Title: MEGANUCLEASE VARIANTS CLEAVING THE GENOME OF A NON-GENOMICALLY INTEGRATING VIRUS AND USES THEREOF

Abstract: The present invention relates to variants of the meganuclease l-Crel which are characterized in that they are able to cleave a DNA target sequen from a non-integrating virus (NIV) genome. The present invention also relates to rational methods to generate such l-Crel variants and the use of the variants as research and diagnostic reagents, as well as medicaments comprising these l-Crel variants for the treatment of NIV infections.
MEGANUCLEASE VARIANTS CLEAVING THE GENOME OF A NON-
GENOMICALLY INTEGRATING VIRUS AND USES THEREOF

The invention relates to a meganuclease variant cleaving the genome of viruses which do not integrate into the genome of the cells which they infect and hence exist as episomal extra genomic DNA molecule(s). In particular the present invention relates to a meganuclease variant which cleaves a target present in the genome of a *Hepadnaviridae* virus and most particularly a hepatitis B Virus. The present invention also relates to a vector encoding such a meganuclease variant, to a cell, an animal or a plant modified by said vector and to the use of such a meganuclease variant and products derived there from, for genome engineering and for *in vivo* and *ex vivo* genome therapy.

Viral infections of various sorts are a serious and continuing health, agricultural and economic problem worldwide. In particular viruses present specific treatment and control problems as they always comprise an intracellular stage to their life cycle, in which the nucleic acid genome of the virus is inserted into a host cell and normally transported to the nucleus. During this stage of the virus life cycle, the virus genome can enter into a dormant state whilst inside a host cell, in which the production of new virus particles/proteins/copies of the viral genome ceases. These characteristics present a significant problem as most medicaments and treatments for viral infection consist of compounds which affect aspects of virus biology involved in the active stages of the virus life cycle, such as compounds which target a viral enzyme or structural protein. Therefore whilst in a dormant state the viral genome resident in the cytoplasm or nucleus of a host cell can not be affected by most conventional anti-virus medicaments and therefore persists.

The present invention relates to viruses which do not integrate into the host genome following insertion of the viral genomic/genetic material into the host cell. That is the viral genetic material exists as an episomal/separate DNA molecule. Most important diseases exhibit such a life cycle, for example DNA ds (double stranded) viruses like *Herpesviridae, Adenoviridae, Papovaviridae* and *Poxviridae*; DNA ss (single stranded) viruses like *Parvoviridae* and DNA ds viruses that replicate through a single stranded RNA intermediate such as *Hepadnaviridae*.
Hepatitis B, a virus of the family *Hepadnaviridae*, is an example of an epidemiologically important virus which following insertion of the virus genome into a host cell, then exists as an episomal DNA molecule separate from the host cell genome in the nucleus.

Infection with hepatitis B virus (HBV) is a world health problem, leading to more than 1 million deaths per year according to the World Health Organization. HBV is transmitted through infected blood, body fluids and by sexual intercourse.

HBV exhibits genetic variability with an estimated rate of 1.4 to 3.2 x 10^3 nucleotide substitutions per site per year. A large number of virus variants arise during replication as a result of nucleotide misincorporations, due to the absence of any proof reading capacity by the viral polymerase. This variability has resulted in well recognized subtypes of the virus (Schaefer, World J. Gastroenterol., 2007, 13:14-21).

HBV is an enveloped DNA-containing virus that replicates through an RNA intermediate. The infectious ("Dane") particle consists of an inner core plus an outer surface coat (Figure 32). The virus is a spherical particle with a diameter of 42 nm and is composed of an outer shell (or envelope) composed of several proteins known collectively as HBs which surrounds an inner protein shell, composed of HBc protein. Finally the HBc protein surrounds the viral DNA and the viral DNA polymerase.

The HBV virion genome is circular and approximately 3.2 kb in size and consists of DNA that is mostly double stranded. It comprises four overlapping open reading frames running in one direction and no non-coding regions. The four overlapping open reading frames (ORFs) in the genome are responsible for the transcription and expression of seven different HBV proteins. The four ORFs are known as C, S, P and X. The C ORF codes for the viral core protein and the e-antigen, the S ORF codes for three related viral envelope proteins, the P ORF codes for viral DNA polymerase and the X ORF codes for a 16.5 kDa protein whose function is not well defined (Figure 33).

The C ORF is divided into the precore region and the core region by two in-frame initiating ATG codons. The hepatitis B virus core antigen (HBcAg) is
initiated from the second ATG and thus contains only the core region. The virus core antigens associate to form the hepatitis B core that encapsulates HBV DNA and DNA polymerase. This protein has been shown to be essential for viral DNA replication. A second protein, the hepatitis B e-antigen (HBeAg) is initiated from the first ATG in the C ORF and thus consists of the pre-core and core region. This protein is targeted to the endoplasmic reticulum where it is cleaved at the N and C terminus and then secreted as a non-particulate HBeAg. This protein is not essential for viral replication and its function remains unknown. The S ORF encodes for three envelope proteins known as small (S), medium (M), and large (L) hepatitis B surface antigen. All three proteins contain the structural domain. The extra domain in M is known as pre-S2 while L contains the pre-S2 and pre-S1 domains. The pre-S1 domain is thought to be the substrate for the viral receptor on hepatocytes and thus essential for viral attachment and entry. All three envelope proteins are components of the infectious viral particles also referred to as Dane particles. However, the S protein by itself or associated with the larger envelope proteins have been shown to form spheres and filaments that are secreted from infected cells in at least 100-fold excess over infectious viral particles. It is thought that these spheres and filaments may serve to titrate out antibodies that are produced by the immune system and thus aid the infectious viral particles to escape the immune system. The P ORF codes for the viral DNA polymerase. This protein consists of two major domains tethered by an intervening spacer region. The amino-terminal domain plays a critical role in the packaging of pre-genomic RNA and in the priming of minus strand DNA while the carboxy-terminal domain is a reverse transcriptase that also has RNase H activity. This protein is essential for viral DNA replication. The X ORF encodes a protein that has been shown to be essential for virus replication in animals but dispensable for viral DNA synthesis in transfected tissue culture cells. It has been suggested that the X protein may play a role in transcriptional activation as well as stimulation of signal transduction pathways and regulation of apoptosis (Seeger and Mason, Microbiol. Mol. Biol. Rev., 2000, 51-68).

The viral genome consists of two partially overlapping DNA strands, called the - and + strands. The - strand is the larger of the two strands and is approximately 3.02 kb - 3.32 kb in length and has a protein covalently attached to its
5' end. The + strand, is approximately 1.7 - 2.8 kb in length and has an RNA oligonucleotide attached at its 5' end.

The viral DNA is found in the nucleus soon after infection of the cell. The partially double-stranded DNA is rendered fully double-stranded by completion of the (+) sense strand and removal of the protein molecule from the (-) sense strand and a short sequence of RNA from the (+) sense strand and the ends are rejoined. This fully double-stranded DNA, a closed and circular DNA structure is known as cccDNA (covalently closed circular DNA).

HBV is a vaccine-preventable disease. Current vaccines are composed of the surface antigen of HBV and are produced by two different methods: plasma derived or recombinant DNA (Maupas et al, Lancet, 1976, 7974: 1367-1370; Mahoney, Clin. Microbiol. Rev., 1999,12:351-366). However HBV vaccines are not available to all at risk individuals and/or are not always administered in the correct form and so cases of HBV infection persist throughout the world.

HBV infection can result in two distinct disease states, acute and chronic HBV infection. Acute HBV is the initial, rapid onset, short duration illness that results from infection with HBV. About 70% of adults with acute hepatitis B have few or no symptoms, while the remaining 30% develop significant symptoms (Seeger and Mason, Microbiol. Mol. Biol. Rev., 2000, 51-68). Rarely (in less than 1 % of adults), individuals with acute hepatitis B can develop acute liver failure (fulminant hepatitis).

Chronic hepatitis B infection may take one of two forms: chronic persistent hepatitis, a condition characterized by persistence of HBV but in which liver damage is minimal; and chronic active hepatitis, in which there is aggressive destruction of liver tissue leading to cirrhosis and/or cancer such as hepatocellular carcinoma.

The prevalence of chronic HBV infection varies greatly in different parts of the world. Chronic HBV infection is highly endemic in developing regions with large populations such as South East Asia, China, sub-Saharan Africa and the Amazon Basin; moderately endemic in parts of Eastern and Southern Europe, the Middle East, Japan, and part of South America and low in most developed areas, such as North America, Northern and Western Europe and Australia.
When HBV infection results in a chronic disease, this cannot currently be cured. Therefore the goal of therapy is the long-term suppression of viral replication, as this is associated with a reduced risk of the development of advanced liver disease including liver cirrhosis and cancer. There are currently two major families of drugs that have been approved for the treatment of chronic Hepatitis B infections, interferons, which boost the immune system in order to eliminate or diminish the virus, and nucleoside/nucleotide analogues, which inhibit viral replication. As all treatments for Hepatitis B infections are administered over long periods of time, one of the major problems is the development of drug resistance. This is particularly the case for nucleoside/nucleotide analogues for which there are a growing number of documented viral polymerase mutants that result in drug resistance (Tillman, World J. Gastroenterol., 2007, 13:125-140).

Liver transplantation is the only long term treatment available for patients with liver failure. However, liver transplantation is complicated by the risk of recurrent hepatitis B infection in patients where the initial liver failure was due to hepatitis B infection or who have a chronic HV infection or a high risk of HBV reinfection; this problem significantly impairs graft and patient survival. In the absence of treatment, HBV reinfection occurs in 75%-80% of persons who undergo liver transplantation.

Therefore in the prior art significant problems exist with treating patients who are chronically infected with HBV and more specifically with reducing the HBV viral titer as far as possible in a patient who requires a liver transplant.

A promising target for the development of treatments for HBV infection and more generally non-integrating viruses is the intracellular episomal HBV/NIV (Non Integrating Virus) genome. The intracellular HBV genome is the molecular basis of HBV persistence. It has been found in both animal models and clinical investigations that cccDNA persists even after years of antiviral therapy and is responsible for the rapid increases in viral titer following withdrawal of treatment or the development of resistance.

It has been proposed (WO 2008/119000) that the intracellular genome of HBV could be targeted and potentially inactivated using a variety of methods such as via RNA interference (RNAi), short interfering RNA (siRNA) and
engineered polydactyl zinc finger protein domains in combination with cleavage
domains generated against a target(s) in the HBV genome. To date however none of
these methods have been shown able to specifically target and/or affect the HBV
virus.

Potential problems exist with all of these proposed mechanisms, for
instance although it appears that RNAi can suppress virtually all classes of DNA and
RNA virus against which they have been tested (Dykxhoorn, DM and Lieberman, J
532—540.), clinical studies are increasingly showing that viruses are able to elude the
apparent ease.

Likewise in theory zinc finger domains could be generated which
are specific to targets in the HBV genome and therefore Zinc finger nucleases (ZFNs),
which are chimeric proteins composed of a 'specific' zinc finger DNA-binding
domain linked to a non-specific DNA-cleavage domain, could be generated to a target
in the HBV genome. Such ZFNs would not be useful as in general ZFNs are known to
be highly cytotoxic (Porteus MH, Baltimore D (2003) Science 300: 763 and Beumer
2391-2403.) due to their cleavage of non-target sequences, leading to genome
degradation. Although various steps have been attempted to attenuate these cytotoxic
effects thus far ZFNs remain simply too toxic for routine use. The generation of such
ZFNs is also a laborsome endeavour as the combination of a given zinc finger domain,
following its generation, with a nuclease domain requires a substantial amount of
work to ensure firstly that the combination is functional and secondly specific.

The inventors of the present invention have developed a new
molecular medicine approach based on the inactivation of the HBV virus whilst in the
infected cell by targeting the cccDNA with engineered meganucleases, which could
lead to the cleavage and elimination or inactivation of the intracellular copies of the
virus genome that allows the virus to persist.

In the wild, meganucleases are essentially represented by homing
endonucleases. Homing Endonucleases (HEs) are a widespread family of natural
meganucleases including hundreds of proteins families (Chevalier, B.S. and B.L. Stoddard, Nucleic Acids Res., 2001, 29, 3151-211A). These proteins are encoded by mobile genetic elements which propagate by a process called "homing": the endonuclease cleaves a cognate allele from which the mobile element is absent, thereby stimulating a homologous recombination event that duplicates the mobile DNA into the recipient locus. Given their exceptional cleavage properties in terms of efficacy and specificity, they could represent ideal scaffolds to derive novel, highly specific endonucleases.

HEs belong to four major families. The LAGLIDADG family, named after a conserved peptidic motif involved in the catalytic center, is the most widespread and the best characterized group. Seven structures are now available. Whereas most proteins from this family are monomeric and display two LAGLIDADG motifs, a few have only one motif, and thus dimerize to cleave palindromic or pseudo-palindromic target sequences.

Although the LAGLIDADG peptide is the only conserved region among members of the family, these proteins share a very similar architecture (Figure 34). The catalytic core is flanked by two DNA-binding domains with a perfect two-fold symmetry for homodimers such as 1-Oel (Chevalier, et al, Nat. Struct. Biol., 2001, 8, 312-316) , 1-Msol (Chevalier et al, J. Mol. Biol., 2003, 329, 253-269) and 1-Ceul (Spiegel et al., Structure, 2006, 14, 869-880) and with a pseudo symmetry for monomers such as l-Scel (Moure et al, J. Mol. Biol., 2003, 334, 685-69, 1-Dmol (Silva et al, J. Mol. Biol., 1999, 286, 1123-1136) or l-Anil (Bolduc et al, Genes Dev., 2003, 17, 2875-2888). Both monomers and both domains (for monomeric proteins) contribute to the catalytic core, organized around divalent cations. Just above the catalytic core, the two LAGLIDADG peptides also play an essential role in the dimerization interface. DNA binding depends on two typical saddle-shaped $\alpha\beta\beta\alpha\beta\beta\alpha$ folds, sitting on the DNA major groove. Other domains can be found, for example in inteins such as VI-Pfut (Ichiyanagi et al, J. Mol. Biol., 2000, 300, 889-901) and PI-SceI (Moure et al, Nat. Struct Biol., 2002, 9, 764-770), whose protein splicing domain is also involved in DNA binding.

The making of functional chimeric meganucleases, by fusing the N-terminal l-Dmol domain with an 1-Oel monomer (Chevalier et al, Mol. Cell., 2002,


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

- Residues Q44, R68 and R70 or Q44, R68, D75 and 177 of l-Crel were mutagenized and a collection of variants with altered specificity at positions ± 3 to 5 of the DNA target (5NNN DNA target) were identified by screening (International PCT Applications WO 2006/097784 and WO 2006/097853; Arnould et al, J. Mol. Biol, 2006, 355, 443-458; Smith et al, Nucleic Acids Res., 2006, 34, el49).

- Residues K28, N30 and Q38 or N30, Y33 and Q38 or K28, Y33, Q38 and S40 of l-Crel were mutagenized and a collection of variants with altered specificity at positions ± 8 to 10 of the DNA target (IONNN DNA target) were identified by screening (Smith et al, Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156).

Two different variants were combined and assembled in a functional heterodimeric endonuclease able to cleave a chimeric target resulting from the fusion of two different halves of each variant DNA target sequence (Arnould et al, precited; International PCT Applications WO 2006/097854 and WO 2007/034262).

Furthermore, residues 28 to 40 and 44 to 77 of l-Crel were shown to form two separable functional subdomains, able to bind distinct parts of a homing

The combination of mutations from the two subdomains of 1-OeI within the same monomer allowed the design of novel chimeric molecules (homodimers) able to cleave a palindromic combined DNA target sequence comprising the nucleotides at positions ±3 to 5 and ±8 to 10 which are bound by each subdomain (Smith et al, Nucleic Acids Res., 2006, 34, e149; International PCT Applications WO 2007/049095 and WO 2007/057781).


The combination of the two former steps allows a larger combinatorial approach, involving four different subdomains. The different subdomains can be modified separately and combined to obtain an entirely redesigned meganuclease variant (heterodimer or single-chain molecule) with chosen specificity. In a first step, couples of novel meganucleases are combined in new molecules ("half-meganucleases") cleaving palindromic targets derived from the target one wants to cleave. Then, the combination of such "half-meganucleases" can result in a heterodimeric species cleaving the target of interest. The assembly of four sets of mutations into heterodimeric endonucleases cleaving a model target sequence or a sequence from different genes has been described in the following patent applications: XPC gene (WO2007/093918), RAG gene (WO2008/010093), HPRT gene (WO2008/059382), beta-2 microglobulin gene (WO2008/102274), Rosa26 gene (WO2008/152523), Human hemoglobin beta gene (WO2009/13622) and Human interleukin-2 receptor gamma chain gene (WO2009019614).

