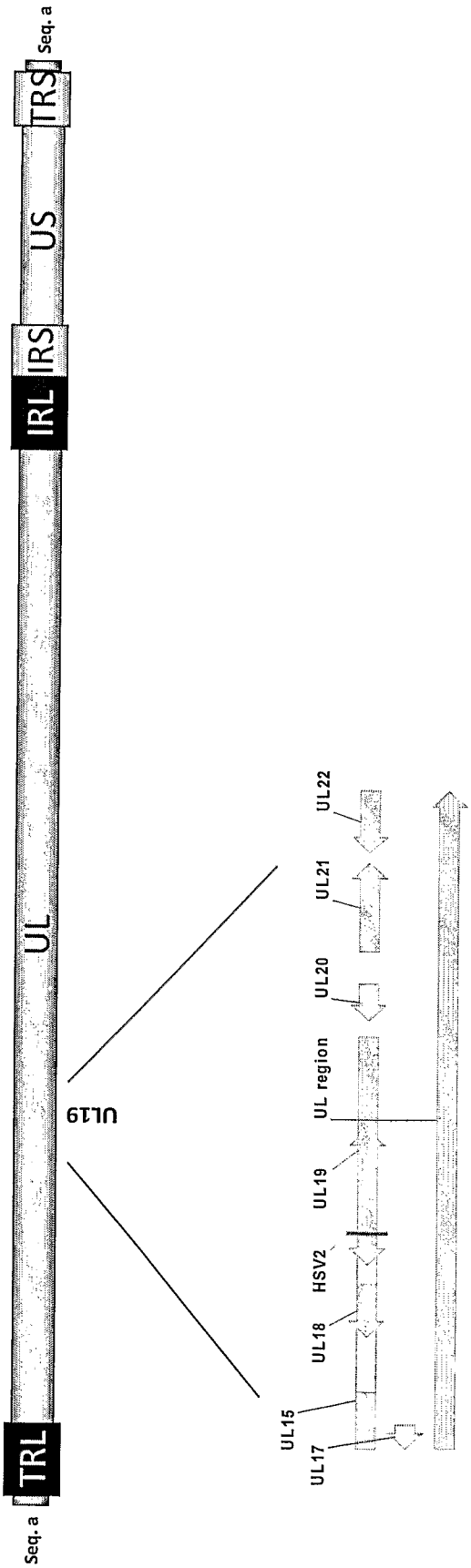

Seq. a
TRS
US
IRS
IRL
UL
UL19
TRL
Seq. a
UL22
UL21
UL20
UL region
UL19
HSV2
UL18
UL15
UL17
NC 001806 UL19
119941 bp

Figure 2

C1221	(T) C-AAA-AC-GTC-GTAC-GAC-GT-TTT-G (A)	(SEQ ID NO:2)
10AAA_P	(T) C-AAA-AC-GTC-GTAC-GAC-GT-TTT-G (A)	(SEQ ID NO:4)
10AGG_P	(T) C-AGG-AC-GTC-GTAC-GAC-GT-CCT-G (A)	(SEQ ID NO:5)
5CAC_P	(T) C-AAA-AC-CAC-GTAC-GTG-GT-TTT-G (A)	(SEQ ID NO:6)
5GCC_P	(T) C-AAA-AC-GCC-GTAC-GGC-GT-TTT-G (A)	(SEQ ID NO:7)
HSV_2.1	(A) T-AAA-CT-CAC-ACAC-GGC-GT-CCT-G (G)	(SEQ ID NO:8)
HSV_2.2	(A) T-AAA-CT-CAC-GTAC-GGC-GT-CCT-G (G)	(SEQ ID NO:9)
HSV_2.3	(A) T-AAA-CT-CAC-GTAC-GTG-AG-TTT-A (T)	(SEQ ID NO:10)
HSV_2.4	(C) C-AGG-AC-GCC-GTAC-GGC-GT-CCT-G (G)	(SEQ ID NO:11)
HSV_2.5	(A) T-AAA-CT-CAC-ACAC-GTG-AG-TTT-A (T)	(SEQ ID NO:12)
HSV_2.6	(C) C-AGG-AC-GCC-ACAC-GGC-GT-CCT-G (G)	(SEQ ID NO:13)

Figure 3

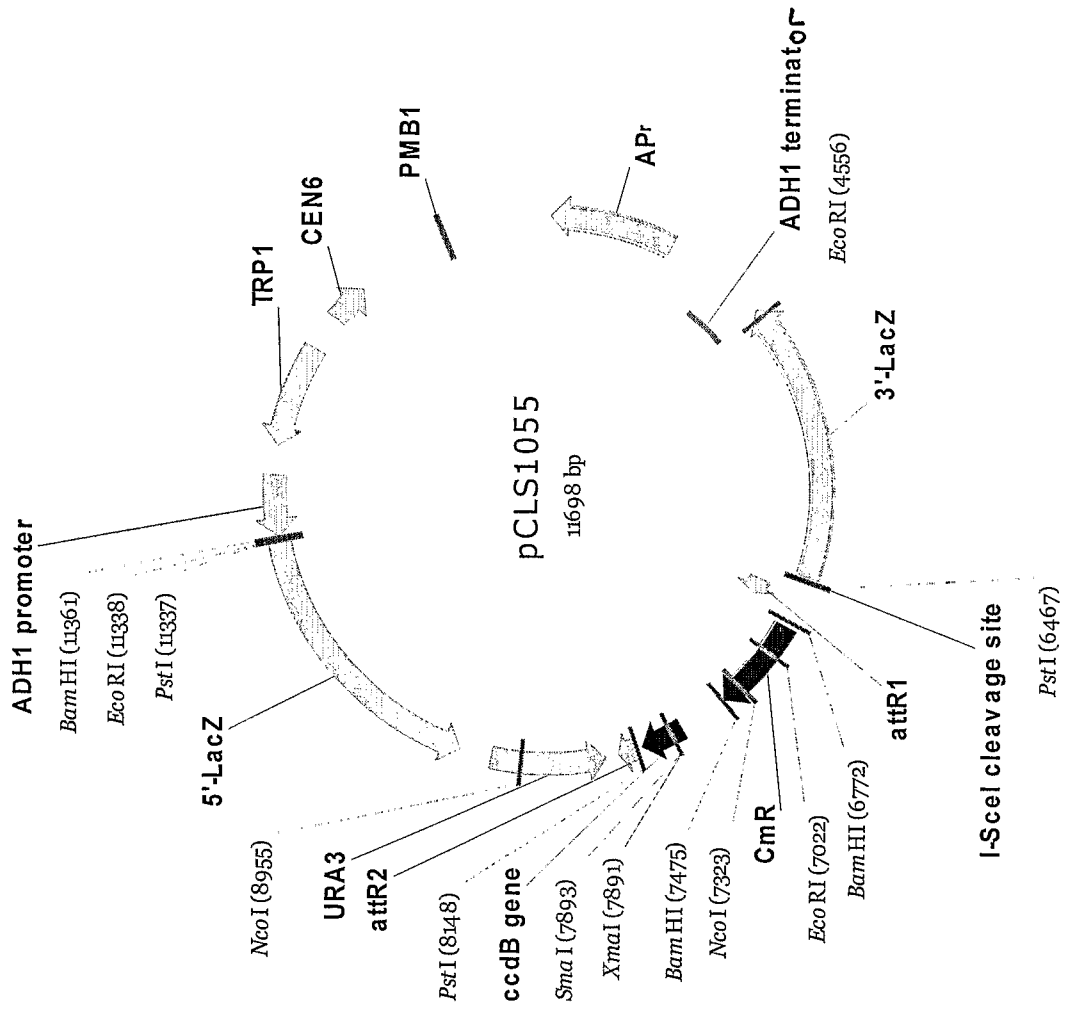

pCLS1055
11698 bp
TRP1
CEN6
PMB1
APr
ADH1 terminator
Eco RI (4556)
3'-LacZ
ADH1 promoter
Bam HI (11361)
Eco RI (11338)
Pst I (11337)
5'-LacZ
Nco I (8955)
URA3
attR2
Pst I (8348)
ccdB gene
Sma I (7893)
Xma I (7891)
Bam HI (7475)
Nco I (7323)
CmR
Eco RI (7022)
Bam HI (6772)
attR1
I-SceI cleavage site
Pst I (6467)

Figure 4

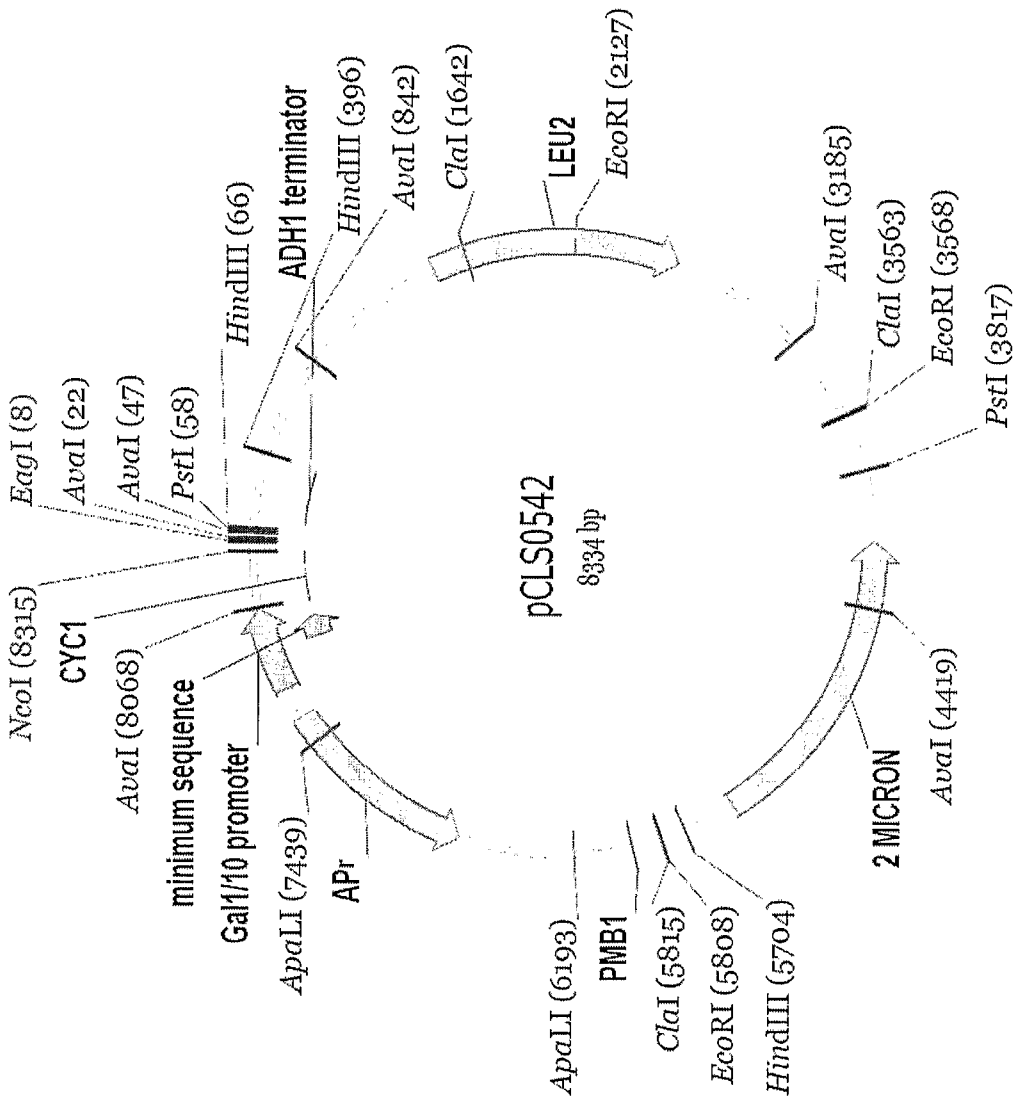

pCLS0542
8334 bp
EagI (8)
AvaI (22)
AvaI (47)
PstI (58)
HindIII (66)
ADH1 terminator
HindIII (396)
AvaI (842)
ClaI (1642)
LEU2
EcoRI (2127)
AvaI (3185)
ClaI (3563)
EcoRI (3568)
PstI (3817)
AvaI (4419)
2 MICRON
HindIII (5704)
EcoRI (5808)
ClaI (5815)
PMB1
ApaLI (6193)
ApaLI (7439)
APr
Gal1/10 promoter
minimum sequence
AvaI (8068)
CYC1
NcoI (8315)

Figure 5

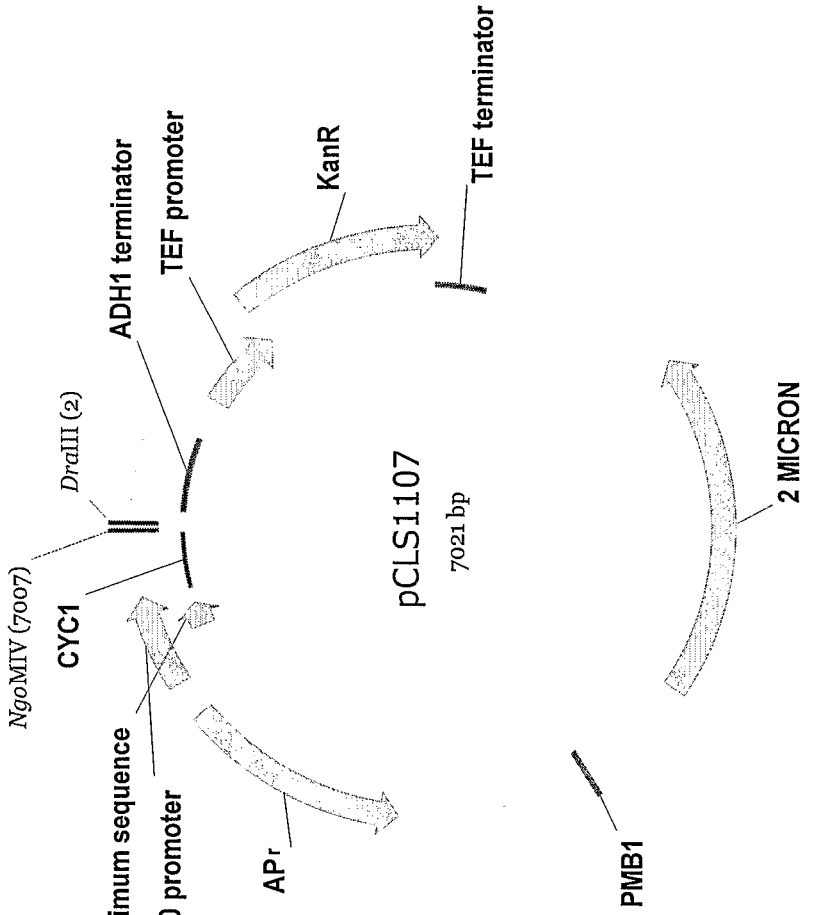

TEF terminator
KanR
TEF promoter
ADH1 terminator
DraIII (2)
NgoMIV (7007)
CYC1
minimum sequence
Gal1/10 promoter
APr
pCLS1107
7021 bp
PMB1
2 MICRON