These variants can be used to cleave genuine chromosomal sequences and have paved the way for novel perspectives in several fields, including gene therapy.
Even though the base-pairs ±1 and ±2 do not display any contact with the protein, it has been shown that these positions are not devoid of content information (Chevalier et al., J. MoL Biol., 2003, 329, 253-269), especially for the base-pair ±1 and could be a source of additional substrate specificity (Argast et al., J. MoL Biol., 1998, 280, 345-353; Jurica et al., MoL. Cell, 1998, 2, 469-476; Chevalier et al., Nucleic Acids Res., 2001, 29, 3757-3774). In vitro selection of cleavable l-Crel targets (Argast et al., precited) randomly mutagenized, revealed the importance of these four base-pairs on protein binding and cleavage activity. It has been suggested that the network of ordered water molecules found in the active site was important for positioning the DNA target (Chevalier et al., Biochemistry, 2004, 43, 14015-14026). In addition, the extensive conformational changes that appear in this region upon I-Crel binding suggest that the four central nucleotides could contribute to the substrate specificity, possibly by sequence dependent conformational preferences (Chevalier et al, 2003, precited).

The inventors seeing these problems with prior art approaches to treating virus infections, in particular the persistence of dormant copies of the virus genome, have now developed a new set of materials which target the otherwise stable episomal virus genome in situ within the nucleus or cytoplasm of an infected cell. The inventors have validated their work using the important disease hepatitis B and have generated several meganuclease variants which can recognize and cleave different targets in the HBV episomal genome.

These materials can be used to manipulate the virus 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-Crel variant, characterized in that at least one of the two l-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 genome of a non-integrating virus (NIV), and being obtainable by a method comprising at least the steps of:
(a) constructing a first series of l-Crel variants having at least one substitution in a first functional subdomain of the LAGLIDADG core domain situated from positions 26 to 40 of l-Crel,

(b) constructing a second series of l-Crel variants having at least one substitution in a second functional subdomain of the LAGLIDADG core domain situated from positions 44 to 77 of 1-OeI,

(c) selecting and/or screening the variants from the first series of step (a) which are able to cleave a mutant l-Crel site wherein at least one of (i) the nucleotide triplet in positions -10 to -8 of the 1-OeI site has been replaced with the nucleotide triplet which is present in positions -10 to -8 of said DNA target sequence from the NIV 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 NIV genome,

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

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

(f) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant l-Crel site wherein at least one of (i) the nucleotide triplet in positions +3 to +5 of the I-Crel site has been replaced with the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the NIV 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 NIV 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-Crel 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 NIV 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 NIV 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 NIV 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 NIV 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-Crel variant which cleaves a sequence wherein (i) the nucleotide triplet in positions +8 to +10 of the I-Crel site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the NIV 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 NIV 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 NIV 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 NIV genome,

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

Q) selecting and/or screening the heterodimers from step (i) which are able to cleave said DNA target sequence from the NIV 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 NIV *in vitro* and *in vivo*; this class of reagents also represent a potential new class of anti-NIV medicament, which instead of acting upon the virion or any component thereof, acts upon the intracellular genomic of the virus.

To validate their invention, the Inventors have identified a series of DNA targets in the genome of the NIV hepatitis B virus (HBV), that are cleaved by I-Crel variants (Figures 1, 11, 21 and 35).

Target sequences can be chosen in any region of the NIV genome, for instance in the coding sequence of a virus gene and in particular in a gene (s) which is essential for the virus include. 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 most viruses all 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-Crel 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-Crel 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 NIV genome.

The inventors have found that although specific meganucleases can be generated to a particular target in the NIV 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 said 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 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 NIV genome.
The inventors have found that the meganucleases can be further improved by using multiple iterations of the additional steps (k) and (1).

In particular said substitution(s) in the subdomain situated from positions 26 to 40 of I-Oel 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-Crel enzymes specificity.

In particular said substitution(s) in the subdomain situated from positions 44 to 77 ofl-Oel 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-Crel enzymes specificity.

In particular the variant comprises one or more substitutions in positions 137 to 143 of I-Crel 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 NIV genome.

In particular the variant comprises one or more substitutions on the entire 1-Crel sequence that improve the binding and/or the cleavage properties of the variant towards said DNA target sequence from the NIV 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-Crel 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-Crel, said heterodimer being able to cleave a non-palindromic DNA target sequence from the HBV genome.

As explained above the I-Crel enzyme acts as a dimer, by ensuring that the variant is a heterodimer this allows a specific combination of two different I-Crel 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-OeI 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 NIV genome.

The inventors have previously established a number of residue changes which can ensure an I-Crel 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.

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 peptidic 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 NIV genome.

In particular the NIV is selected from the group consisting of Zoophaginae or Phytophaginae viruses.

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

In particular the NIV is a virus from a family selected from the group comprising: Herpesviridae, Adenoviridae, Papovaviridae, Poxviridae, Parvoviridae, Hepadnaviridae.

In particular the NIV is selected from the group comprising: herpes simplex virus 1, herpes simplex virus 2, herpes simplex virus 3, Varicella zoster virus, Epstein-Barr virus, Cytomegalovirus, Herpes lymphotropic virus, Roseolovirus,
Rhadinovirus, Adenovirus, Papillomavirus, Polyomavirus, variola virus, vaccinia virus, cowpox virus, monkeypox virus, camel pox, variola virus, vaccinia virus, cowpox virus, monkeypox virus, tanapox virus, yaba monkey tumor virus, molluscum contagiosum virus, Parvovirus B19, hepatitis B.

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_databaseA^IDA.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 NIV is hepatitis B.

In particular the DNA target sequence is from a hepatitis B virus of genotype A.

As indicated above HBV exhibits genetic variability with an estimated rate of 1.4 - 3.2 × 10⁻⁵ nucleotide substitutions per site per year. A large number of virus variants arise during replication as a result of nucleotide misincorporations in the absence of any proof reading capacity by the viral polymerase. This variability has resulted in well recognized subtypes of the virus: HBV has been classified into 8 well defined genotypes on the basis of an inter-group divergence of 8% or more in the complete genomic sequence, each having a distinct geographical distribution. Genotype A is most commonly found in Northern Europe, North America and Central Africa, while genotype B predominates in Asia (China, Indonesia and Vietnam). Genotype C is found in the Far East in Korea, China, Japan and Vietnam as well as the Pacific and Island Countries, while genotype D is found in the Mediterranean countries, the Middle East extending to India, North America and parts of the Asia-Pacific region. Genotype E is related to Africa while genotype F is found predominately in South America, including among Amerindian populations,
and also Polynesia. Genotype G has been found in North America and Europe while the most recently identified genotype H has been reported from America (Schaefer, World J. Gastroenterol., 2007, 13:14-21).

In the present Patent Application the inventors have generated meganuclease variants to targets present in the genome of hepatitis B virus either in genotype A subtype adw2 (Preisler-Adams et al., Nucleic Acids Research, 1993, Vol. 21, No.9), which corresponds to Genbank accession number X70185 or in subtype adr, which corresponds to Genbank accession number M38636.

In particular the variants may be selected from the group consisting of SEQ ID NO: 15 to 20, 22 to 27, 29, 32 to 48, 61 to 75, 84 to 89, 91, 93 to 96, 98 to 109, 124 to 128, 130 to 134, 137 to 144, 146 to 153, 155 to 192.

In particular said DNA target is selected from the group consisting of the sequences SEQ ID NO: 6 to 9, 79 to 82, and 117 to 122.

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

In particular the DNA target is within an open reading frame of the HBV genome, selected from the group: C ORF, S ORF, P ORF and X ORF.

The HBV virion genome contains four overlapping open reading frames (ORFs) in the genome which are responsible for the transcription and expression of seven different hepatitis B proteins. The transcription and translation of these proteins is through the use of multiple in-frame start codons. The HBV genome also contains parts that regulate transcription, determine the site of polyadenylation and a specific transcript for encapsidation into the capsid.

Details concerning the four overlapping open reading frames of the HBV genome are detailed in the introduction above. In particular the DNA target is located in one of the HBV genomic genes selected from the group: viral core protein, e-antigen, small (S) hepatitis B surface antigen, medium (M) hepatitis B surface antigen, large (L) hepatitis B surface antigen, viral DNA polymerase, X protein.

In the present Patent Application the inventors provide meganuclease variants which can cleave targets in the S ORF and P ORF (target HBV12, SEQ ID NO: 6) and two independent targets in the C ORF (target HBV 8, SEQ ID NO: 79 and target HBV3, SEQ ID NO: 117). The cleavage of these sites in
the HBV genome \emph{in vivo} would therefore disrupt the sequence encoding the small (S) hepatitis B surface antigen, medium (M) hepatitis B surface antigen, large (L) hepatitis B surface antigen, viral DNA polymerase, viral core protein and e-antigen of the virus and thereby following a disruption and/or alteration of these gene sequences inactivate the HBV genome.

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 NIV 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 NIV genome surrounding the variant DNA target site. Following cleavage of the target site by the variant these homologous portions can act as a complimentary sequences in a homologous recombination reactions with the NIV genome replacing the existing NIV 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 NIV genome and/or a sequence of an exogenous 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; TRPl 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 NIV genome.

In particular, wherein the sequence which inactivates the NIV 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 NIV genome is a fragment of the NIV 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 NIV genome comprising a DNA recognition and cleavage site of said meganuclease, and thereby generate a cell comprising at least one modified NIV 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 HBV 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 NIV genomes, the present invention also relates to modifying a copy (ies) of the NIV 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 NIV genome as well as a model for screening compounds and other substances for therapeutic effects against cells comprising the modified NIV 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. For example in hepatitis B, chronic infection with this virus is associated with a greatly elevated risk of hepatocellular carcinoma. In the past this association has been explained as a side effect of the episomal hepatitis B genome
upon the hepatocyte host cells. Although this is doubtless true, recently the random genomic insertion of copies of the hepatitis B genome into the host cell genome has also been shown to be a causative factor in hepatocyte carcinoma (Goodarzi et al., 2008, Hep. Mon; 8 (2): 129-133).

Hepatocellular carcinoma is one of the most common cancers in the world and hence a treatment for this condition, using a meganuclease variant which can cleave the randomly integrated hepatitis B genome and have a therapeutic affect upon hepatocytes via one or more of mechanisms detailed herein is therefore also within the scope of the present invention as are other meganuclease variants to genomically integrated copies of virus genetic material which cause a disease phenotype.

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 NIV 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 NIV 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 NIV 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 NIV 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 hepatitis B 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-hepatitis B 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 NIV genome, as defined above, so as to generate genomically modified cells (animal precursor cell or embryo/animal or human cell) having replaced the hepatitis B 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). Meganuclease 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 NIV 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 NIV genome comprising a NIV 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 NIV genome, so as to induce cessation of a NIV genome function such as replication, attenuating or activating the NIV genome or a gene therein, introducing a mutation into a site of interest of a NIV gene, introducing an exogenous gene or a part thereof, inactivating or deleting the NIV 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 NIV 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 NIV 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 a 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 NIV 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 NIV genomes, the sequence to be introduced is a sequence which inactivates the NIV genome. All NIV 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 NIV genome or part thereof, and introduce an exogenous gene or part thereof (knock-in/gene replacement). For making knock-in NIV 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 NIV genome or to introduce a mutation into a site of interest in the NIV genome.

Inactivation of the NIV 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 NIV genome may also occur by insertion of a marker gene within an essential gene of NIV, 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 NIV 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 NIV genome, flanked by sequences sharing homologies with the NIV 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 NIV and in particular a hepatitis B 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 NIV 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 NIV 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 NIV genome contained in an at least one cell of infected
individual a double stranded cleavage at a site of interest of the NIV 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 NIV genome upon recombination between the targeting DNA and the NIV genome, as defined above. The targeting DNA is introduced into the NIV genome under conditions appropriate for introduction of the targeting DNA into the site of interest. The targeting construct may comprise sequences for deleting the NIV genome or a portion thereof and introducing the sequence of an exogenous gene of interest (gene replacement).

Alternatively, the NIV 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 NIV 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 NIV 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 hepatitis B 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 NIV and in particular hepatitis B 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 Glu 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 (β₁β₂ or β₃β₄) 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, el78; 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-Crel (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-Crel, or a variant, or a single-chain
chimeric meganuclease derived from l-Crel. 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 1, SEQ ID NO: 1). 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 l-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 HBV genome" it is intended a 20 to 24 bp sequence of the HBV genome which is recognized and cleaved by a meganuclease variant. In particular the DNA target sequence from then HBV genome is in an essential gene sequence and/or within an essential regulatory sequence and/or within an essential structural sequence of the HBV genome.

- by "first/second/third/nth 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-CreF" it is intended the wild-type I-CreI having the sequence of pdb accession code Ig9y, corresponding to the sequence SEQ ID NO: 193 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. 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: 193) 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-aa_ioa_{9}a_{8}a_{7}c_{6}-g_{5}-c_{4}c_{3}g_{2}t_{1}a_{1}a+iC{i}2g_{3}a_{4}+_{5}a_{4}+_{6}+_{7}h_{8}+_{9}+iog_{10}+ii_{1}a_{12} (SEQ ID NO: 1), 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-Crel gene or a variant thereof, which when allowed to dimerise forms a functional I-Crel 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/амino 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.

- by non-integrating virus, it is intended to mean any virus which following insertion of the virus genome into a host cell, the virus genome exists as a DNA molecule either in the nucleus or cytoplasm.

- 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 repre-
sents 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 "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 NIV 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: The HBV12 target sequences and its derivatives. 10ATT_P, 10TAG_P, 5TGG_P and 5CTT_P are close derivatives cleaved by previously obtained l-Crel variants. They differ from C1221 by the boxed motives. C1221, 10ATT_P, 10TAG_P, 5TGG_P and 5CTT_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. HBV12 is the DNA sequence located at positions 2828-2850 of the Hepatitis B genome (accession number X70185). In the HBV12.2 target, the GAAC sequence in the middle of the target is replaced with GTAC, the bases found in C1221. HBV12.3 is the palindromic sequence derived from the left part of HBV12.2, and HBV12.4 is the palindromic sequence derived from the right part of HBV12.2. As shown in the Figure, the boxed motives from 1OATTJP, 10TAG_P, 5TGG_P and 5CTT_P are found in the HBV12 series of targets.

- Figure 2: pCLS1055 plasmid map.
- Figure 3: pCLS0542 plasmid map.
- Figure 4: Cleavage of HBV12.3 target by combinatorial variants. The figure displays an example of screening of l-Crel combinatorial variants with the HBV12.3 target. Each cluster contains 6 spots: In the 4 left spots, the yeast strain containing the HBV12.3 target mated with a variant from the combinatorial library described in example 2. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions A2 and A4 are KSRSQS/DYSSR and KSSNQS/DYSSR +66H, respectively, (according to the nomenclature of Table 1).
- Figure 5: pCLSl 107 plasmid map.
- Figure 6: Cleavage of HBV12.4 target by combinatorial variants. The figure displays an example of screening of \( l-Cre \) combinatorial variants with the HBV12.4 target. Each cluster contains 6 spots: In the 4 left spots, the yeast strain containing the HBV12.4 target mated with a variant from the combinatorial library described in example 3. The right 2 spots are an internal control. H10, H11 and H12 are negative and positive controls of different strength. On the filter, the sequence of the positive variants at positions A7, D1 and G11 are KNHCQS/RYSN, KNHCQS/RYSNQ and KNHCQS/RYSYN, respectively, (according to the nomenclature of Table III and Table IV).

- Figure 7: Cleavage of the HBV12 target sequences by heterodimeric combinatorial variants. The figure displays an example of screening of combinations of \( l-CreI \) variants against the HBV12 target. Each cluster contains 4 spots: In the 2 left spots, a yeast strain co-expressing an HBV12.3 and an HBV12.4 variant mated with a yeast strain containing the HBV12 target. The right 2 spots are an internal control. The heterodimers displaying the strongest signal with the HBV12 target are observed at positions D2 and D4, corresponding to yeast co-expressing the HBV12.3 variant KSSNQS/DYSSR +66H with the HBV12.4 variants KNHCQS/RYSYQ or KNHCQS/RYSYN, respectively (according to the nomenclature of Table V).

- Figure 8: Cleavage of the HBV12 target. Example of screening against the HBV12 target of \( l-CreI \) refined variants obtained by random mutagenesis of initial variants cleaving HBV12.3 and co-expressed with a variant cutting HBV12.4. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV12 target and the HBV12.4 variant KNHCQS/RYSYN mated with a different clone from the random mutagenesis library described in example 5 (except for H10, H11 and H12: negative and positive controls of different strength). The top right spot is the HBV12.4 variant / HBV12 target strain mated with one of the initial HBV12.3 variants KSSNQS/DYSSR +66H (according to the nomenclature of Table II); the lower right spot is an internal control. On the filter, the sequence of the positive variants at positions A8 and BIO are 32Q,38C,44D,68Y,70S,75S,77R,80A (SEQ ID NO: 46) and 24F,32Q,38C,44D,68Y,70S,75S,77R (SEQ ID NO: 32) respectively.
- Figure 9: Cleavage of the HBV12 target. Example of screening against the HBV12 target of I-Crel refined variants obtained by site-directed mutagenesis of variants cleaving the HBV12.3 target and co-expressed with a variant cutting HBV12.4. Each cluster contains 6 spots: For the 4 left spots, each spot represents the yeast strain containing the HBV12 target and the HBV12.4 variant KNHCQ S/RYSNQ mated with a different clone from the site-directed mutagenesis library described in example 6. The top right spot is the HBV12.4 variant / HBV12 target strain mated with one of the HBV12.3 optimized variants 32Q,38C,44D,68Y,70S,75S,77R,80A (SEQ ID NO: 46, according to the nomenclature of Table VI); the lower right spot is an internal control. H10, H11 and H12 are negative and positive controls of different strength. The sequence of the positive variants at positions A1, A8 and Cl0 are 24F,32Q,38C,44D,68Y,70S,75S,77R,80K (SEQ ID NO: 65); 24F,32Q,38C,44D,68Y,70S,75S,77R,87L 5153G (SEQ ID NO: 61) and 24F,32Q,38C,44D,68Y,70S,75S,77R,105A,132V (SEQ ID NO: 67), respectively.

- Figure 10: Cleavage of the HBV12 target. Example of screening against the HBV12 target of 1-OeI refined variants obtained by site-directed mutagenesis of variants cleaving the HBV12.4 target and co-expressed with a variant cutting HBV12.3. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV12 target and the HBV12.3 variant KSRSQS/DYSSR mated with a different clone from the site-directed mutagenesis library described in example 7. The top right spot is the HBV12.3 variant / HBV12 target strain mated with one of the initial HBV12.4 variants KNHCQS/RYSNQ (according to the nomenclature of Table IV; the lower right spot is an internal control. H10, H11 and H12 are negative and positive controls of different strength. The sequence of the positive variants at positions A12, F9, and G1 are 32H,33C,40R,44R,68Y,70S,75N,77Q (SEQ ID NO: 69); 32H,33C,44R,68Y,70S,75Y,77Q,87L (SEQ ID NO: 71) and 19S,32H,33C,44R,68Y,70S,75D77R (SEQ ID NO: 70), respectively.

- Figure 11: The HBV8 target sequences and its derivatives. 10TGA_P, 10CAA_P, 5CTT_P and 5TCT_P are close derivatives cleaved by previously obtained 1-OeI variants. They differ from C1221 by the boxed motives. C1221, 10TGA_P, 10CAA_P, 5CTT_P and 5TCT_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. HBV8 is the DNA sequence located at positions 1908-1929 of the Hepatitis B genome (accession number X70185). In the HBV8.2 target, the ATAA sequence in the middle of the target is replaced with GTAC, the bases found in C1221. HBV8.3 is the palindromic sequence derived from the left part of HBV8.2, and HBV8.4 is the palindromic sequence derived from the right part of HBV8.2. As shown in the Figure, the boxed motives from 10TGA_P, 10CAA_P, 5CTTjP and 5TCTjP are found in the HBV8 series of targets.

- Figure 12: Cleavage of HBV8.3 target by combinatorial variants.

The figure displays an example of screening of l-Crel combinatorial variants with the HBV8.3 target. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8.3 target mated with a variant from the combinatorial library described in example 9. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions A3, A12 and F9 are KNSCRS/RYSDN, KHSCHS/RYSYN and KNSARS/RYSDN, respectively, (according to the nomenclature of Table X).

- Figure 13: Cleavage of HBV8.4 target by combinatorial variants.

The figure displays an example of screening of I-Crel combinatorial variants with the HBV8.4 target. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8.4 target mated with a variant from the combinatorial library described in example 10. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions A1, A2 are KNSHQK/QRSNK and KNSHQK/QRSNK + 163Q, respectively, (according to the nomenclature of Table XI and Table XII).

- Figure 14: pCLS1884 plasmid map.

- Figure 15: Cleavage of HBV8.4 target by combinatorial variants containing 105A and 132V mutations. The figure displays an example of screening of l-Crel combinatorial variants with the HBV8.4 target. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8.4 target mated with a variant from the combinatorial library containing the 105A and 132V substitutions described in example 11. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions A1, A2, A3 and A4 are KNSHQK/KASNI
+105A132V, KNEYQS/QSSNR + 105A132V, KNEYQS/QASNR + 105A132V and KNSHQQ/KNANI + 105A 32V respectively, (according to the nomenclature of Table XIII).

- Figure 16: Cleavage of the HBV8 target. Example of screening against the HBV8 target of I-Crel refined variants obtained by random mutagenesis of initial variants cleaving HBV8.4 and co-expressed with a variant cutting HBV8.3. Each cluster contains 6 spots: In the 4 left spots, the yeast strain containing the HBV8 target and the HBV8.3 variant KNSCRS/RYSDN mated with two different clones from the random mutagenesis library (clone 1, upper left and middle spots; clone 2, lower left and middle spots) described in example 12. H10, H11 and H12: negative and positive controls of different strength. The 2 right spots are an internal control. On the filter, the sequence of the positive variants at positions A3 and A9 are 33H,40Q,70S,75N,77K,105A,132V (SEQ ID NO: 99) and 33H,40Q,68A,70S,75N,77R,105A,132V (SEQ ID NO: 100), respectively.

- Figure 17: Cleavage of the HBV8 target. Example of screening against the HBV8 target of 1-OeI refined variants obtained by site-directed mutagenesis of variants cleaving the HBV8.4 target and co-expressed with a variant cutting HBV8.3. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8 target and the HBV8.3 variant KNSCRS/RYSDN mated with a different clone from the site-directed mutagenesis library described in example 13. The top right spot is the HBV8.3 variant / HBV8 target strain mated with one of the optimized HBV8.4 variants 33H,40Q,70S,75N,77K,105A,132V (SEQ ID NO: 99, according to the nomenclature of Table XIV); the lower right spot is an internal control. H1O, H11 and H12 are negative and positive controls of different strength. The sequence of the positive variants at positions CI1, D1O, and G8 are 19S,33H,40Q,70S,75N,77K,105A,132V (SEQ ID NO: 107); 19S,33H,40Q,70S,75N,77K,105A (SEQ ID NO: 108) and 19S,33H,40Q,43I,70S,75N,77K,105A,132V (SEQ ID NO: 106), respectively.

- Figure 18: pCLS1058 plasmid map.

- Figure 19: pCLS1768 plasmid map.

- Figure 20: HBV8 target cleavage in CHO cells. Extrachromosomal assay in CHO cells for heterodimers displaying strong cleavage activity against the
HBV8 target as described in example 14. OD values indicated were observed 3 hours after lysis/revelation buffer addition. HD1 represents the results obtained with co-expression of the HBV8.3 variant 33C,38R,44R,68Y,70S,77N (SEQ ID NO: 85) with HBV8.4 variant 19S,33H,40Q,43I,70S,75N,77K,105A,132V (SEQ ID NO: 106). HD2 represents the results obtained with co-expression of the HBV8.3 variant 33C,38R,44R,68Y,70S,77N (SEQ ID NO: 85) and HBV8.4 variant 19S,33H,40Q,70S,75N,77K,105A,132V (SEQ ID NO: 107). I-SceI and empty vector are presented as positive and negative controls, respectively.

- Figure 21: The HBV3 target sequences and its derivatives.

10TGC_P, 10TCTJP, 5TAC_P and 5TCC_P are close derivatives cleaved by previously obtained I-CreI variants. They differ from C1221 by the boxed motives. C1221, 10TGC_P, 10TCTJ\ 5TACJP and 5TCC_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. HBV3 is the DNA sequence located at positions 2216-2237 of the Hepatitis B genome (accession number M38636). In the HBV3.2 target, the TTTT sequence in the middle of the target is replaced with GTAC, the bases found in C1221. HBV3.3 is the palindromic sequence derived from the left part of HBV3.2, and HBV3.4 is the palindromic sequence derived from the right part of HBV3.2. HBV3.5 and HBV3.6 are pseudo-palindromic targets similar to HBHV3.3 and HBV3.4 except that they contain the tttt sequence at positions -2 to 2. As shown in the Figure, the boxed motives from 10TGC_P, 10TCT_P, 5TAC_P and 5TCCJP are found in the HBV3 series of targets.

- Figure 22: Cleavage of HBV3.3 target by combinatorial variants. The figure displays an example of screening of I-CreI combinatorial variants with the HBV3.3 target. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV3.3 target mated with a variant from the combinatorial library described in example 16. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions C9, D8 and H8 are KNSCRS/AYSRT, KNSSRQ/AYSRI and KNSCSS/NYSRY, respectively, (according to the nomenclature of Table XVI and XVII).

- Figure 23: Cleavage of HBV3.4 target by combinatorial variants. The figure displays an example of screening of I-CreI combinatorial variants with the
HBV3.4 target. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV3.4 target mated with a variant from the combinatorial library described in example 17. The right 2 spots are an internal control. On the filter, the sequence of the positive variants at positions Cl, E3 and G8 are KNSCYS/KYSNV +45M, KNSSYS/KHNNI and KNSGYS/KYSNV +45M, respectively, (according to the nomenclature of Table XVIII and Table XIX).

- Figure 24: Cleavage of the HBV3.2 target sequences by heterodimeric combinatorial variants. The figure displays an example of screening of combinations of I-Crel variants against the HBV3.2 target. Each cluster contains 4 spots: In the 2 left spots; a yeast strain co-expressing the HBV3.3 and HBV3.4 combinatorial variants was mated with a yeast strain containing the HBV3 target as described in example 18. The right 2 spots are an internal control. All heterodimers tested resulted in strong cleavage of the HBV3.2 target.

- Figure 25: Cleavage of the HBV3.5 target. Example of screening against the HBV3.5 target of I-Crel refined variants obtained by random mutagenesis of initial variants cleaving HBV8.3. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8.5 target mated with a clone from the random mutagenesis library described in example 19. H10, H11 and H12: negative and positive controls of different strength. The top right spot is the HBV3.5 target strain mated with one of the initial HBV3.3 variants KNSCRS/AYSRT (according to the nomenclature of Table XVII). The right lower spot is an internal control. On the filter, the sequence of the positive variants at positions A4 and F12 are 26R,33C,38S,44N,68Y,70S,75R,77Y,81T (SEQ ID NO: 146) and 33C,38R,44A,68Y,70S,75R,77T,132V (SEQ ID NO: 147), respectively.

- Figure 26: Cleavage of the HBV3.6 target. Example of screening against the HBV3.6 target of I-Crel refined variants obtained by random mutagenesis of initial variants cleaving HBV8.4. Each cluster contains 4 spots: In the 2 left spots, the yeast strain containing the HBV8.6 target mated with two different clones from the random mutagenesis library described in example 20. H10, H11 and H12: negative and positive controls of different strength. The top right spot is the HBV3.6 target strain mated with one of the initial HBV3.4 variants KNSGYS/KYSNY (according to the nomenclature of Table XVIII). The right lower spot is an internal control. On the
filter, the sequence of the positive variants at positions B1 and G4 are 33C, 38Y, 44K, 64I, 68Y, 70S, 75N, 77Y, 85R (SEQ ID NO: 157) and 33S, 38Y, 44K, 45M, 68Y, 70S, 75N, 77V, 86T (SEQ ID NO: 161), respectively.

- Figure 27: Cleavage of the HBV3 target sequences by optimized heterodimeric variants. The figure displays an example of screening of I-OeI variants against the HBV3 target. Each cluster contains 4 spots: In the 2 left spots and the upper right spot, a yeast strain co-expressing an HBV3.3 and an HBV3.4 variant mated with a yeast strain containing the HBV3 target. The lower right spot is an internal control. The heterodimers displaying the strongest signal with the HBV3 target are observed at positions A1 and A11, corresponding to yeast co-expressing the HBV3.3 variant 26R, 33C, 38S, 44N, 68Y, 70S, 75R, 77Y, 81T (SEQ ID NO: 146) with the HBV3.4 variants 33S, 38Y, 44K, 68Y, 70S, 75N, 77L (SEQ ID NO: 162) and 2D, 33S, 38Y, 44K, 68Y, 70S, 75N, 77Y, 140M (SEQ ID NO: 155), respectively.

- Figure 28: Cleavage of the HBV3 target. Example of secondary screening against the HBV3 target of l-Crel refined variants obtained by random mutagenesis of variants cleaving the HBV3.4 target and co-expressed with a variant cutting HBV3.3. Extrachromosomal assay in CHO cells for heterodimers displaying cleavage activity against the HBV3 target as described in example 22. HBV3.4 variants, both the initial (33S, 38Y, 44K, 68Y/70S, 75N, 77L (SEQ ID NO: 162)) and optimized (see Table XXIV) variants, were co-expressed with the HBV3.3 variant 26R, 33C, 38S, 44N, 68Y, 70S, 75R, 77Y, 81T (SEQ ID NO: 146) and examined for their ability to cleave the HBV3 target. OD values indicated were observed 3 hours after lysis/revelation buffer addition. l-Scel is presented as a positive control.

- Figure 29: Cleavage of the HBV3 target. Example of secondary screening against the HBV3 target of I-Crel refined variants obtained by random mutagenesis of variants cleaving the HBV3.3 target and co-expressed with a variant cutting HBV3.4. Extrachromosomal assay in CHO cells for heterodimers displaying cleavage activity against the HBV3 target as described in example 23. HBV3.3 variants, both the initial (26R, 33C, 38S, 44N, 68Y, 70S, 75R, 77Q, 81T (SEQ ID NO: 146)) and optimized (see Table XXV) variants, were co-expressed with the HBV3.4 variant 19S, 33C, 38Y, 44K, 68Y, 70S, 75N, 77Q (SEQ ID NO: 84) and examined for their ability to cleave the HBV3 target. OD values indicated were observed 3 hours
after lysis/revelation buffer addition. *l-Scel* is presented as a positive control.

- Figure 30: Cleavage of the HBV3 target. Example of secondary screening against the HBV3 target of *l-Crel* refined variants obtained by site-directed mutagenesis. Extrachromosomal assay in CHO cells for heterodimers displaying cleavage activity against the HBV3 target as described in example 24. HBV3.3 variant containing site-directed mutations (3.3_R5) was co-expressed with either the initial HBV3.4 variant (3.4_A7) or one of four HBV3.4 variants containing site-directed mutations (3.4_R2, 3.4_R4, 3.4_R5, 3.4_R6; see Table XXVI) variants, and examined for their ability to cleave the HBV3 target in comparison to the original HBV3 heterodimer (3.3_F1/ 3.4_A7). OD values indicated were observed 3 hours after lysis/revelation buffer addition. *l-Scel* is presented as a positive control.

- Figure 31: Cleavage of the HBV3 target. Example of screening of *l-Crel* single chain molecules for cleavage activity against the HBV3 target. Extrachromosomal assay in CHO cells for single chain molecules displaying cleavage activity against the HBV3 target as described in example 25. Two single-chain molecules (SC_34 and SC_OH_34) were examined for their ability to cleave the HBV3 target in comparison to the HBV3 heterodimer (3.3JR.5/ 3.4_R4). OD values indicated were observed 3 hours after lysis/revelation buffer addition. *l-Scel* and empty vector are presented as positive and negative controls, respectively.

- Figure 32: shows schematic representation of HBV as an enveloped DNA-containing virus. The viral particle consists of an inner core plus an outer surface coat.

- Figure 33: shows a schematic representation of the HBV genome.

- Figure 34: shows a structural representation of a LAGLIDADG enzyme in combination with its DNA target.

- Figure 35: shows a schematic representation of the coding sequences present in the HBV genome and the HBV3, 8 and 12 targets identified in the HBV genome for which meganuclease variants according to the present invention have been made.

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

**Example 1:** Strategy for engineering novel meganucleases

**5** cleaving the HBV12 target from the Hepatitis B genome

HBV12 is a 22 bp (non-palindromic) target located in the coding sequence of the RNA dependent DNA polymerase gene in the Hepatitis B genome. The target sequence corresponds to positions 2828-2850 of the Hepatitis B genome (accession number X70185, Figure 35).

The HBV12 sequence is partly a patchwork of the 10ATTJP, 10TAGJP, 5TGG_P and 5_CTTJP targets (Figure 1) which are cleaved by previously identified meganucleases, obtained as described in International PCT Applications WO 2006/097784 and WO 2006/097853; Arnould *et al.*, J. Mol. Biol., 2006, 355, 443-458; Smith *et al.*, Nucleic Acids Res., 2006. Thus the inventors set out to determine whether HBV12 could be cleaved by combinatorial variants resulting from these previously identified meganucleases.