Figure 6

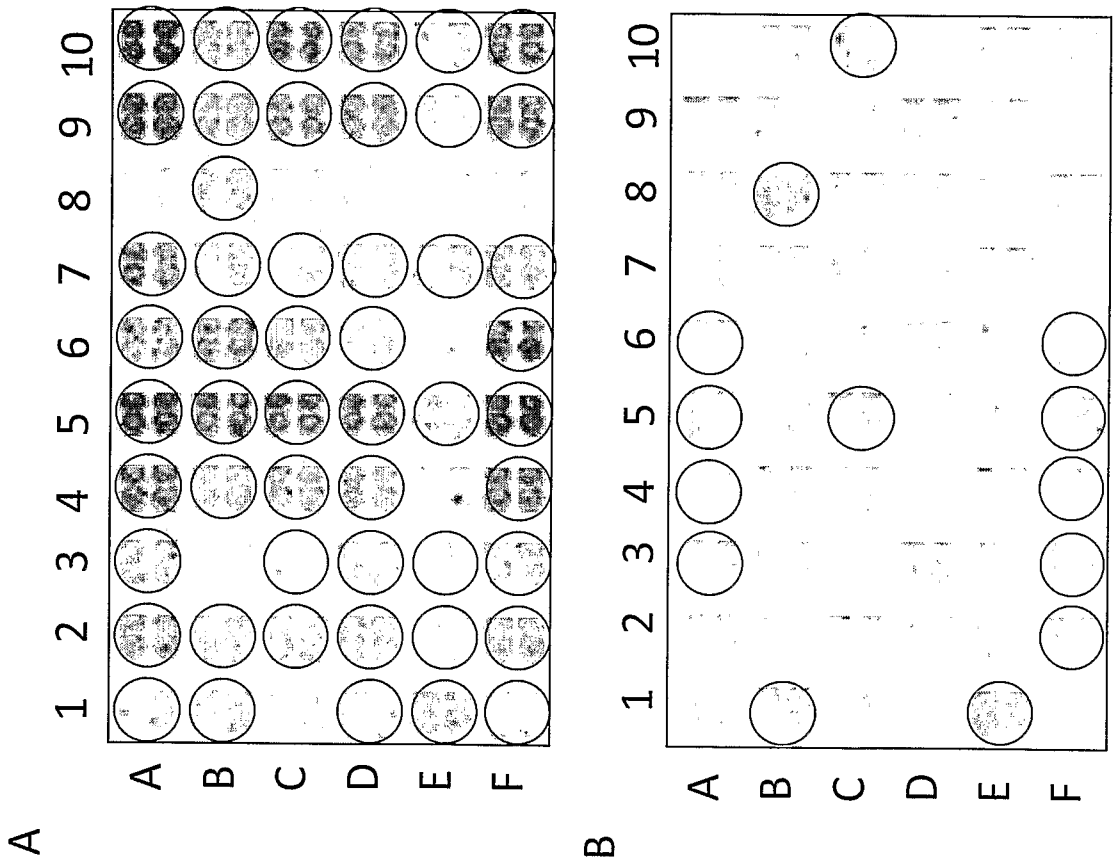


Figure 7

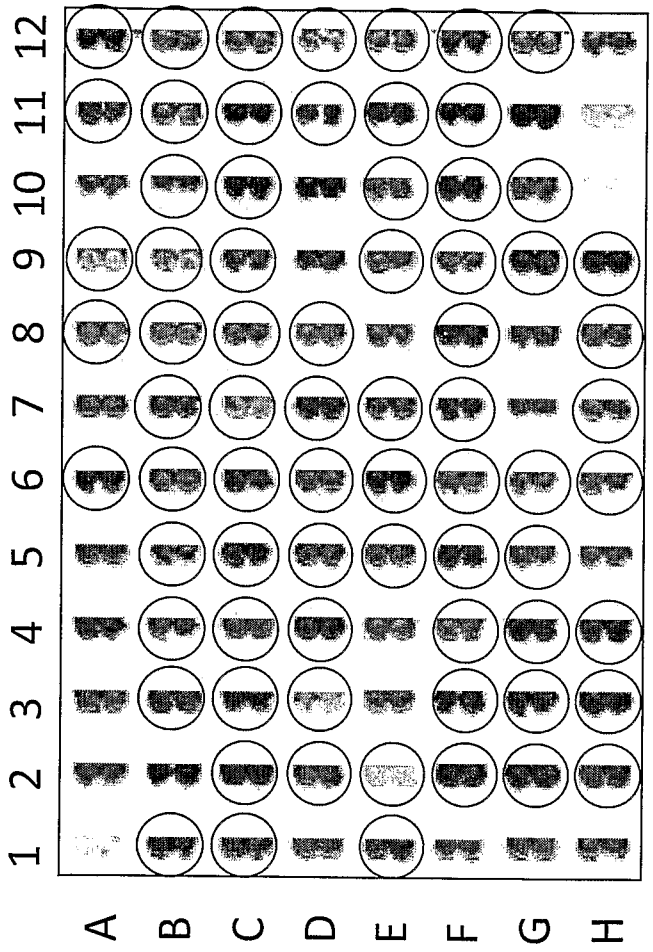


Figure 8

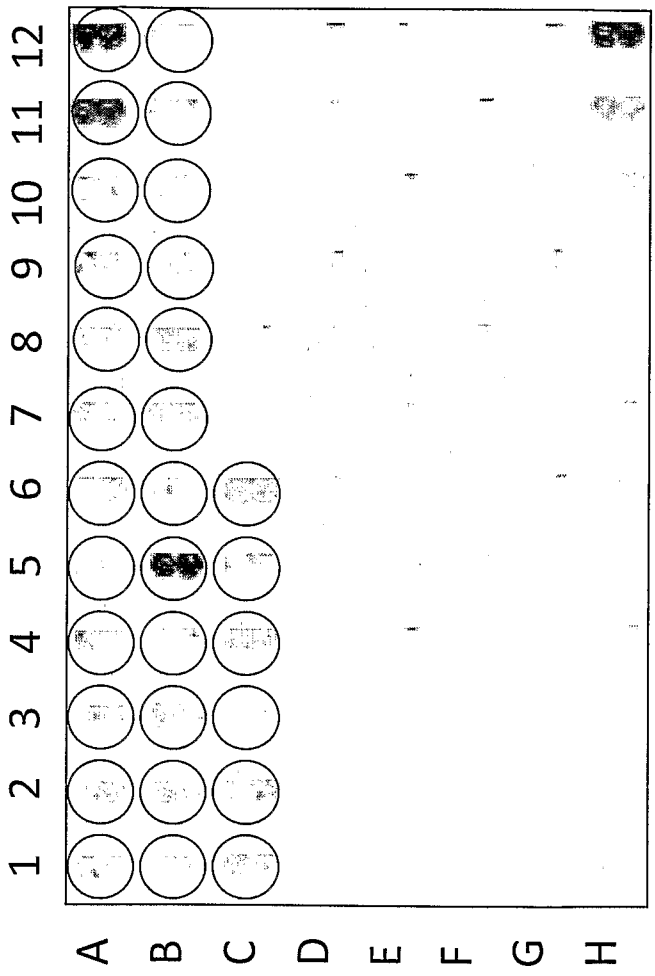


Figure 9

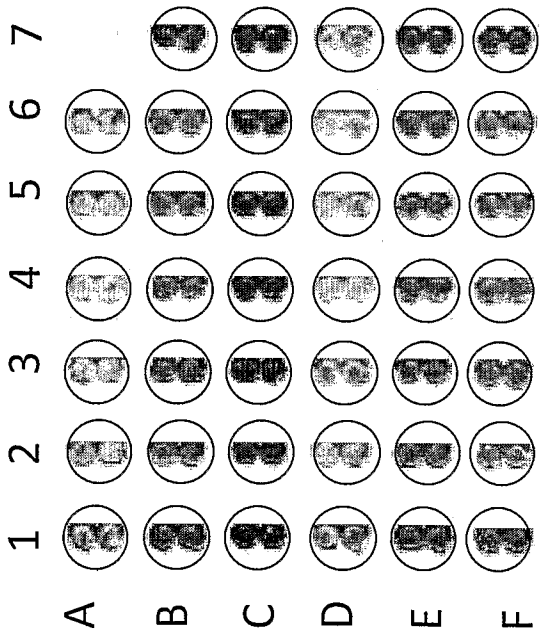


Figure 10

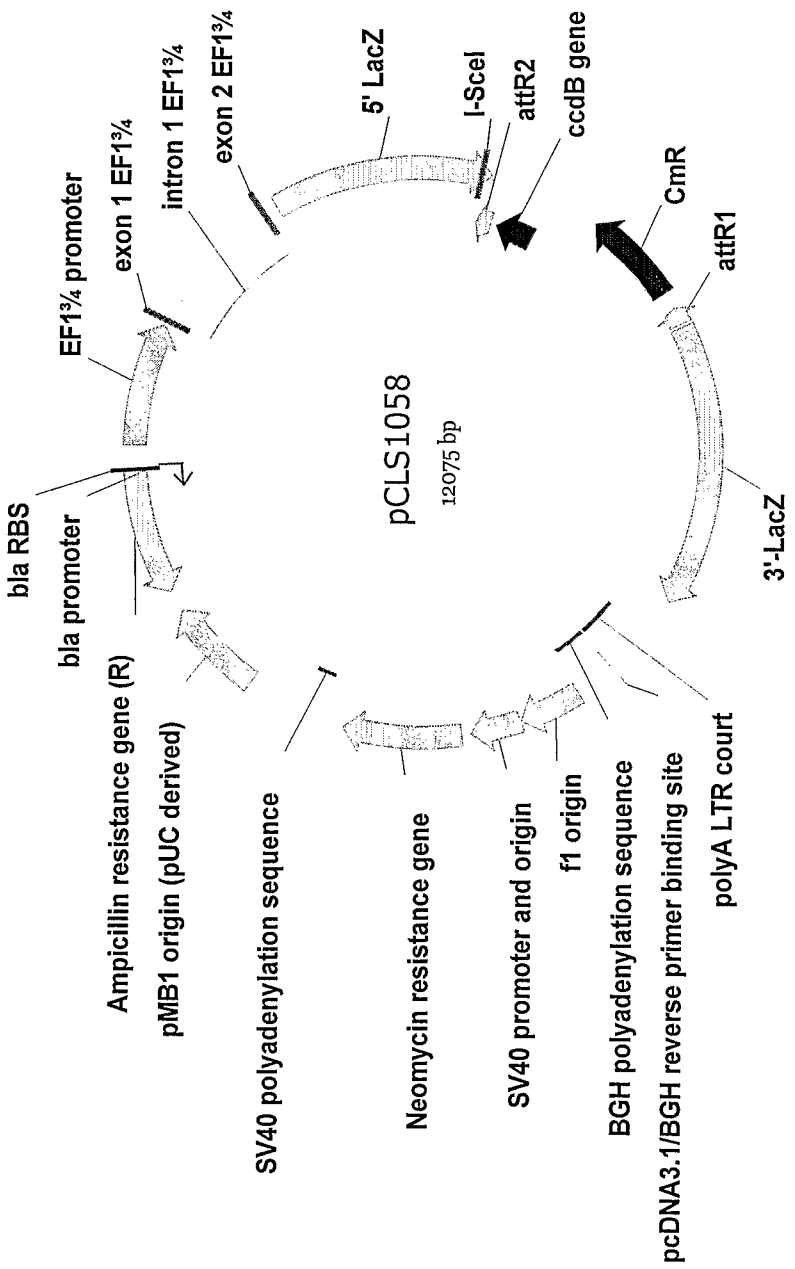

pCLS1058
12075 bp
bla RBS
bla promoter
Ampicillin resistance gene (R)
pMB1 origin (pUC derived)
SV40 polyadenylation sequence
Neomycin resistance gene
SV40 promoter and origin
f1 origin
BGH polyadenylation sequence
pcDNA3.1/BGH reverse primer binding site
polyA LTR court
3'-LacZ
attR1
CmR
ccdB gene
attR2
I-SceI
5' LacZ
exon 2 EF1¾
intron 1 EF1¾
exon 1 EF1¾
EF1¾ promoter

Figure 11

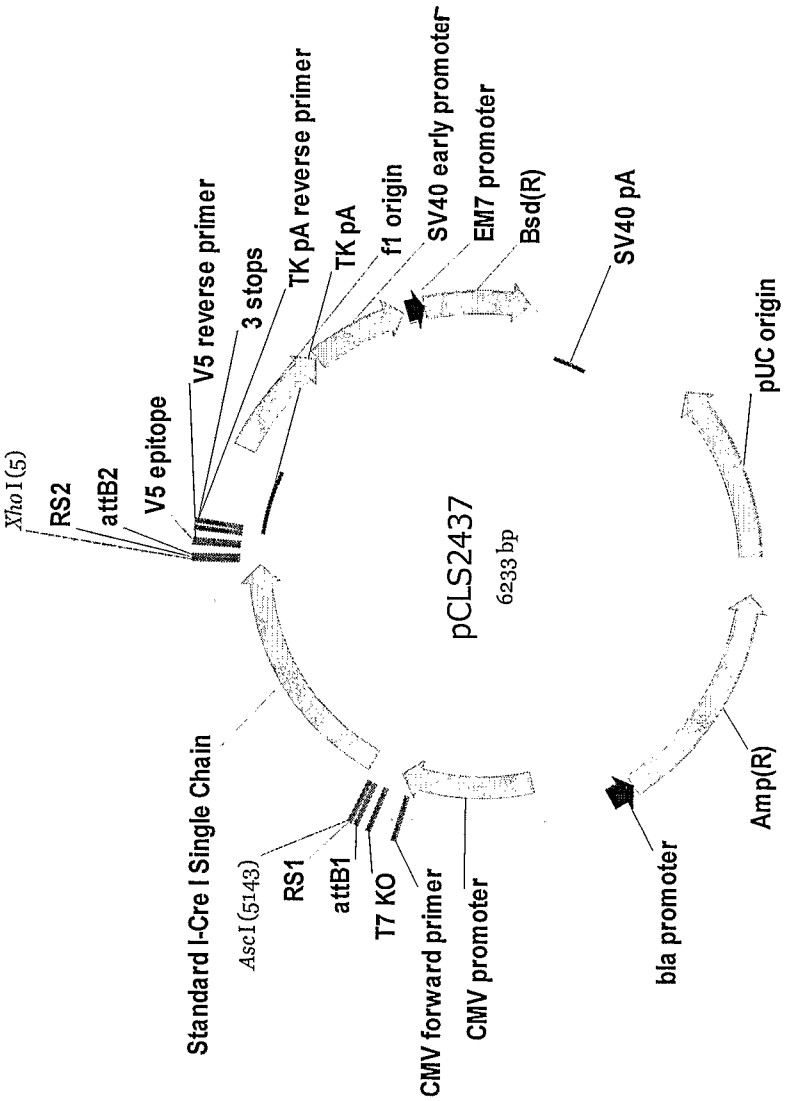
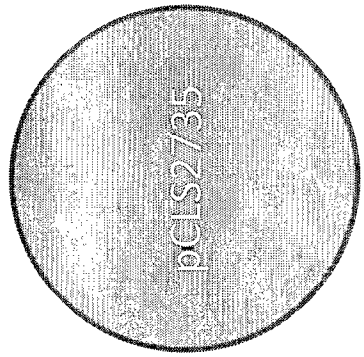

pCLS2437
6233 bp
XhoI(5)
RS2
attB2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
Amp(R)
bla promoter
CMV forward primer
CMV promoter
T7 KO
attB1
RS1
AscI(5143)
Standard I-Cre I Single Chain