The 10ATTJP, 10TAG_P, 5TGG_P and 5_CTT_P target sequences are 24 bp derivatives of C1221, a palindromic sequence cleaved by l-Crel (*Arnould et al.*, precited). However, the structure of l-Crel 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 HBV12 series of targets were defined as 22 bp sequences instead of 24 bp. HBV12 differs from C1221 in the 4 bp central region. According to the structure of the l-Crel protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the l-Crel 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 gaac sequence in -2 to 2 was first substituted with the gtac sequence from C1221, resulting
in target HBV12.2 (Figure 1). Then, two palindromic targets, HBV12.3 and HBV12.4, were derived from HBV12.2 (Figure 1). Since HBV12.3 and HBV12.4 are palindromic, they should be cleaved by homodimeric proteins. Thus, proteins able to cleave the HBV12.3 and HBV12.4 sequences as homodimers were first designed (examples 2 and 3) and then co-expressed to obtain heterodimers cleaving HBV12 (example 4). Heterodimers cleaving the HBV12 target could be identified. In order to improve cleavage activity for the HBV12 target, a series of variants cleaving HBV12.3 and HBV12.4 was chosen, and then refined. The chosen variants were subjected to random or site-directed mutagenesis, and used to form novel heterodimers that were screened against the HBV12 target (examples 5, 6 and 7). Strong cleavage activity of the HBV12 target could be observed for these heterodimers.

**Example 2: Identification of meganucleases cleaving HBV12.3**

This example shows that I-Crel variants can cut the HBV12.3 DNA target sequence derived from the left part of the HBV12.2 target in a palindromic form (Figure 1). 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 HBV12.3 will be noted tattctggtJP).

HBV12.3 is similar to 10ATT_P at positions ±1, ±2, ±8, ±9, and ±10 and to 5TGG_P at positions ±1, ±2, ±3, ±4, ±5 and ±10. It was hypothesized that positions ±6, ±7 and ±11 would have little effect on the binding and cleavage activity. Variants able to cleave the 10ATT_P target were obtained by mutagenesis of *I-Crel* N75 or D75, at positions 28, 30, 32, 33, 38, 40 and 70, as described previously in Smith *et al.* Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156. Variants able to cleave 5TGG_P were obtained by mutagenesis on I-Crel N75 at positions 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, el49; International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156.

Both sets of proteins are mutated at position 70. However, the existence of two separable functional subdomains was hypothesized. This implies that this position has little impact on the specificity at bases 10 to 8 of the target.
Therefore, to check whether combined variants could cleave the HBV12.3 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TGG_P were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 1OATTJP.

A) Material and Methods

a) Construction of target vector

The target was cloned as follows: an oligonucleotide corresponding to the HBV12.3 target sequence flanked by gateway cloning sequences was ordered from PROLIGO: 5’ tggcatataagttatatcttgggtacccaagaatatactgtgtca 3’ (SEQ ID NO: 10). 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 2). Yeast reporter vector was transformed into Saccharomyces cerevisiae strain FYBL2-7B (MAT a, ura3Δ 851, trpl Δ 63, leu2Δ l, lys2Δ 202), resulting in a reporter strain.

b) Mating of meganuclease expressing clones and screening in yeast

I-Crel variants cleaving 1OATTJP or 5TGG_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 10ATT_P and 5TGG_P targets. In order to generate I-Crel 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 l-Crel coding sequence. For both the 5’ and 3’ end, PCR amplification is carried out using primers (GalIOF 5’-gcaacctttagtgctgacacatacaggg-3’ (SEQ ID NO: 11) or GalIOR 5’-accaacctttagtgctgacacatacaggg-3’ (SEQ ID NO: 12)) specific to the vector (pCLS0542, Figure 3) and primers (assF 5’-ctannttgagttctctcagacctactatg-3’ (SEQ ID NO: 13) or assR 5’-aatccctaactggtgctgacacatacaggg-3’ (SEQ ID NO: 14)), where nnn codes for residue 40, specific to the l-Crel 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 GalIOF and assR or assF and GalIOR 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 3) linearized by digestion with Ncol and Eagl were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3Δ 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 gridded (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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-Crel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TGG_P with the
mutations at 28, 30, 32, 33, 38 and 40 from proteins cleaving 10ATT_P on the I-Crel scaffold, resulting in a library of complexity 94. Examples of combinatorial variants are displayed in Table I. This library was transformed into yeast and 2232 clones (23.7 times the diversity) were screened for cleavage against the HBV 12.3 DNA target (tattt_geg_P, SEQ ID NO: 8). Six positive clones were found, which after sequencing turned out to correspond to six different novel endonuclease variants (Table II). Examples of positives are shown in Figure 4. All six variants display non parental combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77. Such combinations likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be I-Oel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.

**Table I:** Panel of variants* theoretically present in the combinatorial library

<table>
<thead>
<tr>
<th>Amino acids at positions 44, 68, 70, 75 and 77</th>
<th>KDSRQS</th>
<th>KSSMQS</th>
<th>KSSCQS</th>
<th>KNNQGS</th>
<th>KNYYS</th>
<th>KNQYG</th>
<th>KQSTQS</th>
<th>KNDYCS</th>
<th>KNSNTS</th>
<th>KCSQQS</th>
<th>KNSCAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDSRQS stands for D44, Y68, S70, S75 and R77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYSER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YRSIV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Only 22 out of the 94 combinations are displayed. None of them were identified in the positive clones.

**Table II:** I-Crel variants capable of cleaving the HBV12.3 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40, 44, 68, 70, 75 and 77 of the I-Crel variants (ex: KSRSQS/DYSSR stands for K28, S30, R32, S33, Q38, S40, D44, Y68, S70, S75 and R77)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KSRSQS/DYSSR</td>
<td>15</td>
</tr>
<tr>
<td>KQNYCS/DYSSR</td>
<td>16</td>
</tr>
<tr>
<td>KNKTQS/DYSSR</td>
<td>17</td>
</tr>
<tr>
<td>KRYSQS/DYSSR</td>
<td>18</td>
</tr>
<tr>
<td>KSSNQS/DYSSR +66H</td>
<td>19</td>
</tr>
<tr>
<td>KNAAGS/DYSSR +89A</td>
<td>20</td>
</tr>
</tbody>
</table>
Example 3: Identification of meganucleases cleaving HBV12.4

This example shows that 1-Oel variants can cleave the HBV 12.4 DNA target sequence derived from the right part of the HBV 12.2 target in a palindromic form (Figure 1). 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, HBV 12.4 will be called tgtgctcttttgt_P).

HBV 12.4 is similar to 5CTTJP at positions ±1, ±2, ±3, ±4, ±5 and ±9 and to 1OTAGJP at positions ±1, ±2, ±4, ±8, ±9 and ±10. It was hypothesized that positions ±6, ±7 and ±11 would have little effect on the binding and cleavage activity. Variants able to cleave 5CTT_P were obtained by mutagenesis of I-Crel N75 at positions 24, 44, 68, 70, 75 and 77, as described previously (Arnould et al., J. Mol. Biol., 2006, 355, 443-458; Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156). Variants able to cleave the 10TAG_P target were obtained by mutagenesis of I-Oel N75 or D75, at positions 28, 30, 32, 33, 38 and 40, as described previously in Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156.

Mutations at positions 24 found in variants cleaving the 5CTT_P target will be lost during the combinatorial process. But it was hypothesized that this will have little impact on the capacity of the combined variants to cleave the HBV 12.4 target.

Therefore, to check whether combined variants could cleave the HBV12.4 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5CTT_P were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TAG_P.

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 2, with the exception that an oligonucleotide corresponding to the HBV12.4 target sequence was used: 5’ tgcatacaagttttgtgctcttgtacaagagctacacaatcgtctgtca 3’ (SEQ ID NO: 21).
b) Construction of combinatorial variants

l-Cre\textsuperscript{i} variants cleaving 10TAG\_P or 5CTT\_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. MoI. Biol., 2006, 355, 443-458; International PCT Applications WO 2006/097784 and WO 2006/097853, respectively for the 1OTAGJP and 5CTTJP targets. In order to generate l-Crel derived coding sequences containing mutations from both series, separate overlapping PCR reactions were carried out that amplify the 5\textsuperscript{'} end (aa positions 1-43) or the 3\textsuperscript{'} end (positions 39-167) of the l-Crel coding sequence. For both the 5\textsuperscript{'} and 3\textsuperscript{'} end, PCR amplification is carried out using primers (GalIOF 5\textsuperscript{-}gcaaccttgattggagacttgacc-3\textsuperscript{'} (SEQ ID NO: 11) or GalIOR 5\textsuperscript{-}acaaccttgattggagacttgacc-3\textsuperscript{'} (SEQ ID NO: 12)) specific to the vector (pCLS1107, Figure 5) and primers (assF 5\textsuperscript{-}ctannnttgaccttt-3\textsuperscript{'} (SEQ ID NO: 13) or assR 5\textsuperscript{-}aaaggtcaannntag-3\textsuperscript{'} (SEQ ID NO: 14)), where nnn codes for residue 40, specific to the l-Cre\textsuperscript{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 GalIOF and assR or assF and GalIOR 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 5) linearized by digestion with DraIII and NgoMLV were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MAT\textalpha, trpl \Delta 63, leu2\Delta 1, his3\Delta 200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymiol., 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. MoI. Biol., 2006, 355, 443-458). Mating was performed using a colony griddler (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm\textsuperscript{2}). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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, including 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.

d) Sequencing of variants

Experimental procedure is as described in example 2.

B) Results

1-OeI combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5CTT_P with the mutations 28, 30, 32, 33, 38 and 40 from proteins cleaving 10TAG_P on the I-Crel scaffold, resulting in a library of complexity 720. Examples of combinatorial variants are displayed in Table III. This library was transformed into yeast and 2232 clones (3.1 times the diversity) were screened for cleavage against the HBV12.4 DNA target (gtagctcttgt_P, SEQ ID NO: 9). A total of 664 positive clones were found to cleave HBV12.4. Sequencing and validation by secondary screening of 93 of the I-Crel variants resulted in the identification of 51 different novel endonucleases. Examples of positives are shown in Figure 6. The sequence of several 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 Table IV). Such variants likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be I-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.
Table III: Panel of variants* theoretically present in the combinatorial library

<table>
<thead>
<tr>
<th>Amino acids at positions 44, 68, 70, 75 and 77 (ex: RYSDN stands for R44, Y68, S70, D75 and N77)</th>
<th>Amino acids at positions 28, 30, 32, 33, 38 and 40 (ex: KNHCQS stands for K28, N30, H32, C33, Q38 and S40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RYSDN</td>
<td>+</td>
</tr>
<tr>
<td>KTSDR</td>
<td>+</td>
</tr>
<tr>
<td>RYSYN</td>
<td>+</td>
</tr>
<tr>
<td>KESNR</td>
<td></td>
</tr>
<tr>
<td>RYSDQ</td>
<td>+</td>
</tr>
<tr>
<td>QASQR</td>
<td></td>
</tr>
<tr>
<td>RNSNN</td>
<td></td>
</tr>
<tr>
<td>RYSYN</td>
<td>+</td>
</tr>
<tr>
<td>RINSNQ</td>
<td></td>
</tr>
<tr>
<td>RYATQ</td>
<td></td>
</tr>
<tr>
<td>KASDV</td>
<td></td>
</tr>
</tbody>
</table>

* Only 220 out of the 720 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HBV 12.4 target was found among the identified positives.

Table IV: I-Crel variants with additional mutations capable of cleaving the HBV12.4 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-Crel variants (ex: KRGYQS/KYSNI stands for K28, R30, G32, Y33, S38, K44, Y68, S70, N75 and I77)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNHCQS/RYSYN</td>
<td>22</td>
</tr>
<tr>
<td>KNHCQS/RYSNO +117K</td>
<td>23</td>
</tr>
<tr>
<td>KNHCQS/RYSNO</td>
<td>24</td>
</tr>
<tr>
<td>KNHCQS/RYSNN</td>
<td>25</td>
</tr>
<tr>
<td>KNHCQS/RYSNO +151A</td>
<td>26</td>
</tr>
<tr>
<td>KNHCQS/RYSNO</td>
<td>27</td>
</tr>
</tbody>
</table>

Example 4: Making of meganucleases cleaving HBV12

1-Oel variants able to cleave each of the palindromic HBV 12.2 derived targets (HBV12.3 and HBV12.4) were identified in example 2 and example 3. Pairs of such variants (one cutting HBV12.3 and one cutting HBV12.4) 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 HBV12 target, which differs from the HBV 12.2 sequence by 2 bp at positions 1 and 2.

**A) Materials and Methods**

a) **Construction of target vector**

The experimental procedure is as described in example 2, with the exception that an oligonucleotide corresponding to the HBV12 target sequence: 5’ tggcataaagtttatattcttgggaacaagagctacacaatcgtctgtca3’ (SEQ ID NO: 28) was used.

b) **Co-expression of variants**

Yeast DNA was extracted from variants cleaving the HBV 12.3 target (pCLS542 expression vector) as well as those cleaving the HBV 12.4 target (pCLSI 107 expression vector) using standard protocols and were used to transform *E. coli*. Plasmid DNA derived from a HBV12.3 variant and a HBV12.4 variant was then co-transformed into the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3 Δ 200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Transformants were selected on synthetic medium lacking leucine and containing G418.

c) **Mating of meganuclease co-expressing clones and screening in yeast**

Mating was performed using a colony griddler (QpixII, Genetix). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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, including 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 HBV 12.4 target (7 variants chosen among those described in Table III and Table IV) and the six variants cleaving the HBV12.3 target (described in Table II) resulted in weak cleavage of the HBV12 target in certain cases (Figure 7). Functional combinations are summarized in Table V.

Table V: Cleavage of the HBV12 target by the heterodimeric variants

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-Crel variants cleaving the HBV12.3 target (ex: KRYSQS/DYSSR stands for K28, R30, Y32, S33, Q38, S40/D44, Y68, S70, S75 and R77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRYSQS/ DYSSR (SEQ ID NO: 18)</td>
</tr>
<tr>
<td>KNCQCS/ RYSYN (SEQ ID NO: 39)</td>
</tr>
<tr>
<td>KNCQCS/ RYSYN (SEQ ID NO: 22)</td>
</tr>
<tr>
<td>KNCQCS/ RYSNQ +117K (SEQ ID NO: 23)</td>
</tr>
<tr>
<td>KNCQCS/ RYSNQ (SEQ ID NO: 24)</td>
</tr>
<tr>
<td>KNCQCS/ RYSNN (SEQ ID NO: 25)</td>
</tr>
<tr>
<td>KNCQCS/ RYSDQ +151A (SEQ ID NO: 26)</td>
</tr>
<tr>
<td>KNCQCS/ RYSDQ (SEQ ID NO: 27)</td>
</tr>
</tbody>
</table>

+ indicates a functional combination

*indicates that the combination weakly cuts the HBV12 target.

Example 5: Improvement of meganucleases cleaving HBV12 by random mutagenesis of proteins cleaving HBV12.3 and assembly with proteins cleaving HBV12.4

I-Crel variants able to cleave the HBV12 target by assembly of variants cleaving the palindromic HBV12.3 and HBV12.4 target have been previously identified in example 4. However, these variants display weak activity with the HBV12 target.

Therefore five combinatorial variants cleaving HBV12.3 were mutagenized, and variants were screened for cleavage activity of HBV12 when co-expressed with a variant cleaving HBV12.4. According to the structure of the I-Oel
protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the l-Cre 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, only one of the two components of the heterodimers cleaving HBV12 was mutagenized.

Thus, in a first step, proteins cleaving HBV12.3 were mutagenized, and in a second step, it was assessed whether they could cleave HBV12 when co-expressed with a protein cleaving HBV12.4.

A) Material and Methods

a) Construction of libraries by random mutagenesis

Random mutagenesis was performed on a pool of chosen variants, by PCR using Mn²⁺. PCR reactions were carried out that amplify the I-Crel coding sequence using the primers preATGCreFor (5'-gcataaatcattacttcatagacgcaaaacaaatatacagcggccttgccacc-3'; SEQ ID NO: 30) and ICrelpostRev (5'-ggctcgaggagctcgtctagaggatcgctcgagttatcagtcggccgc-3'; SEQ ID NO: 31), which are common to the pCLS0542 (Figure 3) and pCLS1107 (Figure 5) vectors. Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLS0542) linearized by digestion with Ncol and Eagl were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATa, trpl Δ 63, leu2Δ l, his3Δ 200) 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-Crel variant were generated by in vivo homologous recombination in yeast.

b) Variant-target yeast strains, screening and sequencing

The yeast strain FYBL2-7B [MATa, ura3Δ 851, trpl Δ 63, leu2Δ l, lys2Δ 202] containing the HBV12 target in the yeast reporter vector (pCLS1055, Figure 2) was transformed with variants, in the kanamycin vector (pCLSI 107), cutting the HBV12.4 target, using a high efficiency LiAc transformation protocol. Variant-target yeast strains were used as target strains for mating assays as described in
example 4. Positives resulting clones were verified by sequencing (MILLEGEN) as described in example 2.

**B) Results**

Five variants cleaving HBV12.3 (1-OeI 30R,32Y,33S,44D,68Y,70S,75S,77R, l-Crel 30S,32R,33S,44D,68Y,70S,75S,77R, 1-Oel 32A,33A,38G,44D,68Y,70S,75S,77R,89A, I-Crel 32Q,38C,44D,68Y,70S,75S,77R and l-Crel 32K,33T,44D,68Y,70S,75S,77R, also called KRYSQS/DYSSR (SEQ ID NO: 18), KSRSQS/DYSSR (SEQ ID NO: 15), KNAAGS/DYSSR +89A (SEQ ID NO: 20), KNQYCS/DYSSR (SEQ ID NO: 16) and KNKTQS/DYSSR (SEQ ID NO: 17) respectively, according to the nomenclature of Table II, were pooled, randomly mutagenized and transformed into yeast. 2232 transformed clones were then mated with a yeast strain that contains (i) the HBV12 target in a reporter plasmid (ii) an expression plasmid containing a variant that cleaves the HBV1 2.4 target (I-Crel 32H,33C,44R,68Y,70S,75N,77Q or KNHCQS/RYSNQ (SEQ ID NO: 24) according to the nomenclature of Table IV. After mating with this yeast strain, 156 clones were found to cleave the HBV12 target more efficiently than the original variant. Thus, 156 positives contained proteins able to form heterodimers with KNHCQS/RYSNQ with an improved cleavage activity for the HBV1 2 target. An example of positives is shown in Figure 8. Sequencing of the strongest 93 positive clones indicates that 29 distinct variants were identified (examples listed in Table VI).
Table VI: Functional variant combinations displaying improved cleavage activity for HBV12.

<table>
<thead>
<tr>
<th>Variant HBV12.4</th>
<th>Optimized* Variants HBV12.3 (SEQ ID NO: 32 to 48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-Crel 28K 30N 32H 33C 38Q 40S 44R 68Y 70S 75N 77Q</td>
<td>24F 32Q 38C 44D 68Y 70S 75S 77R</td>
</tr>
<tr>
<td></td>
<td>30R 32Y 33S 44D 64A 68Y 70S 75S 77R</td>
</tr>
<tr>
<td></td>
<td>30S 32H 33S 44D 68Y 70S 75S 77R</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 81V</td>
</tr>
<tr>
<td></td>
<td>7E 30S 32R 33S 44D 68Y 70S 75S 77R</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 107R</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 153G</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 81T 160E</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 81V 162P</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 89A</td>
</tr>
<tr>
<td></td>
<td>30S 32R 33S 44D 68Y 70S 75S 77R 99R</td>
</tr>
<tr>
<td></td>
<td>32G 38G 44D 64A 68Y 70S 75S 77R 89A</td>
</tr>
<tr>
<td></td>
<td>32G 38G 44D 68Y 70S 75S 77R</td>
</tr>
<tr>
<td></td>
<td>32G 38G 44D 68Y 70S 75S 77R 81T</td>
</tr>
<tr>
<td></td>
<td>32Q 38C 44D 68Y 70S 75S 77R 80A</td>
</tr>
<tr>
<td></td>
<td>32R 33S 44D 68Y 70S 75S 76F 77R</td>
</tr>
<tr>
<td></td>
<td>32R 38C 44D 68Y 70S 75S 77R 109V 157K</td>
</tr>
</tbody>
</table>

* Mutations resulting from random mutagenesis are in bold.

Example 6: Improvement of meganucleases cleaving HBV12 by site-directed mutagenesis of proteins cleaving HBV12.3 and assembly with proteins cleaving HBV12.4

The optimized I-Crel variants cleaving HBV 12.3 described in Table VI that resulted from random mutagenesis as described in example 5 were further mutagenized by introducing selected amino-acid substitutions in the proteins and screening for more efficient variants cleaving HBVI 2 in combination with a variant cleaving HBV 12.4.

Six amino-acid substitutions have been found in previous studies to enhance the activity of I-Crel derivatives: these mutations correspond to the replacement of Glycine 19 with Serine (G19S), Phenylalanine 54 with Leucine (F54L), Glutamic acid 80 with Lysine (E80K), Phenylalanine 87 with Leucine (F87L), Valine 105 with Alanine (V105A) and Isoleucine 132 with Valine (I132V). These
mutations were introduced into the coding sequence of proteins cleaving HBV12.3, and the resulting proteins were tested for their ability to induce cleavage of the HBV12 target, upon co-expression with a variant cleaving HBV 12.4.

A) **Material and Methods**

a) **Site-directed mutagenesis**

A site-directed mutagenesis library was created by PCR on a pool of chosen variants. For example, to introduce the G19S substitution into the coding sequence of the variants, two separate overlapping PCR reactions were carried out that amplify the 5’ end (residues 1-24) or the 3’ end (residues 14-167) of the I-Crel coding sequence. For both the 5’ and 3’ end, PCR amplification is carried out using a primer with homology to the vector (GalIOF 5’-gcaacctttgtgctgacacattag-3’ (SEQ ID NO: 11) or GalIOR 5’-acaaccttgattggagacttgacc-3’(SEQ ID NO: 12)) and a primer specific to the I-Crel coding sequence for amino acids 14-24 that contains the substitution mutation G19S (G19SF 5’-gccggctttgtggactctgacggtagcatcatc-3’ (SEQ ID NO: 49) or G19SR 5’-gatgatgctaccgtcagagtccacaaagccggc-3’(SEQ ID NO: 50)). The resulting PCR products contain 33bp of homology with each other. The PCR fragments were purified.

The same strategy was used with the following pairs of oligonucleotides to introduce the F54L, E80K, F87L, V105A and 1132V substitutions, respectively:

* F54LF: 5’-accagcgcgtgtctgctgagacaaactagtg-3’ and F54LR: 5’-cactagttttggcagcaccaacggcgtggtg-3’ (SEQ ID NO: 51 and 52);

* E80KF: 5’-ttaagcaaaatcaagccgctgcacaacttcctg-3’ and E80KR: 5’-caggaagttgtgcagcggcttgattttgcttaa-3’ (SEQ IDNO: 53 and 54);

* F87LF: 5’-aacgctgtcacaacatctgcactcaactgcag-3’ and F87LR: 5’-ctgcaatgactcagcgcttgctggctggtg-3’ (SEQ ID NO: 55 and 56);

* V105AF: 5’-aacccgctgtcacaacatctgcactcaactgcag-3’ and V105AR: 5’-ttgcaatgactcagcgcttgctggctggtg-3’ (SEQ ID NO: 57 and 58);

* I132VF: 5’-acctgggtggatcaggttgcagctctgaacgat-3’ and I132VR: 5’-atcgctagctcacaacatctgcactcaactgcag-3’ (SEQ ID NO: 59 and 60).

The two overlapping PCR fragments for each of the six site-directed mutations were pooled and a total of approximately 25ng was combined with 75ng of
vector DNA (pCLS0542, Figure 3) linearized by digestion with Ncol and Eagl. The DNA was then used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATa, trpl Δ 63, leu2Δ 1, his3Δ 200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods EnzymoL, 2002, 350, 87-96). Intact coding sequences containing the one or more of the above described site directed substitutions are generated by *in vivo* homologous recombination in yeast.

c) **Mating of meganuclease expressing clones and screening in yeast**

The experimental procedure is as described in example 5.

d) **Sequencing of variants**

The experimental procedure is as described in example 2.

B) Results

A library containing site-directed mutations (GI9S, F54L, E80K, F87L, V105A, I132V) was constructed from a pool of 7 variants cleaving HBV12.3 (32Q,38C,44D,68Y,70S,75S,77R,80A (SEQ ID NO: 46), 24F,32Q,38C,44D,68Y,70S,75S,77R (SEQ ID NO: 32), 30S,32R,33S,44D,68Y,70S,75S,77R,81V,162P (SEQ ID NO: 40), 30S,32R,33S,44D,68Y,70S,75S,77R,153G (SEQ ID NO: 38), 32R,33S,44D,68Y,70S,75S,76F,77R (SEQ ID NO: 47), 7E,30S,32R,33S,44D,68Y,70S,75S,77R (SEQ ID NO: 36) and 30S,32H,33S,44D,68Y,70S,75S,77R (SEQ ID NO: 34) according to the nomenclature of Table VI). The library was transformed into yeast and 1674 individual clones were picked and mated with a yeast strain that contains (i) the HBV12.2 target in a reporter plasmid (ii) an expression plasmid containing a variant that cleaves the HBV12.4 target (32H,33C,44R,68Y,70S,75N,77Q or KNHCQS/RYSNQ (SEQ ID NO: 24) according to the nomenclature of Table IV).

After mating with this yeast strain, 122 clones were found to cleave the HBV12 target more efficiently than the original variants. An example of positives is shown in Figure 9. The sequence of eight of the best I-Crel variants cleaving the HBV12 target when forming a heterodimer with the KNHCQS/RYSNQ variant are listed in Table VII.
Table VII: Functional variant combinations displaying strong cleavage activity for HBV12.

<table>
<thead>
<tr>
<th>VARIANT HBV12.4</th>
<th>Optimized* Variants HBV12.3 (SEQ ID NO: 61 to 68)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24F 32Q 38C 44D 68Y 70S 75S 77R 87L 153G</td>
</tr>
<tr>
<td></td>
<td>24F 32Q 38C 44D 68Y 70S 75S 77R 132V</td>
</tr>
<tr>
<td></td>
<td>19S 32Q 38C 44D 68Y 70S 75S 77R 81V</td>
</tr>
<tr>
<td></td>
<td>7E 24F 32Q 38C 44D 68Y 70S 75S 77R 80K</td>
</tr>
<tr>
<td></td>
<td>24F 32Q 38C 44D 68Y 70S 75S 77R 80K</td>
</tr>
<tr>
<td></td>
<td>24F 32Q 38C 44D 68Y 70S 75S 77R 105A 132V</td>
</tr>
<tr>
<td></td>
<td>2F 32Q 38C 44D 68Y 70S 75S 77R 80A 153G</td>
</tr>
</tbody>
</table>

* Mutations resulting from site-directed mutagenesis are in bold.

Example 7: Improvement of meganucleases cleaving HBV12 by site-directed mutagenesis of proteins cleaving HBV12.4 and assembly with proteins cleaving HBV12.3

The initial I-Crel variants cleaving HBV12.4 described in Tables III and IV were mutagenized by introducing selected amino-acid substitutions in the proteins and screening for more efficient variants cleaving HBV12 in combination with a variant cleaving HBV12.3.

Six amino-acid substitutions have been found in previous studies to enhance the activity of I-Crel derivatives: these mutations correspond to the replacement of Glycine 19 with Serine (G19S), Phenylalanine 54 with Leucine (F54L), Glutamic acid 80 with Lysine (E80K), Phenylalanine 87 with Leucine (F87L), Valine 105 with Alanine (V105A) and Isoleucine 132 with Valine (I132V). These mutations were introduced into the coding sequence of proteins cleaving HBV12.4, and the resulting proteins were tested for their ability to induce cleavage of the HBV12 target, upon co-expression with a variant cleaving HBV12.3.

A) Material and Methods

a) Site-directed mutagenesis

A site-directed mutagenesis library was created by PCR on a pool of chosen variants. For example, to introduce the G19S substitution into the coding
sequence of the variants, two separate overlapping PCR reactions were carried out that amplify the 5' end (residues 1-24) or the 3' end (residues 14-167) of the I-Crel coding sequence. For both the 5' and 3' end, PCR amplification is carried out using a primer with homology to the vector (GalIIF 5'-gcaaccttagtgctgcagcatacagg-3' (SEQ ID NO: 11) or GalIIR 5'-acaaccttgattggagactgacc-3'(SEQ ID NO: 12) and a primer specific to the I-Crel coding sequence for amino acids 14-24 that contains the substitution mutation G19S (G19SF 5'-gccggctttgtggactctgacggtagcato-3' (SEQ ID NO: 49) or G19SR 5'-gatgatgctaccgtcagagtccacaaagccggc-3'(SEQ ID NO: 50)). The resulting PCR products contain 33bp of homology with each other. The PCR fragments were purified.

The same strategy was used with the following pairs of oligonucleotides to introduce the F54L, E80K, F87L, V105A and I132V substitutions, respectively:

* F54LF: 5'-accagccgttggctgcagcatacagg-3' and F54LR: 5'-cactagttgctgcagcatacagg-3' (SEQ ID NO: 51 and 52);

* E80KF: 5'-ttaagcaaaatcaagccgctgcacaacttcctg-3' and E80KR: 5'-caggaagttgtgcagcggcttgattttgcttaa-3' (SEQ ID NO: 53 and 54);

* F87LF: 5'-aagccgctgcacacccgtcactcaactgcag-3' and F87LR: 5'-ctgcagttgagtcagcaggttgtgcagcggctt-3' (SEQ ID NO: 55 and 56);

* V105AF: 5'-aaacaggcaaacctggctctgaaaattatcgaa-3' and V105AR: 5'-ttcgataattttcagagccaggtttgcctgttt-3' (SEQ ID NO: 57 and 58);

* I132VF: 5'-acctgggtggatcaggttgcagctctgaacgat-3' and I132VR: 5'-atcgttcagagctgcaacctgatccacccaggt-3' (SEQ ID NO: 59 and 60).

The two overlapping PCR fragments for each of the six site-directed mutations were pooled and a total of approximately 25ng was combined with 75ng of vector DNA (pCLS1107, Figure 5) linearized by digestion with DraIII and AgeMV. The DNA was then used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MATa, trpl Δ 63, leu2Δ 1, Ms3Δ 200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Intact coding sequences containing one or more of the above described site directed substitutions are generated by in vivo homologous recombination in yeast.
c) Mating of meganuclease expressing clones and screening in yeast

The experimental procedure is as described in example 5.

d) Sequencing of variants

The experimental procedure is as described in example 2.

B) Results

A library containing site-directed mutations (G19S, F54L, E80K, F87L, V105A, 1132V) was constructed from a pool of 7 variants cleaving HBV 12.4 (32N,33C,44R,68Y,70S,75Y,77N, 32H,33C,44R,68Y,70S,75Y,77N, 32H,33C,44R,68Y,70S,75N,77Q,17K, 32H,33C,44R,68Y,70S,75N,77Q, 32H,33C,44R,68Y,70S,75N,77Q, also called KNHCQS/RYSYN (SEQ ID NO: 29), KNHCQS/RYSYN (SEQ ID NO: 22), KNHCQS/RYSNQ +117K (SEQ ID NO: 23), KNHCQS/RYSNQ (SEQ ID NO: 24), KNHCQS/RYSNN (SEQ ID NO: 25), KNHCQS/RYSNDQ +151A (SEQ ID NO: 26) and KNHCQS/RYSNDQ (SEQ ID NO: 27), respectively, according to the nomenclature of Table III and Table IV). The library was transformed into yeast and 1116 individual clones were picked and mated with a yeast strain that contains (i) the HBV12 target in a reporter plasmid (ii) an expression plasmid containing a variant that cleaves the HBV 12.3 target (30S,32R,33S,44D,68Y,70S,75S,77R or KSRSQS/DYSSR (SEQ ID NO: 15) according to the nomenclature of Table II).

After mating with this yeast strain, >200 clones were found to cleave the HBV12 target more efficiently than the original variants. An example of positives is shown in Figure 10. The sequence of seven of the best I-Crel variants cleaving the HBV12 target when forming a heterodimer with the KSRSQS/DYSSR variant are listed in Table VIII.
Table VIII: Functional variant combinations displaying strong cleavage activity for HBV12.

<table>
<thead>
<tr>
<th>VARIANT HBV12.3</th>
<th>Optimized* Variants HBV12.4 (SEQ ID NO: 69 to 75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Crel 28K 368 32R 35S 38O 46S 44D 68Y 70S 75S 77R(KSBRSQDYSRR)</td>
<td>32H 33C 40R 44R 68Y 70S 75N 77Q 19S</td>
</tr>
<tr>
<td></td>
<td>2H 32C 44R 68Y 70S 75D 77R</td>
</tr>
<tr>
<td></td>
<td>32H 33C 44R 68Y 70S 75Y 77Q 87L</td>
</tr>
<tr>
<td></td>
<td>32H 33C 44R 68Y 70S 75D 77N 80K</td>
</tr>
<tr>
<td></td>
<td>32H 33C 44R 68Y 70S 75D 77Q 87L 105A 151A</td>
</tr>
<tr>
<td></td>
<td>32H 33C 44R 54L 68Y 70S 75D 77Q</td>
</tr>
<tr>
<td></td>
<td>32H 33C 44R 68Y 70S 75D 77Q 87L 117K</td>
</tr>
</tbody>
</table>

* Mutations resulting from site-directed mutagenesis are in bold.

Example 8: Strategy for engineering novel meganucleases cleaving the HBV8 target from the Hepatitis B genome

HBV8 is a 22 bp (non-palindromic) target located in the coding sequence of the core protein gene in the Hepatitis B genome. The target sequence corresponds to positions 1908-1929 of the Hepatitis B genome (accession number X70185, Figure 35).

The HBV8 sequence is partly a patchwork of the 10TGA_P, 10CAAJP, 5CTTJP and 5_TCT_P targets (Figure 11) which are cleaved by previously identified meganucleases, obtained as described in International PCT Applications WO 2006/097784 and WO 2006/097853; Arnould et al, J. Mol. Biol., 2006, 355, 443-458; Smith et al, Nucleic Acids Res., 2006. Thus the inventors set out to determine whether HBV8 could be cleaved by combinatorial variants resulting from these previously identified meganucleases.