Figure 12



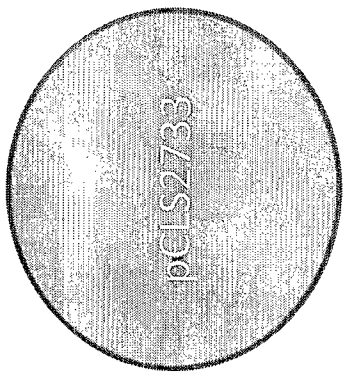


Figure 13

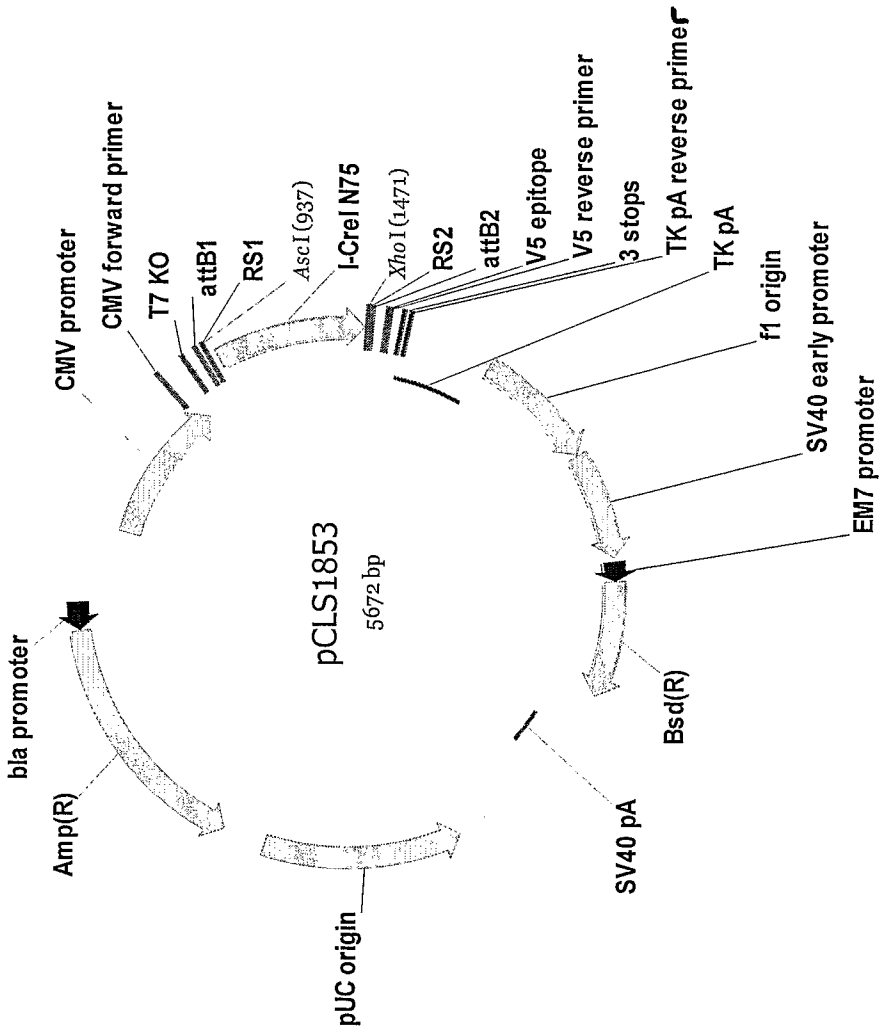

CMV promoter
CMV forward primer
T7 KO
attB1
RS1
AscI (937)
I-CreI N75
XhoI (1471)
RS2
attB2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
Amp(R)
bla promoter
pCLS1853
5672 bp

Figure 14

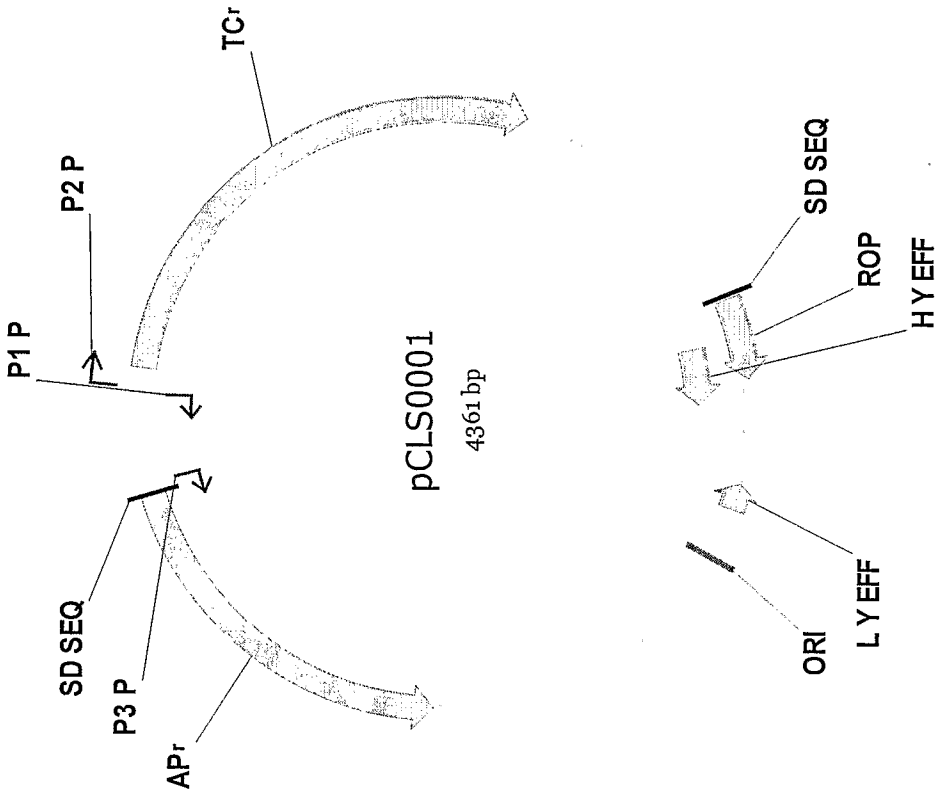

TCr
P2 P
P1 P
pCLS0001
4361 bp
SD SEQ
P3 P
APr
SD SEQ
ROP
H Y EFF
ORI
L Y EFF

Figure 15

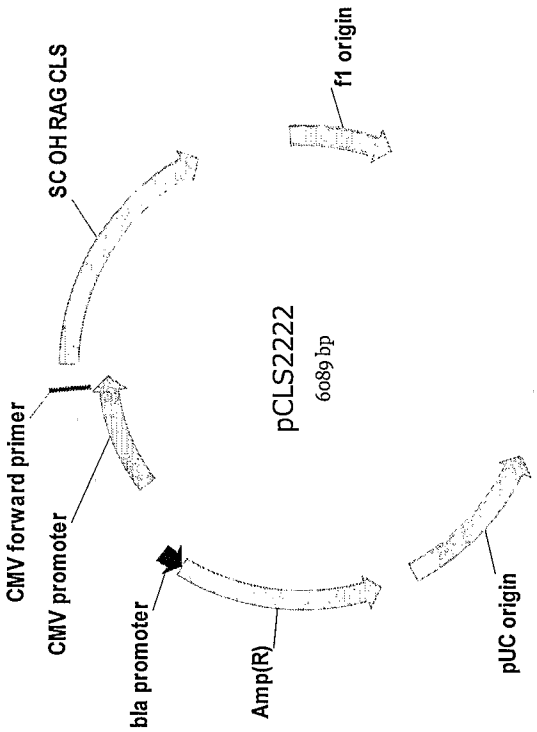

SC OH RAG CLS
f1 origin
CMV forward primer
CMV promoter
pCLS2222
6089 bp
bla promoter
Amp(R)
pUC origin

Figure 16

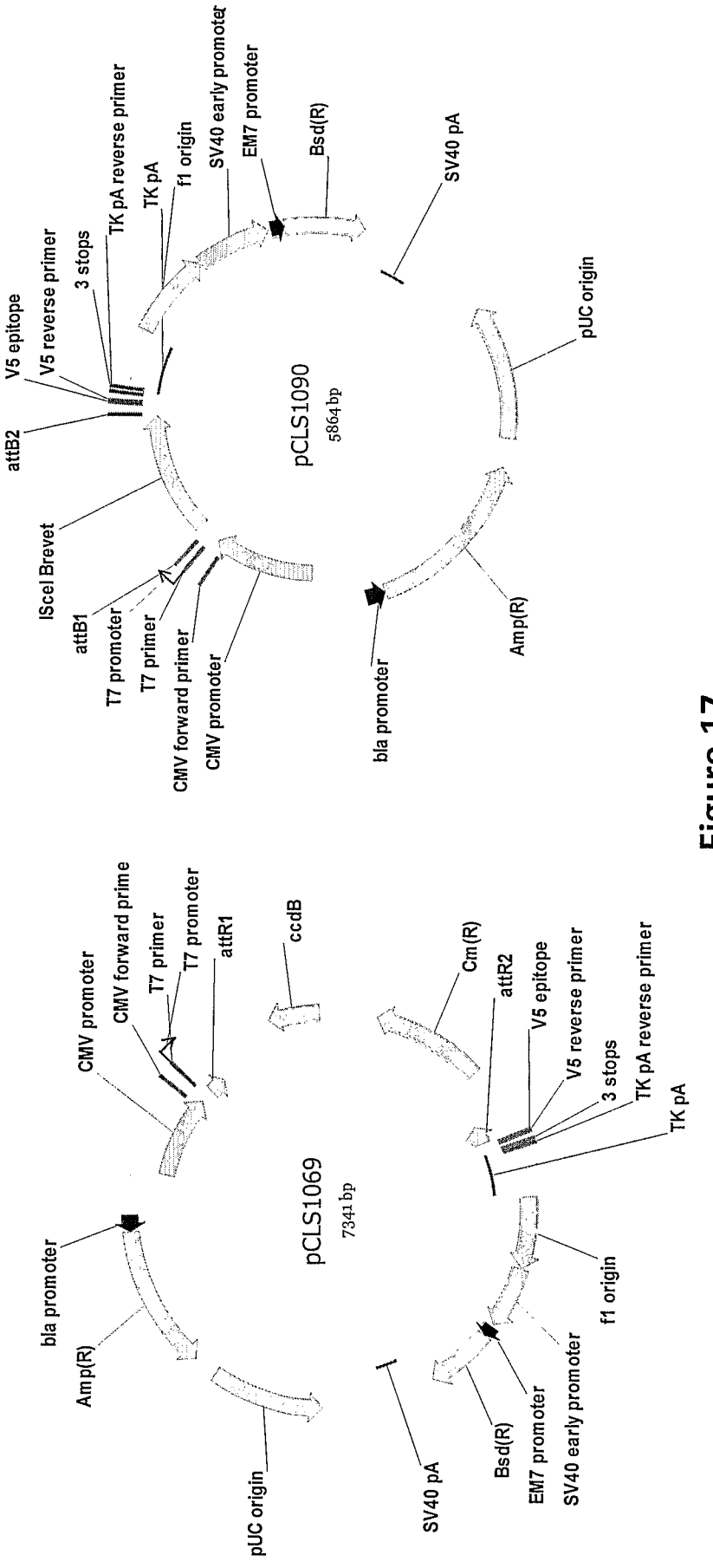

pCLS1090
5864 bp
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
Amp(R)
bla promoter
ISceI Brevet
attB1
T7 promoter
T7 primer
CMV forward primer
CMV promoter
attB2
pCLS1069
7341 bp
CMV promoter
CMV forward prime
T7 primer
T7 promoter
attR1
ccdB
Cm(R)
attR2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
Amp(R)
bla promoter

Figure 17

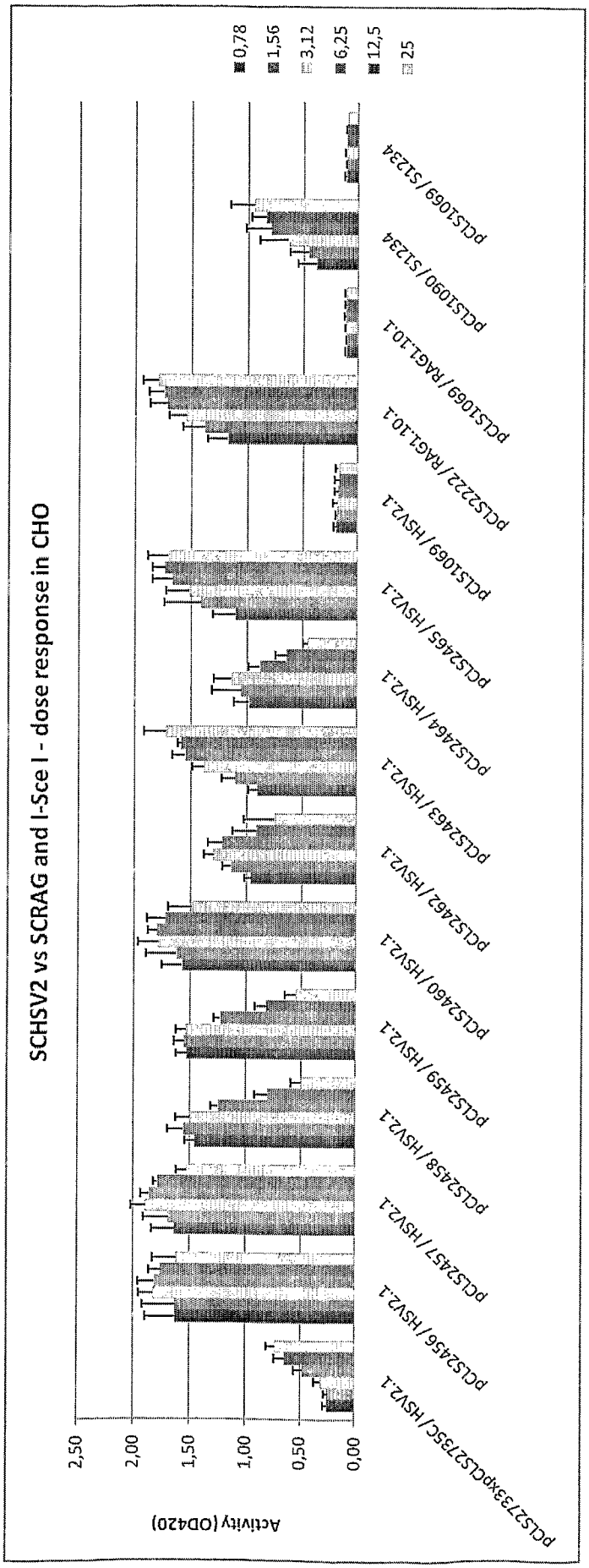

SCHSV2 vs SCRAG and I-Sce I - dose response in CHO
0,78
1,56
3,12
6,25
12,5
25
Activity (OD420)
2,50
2,00
1,50
1,00
0,50
0,00
pCLS2733xpCLS2735C / HSV2.1
pCLS2456 / HSV2.1
pCLS2457 / HSV2.1
pCLS2458 / HSV2.1
pCLS2459 / HSV2.1
pCLS2460 / HSV2.1
pCLS2462 / HSV2.1
pCLS2463 / HSV2.1
pCLS2464 / HSV2.1
pCLS2465 / HSV2.1
pCLS1069 / HSV2.1
pCLS2222 / RAG1.10.1
pCLS1069 / RAG1.10.1
pCLS1090 / S1234
pCLS1069 / S1234