The 10TGA_P, 10CAA_P, 5CTTJP and 5_TCT_P target sequences are 24 bp derivatives of C1221, a palindromic sequence cleaved by I-Crel (Arnould et al, precited). However, the structure of l-Crel 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 HBV8 series of targets were defined as 22 bp sequences instead of 24 bp. HBV8 differs from C1221 in the 4 bp central region.
According to the structure of the I-Oel protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the I-Crel 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 ataa sequence in -2 to 2 was first substituted with the gtac sequence from C1221, resulting in target HBV8.2 (Figure 11). Then, two palindromic targets, HBV8.3 and HBV8.4, were derived from HBV8.2 (Figure 11). Since HBV8.3 and HBV8.4 are palindromic, they should be cleaved by homodimeric proteins. Thus, proteins able to cleave the HBV8.3 and HBV8.4 sequences as homodimers were first designed (examples 9, 10 and 11). In order to improve the weak cleavage activity of HBV8.4 variants, a series of variants cleaving HBV8.4 was subjected to random mutagenesis and screened for cleavage activity of the HBV8 target when co-expressed with a protein cleaving HBV8.3 (example 12). Cleavage activity of the HBV8 target could be observed for these heterodimers. To further improve cleavage activity for the HBV8 target, HBV8.4 variants were optimized by site-directed mutagenesis and used to form novel heterodimers that were screened against the HBV8 target (example 13). Improved cleavage activity of the HBV8 target could be observed for these heterodimers. Chosen heterodimers were then cloned into mammalian expression vectors and screened against the HBV8 target in CHO cells (example 14). Strong cleavage activity for the HBV8 target could be observed for these heterodimers in mammalian cells.

**Example 9: Identification of meganucleases cleaving HBV8.3**

This example shows that I-Oel variants can cut the HBV8.3 DNA target sequence derived from the left part of the HBV8.2 target in a palindromic form (Figure 11). 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 HBV8.3 will be noted tggacccttgt_P).

HBV8.3 is similar to 10TGA_P at positions ±1, ±2, ±4, ±6, ±8, ±9 and ±10 and to 5CTT_P at positions ±1, ±2, ±3, ±4, ±5, ±6 and ±8. It was hypothesized that positions ±7 and ±11 would have little effect on the binding and cleavage activity. Variants able to cleave the 10TGA_P target were obtained by

Both sets of proteins are mutated at position 70. However, the existence of two separable functional subdomains was hypothesized. This implies that this position has little impact on the specificity at bases 10 to 8 of the target. Mutations at positions 24 found in variants cleaving the 5CTTJP target will be lost during the combinatorial process. But it was hypothesized that this will have little impact on the capacity of the combined variants to cleave the HBV8.3 target.

Therefore, to check whether combined variants could cleave the HBV8.3 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5CTTJP were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TGA_P.

**A) Material and Methods**

a) **Construction of target vector**

The target was cloned as follows: an oligonucleotide corresponding to the HBV8.3 target sequence flanked by gateway cloning sequences was ordered from PROLIGO: 5′ tgtgatacagttttgacctgctagtctgta 3′ (SEQ ID NO: 83). 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 2). Yeast reporter vector was transformed into *Saccharomyces cerevisiae* strain FYBL2-7B (*MAT a, ura3Δ 851, trpl Δ 63, leu2Δ 1, lys2Δ 202*), resulting in a reporter strain.

b) **Mating of meangnuclease expressing clones and screening in yeast**

and WO 2006/097853, respectively for the IOTGA? and 5CTT_P targets. In order to generate I-CreI 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 1-OeI coding sequence. For both the 5' and 3' end, PCR amplification is carried out using primers (GalLOF 5'-gcaactttagtgacacatacagg-3' (SEQ ID NO: 11) or GalIOR 5'-acaaccttgatggagactgcc-3'(SEQ ID NO: 12)) specific to the vector (pCLS0542, Figure 3) and primers (assF 5'-ctannttgaccttt-3' (SEQ ID NO: 13) or assR 5'-aaaggtcaannntag-3'(SEQ ID NO: 14)), where nnn codes for residue 40, specific to the I-CreI 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 GalLOF and assR or assF and GalIOR 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 3) linearized by digestion with Ncol and Eagl were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATa, trpl Δ 63, leu2Δ 1, his3Δ 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 ah, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony gridded (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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-GaI 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

l-Crel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5CTTJP with the mutations at 28, 30, 32, 33, 38 and 40 from proteins cleaving 10TGA_P on the I-Crel scaffold, resulting in a library of complexity 1600. Examples of combinatorial variants are displayed in Table IX. This library was transformed into yeast and 2304 clones (1.4 times the diversity) were screened for cleavage against the HBV8.3 DNA target (ttgacccttgt_P, SEQ ID NO: 81). A total of 160 positive clones were found to cleave HBV8.3. Sequencing and validation by secondary screening of 79 of the best l-Crel variants resulted in the identification of 55 different novel endonucleases. Examples of positives are shown in Figure 12. The sequence of several 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 Table X). Such combinations likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be l-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.
Table IX: Panel of variants* theoretically present in the combinatorial library

*Only 200 out of the 1600 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HBV8.3 target was found among the identified positives.

Table X: l-Cre variants capable of cleaving the HBV8.3 DNA target.

Example 10: Identification of meganucleases cleaving HBV8.4

This example shows that l-Crel variants can cleave the HBV8.4 DNA target sequence derived from the right part of the HBV8.2 target in a palindromic form (Figure 11). 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, HBV8.4 will be called ccaattctgt_P).

HBV8.4 is similar to 5TCT_P at positions ±1, ±2, ±3, ±4, ±5, ±7, ±8, ±9 and ±11 and to 10CAA_P at positions ±1, ±2, ±7, ±8, ±9, ±10 and ±11. It was hypothesized that position ±6 would have little effect on the binding and cleavage activity. Variants able to cleave 5TCT_P were obtained by mutagenesis of \( l\text{-Crel} \) N75 at positions 24, 44, 68, 70, 75 and 77, as described previously (Arnould \textit{et al.}, J. Mol. Biol., 2006, 355, 443-458; Smith \textit{et al.} Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156). Variants able to cleave the 10CAA_P target were obtained by mutagenesis of \( l\text{-Crel} \) N75 or D75, at positions 28, 30, 32, 33, 38 and 40, as described previously in Smith \textit{et al.} Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156.

Mutations at position 24 found in variants cleaving the 5TCTJP target will be lost during the combinatorial process. But it was hypothesized that this will have little impact on the capacity of the combined variants to cleave the HBV8.4 target.

Therefore, to check whether combined variants could cleave the HBV8.4 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCTJP were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10CAA_P.

**A) Material and Methods**

a) \textbf{Construction of target vector}

The experimental procedure is as described in example 9, with the exception that an oligonucleotide corresponding to the HBV8.4 target sequence was used: 5′ tgccatacagttttccaaatctgtacagaaattggacatgtcgta 3′ (SEQ ID NO: 90).

b) \textbf{Construction of combinatorial variants}

\( l\text{-Crel} \) variants cleaving 10CAA_P or 5TCTJP were previously identified, as described in Smith \textit{et al.} Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156, and Arnould \textit{et al.}, J. Mol. Biol., 2006, 355, 443-458; International PCT Applications WO
2006/097784 and WO 2006/097853, respectively for the 10CAA_P and 5TCT_P targets. In order to generate l-Crel 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 l-OeI coding sequence. For both the 5' and 3' end, PCR amplification is carried out using primers (GalIOF 5'-gcaacctttgattgacatacagg-3' (SEQ ID NO: 11) or GalIOR 5'-acaaccttgagagatggacc-3' (SEQ ID NO: 12) specific to the vector (pCLS1 107, Figure 5) and primers (assF 5'-ctamintgaccttt-3' (SEQ ID NO: 13) or assR 5'-aaaggtcaanntag-3'(SEQ ID NO: 14), where nnn codes for residue 40, specific to the l-Crel 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 GalIOF and assR or assF and GalIOR 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 (pCLS1 107, Figure 5) linearized by digestion with DraIII and NgoMLV were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3Δ 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 (Arnauld et ai, J. Mol. Biol., 2006, 355, 443-458). Mating was performed using a colony grdder (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). 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, including 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 raM 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

The experimental procedure is as described in example 9.

B) Results

I-Crel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCT_P with the mutations 28, 30, 32, 33, 38 and 40 from proteins cleaving 10CAA_P on the I-Crel scaffold, resulting in a library of complexity 1600. Examples of combinatorial variants are displayed in Table XI. This library was transformed into yeast and 2304 clones (1.4 times the diversity) were screened for cleavage against the HBV8.4 DNA target (ccaaattctgt_P SEQ ID NO:82). Two positive clones were found, which after sequencing turned out to correspond to two different novel endonuclease variants (Table XI and Table XII). Examples of positives are shown in Figure 13. One of these two variants display non parental combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77. Such combinations likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be I-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.
Table XI: Panel of variants* theoretically present in the combinatorial library

<table>
<thead>
<tr>
<th>Amino acids at positions 44, 68, 70, 75 and 77 (ex: KGNI stands for K44, Q68, G70, N75 and I77)</th>
<th>Amino acids at positions 28, 30, 32, 33, 38 and 49 (ex: KNSHQ/QRSNK stands for K28, N30, S32, H33, Q38, Q40/Q44, R68, S70, N75 and K77, 163Q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGGNI</td>
<td>KNSHQ/QRSNK +163Q</td>
</tr>
<tr>
<td>KANI</td>
<td></td>
</tr>
<tr>
<td>KADNI</td>
<td></td>
</tr>
<tr>
<td>KNENI</td>
<td></td>
</tr>
<tr>
<td>KGNI</td>
<td></td>
</tr>
<tr>
<td>GRSNK</td>
<td></td>
</tr>
<tr>
<td>KSSNI</td>
<td></td>
</tr>
<tr>
<td>KSNI</td>
<td></td>
</tr>
<tr>
<td>QSSNR</td>
<td></td>
</tr>
<tr>
<td>QSSNR</td>
<td></td>
</tr>
<tr>
<td>HYDRH</td>
<td></td>
</tr>
</tbody>
</table>

* Only 150 out of the 1600 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HBV8.4 target was found among the identified positives.

Table XII: I-CreI variants with additional mutations capable of cleaving the HBV8.4 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-CreI variants (ex: KNSHQ/QRSNK +163Q stands for K28, N30, S32, H33, Q38, Q40/Q44, R68, S70, N75 and K77, 163Q)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNSHQ/QRSNK +163Q</td>
<td>91</td>
</tr>
</tbody>
</table>

Example 11: Identification of meganucleases cleaving HBV8.4 through the generation of combinatorial variants containing 105A and 132V substitutions

A combinatorial library containing selected amino-acid substitutions was produced as an alternative approach to generating I-CreI variants that cleave the HBV8.4 DNA target.

Two amino-acid substitutions have been found in previous studies to enhance the activity of I-CreI derivatives and could easily be incorporated into a combinatorial library: these mutations correspond to the replacement of Valine
105 with Alanine (V105A) and Isoleucine 132 with Valine (I132V). Both of these substitutions were introduced into all variants of a combinatorial library containing mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCT_P combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10CAA_P.

A) Material and Methods

a) Construction of combinatorial variants

1-Crel variants cleaving 10CAA_P or 5TCT_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 10CAA_P and 5TCT_P targets. In order to generate 1-Crel 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 39-104) of the L-OeI coding sequence. The remaining 3’ sequences of I-Crel containing the 105A and 132V substitutions are present in the vector pCLS1884. For both the 5’ and 3’ end amplifications, PCR is carried out using primers (GalIOF 5'-gcaaccttagtgtgacacatacaggg3' (SEQ ID NO: 11) or CreRevBsgl 5’- caggtttgctgttcttctacaaacgtt3' (SEQ ID NO: 92)) containing homology to the vector (pCLS1884, Figure 14) and primers (assF 5'-ctannncttgaccttt3' (SEQ ID NO: 13) or assR 5’-aaggtcaannntag-3' (SEQ ID NO: 14)), where nnn codes for residue 40, specific to the L-OeI 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 GalIOF and assR or assF and CreRevBsgl 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 (pCLS1884, Figure 14) linearized by digestion with Ncol and Bsgl were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3Δ 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 as well as the 105A and 132V substitutions is generated by in vivo homologous recombination in yeast.
b) Mating of meganuclease expressing clones and screening in yeast
The experimental procedure is as described in example 10.

d) Sequencing of variants
The experimental procedure is as described in example 9.

B) Results

1-Oel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCTJP with the mutations 28, 30, 32, 33, 38 and 40 from proteins cleaving 10CAA_P on an I-Crel scaffold containing the amino acid substitutions 105A and 132V, resulting in a library of complexity 1600. This library was transformed into yeast and 2304 clones (1.4 times the diversity) were screened for cleavage against the HBV8.4 DNA target (ccaaattctgtJP). Four positive clones were found, which after sequencing turned out to correspond to four different novel endonuclease variants (Table XIII). Examples of positives are shown in Figure 15. All four variants contain the 105A and 132V substitutions as well as display non parental combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77. Such combinations likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be I-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.

Table XIII: I-Crel variants with additional mutations capable of cleaving the HBV8.4 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40/44, 68, 70, 75 and 77 of the I-Crel variants (ex: KNSHQ/KASNI +105A+132V stands for K28, N30, S32, H33, Q38, Q40/ K44, A68, S70, N75 and 177, 105A, 132V)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNSHQ/KASNI +105A+132V</td>
<td>93</td>
</tr>
<tr>
<td>KNSHQ/KNANI +105A+132V</td>
<td>94</td>
</tr>
<tr>
<td>KNEYQ8/QASNRPK +105A+132V</td>
<td>95</td>
</tr>
<tr>
<td>KNEYQ8/QSSNR +105A+132V</td>
<td>96</td>
</tr>
</tbody>
</table>
Example 12: Improvement of meganucleases cleaving HBV8.4 by random mutagenesis

I-Crel variants able to cleave the palindromic HBV8.4 target have been previously identified in examples 10 and 11. However, the HBV8.4 variants display very weak activity with the HBV8.4 target. In this example, it was determined if the activity of the HBV8.4 meganucleases could be increased and at the same time it was tested whether they could cleave HBV8 efficiently when co-expressed with a protein cleaving HBV8.3. The six combinatorial variants cleaving HBV8.4 were mutagenized by random mutagenesis, and in a second step, it was assessed whether they could cleave HBV8 when co-expressed with a protein cleaving HBV8.3.

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 9, with the exception that an oligonucleotide corresponding to the HBV8 target sequence: 5’
tggcataaattactatacttctatagacacgcaaacacaaatacacagcggccttgccacc-3’ (SEQ ID NO: 97) was used.

b) Construction of libraries by random mutagenesis

Random mutagenesis was performed on a pool of chosen variants, by PCR using Mn²⁺. PCR reactions were carried out that amplify the 1-OeI coding sequence using the primers preATGCreFor (5’- gcataaattactatacttcatagacacgcaaacacaaatacacagcggccttgccacc-3’; SEQ ID NO: 30) and ICrelpostRev (5’-ggctgcggaggctgcgtctagaggctgcgtctaggtatcagtctgcgtgaccgcggc-3’; SEQ ID NO: 31). Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLS1107, Figure 5) linearized by digestion with Drall and NgoMV were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATA, trplΔ6S, 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 1-OeI variant were generated by in vivo homologous recombination in yeast.

c) Variant-target yeast strains

The yeast strain FYBL2-7B (MATA, ura3Δ851, trplΔ63, leu2Δ1, lys2Δ202) containing the HBV8 target in the yeast reporter vector (pCLS1055,
Figure 2) was transformed with variants, in the leucine vector (pCLS0542), cutting the HBV8.3 target, using a high efficiency LiAc transformation protocol.

d) Mating of meganuclease expressing clones and screening in yeast

Mating was performed using a colony gridded (QpixII, Genetix).

Variants were gridded on nylon filters covering YPD plates, using a low gridding density (about 4 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of a variant-target yeast strain for the target of interest. 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, including 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.

e) Sequencing of variants

The experimental procedure is as described in example 9.

B) Results

Six variants cleaving HBV8.4 (I-OeI 33H,40Q,70S,75N,77K, I-OeI 33H,40Q,70S,75N,77K,163Q, I-Crel 33H,40Q,44K,68A,70S,75N,105A,132V, I-Crel 33H,40Q,44K,68A,70S,75N,105A,132V, I-Crel 32E,68A,70S,75N,77R,105A,132V and I-Crel 32E,68A,70S,75N,77R,105A,132V also called KNSHQQ/QRSNK (SEQ ID NO: 98), KNSHQQ/QRSNK +163Q (SEQ ID NO: 91), KNSHQQ/KASNI +105A+132V (SEQ ID NO: 93), KNSHQQ/KNASNI +105A+132V (SEQ ID NO: 94), KNEYQS/QASNR +105A+132V (SEQ ID NO: 95), and KNEYQS/QSSNR +105A+132V (SEQ ID NO: 96), respectively, according to the nomenclature of Table XI, XII and XIII) were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then mated with a yeast strain that contains (i) the HBV8 target in a reporter plasmid (ii) an expression plasmid containing a variant that cleaves the HBV8.3 target (I-Crel 33C,38R,44R,68Y,70S,75D,77N or KNSCRS/RYSND (SEQ ID NO: 85) according to...
the nomenclature of table X). After mating with this yeast strain, 379 clones were found to cleave the HBV8 target. Thus, 379 positives contained proteins able to form heterodimers with KNSCR5/RYSND with cleavage activity for the HBV8 target. An example of positives is shown in Figure 16. Sequencing of the strongest 186 positive clones indicates that 32 distinct variants were identified (examples listed in Table XIV).