Figure 18

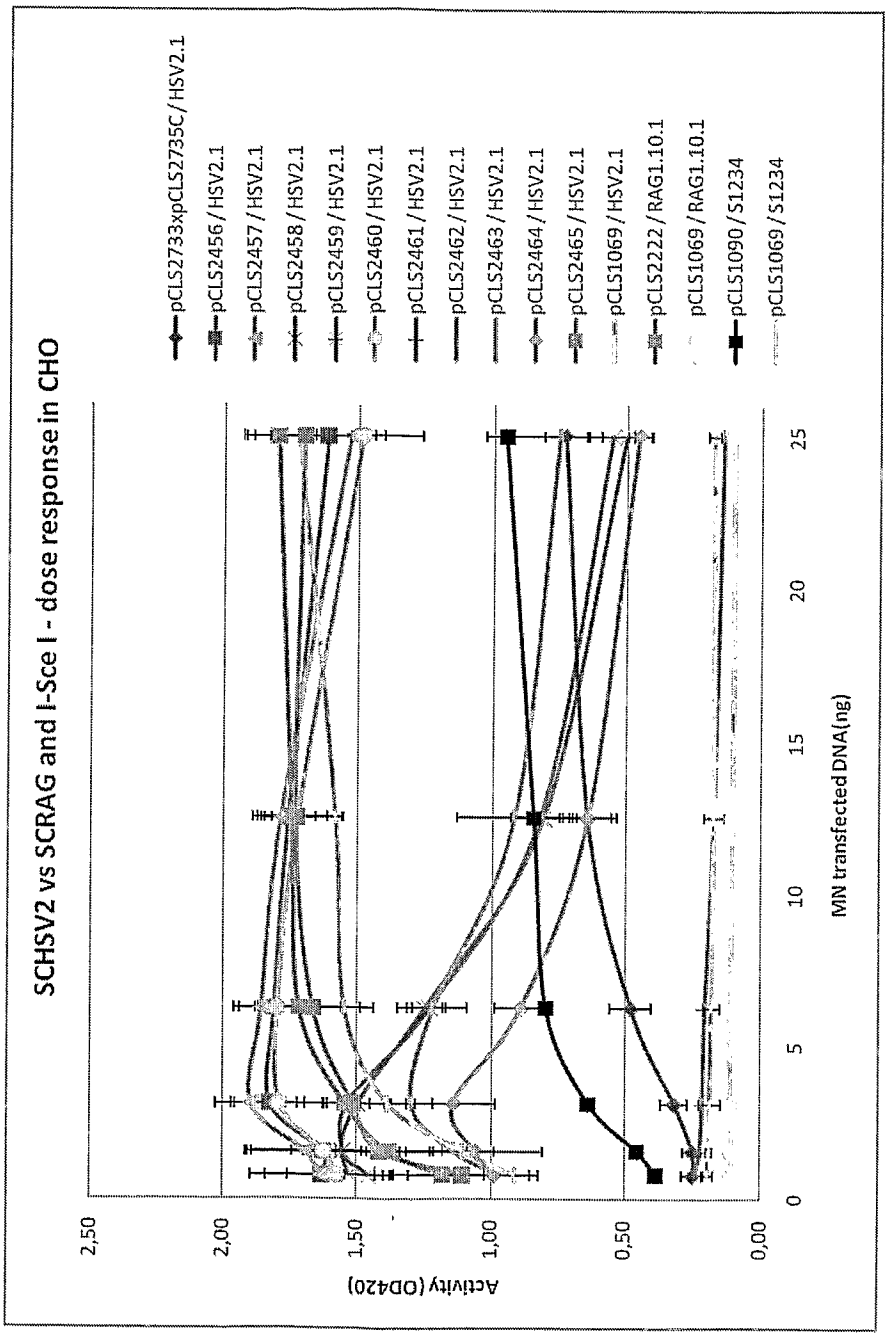

SCHSV2 vs SCRAG and I-Sce I - dose response in CHO
pCLS2733xpCLS2735C/HSV2.1
pCLS2456/HSV2.1
pCLS2457/HSV2.1
pCLS2458/HSV2.1
pCLS2459/HSV2.1
pCLS2460/HSV2.1
pCLS2461/HSV2.1
pCLS2462/HSV2.1
pCLS2463/HSV2.1
pCLS2464/HSV2.1
pCLS2465/HSV2.1
pCLS1069/HSV2.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCLS1090/S1234
pCLS1069/S1234
Activity (OD420)
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
MN transfected DNA(ng)

Figure 19

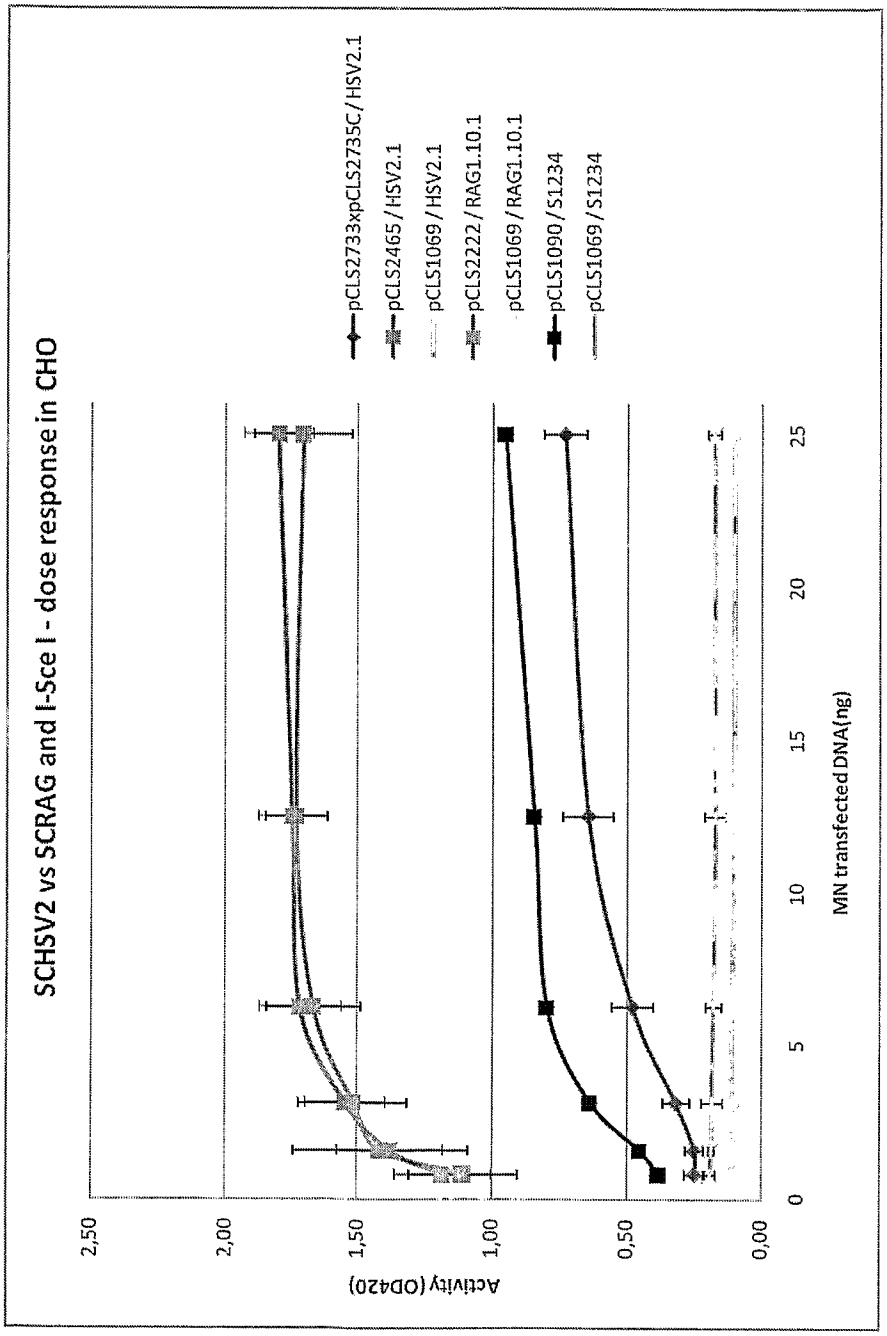

SCHSV2 vs SCRAG and I-Sce I - dose response in CHO
pCLS2733xpCLS2735C/HSV2.1
pCLS2465/HSV2.1
pCLS1069/HSV2.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCLS1090/S1234
pCLS1069/S1234
Activity (OD420)
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
MN transfected DNA(ng)

Figure 20

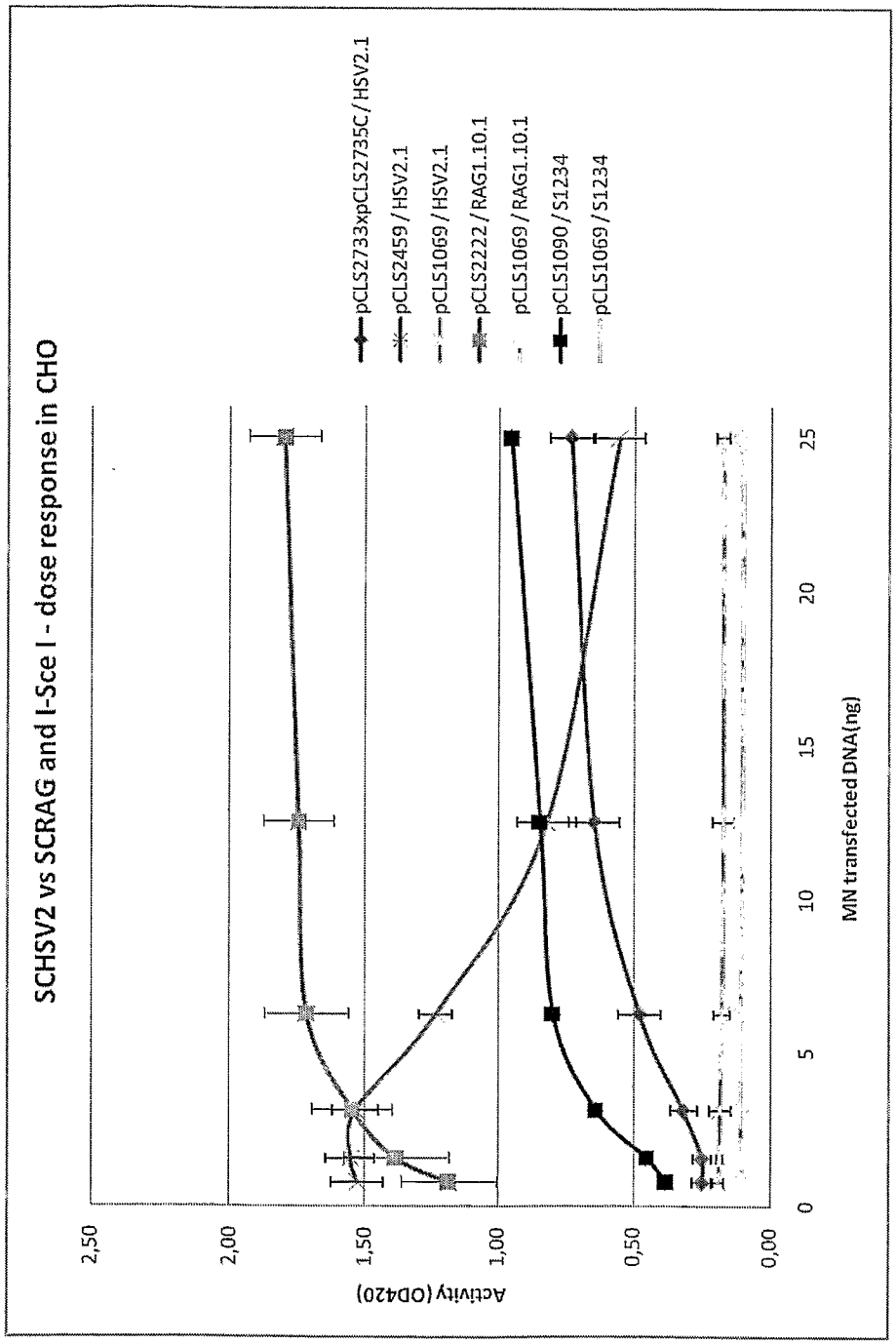

SCHSV2 vs SCRAG and I-Sce I - dose response in CHO
pCLS2733xpCLS2735C / HSV2.1
pCLS2459 / HSV2.1
pCLS1069 / HSV2.1
pCLS2222 / RAG1.10.1
pCLS1069 / RAG1.10.1
pCLS1090 / S1234
pCLS1069 / S1234
Activity (OD420)
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
MN transfected DNA(ng)

Figure 21

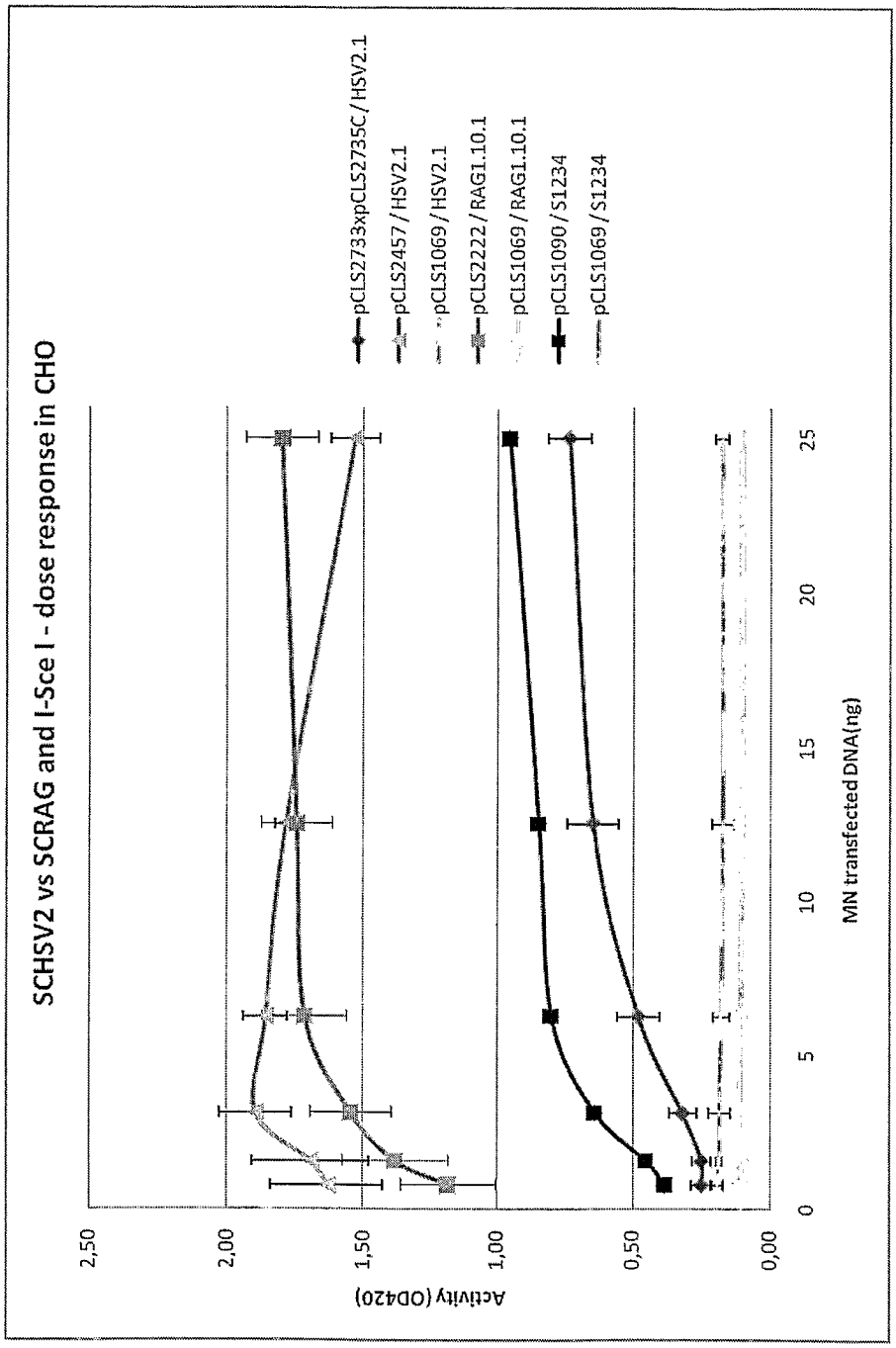

SCHSV2 vs SCRAG and I-Sce I - dose response in CHO
pCLS2733xpCLS2735C/HSV2.1
pCLS2457/HSV2.1
pCLS1069/HSV2.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCLS1090/S1234
pCLS1069/S1234
Activity (OD420)
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
MN transfected DNA(ng)

Figure 22

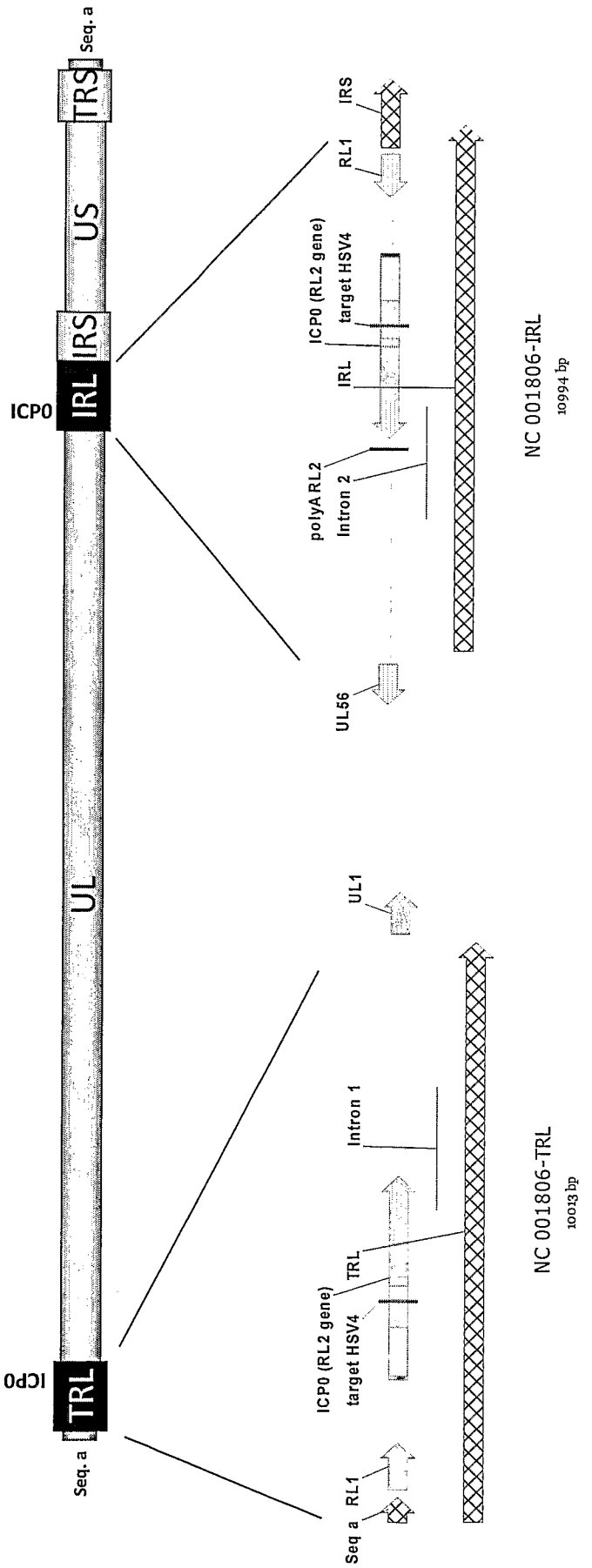

Seq. a
TRS
US
IRS
IRL
ICP0
UL
TRL
ICP0
Seq. a
IRS
RL1
ICP0 (RL2 gene)
target HSV4
IRL
polyA RL2
Intron 2
UL56
NC 001806-IRL
10994 bp
UL1
Intron 1
TRL
ICP0 (RL2 gene)
target HSV4
Seq a
RL1
NC 001806-TRL
10013 bp

Figure 23

C1221	(T) C-AAA-AC-GTC-GTAC-GAC-GT-TTT-G (A)	(SEQ ID NO:2)
10AAG_P	(T) C-AAG-AC-GTC-GTAC-GAC-GT-CTT-G (A)	(SEQ ID NO:14)
10ACT_P	(T) C-ACT-AC-GTC-GTAC-GAC-GT-AGT-G (A)	(SEQ ID NO:15)
5GGT_P	(T) C-AAA-AC-GGT-GTAC-ACC-GT-TTT-G (A)	(SEQ ID NO:37)
5CAG_P	(T) C-AAA-AC-CAG-GTAC-CTG-GT-TTT-G (A)	(SEQ ID NO:16)
HSV_4	(C) C-AAG-CT-GGT-GTAC-CTG-AT-AGT-G (G)	(SEQ ID NO:17)
HSV_4.3	(C) C-AAG-CT-GGT-GTAC-ACC-AG-CTT-G (G)	(SEQ ID NO:18)
HSV_4.4	(C) C-ACT-AT-CAG-GTAC-CTG-AT-AGT-G (G)	(SEQ ID NO:19)

Figure 24

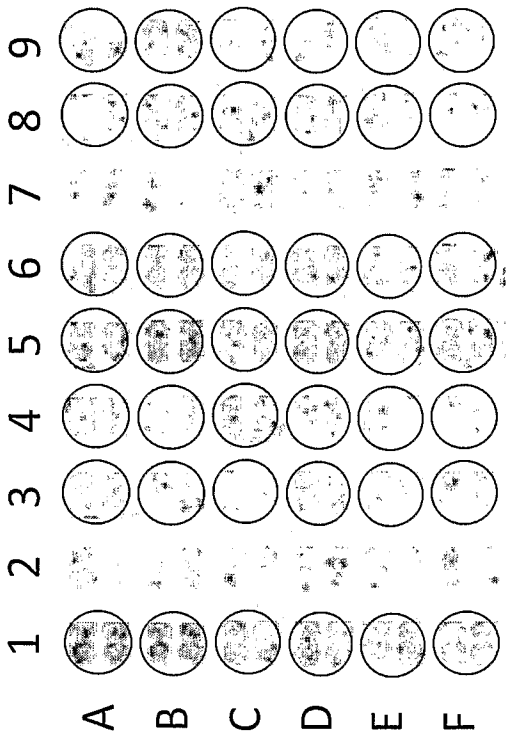


Figure 25

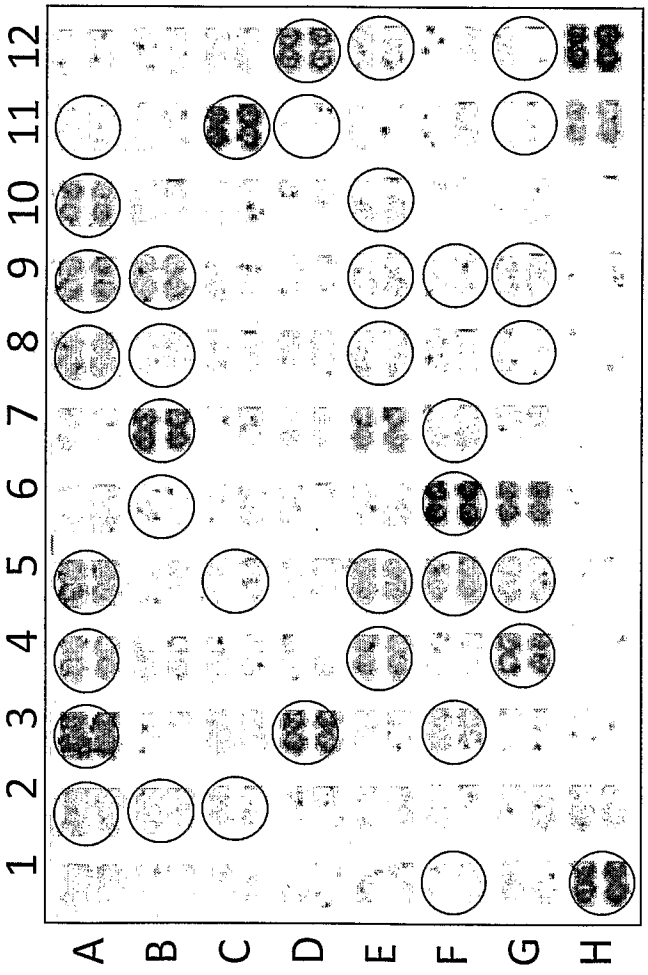


Figure 26

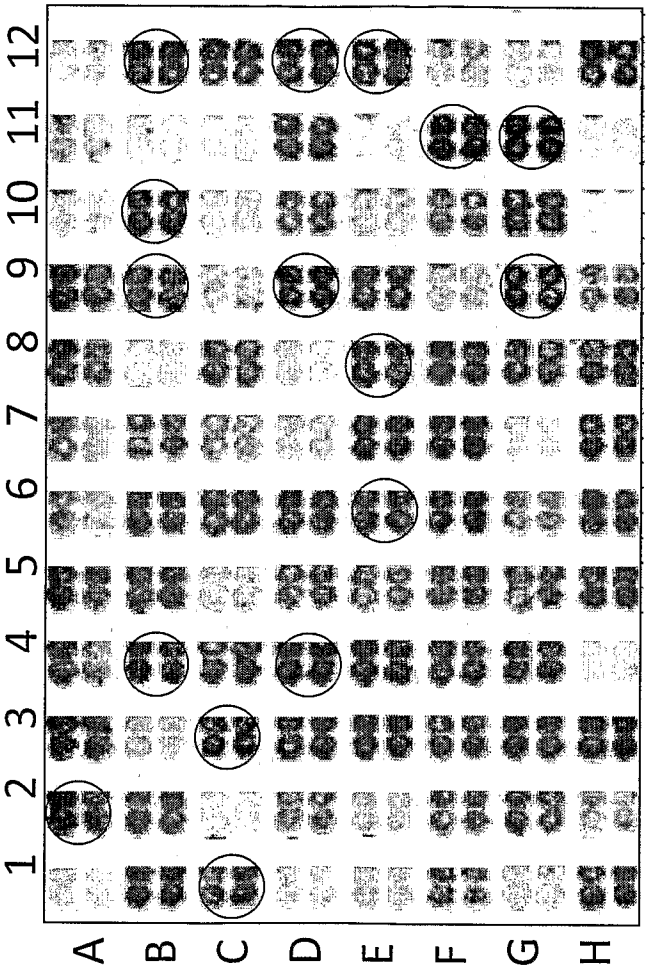


Figure 27

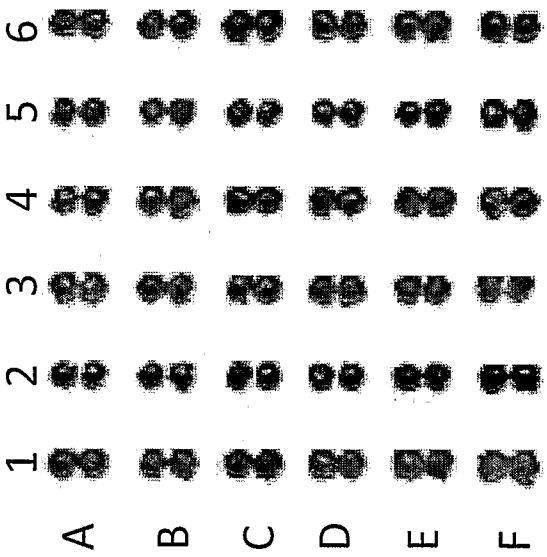


Figure 28

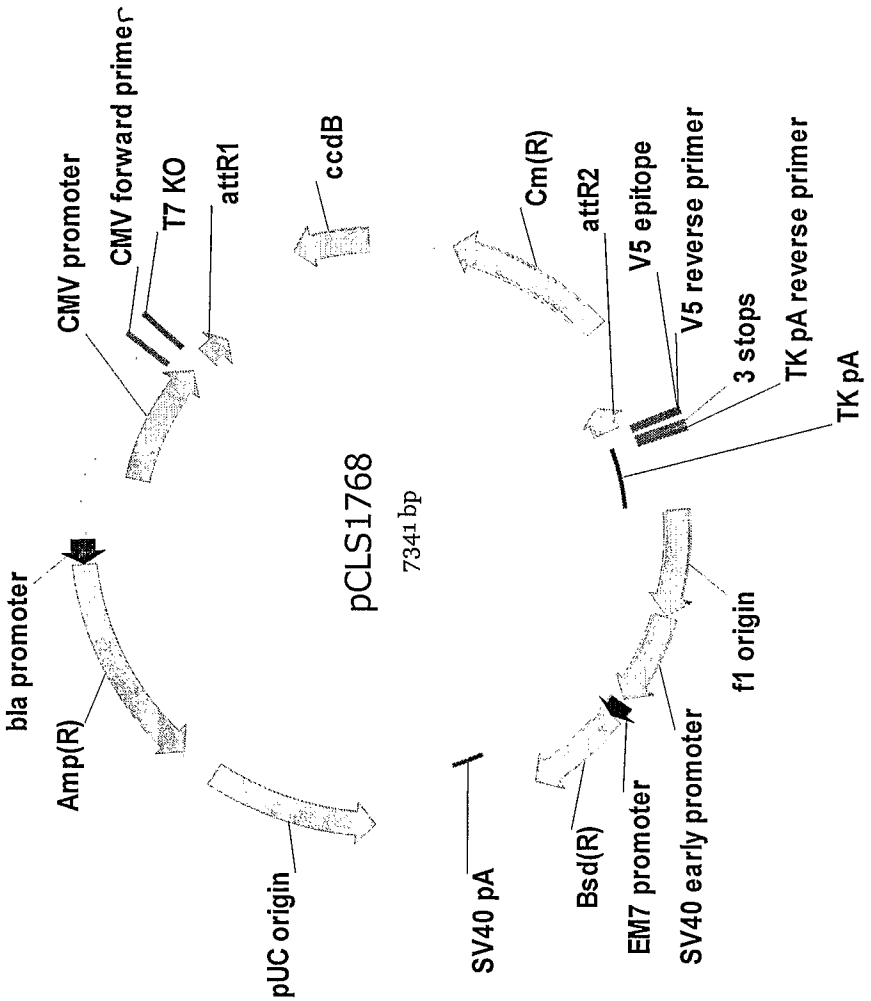

pCLS1768
7341 bp
CMV promoter
CMV forward primer
T7 KO
attR1
ccdB
Cm(R)
attR2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
Amp(R)
bla promoter