**Table XIV: Functional variant combinations displaying cleavage activity for HBV8.**

<table>
<thead>
<tr>
<th>VARIANT: HBV8.3</th>
<th>Optimized Variants HBV8.4 (SEQ ID NO: 99 to 105)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l-Crel 33C 38R 44R 68Y 70S 75D 77N (KNSCR5/RYSND)</td>
<td>33H 40Q 70S 75N 77K 105A 132V</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 68A 70S 75N 77R 105A 132V</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 68A 70S 75N 77R 132V</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 70S 75N 77K 132V</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 68S 70S 75N 77R 105A 132V</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 70S 75N 77K 105A</td>
</tr>
<tr>
<td></td>
<td>33H 40Q 68A 70S 75N 77R 105A</td>
</tr>
</tbody>
</table>

* Mutations resulting from mutagenesis are in bold.

**Example 13: Improvement of meganucleases cleaving HBV8 by site-directed mutagenesis of proteins cleaving HBV8.4 and assembly with proteins cleaving HBV8.3**

The l-Crel optimized variants cleaving HBV8.4 described in example 12 were further mutagenized by introducing selected amino-acid substitutions in the proteins and screening for more efficient variants cleaving HBV8 in combination with a variant cleaving HBV8.3.

Two amino-acid substitutions found in previous studies to enhance the activity of l-Crel derivatives were introduced into HBV8.4 variants: these mutations correspond to the replacement of Glycine 19 with Serine (G19S) and Phenylalanine 54 with Leucine (F54L). These mutations were individually introduced into the coding sequence of proteins cleaving HBV8.4, and the resulting proteins were
tested for their ability to induce cleavage of the HBV8 target, upon co-expression with a variant cleaving HBV8.3.

A) Material and Methods

a) Site-directed mutagenesis

Site-directed mutagenesis libraries were created by PCR on a pool of chosen variants. For example, to introduce the G19S substitution into the coding sequence of the variants, two separate overlapping PCR reactions were carried out that amplify the 5’ end (residues 1-24) or the 3’ end (residues 14-167) of the I-Crel coding sequence. For both the 5’ and 3’ end, PCR amplification is carried out using a primer with homology to the vector (GalIIOF 5’-gcaacctttagtgcagcatacagg-3’ (SEQ ID NO: 11) or GalIOR 5’-aacaccttgattggagacttgacc-3’ (SEQ ID NO: 12)) and a primer specific to the I-Crel coding sequence for amino acids 14-24 that contains the substitution mutation G19S (G19SF 5’-gccggctttgtggactctgacggtagcatcatc-3’ (SEQ ID NO: 49) or G19SR 5’-gatgatgctaccgtcagagtccacaaagccggo-3’ (SEQ ID NO: 50). The resulting PCR products contain 33bp of homology with each other. The PCR fragments were purified. Approximately 25ng of each of the two overlapping PCR fragments and 75ng of vector DNA (pCLS1 107, Figure 5) linearized by digestion with DraIII and NgoMIV were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATa, trpl Δ 63, leu2Δ l, his3A200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Intact coding sequences containing the G19S substitution are generated in vivo homologous recombination in yeast.

The same strategy is used with the following pair of oligonucleotides to create the library containing the F54L substitution:

* F54LF: 5’-accacgcgcgtggctgtggcagaaacctagtg-3’ and F54LR: 5’-cactagtgtgccgagccagctgg-3’ (SEQ ID NO: 51 and 52);

b) Mating of meganuclease expressing clones and screening in yeast

The experimental procedure is as described in example 12.

c) Sequencing of variants

The experimental procedure is as described in example 9.
B) Results

Libraries containing one of two amino-acid substitutions (G19S or F54L) were constructed on a pool of five variants cleaving HBV8.4 (33H,40Q,70S,75N,77K,105A,132V (SEQ ID NO: 99); 33H,40Q,68A,70S,75N,77R,105A,132V (SEQ ID NO: 100); 33H,40Q,68A,70S,75N,77R,77K,105A,132V (SEQ ID NO: 102) and 33H,40Q,68S,70S,75N,77R,105A,132V (SEQ ID NO: 103), according to the nomenclature of Table XIV). 576 transformed clones for each library were then mated with a yeast strain that contains (i) the HBV8 target in a reporter plasmid (ii) an expression plasmid containing a variant that cleaves the HBV8.3 target (1-OeI 33C,38R,44R,68Y,70S,75D,77N or KNSCR5/RYSDN (SEQ ID NO: 85) according to the nomenclature of Table X).

After mating with this yeast strain, a large number of clones (>100) in the library containing amino-acid substitution Glycine 19 with Serine (G19S), were found to cleave the HBV8 target more efficiently than the original variants. An example of positives is shown in Figure 17. The sequence of the four best L-OeI variants cleaving the HBV8 target when forming a heterodimer with the KNSCR5/RYSDN variant are listed in Table XV.

**Table XV: Functional variant combinations displaying strong cleavage activity for HBV8.**

<table>
<thead>
<tr>
<th>VARIANT HBV8.3</th>
<th>Optimized* Variants HBV8.4 (SEQ ID NO: 106 to 109)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19S 33H 40Q 43I 70S 75N 77K 105A 132V</td>
</tr>
<tr>
<td></td>
<td>19S 33H 40Q 70S 75N 77K 105A 132V</td>
</tr>
<tr>
<td></td>
<td>19S 33H 40Q 70S 75N 77K 105A</td>
</tr>
<tr>
<td></td>
<td>19S 33H 40Q 70S 75N 77K 132V</td>
</tr>
</tbody>
</table>

*Mutations resulting from mutagenesis are in bold.

**Example 14: Validation of HBV8 target cleavage in an extrachromosomal model in CHO cells**

L-Crel variants able to efficiently cleave the HBV8 target in yeast when forming heterodimers were described in examples 12 and 13. In order to further validate heterodimers displaying strong cleavage activity for the HBV8 target in yeast cells, the efficiency of chosen combinations of variants to cut the HBV8 target was...
analyzed, 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 HBV8 target in a vector for CHO screen

The target was cloned as follows: oligonucleotide corresponding to the HBV8 target sequence flanked by gateway cloning sequence was ordered from PROLIGO: 5' tggcataacaagttgtacaaaaagcaggcttcgaaggagatagaaccatggccaatacaatataaacaagagtctcG-3' (SEQ ID NO: 97). 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 18). Cloned target was verified by sequencing (MILLEGEN).

b) Re-cloning of meganucleases

The ORF of 1-OeI variants cleaving the HBV8.3 and HBV8.4 targets identified in examples 9 and 13 were re-cloned in pCLS1768 (Figure 19). ORFs were amplified by PCR on yeast DNA using the attBl-ICreIFor (5' ggggacagttgtacaaaaagcaggcttcgaaggagatagaaccatggccaatacaatataaacaagagtctcG-3'; SEQ ID NO: 110) and attB2-ICreIRev (5' ggggacacgttggatcaaaaaagcaggcttcgaaggagatagaaccatggccaatacaatataaacaagagtctcG-3'; SEQ ID NO: 111) primers. PCR products were cloned in the CHO expression vector pCLS1768 (Figure 19) using the Gateway protocol (INVITROGEN). Resulting clones were verified by sequencing (MILLEGEN).

c) Extrachromosomal assay in mammalian cells

CHO 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 IIOX 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 run. The entire process is performed on an automated Velocity 11 BioCel platform.

Per assay, 150 ng of target vector was co-transfected with 12.5 ng of each one of both variants (12.5 ng of variant cleaving palindromic HBV8.3 target and 12.5 ng of variant cleaving palindromic HBV8.4 target).

2) Results

One HBV8.3 variant (I-Crel 33C,38R,44R,68Y,70S,75D, 77N, SEQ ID NO: 85) and two HBV8.4 variants (I-Oel 19S 33H 40Q 431 70S 75N 77K 105A 132V, SEQ ID NO: 106 and I-Oel 19S 33H 40Q 70S 75N 77K 105A 132V, SEQ ID NO:107) described in examples 9 and 13 were first re-cloned in pCLS1768 (Figure 19). Then, in order to validate the cleavage activity of the heterodimers with the HBV8 target, the I-Crel variants cleaving the HBV8.3 or HBV8.4 targets were tested together as heterodimers against the HBV8 target in the CHO extrachromosomal assay.

Figure 20 shows the results obtained for the two heterodimers against the HBV8 target in CHO cells assay, compared to the activity of l-Scel against its target (tagggataacaggttaat, SEQ ID NO: 112). Analysis of the efficiencies of cleavage of the HBV8 sequence demonstrates that both combinations of I-Crel variants are able to cut the HBV8 target in CHO cells with an activity similar to that of I-Scel against the I-Scel target.

Example 15: Strategy for engineering novel meganucleases cleaving the HBV3 target from the HBV genome

HBV3 is a 22 bp (non-palindromic) target located in the coding sequence of the core protein gene in the Hepatitis B genome. The target sequence corresponds to positions 2216-2237 of the Hepatitis B genome (accession number M38636, Figure 35).

The HBV3 sequence is partly a patchwork of the 1OTGCJP, 1OTCTJP, 5TAC_P and 5TCC_P targets (Figure 21) which are cleaved by previously identified meganucleases, obtained as described in International PCT Applications WO 2006/097784 and WO 2006/097853; Arnould et al, J. MoL Biol, 2006, 355, 443-458; Smith et al., Nucleic Acids Res., 2006. Thus the inventors set out to
determine whether HBV3 could be cleaved by combinatorial variants resulting from these previously identified meganucleases.

The 10TGC_P, 1OTCTJP, 5TAC_P and 5TCC_P target sequences are 24 bp derivatives of C1221, a palindromic sequence cleaved by l-Crel (Arnould et al., precited). However, the structure of 1-OeI 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 HBV3 series of targets were defined as 22 bp sequences instead of 24 bp. HBV3 differs from C1221 in the 4 bp central region. According to the structure of the l-Crel protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the l-Crel protein (Chevalier et al, Nat. Struct. Biol., 2001, 8, 312-316; Chevalier and Stoddard, Nucleic Acids Res., 2001, 29, 3151-311A; 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 tttt sequence in -2 to 2 was first substituted with the gtac sequence from C1221, resulting in target HBV3.2 (Figure 21). Then, two palindromic targets, HBV3.3 and HBV3.4, were derived from HBV3.2 (Figure 21). Since HBV3.3 and HBV3.4 are palindromic, they should be cleaved by homodimeric proteins. In addition, to test the influence of the tttt sequence on the activity of homodimeric proteins two pseudo-palindromic targets were created, containing the tttt sequence at positions -2 to 2 (targets HBV3.5 and HBV3.6, Figure 21). Thus, proteins able to cleave the HBV3.3 and HBV3.4 sequences as homodimers were first designed (examples 16 and 17) and then co-expressed to obtain heterodimers cleaving HBV3.2 (example 18). In order to obtain cleavage activity for the HBV3 target, a series of variants cleaving HBV3.3 and HBV3.4 was chosen and refined. The chosen variants were subjected to random mutagenesis, screened for activity with the HBV3.5 and HBV3.6 targets (examples 19 and 20) and were subsequently used to form novel heterodimers that were screened against the HBV3 target (example 21). Heterodimers could be identified with cleavage activity for the HBV3 target. To further improve the cleavage activity for the HBV3
target, a series of variants cleaving HBV3.3 and HBV3.4 was chosen, refined, cloned into mammalian expression vectors and screened against the HBV3 target in CHO cells (examples 22, 23 and 24). Heterodimers could be identified with strong cleavage activity for the HBV3 target in mammalian cells. Finally, a single-chain construct was assembled and screened against the HBV3 target in CHO cells (example 25). The single-chain construct displayed cleavage activity for the HBV3 target in mammalian cells that was comparable to the HBV3.3/HBV3.4 heterodimer.

Example 16: Identification of meganucleases cleaving HBV3.3

This example shows that I-Crel variants can cut the HBV3.3 DNA target sequence derived from the left part of the HBV3.2 target in a palindromic form (Figure 21). 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 HBV3.3 will be noted ctgcctacgtp).

HBV3.3 is similar to 10TGC_P at positions ±1, ±2, ±3, ±8, ±9, ±10 and ±11 and to 5TACJP at positions ±1, ±2, ±3, ±4, ±5 and ±11. It was hypothesized that positions ±6 and ±7 would have little effect on the binding and cleavage activity. Variants able to cleave the 10TGC=_P target were obtained by mutagenesis of I-Crel N75 or D75, at positions 28, 30, 32, 33, 38 and 40, as described previously in Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156. Variants able to cleave 5TACJP were obtained by mutagenesis on I-Crel 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, el49; International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156.

Mutations at positions 24 found in variants cleaving the 5TAC_P target will be lost during the combinatorial process. But it was hypothesized that this will have little impact on the capacity of the combined variants to cleave the HBV3.3 target.

Therefore, to check whether combined variants could cleave the HBV3.3 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TAC_P were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TGC_P.
A) Material and Methods

a) Construction of target vector

The target was cloned as follows: an oligonucleotide corresponding to the HBV3.3 target sequence flanked by gateway cloning sequences was ordered from PROLIGO: 5'-tggtataacatgttctgcctagtaagctagcatacagtctgtcgtg 3' (SEQ ID NO: 123). 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 2). Yeast reporter vector was transformed into *Saccharomyces cerevisiae* strain FYC2-6A (*MAT*α, *ura3Δ 851, trpl Δ 63, leu2Δ I*, *his3Δ 200*), resulting in a reporter strain.

b) Construction of combinatorial variants

*I-CreI* variants cleaving 10TGC_P or 5TAC_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. MoI. Biol., 2006, 355, 443-458; International PCT Applications WO 2006/097784 and WO 2006/097853, respectively for the 10TGC_P and 5TAC_P targets. In order to generate 1-Oei 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 1-Oei coding sequence. For both the 5' and 3' end, PCR amplification is carried out using primers (GalloF 5'-gcaacttagtgctgacatagcagg-3' (SEQ ID NO: 11) or GallOR 5'-acaacttggtggactagcaggc-3'(SEQ ID NO: 12)) specific to the vector (pCLS0542, Figure 3) and primers (assF 5'-ctanmnttgaccttt-3' (SEQ ID NO: 13) or assR 5'-aaaggtcaanntag-3'(SEQ ID NO: 14)), where nnn codes for residue 40, specific to the I-CreI 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 GalloF and assR or assF and GallOR 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 3) linearized by digestion with *NcoI* and *EagI* were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (*MAT*α, *trpl Δ 63, leu2Δ I*, *his3Δ 200*)

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 gridded (QpixII, GENETIX). Variants were gridded on nylon Filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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-Crel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TAC_P with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TGC_P on the I-Crel scaffold, resulting in a library of complexity 1150. Examples of combinatorial variants are displayed in Table XVI. This library was transformed into yeast and 4608 clones (4 times the diversity) were screened for cleavage against the HBV3.3 DNA target.
(ctgcctacgt_P, SEQ ID NO: 119). A total of 550 positive clones were found to cleave HBV3.3. Sequencing and validation by secondary screening of 186 of the best l-Crel variants resulted in the identification of 120 different novel endonucleases. Examples of positives are shown in Figure 22. The sequence of several of the variants identified display non parenteral combinations at positions 28, 30, 32, 33, 38, 40 or 44, 68, 70, 75, 77 as well as additional mutations (see examples Table XVII). Such variants likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be l-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.