Figure 29

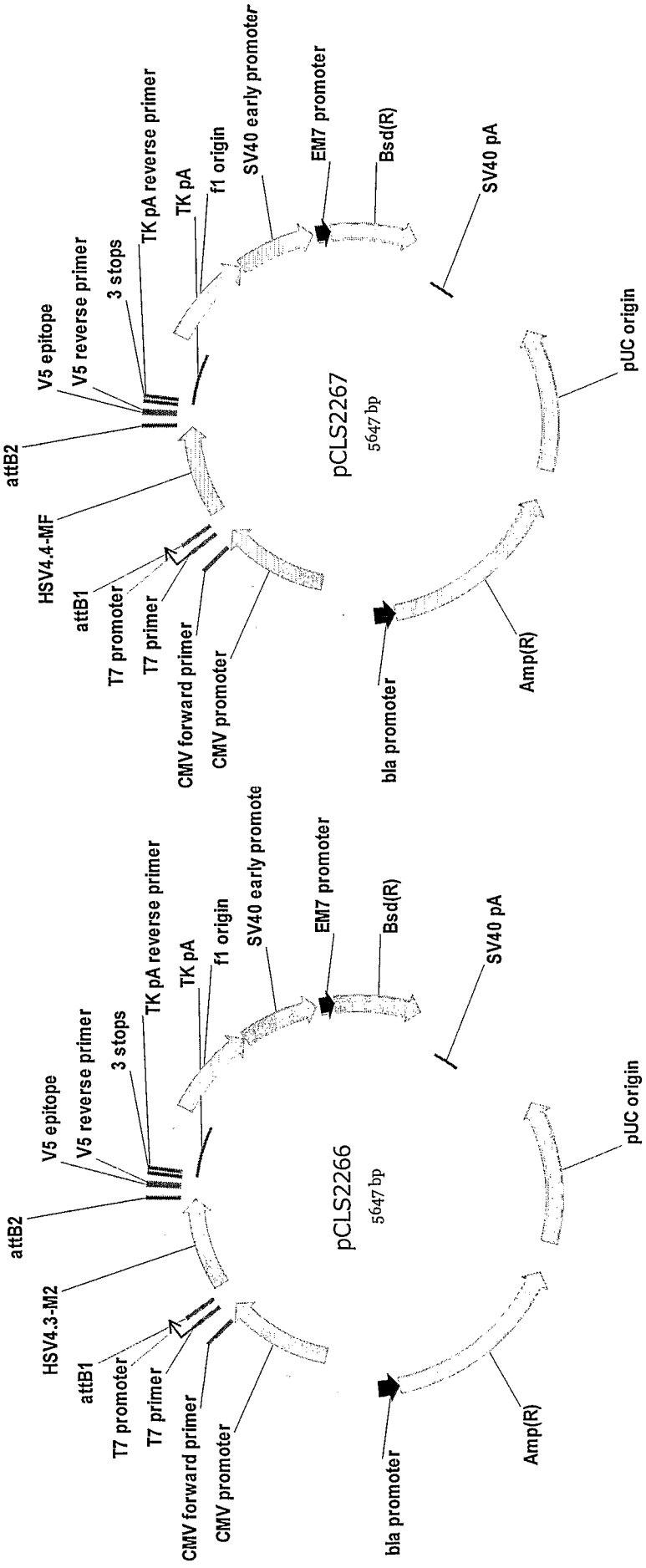

pCLS2267
5647 bp
attB2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promoter
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
HSV4.4-MF
attB1
T7 promoter
T7 primer
CMV forward primer
CMV promoter
bla promoter
Amp(R)
pCLS2266
5647 bp
attB2
V5 epitope
V5 reverse primer
3 stops
TK pA reverse primer
TK pA
f1 origin
SV40 early promote
EM7 promoter
Bsd(R)
SV40 pA
pUC origin
HSV4.3-M2
attB1
T7 promoter
T7 primer
CMV forward primer
CMV promoter
bla promoter
Amp(R)

Figure 30

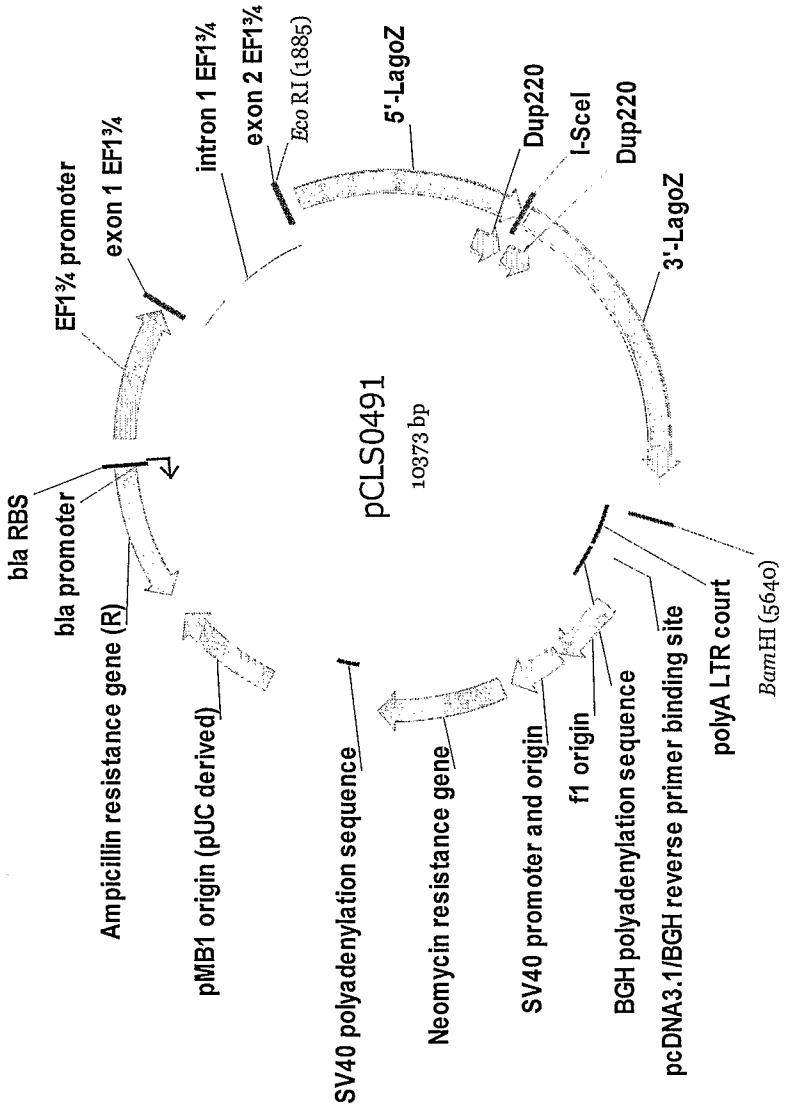

pCLS0491
10373 bp
EF1¾ promoter
exon 1 EF1¾
intron 1 EF1¾
exon 2 EF1¾
EcoRI (1885)
5'-LagoZ
Dup220
I-SceI
Dup220
3'-LagoZ
BamHI (5640)
polyA LTR court
pcDNA3.1/BGH reverse primer binding site
BGH polyadenylation sequence
f1 origin
SV40 promoter and origin
Neomycin resistance gene
SV40 polyadenylation sequence
pMB1 origin (pUC derived)
Ampicillin resistance gene (R)
bla promoter
bla RBS

Figure 31

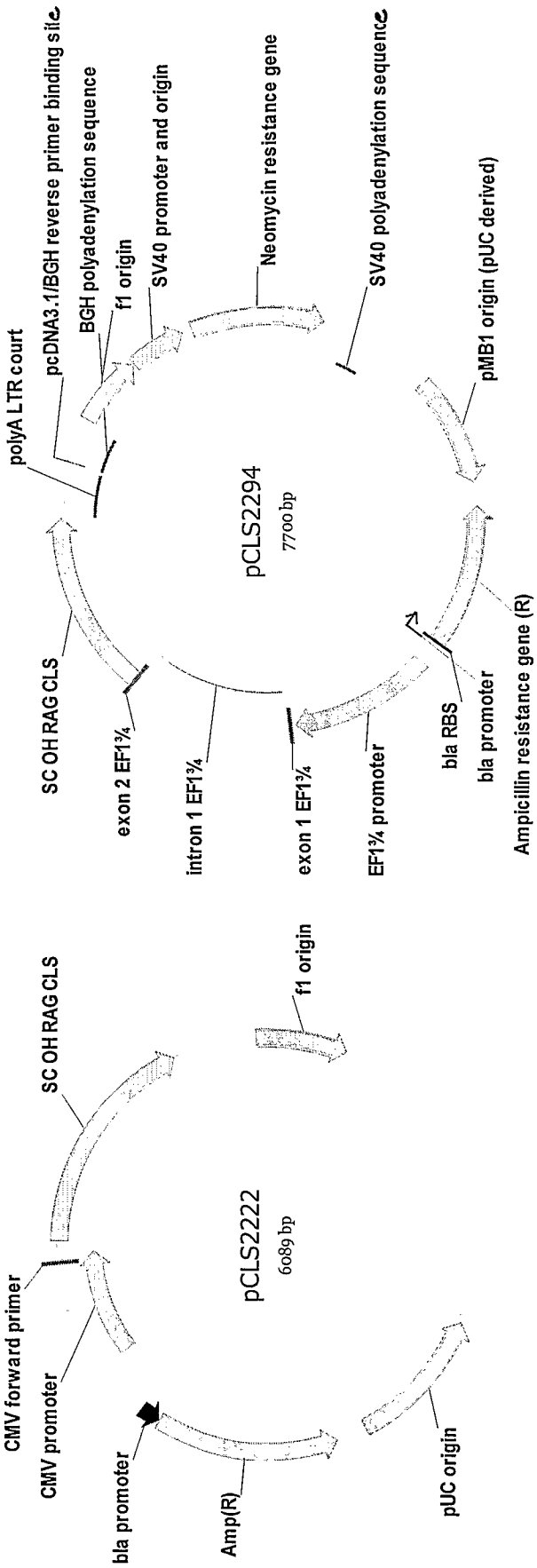

pCLS2294
7700 bp
polyA LTR court
pcDNA3.1/BGH reverse primer binding site
BGH polyadenylation sequence
f1 origin
SV40 promoter and origin
Neomycin resistance gene
SV40 polyadenylation sequence
pMB1 origin (pUC derived)
Ampicillin resistance gene (R)
bla promoter
bla RBS
EF1¾ promoter
exon 1 EF1¾
intron 1 EF1¾
exon 2 EF1¾
SC OH RAG CLS
pCLS2222
6089 bp
SC OH RAG CLS
f1 origin
CMV forward primer
CMV promoter
bla promoter
Amp(R)
pUC origin

Figure 32

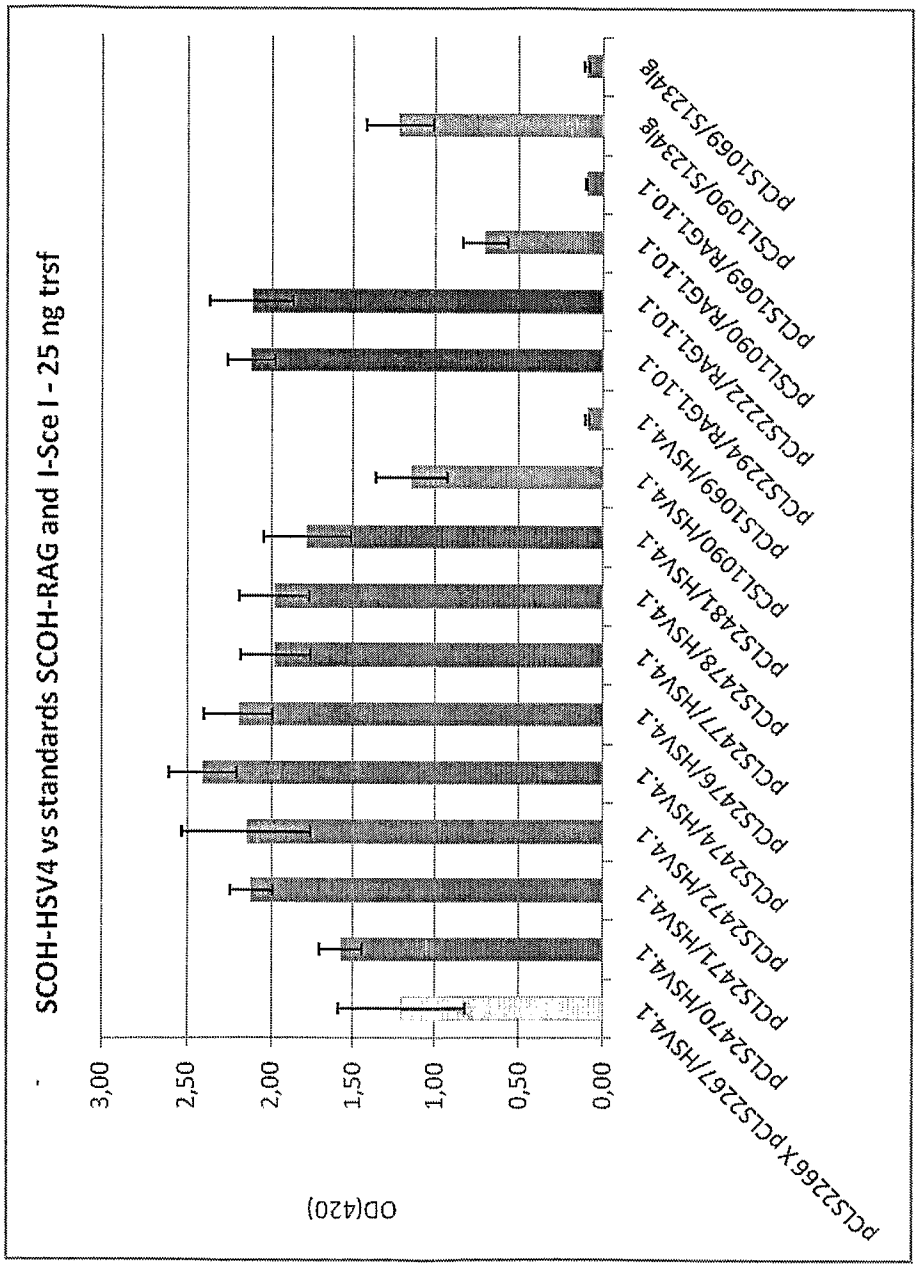

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - 25 ng trsf
OD(420)
3,00
2,50
2,00
1,50
1,00
0,50
0,00
pCLS2266 X pCLS2267/HSV4.1
pCLS2470/HSV4.1
pCLS2471/HSV4.1
pCLS2472/HSV4.1
pCLS2474/HSV4.1
pCLS2476/HSV4.1
pCLS2477/HSV4.1
pCLS2478/HSV4.1
pCLS2481/HSV4.1
pCLS1090/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1090/RAG1.10.1
pCLS1069/RAG1.10.1
pCLS1090/S1234IE
pCLS1069/S1234IE

Figure 33

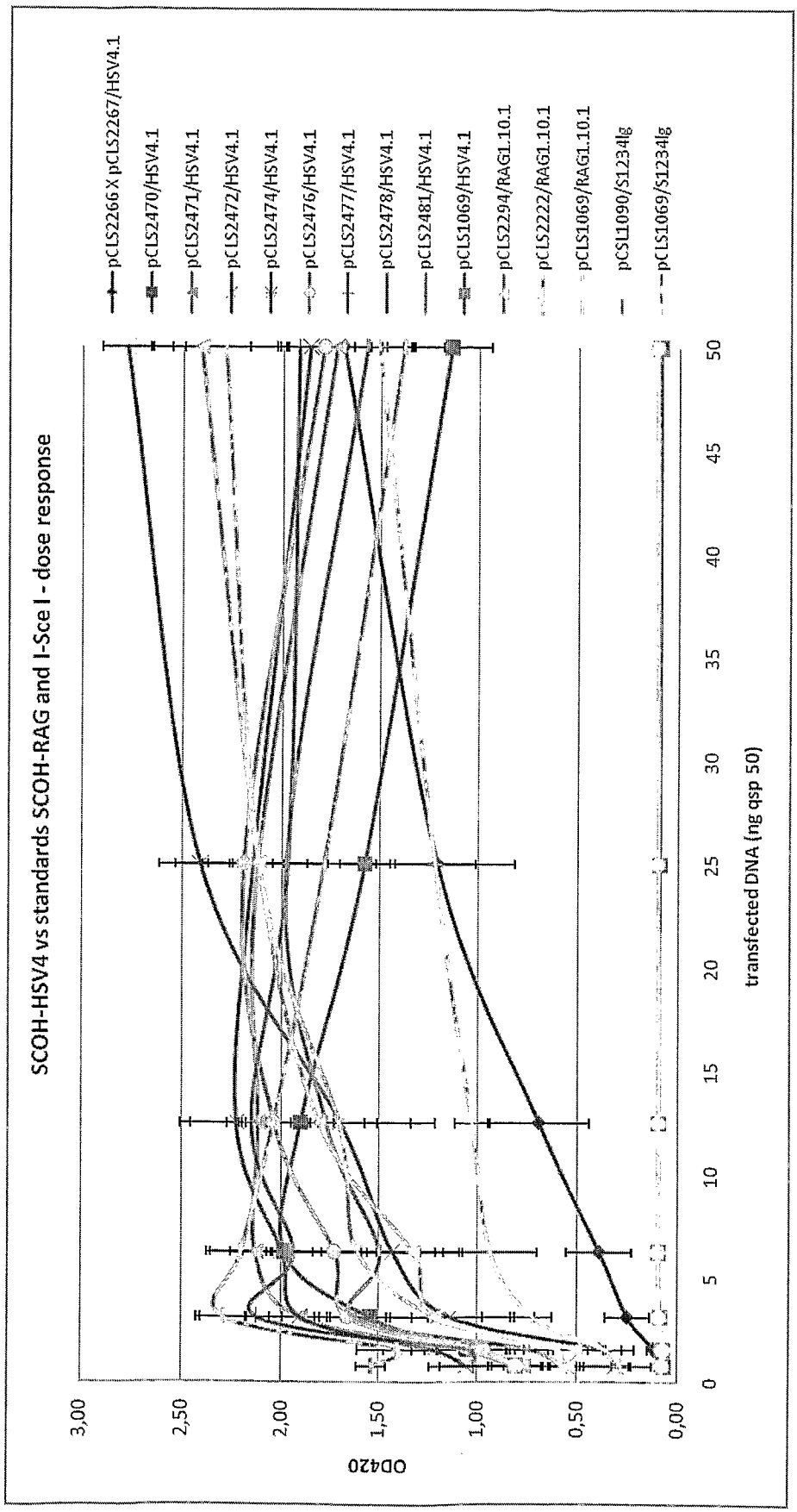

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - dose response
pCLS2266 X pCLS2267/HSV4.1
pCLS2470/HSV4.1
pCLS2471/HSV4.1
pCLS2472/HSV4.1
pCLS2474/HSV4.1
pCLS2476/HSV4.1
pCLS2477/HSV4.1
pCLS2478/HSV4.1
pCLS2481/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCSL1090/S1234lg
pCLS1069/S1234lg
OD420
3,00
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
30
35
40
45
50
transfected DNA (ng qsp 50)

Figure 34

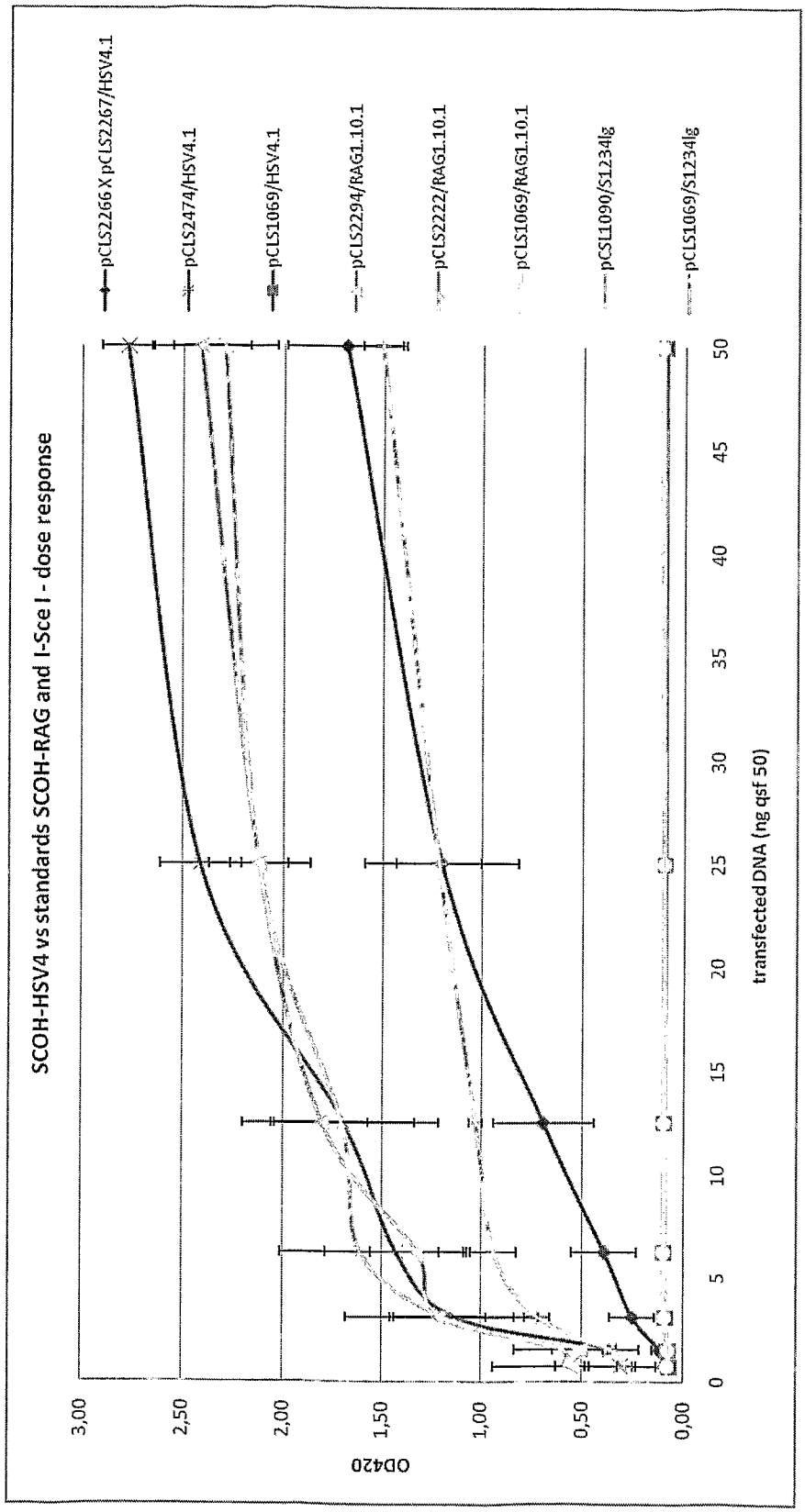

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - dose response
pCLS2266 X pCLS2267/HSV4.1
pCLS2474/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCSL1090/S1234Ig
pCLS1069/S1234Ig
OD420
3,00
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
30
35
40
45
50
transfected DNA (ng qsf 50)

Figure 35

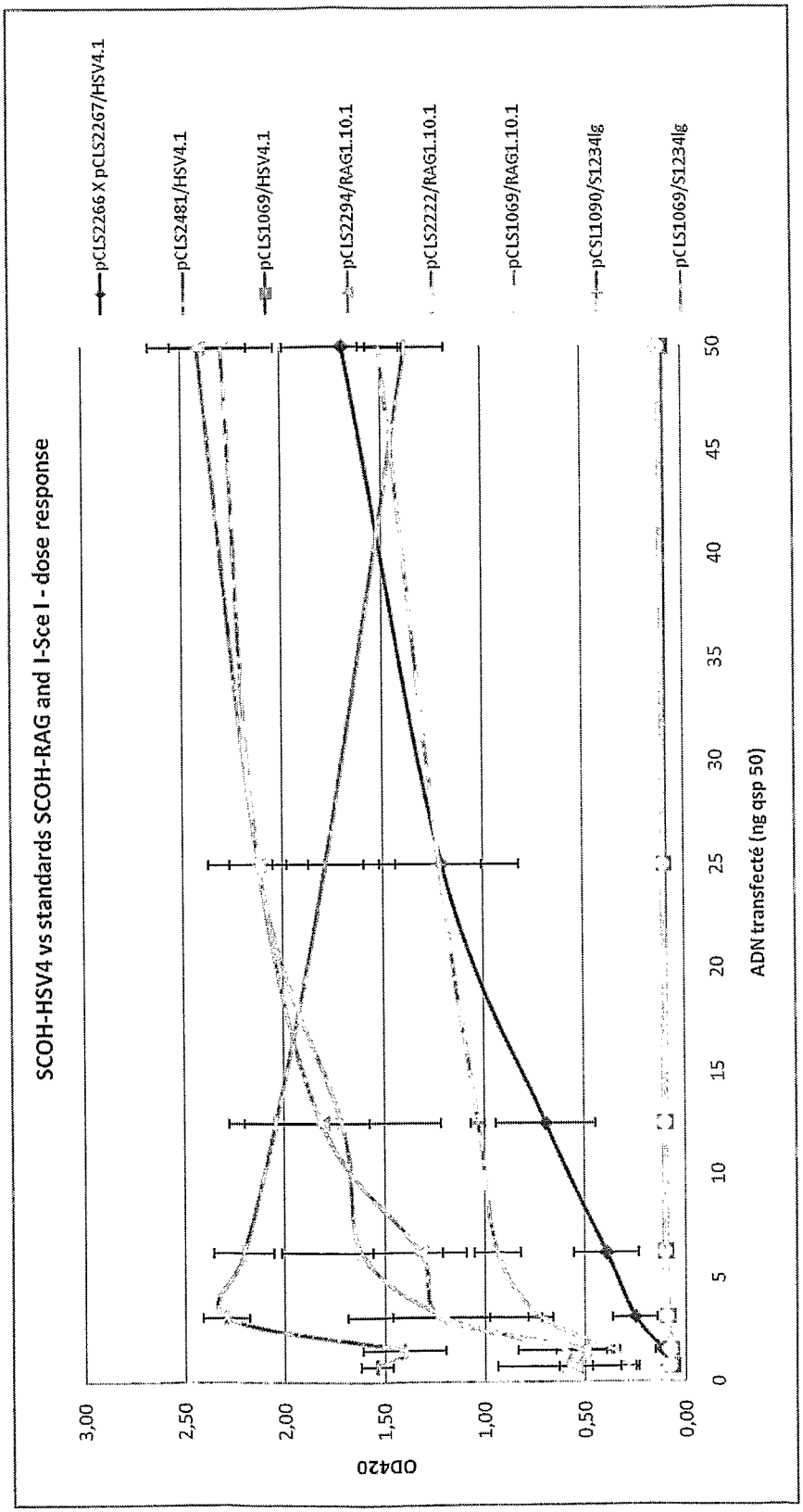

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - dose response
pCLS2266 X pCLS2267/HSV4.1
pCLS2481/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCSL1090/S1234Ig
pCLS1069/S1234Ig
OD420
3,00
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
30
35
40
45
50
ADN transfecté (ng qsp 50)

Figure 36

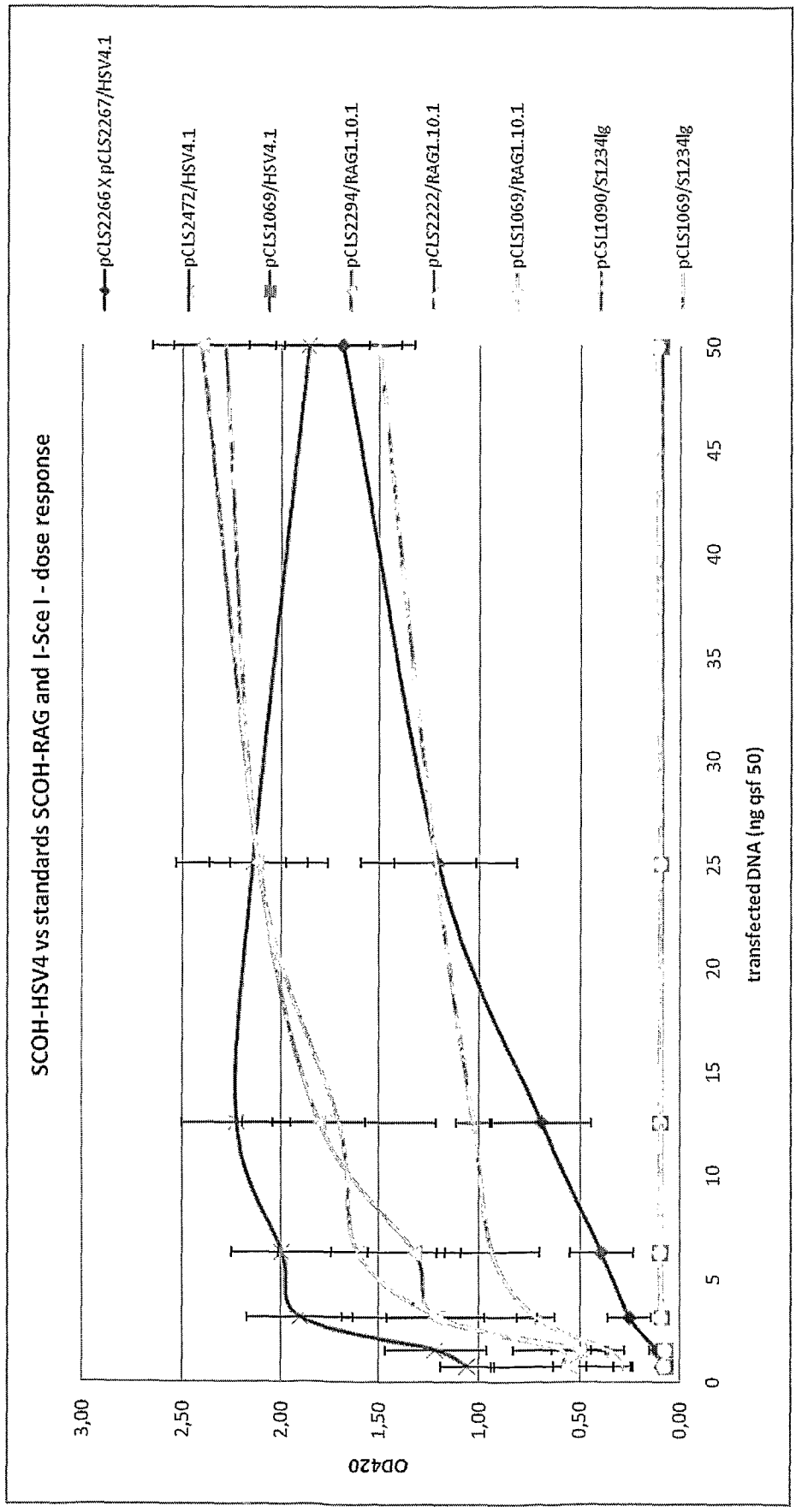

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - dose response
pCLS2266 X pCLS2267/HSV4.1
pCLS2472/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCLS1090/S1234lg
pCLS1069/S1234lg
OD420
3,00
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
30
35
40
45
50
transfected DNA (ng qsf 50)

Figure 37

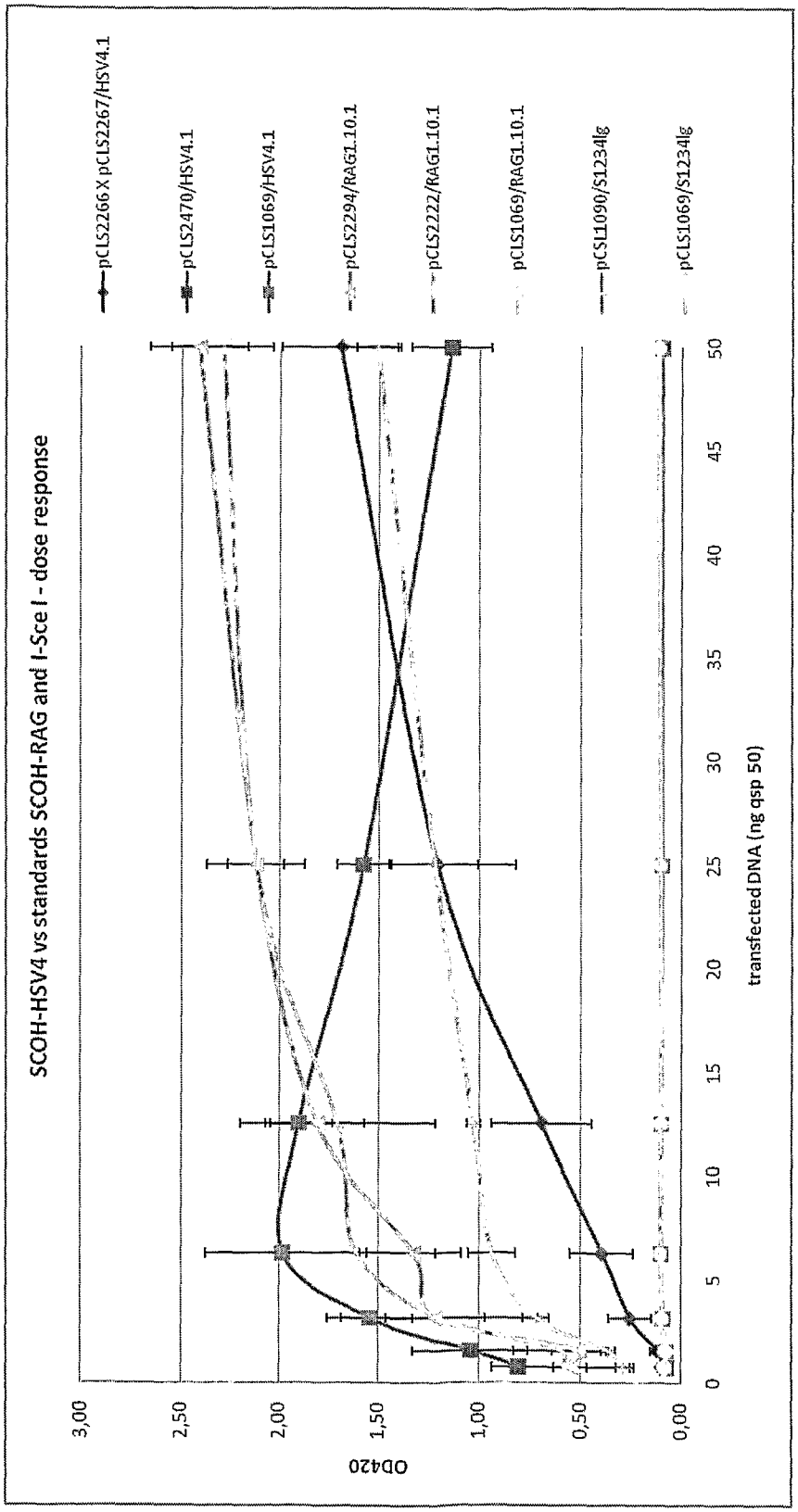

SCOH-HSV4 vs standards SCOH-RAG and I-Sce I - dose response
pCLS2266 X pCLS2267/HSV4.1
pCLS2470/HSV4.1
pCLS1069/HSV4.1
pCLS2294/RAG1.10.1
pCLS2222/RAG1.10.1
pCLS1069/RAG1.10.1
pCSL1090/S1234lg
pCLS1069/S1234lg
OD420
3,00
2,50
2,00
1,50
1,00
0,50
0,00
0
5
10
15
20
25
30
35
40
45
50
transfected DNA (ng qsp 50)

Figure 38

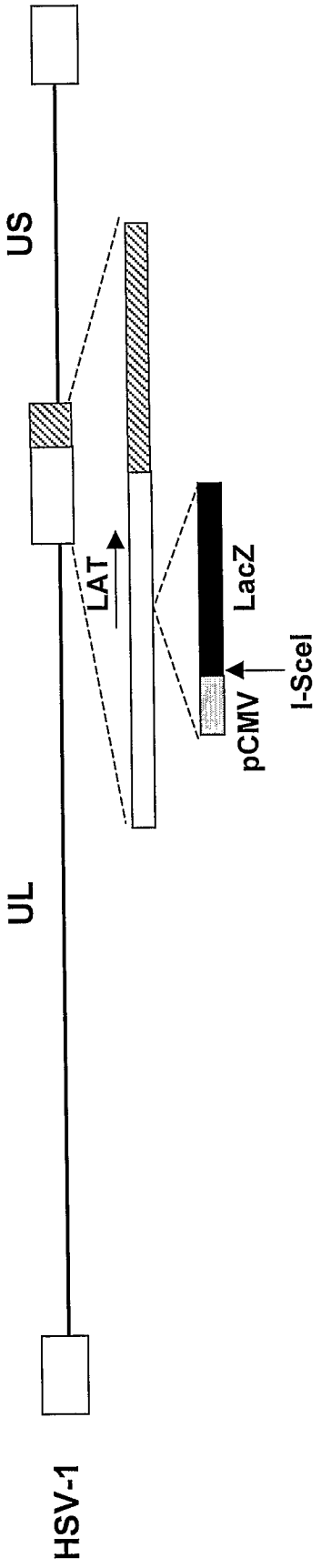


Figure 39

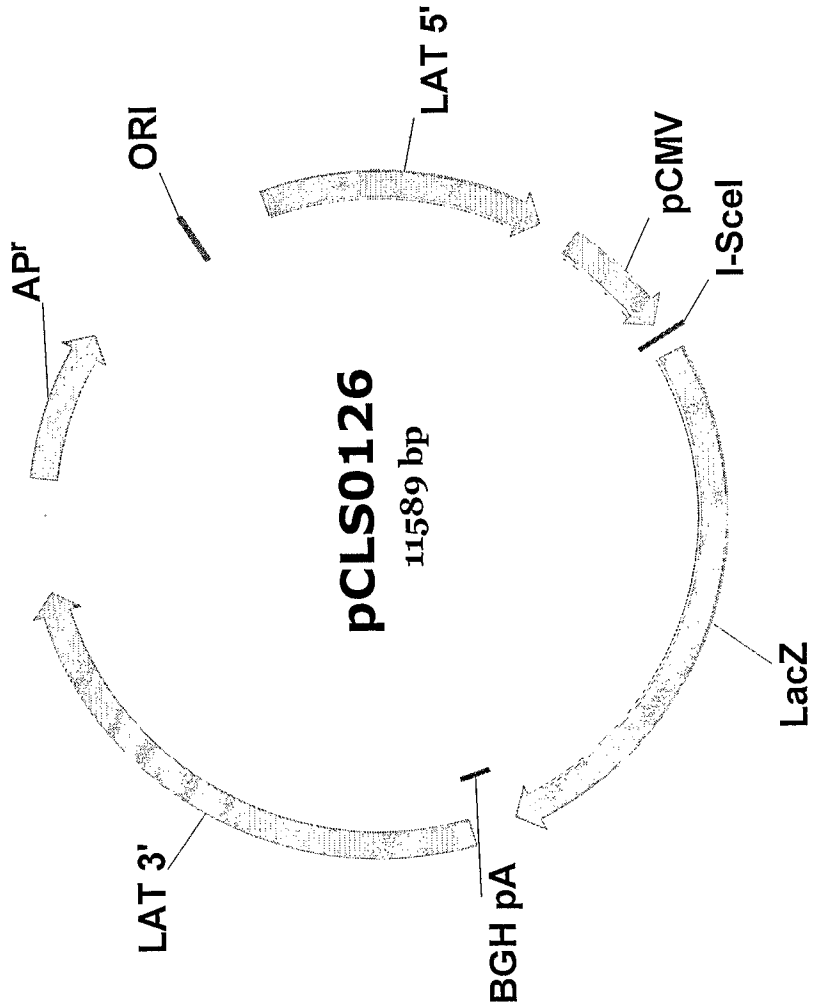


Figure 40

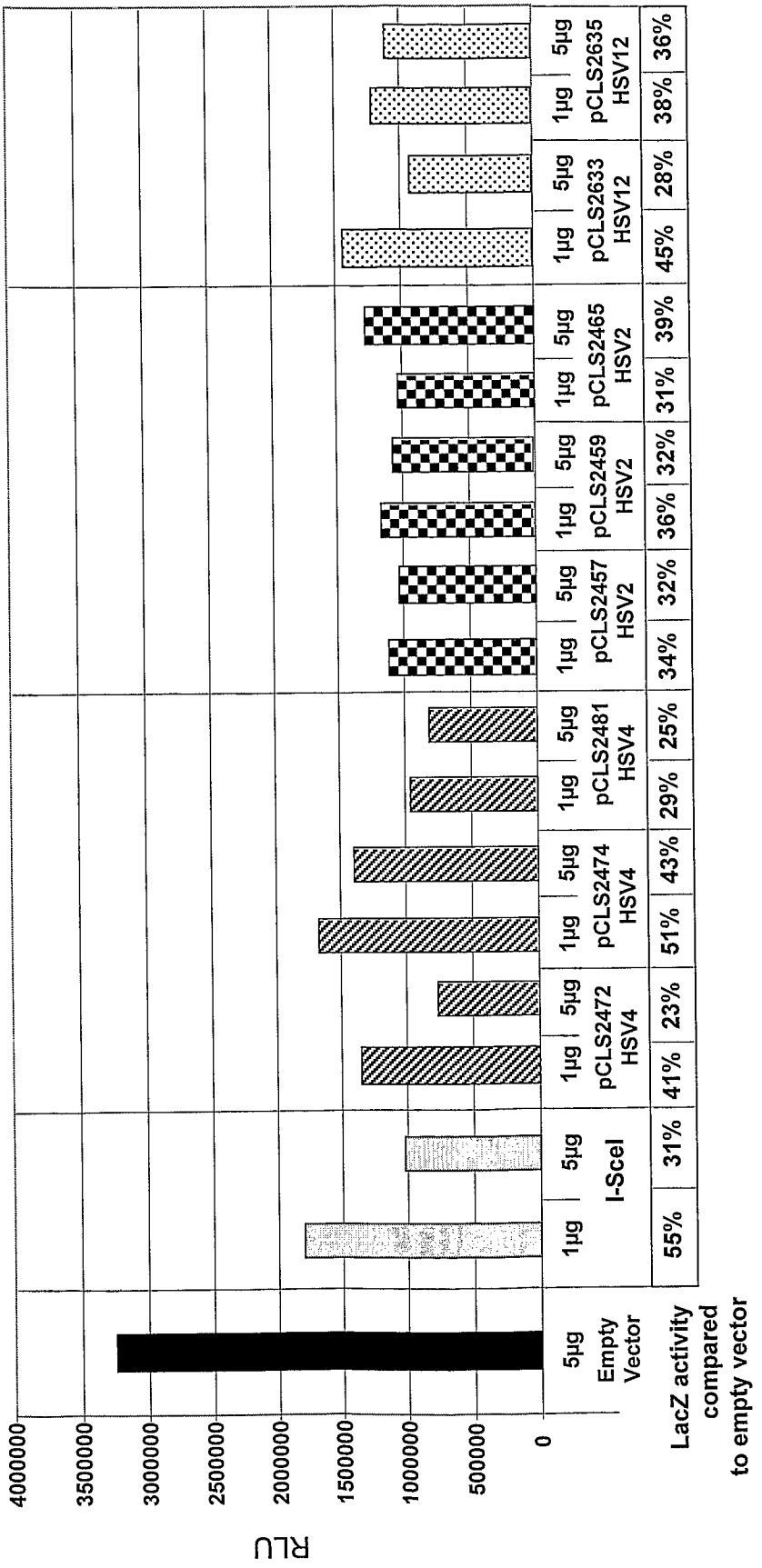

RLU
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
5µg Empty Vector
1µg I-SceI 5µg
1µg pCLS2472 HSV4 5µg
1µg pCLS2474 HSV4 5µg
1µg pCLS2481 HSV4 5µg
1µg pCLS2457 HSV2 5µg
1µg pCLS2459 HSV2 5µg
1µg pCLS2465 HSV2 5µg
1µg pCLS2633 HSV12 5µg
1µg pCLS2635 HSV12 5µg
LacZ activity compared to empty vector
55% 31% 41% 23% 51% 43% 29% 25% 34% 32% 36% 32% 31% 39% 45% 28% 38% 36%

Figure 41

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2009/007171

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/22

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

B. FIELDS SEARCHED

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

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, Sequence Search, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/093152 A (CELLECTIS [FR]; FAJARDO SANCHEZ EMMANUEL [ES]; GRIZOT SYLVESTRE [FR];) 7 August 2008 (2008-08-07) abstract; sequences 16,18 page 3, lines 1-24 page 4, line 12 - page 5, line 30 page 7, line 6 - page 8, line 13 page 12, lines 4-28 page 14, line 7 - page 15, line 4 page 21, lines 16-25 ----- -/--	1-15, 19-29

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 January 2010	29/06/2010

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Montrone, Marco

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2009/007171

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2007/049095 A (CELLECTIS [FR]; PAQUES FREDERIC [FR]) 3 May 2007 (2007-05-03) abstract page 5, line 15 - page 6, line 16 page 7, lines 11,12 page 11, line 2 page 12, line 29 - page 15, line 31 page 19, lines 16-20 page 22, lines 14-26	1-15, 19-29
Y	WO 2007/093918 A (CELLECTIS [FR]; ARNOULD SYLVAIN [FR]; PEREZ-MICHAUT CHRISTOPHE [FR]; S) 23 August 2007 (2007-08-23) abstract claims 1-12	1-29
Y	ARNOULD S ET AL: "Engineering of Large Numbers of Highly Specific Homing Endonucleases that Induce Recombination on Novel DNA Targets" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 355, no. 3, 20 January 2006 (2006-01-20), pages 443-458, XP024950505 ISSN: 0022-2836 [retrieved on 2006-01-20] cited in the application abstract	1-29
Y	SUSSMAN D ET AL: "Isolation and Characterization of New Homing Endonuclease Specificities at Individual Target Site Positions" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 342, no. 1, 3 September 2004 (2004-09-03), pages 31-41, XP004844889 ISSN: 0022-2836 abstract	1-29
Y	SMITH JULIANNE ET AL: "A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences." NUCLEIC ACIDS RESEARCH 2006, vol. 34, no. 22, 2006, pages E149-1-E149-12, XP002559417 ISSN: 1362-4962 cited in the application abstract page 2, column 1, paragraphs 3,4; figure 1 figure 2 page 6, column 1, paragraph 2 - page 7, column 1, paragraph 1 page 11, column 1, paragraph 2	1-29

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2009/007171

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Invention 1 claims 1-29(all partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1 claims 1-29(all partially)

An I-CreI variant, characterized in that at least one of the two I-CreI monomers has at least two substitutions, one in each of the two functional subdomains of the LAGLIDADG core domain situated from positions 26 to 40 and 44 to 77 of I-CreI, said variant being able to cleave a DNA target sequence from the herpes simplex virus (HSV) genome, wherein said variant is characterised by SEQIDNO: 25.

Invention 2-422 claims 1-29(all partial)

An I-CreI variant, characterized in that at least one of the two I-CreI monomers has at least two substitutions, one in each of the two functional subdomains of the LAGLIDADG core domain situated from positions 26 to 40 and 44 to 77 of I-CreI, said variant being able to cleave a DNA target sequence from the herpes simplex virus (HSV) genome, wherein said variant is characterised by SEQIDNO: 26 to 36, 40 to 90, 93 to 151, 153 to 168, 171 to 246, 249 to 261, 267 to 273, 275 to 288, 290 to 433, 436 to 463 and 470 to 471.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IB2009/007171

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008093152 A	07-08-2008	AU 2008211635 A1	07-08-2008
		CA 2677177 A1	07-08-2008
		CN 101678087 A	24-03-2010
		EP 2121004 A2	25-11-2009
		WO 2008093249 A2	07-08-2008
WO 2007049095 A	03-05-2007	AU 2006314220 A1	24-05-2007
		CA 2625774 A1	24-05-2007
		CN 101310015 A	19-11-2008
		EP 1979478 A2	15-10-2008
		WO 2007057781 A2	24-05-2007
		JP 2009513130 T	02-04-2009
		US 2009220476 A1	03-09-2009
WO 2007093918 A	23-08-2007	AU 2007216264 A1	23-08-2007
		CA 2639149 A1	23-08-2007
		CN 101384712 A	11-03-2009
		EP 1989299 A2	12-11-2008
		WO 2007093836 A1	23-08-2007
		JP 2009526521 T	23-07-2009
		US 2009222937 A1	03-09-2009