**Table XVI: Panel of variants** theoretically present in the combinatorial library

<table>
<thead>
<tr>
<th>Amino acids at positions 44, 68, 76, 75 and 77 (ex: AYSRT stands for A44, Y68, S76, T75 and T77)</th>
<th>KNSSS</th>
<th>KNQTS</th>
<th>KNSTAS</th>
<th>KNSTSS</th>
<th>KNQCS</th>
<th>KNCSS</th>
<th>KNSCAS</th>
<th>KNQVS</th>
<th>KNSCS</th>
<th>KNSCTS</th>
<th>KNQCS</th>
<th>KNSTTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AYSRT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARNHI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATRNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARRNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKGHI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHNRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKRHI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NARHNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NKRRNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNSRY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AYRSR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRSRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRSRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSNRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARSRR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARSRL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARSRK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NARSR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYSRV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYSRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYSRR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNSRY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AYRSR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Only 264 out of the 1150 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HBV3.4 target was found among the identified positives.
Table XVII: l-Crel variants capable of cleaving the HBV3.3 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40 / 44, 68, 70, 75 and 77 of the I-Crel variants (ex: KNSSRH/NYSRY stands for K28, N30, S32, S33, R38, H40 / N44, Y68, S70, R75 and Y77)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNSSRH/NYSRY</td>
<td>124</td>
</tr>
<tr>
<td>KNSSRQ/AYSRI</td>
<td>125</td>
</tr>
<tr>
<td>KNSCR/SAYSRT</td>
<td>126</td>
</tr>
<tr>
<td>KNSSRE/NYSRY</td>
<td>127</td>
</tr>
<tr>
<td>KNSSRQ/NYSRV</td>
<td>128</td>
</tr>
</tbody>
</table>

Example 17: Making of meganucleases cleaving HBV3.4

This example shows that I-Crel variants can cleave the HBV3.4 DNA target sequence derived from the right part of the HBV3.2 target in a palindromic form (Figure 21). 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, HBV3.4 will be called ttcttccgt_P).

HBV3.4 is similar to 5TCC_P at positions ±1, ±2, ±3, ±4, ±5 and to 10TCTJP at positions ±1, ±2, ±3, ±8, ±9 and ±10. It was hypothesized that positions ±6, ±7 and ±11 would have little effect on the binding and cleavage activity. Variants able to cleave 5TCCJP were obtained by mutagenesis of I-Crel N75 at positions 24, 44, 68, 70, 75 and 77, as described previously (Arnould et al., J. Mol. Biol., 2006, 355, 443-458; Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2006/097784, WO 2006/097854, WO 2007/060495 and WO 2007/049156). Variants able to cleave the 10TCT_P target were obtained by mutagenesis of I-Crel N75 or D75, at positions 28, 30, 32, 33, 38, 40 and 70, as described previously in Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049156.

Both sets of proteins are mutated at position 70. However, the existence of two separable functional subdomains was hypothesized. This implies that this position has little impact on the specificity at bases 10 to 8 of the target. Mutations at positions 24 found in variants cleaving the STCC_P target will be lost.
during the combinatorial process. But it was hypothesized that this will have little impact on the capacity of the combined variants to cleave the HBV3.4 target.

Therefore, to check whether combined variants could cleave the HBV3.4 target, mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCC_P were combined with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TCT_P.

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 16, with the exception that an oligonucleotide corresponding to the HBV3.4 target sequence was used: 5’ tggcatacaagttttttctttcgtagcgaatcaagcaatgtgca 3’ (SEQ ID NO: 129).

b) Construction of combinatorial variants

l-Crel variants cleaving 10TCT_P or 5TCCJP were previously identified, as described in Smith et al. Nucleic Acids Res., 2006, 34, el49; International PCT Applications WO 2007/060495 and WO 2007/049 156, and Arnould et al, J. Mol. Biol., 2006, 355, 443-458; International PCT Applications WO 2006/097784 and WO 2006/097853, respectively for the 1OTCTJP and 5TCCJP targets. In order to generate l-Crel 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-Crel coding sequence. For both the 5’ and 3’ end, PCR amplification is carried out using primers (GalIOF 5’-gcaacttttagctgacgatcaccataggg-3’ (SEQ ID NO: 11) or GalIOR 5’-acaacttttagctgacgatcaccataggg-3’ (SEQ ID NO: 12)) specific to the vector (pCLS1107, Figure 5) and primers (assF 5’-ctannncttgaccttggacctgacc-3’ (SEQ ID NO: 13) or assR 5’-aaaggtcaannntag-3’(SEQ ID NO: 14)), where nnn codes for residue 40, specific to the 1-OeI coding sequence for amino acids 39-43. The PCR fragments resulting from the amplification reaction were cloned by using 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 GalIOF and assR or assF and GalIOR 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 5) linearized by digestion with DraIII and NgoMLV were used to transform the yeast
Saccharomyces cerevisiae strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3Δ 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 gridded plate (QpixII, GENETIX). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of the reporter-harboring yeast strain for the target of interest. 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, including 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 nM 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

The experimental procedure is as described in example 16.

B) Results

I-Crel combinatorial variants were constructed by associating mutations at positions 44, 68, 70, 75 and 77 from proteins cleaving 5TCC_P with the 28, 30, 32, 33, 38 and 40 mutations from proteins cleaving 10TCT_P on the I-Crel scaffold, resulting in a library of complexity 1196. Examples of combinatorial variants are displayed in Table XVIII. This library was transformed into yeast and 4608 clones (3.8 times the diversity) were screened for cleavage against the HBV3.4 DNA target (ttctttccgt_P, SEQ ID NO: 120). A total of 257 positive clones were found to cleave HBV3.4. Sequencing and validation by secondary screening of 178 of the best I-Oel variants resulted in the identification of 98 different novel endonucleases. Examples of positives are shown in Figure 23. The sequence of several 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 Table XIX). Such variants likely result from PCR artifacts during the combinatorial process. Alternatively, the variants may be l-Crel combined variants resulting from micro-recombination between two original variants during in vivo homologous recombination in yeast.

Table XVIII: Panel of variants* theoretically present in the combinatorial library

<table>
<thead>
<tr>
<th>Amino acids at positions 44, 68, 70, 75 and 77 (ex: KNENI stands for K44, H68, E70, N75 and I77)</th>
<th>Amino acids at positions 28, 30, 32, 33, 38 and 40 (ex: KNSSYS stands for K28, N30, S32, C33, Y38 and S40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNENI +</td>
<td>KNSSYS +</td>
</tr>
<tr>
<td>KQSEI +</td>
<td>KSSQK +</td>
</tr>
<tr>
<td>KQSNQ +</td>
<td>KSSQK +</td>
</tr>
<tr>
<td>QGSR +</td>
<td>KNSSYS +</td>
</tr>
<tr>
<td>KQSIM +</td>
<td>KQSIM +</td>
</tr>
<tr>
<td>KGNN +</td>
<td>KGNN +</td>
</tr>
<tr>
<td>QRSYR +</td>
<td>QRSYR +</td>
</tr>
<tr>
<td>KSYIN +</td>
<td>KSYIN +</td>
</tr>
</tbody>
</table>

* Only 220 out of the 1,196 combinations are displayed.

+ indicates that a functional combinatorial variant cleaving the HBV3.4 target was found among the identified positives.

Table XIX: l-Cre variants with additional mutations capable of cleaving the HBV3.4 DNA target.

<table>
<thead>
<tr>
<th>Amino acids at positions 28, 30, 32, 33, 38, 40, 68, 70, 75 and 77 of the l-CreI variants (ex: KNSSYS/KHNNI stands for K28, N30, S32, C33, Y38, S40, K44, H68, N70, N75 and I77)</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNSSYS/KHNNI</td>
<td>130</td>
</tr>
<tr>
<td>KNSSWS/KYSNH</td>
<td>131</td>
</tr>
<tr>
<td>KQSAQS/QYSYR</td>
<td>132</td>
</tr>
<tr>
<td>KNSSCAS/KYSNY</td>
<td>133</td>
</tr>
<tr>
<td>KNSSCCS/KYSNH</td>
<td>134</td>
</tr>
</tbody>
</table>
Example 18: Making of meganucleases cleaving HBV3.2

1-Crel variants able to cleave each of the palindromic HBV3.2 derived targets (HBV3.3 and HBV3.4) were identified in example 16 and example 17. Pairs of such variants (one cutting HBV3.3 and one cutting HBV3.4) 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 HBV3.2 and the non palindromic HBV3 targets.

A) Materials and Methods

a) Construction of target vector

The experimental procedure is as described in example 2, with the exception that an oligonucleotide corresponding to the HBV3.2 target sequence: 5’ tggcatacaagtttctgctactgtaaacggaagaacaatcgtctga 3’ (SEQ ID NO: 135) or the HBV3 target sequence: 5’ tggcatacaagtttctgctacttttggaagagaaacaatcgtctgtca 3’ (SEQ ID NO: 136) was used.

b) Co-expression of variants

Yeast DNA was extracted from variants cleaving the HBV3.4 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 HBV3.3 target in the pCLS0542 expression vector. Transformants were selected on synthetic medium lacking leucine and containing G418.

c) Mating of meganucleases co-expressing clones and screening in yeast

Mating was performed using a colony griddler (QpixII, Genetix). Variants were gridded on nylon filters covering YPD plates, using a low gridding density (4-6 spots/cm²). A second gridding process was performed on the same filters to spot a second layer consisting of reporter-harboring yeast strain for the target of interest. 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, including 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 HBV3.4 target (7 variants chosen among those described in Table XVIII and Table XIX) and four variants cleaving the HBV3.3 target (described in Table XVI and Table XVII) resulted in efficient cleavage of the HBV3.2 target in all cases (Figure 24). However, none of these combinations were able to cut the HBV3 natural target that differs from the HBV3.2 sequence by 2 bp at positions 1 and 2 (Figure 21). Functional combinations cleaving HBV3.2 are summarized in Table XX.

Table XX: Cleavage of the HBV3.2 target by the heterodimeric variants

<table>
<thead>
<tr>
<th>Variant HBV3.4, amino acids at positions 28, 30, 32, 33, 36, 40 / 44</th>
<th>Variant HBV3.3, amino acids at positions 28, 30, 32, 33, 36, 40 / 44 (ex: KNSSCSK/NSRY stands for K28, N30, S32, S33, S36, S44/44, Y68, Y70, Y75 and Y77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNSSCSK/NSRY (SEQ ID NO: 137)</td>
<td>KNSSCSK/NSRY (SEQ ID NO: 137)</td>
</tr>
<tr>
<td>KNSSCSK/NSRY +45M (SEQ ID NO: 140)</td>
<td>KNSSCSK/NSRY +45M (SEQ ID NO: 141)</td>
</tr>
<tr>
<td>KNSSCSK/KHHI (SEQ ID NO: 130)</td>
<td>KNSSCSK/KHHI (SEQ ID NO: 130)</td>
</tr>
<tr>
<td>KNSSCSK/KHY (SEQ ID NO: 142)</td>
<td>KNSSCSK/KHY (SEQ ID NO: 142)</td>
</tr>
<tr>
<td>KNSSCSK/KYNS (SEQ ID NO: 143)</td>
<td>KNSSCSK/KYNS (SEQ ID NO: 143)</td>
</tr>
<tr>
<td>KNSSCSK/KYSQ (SEQ ID NO: 144)</td>
<td>KNSSCSK/KYSQ (SEQ ID NO: 144)</td>
</tr>
</tbody>
</table>

+ indicates a functional combination

Example 19: Improvement of meganucleases cleaving HBV3.3 by random mutagenesis
l-Crel variants able to cleave the HBV3.2 target by assembly of variants cleaving the palindromic HBV3.3 and HBV3.4 target have been previously identified in example 18. However, none of these variants were able to cleave the HBV3 target.

Therefore, four combinatorial variants cleaving HBV3.3 were mutagenized, and variants were screened for cleavage activity of the HBV3.5. HBV3.5 is a pseudo-palindromic target similar to HBHV3.3 except that it contains the tttt sequence at positions -2 to 2 (Figure 21). The Inventors have previously observed that the association of a variant cleaving a pseudo-palindromic target with a wild-type sequence at positions -2 to 2 with a variant cleaving the other pseudo-palindromic target will increase the probability of cleavage of the target of interest. According to the structure of the l-Crel protein bound to its target, there is no contact between the 4 central base pairs (positions -2 to 2) and the l-OeI protein (Chevalier et at, 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.

Thus, in a first step, proteins cleaving HBV3.3 were mutagenized, and in a second step, it was assessed whether they displayed increased activity with the target HBV3.5.

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 16, with the exception that an oligonucleotide corresponding to the HBV3.5 target sequence was used: 5' tggcatacagtacctgcctagaagggagcaatctgta 3' (SEQ ID NO: 145).

b) Construction of libraries by random mutagenesis

Random mutagenesis was performed on a pool of chosen variants, by PCR using Mn²⁺. PCR reactions were carried out that amplify the l-Crel coding sequence using the primers preATGCreFor (5'-gctgatactactatactagagcacaacacaaatatacagctggtgccac-3'; SEQ ID NO: 30) and lCrepostRev (5'-ggctcgagagctgtcagatcagttatcagctggtgccac-3'; SEQ ID NO: 31), which are common to the pCLS0542 (Figure 3) and pCLS1107 (Figure 5)
vectors. Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLS542, Figure 3) by digestion with Ncol and Eagl were used to transform the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATa, 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 *l-Crel* variant were generated by *in vivo* homologous recombination in yeast.

c) **Mating of meganuclease expressing clones and screening in yeast**

The experimental procedure is as described in example 16.

d) **Sequencing of variants**

The experimental procedure is as described in example 16.

**B) Results**

Four variants cleaving HBV3.3, *l-Crel* 33C,38S,44N,68Y,70S,75R,77Y; *l-Crel* 33C,38S,44N,70S,75R77N; *l-Crel* 33C,38R,44A,68Y,70S,75R,77T and *l-Crel* 33S,38R,40E,44N,68Y,70S,75R,77Y also called KNSSCSS/NYSRY (SEQ ID NO: 137), KNSSCSS/NRSRN (SEQ ID NO: 138), KNSCRs/AYSRT (SEQ ID NO: 126) and KNSSRE/NYSRY (SEQ ID NO: 127) according to the nomenclature of Table XVI and Table XVII were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then screened for cleavage against the HBV3.5 DNA target. 58 clones were found to cleave the HBV3.5 target more efficiently than the original variant. An example of positives is shown in Figure 25. Sequencing of these positive clones indicates that 28 distinct variants were identified (see examples Table XXI).
Table XXI: Functional variant combinations displaying strong cleavage activity for HBV3.5.

<table>
<thead>
<tr>
<th>Optimized* Variants HBV3.3 (SEQ ID NO: 146 to 153)</th>
</tr>
</thead>
<tbody>
<tr>
<td>26R 33C 38S 44N 68Y 70S 75R 77Y 81T</td>
</tr>
<tr>
<td>33C 38R 44A 68Y 70S 75R 77T 132V</td>
</tr>
<tr>
<td>33C 38R 44A 68Y 70S 75R 77Y</td>
</tr>
<tr>
<td>33C 38R 44A 57N 68Y 70S 75R 77Y 80G</td>
</tr>
<tr>
<td>33C 38R 44A 68Y 70S 75R 77T 83S</td>
</tr>
<tr>
<td>17A 33C 38R 44A 68Y 70S 75R 77T</td>
</tr>
<tr>
<td>33C 38R 44A 62L 68Y 70S 75R 77Y</td>
</tr>
<tr>
<td>33C 38R 44N 68Y 70S 72P 75R 77Y</td>
</tr>
</tbody>
</table>

* Mutations resulting from random mutagenesis are in bold.

Example 20: Improvement of meganucleases cleaving HBV3.4

As a complement to example 19 we also decided to perform random mutagenesis with variants that cleave HBV3.4. The mutagenized proteins cleaving HBV3.4 were then tested to determine if they could efficiently cleave the pseudopalindromic target HBV3.6

A) Material and Methods

a) Construction of target vector

The experimental procedure is as described in example 16, with the exception that an oligonucleotide corresponding to the HBV3.6 target sequence was used: 5' tggcatacaagttttctcttttgaagagaaatcgtctgtca 3' (SEQ ID NO: 154).

b) 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 1-Oel coding sequence using the primers preATGCreFor (5' gcatatacattaactctcttagacaagctcacaacactacagcggcttgcacc-3'; SEQ ID NO: 30) and ICrelpostRev (5' ggctcgaggagctcagacgcgttctgatacgcgttacagcggccgcacc-3'; SEQ ID NO: 31). Approximately 25 ng of the PCR product and 75 ng of vector DNA (pCLSI 107, Figure 5) linearized by digestion with DraIII and NgoMN were used to transform the yeast Saccharomyces cerevisiae strain FYC2-6A (MATa, trplΔ 63,
leu2Δ l, his3Δ 200) 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 l-Crel variant were generated by in vivo homologous recombination in yeast.

c) Mating of meganuclease expressing clones and screening in yeast
The experimental procedure is as described in example 16.

d) Sequencing of variants
The experimental procedure is as described in example 16.

B) Results

Seven variants cleaving HBV3.4 (1-OeI 33C,38Y,44K,45M,68Y,70S,75N,77V l-Crel 33C,38Y,44K,68Y,70S,75N,77Y, 1-OeI 33G,38Y,44K,68Y,70S,75N,77Q, l-Crel 33G,38Y,44K,68Y,70S,75N,77Y, 1-OeI 33S,38Y,44K,68H,70N,75N, l-Creï 33S,38Y,44K,45M,68Y,70S,75N,77V, and l-Crel 33C,38Y,44K,68Y,70S,75N,77Q) also called KNSCYS/KYSNV +45M (SEQ ID NO: 140), KNSCYS/KYSNY (SEQ ID NO: 142), KNSGYS/KYSNQ (SEQ ID NO: 143), KNSGYS/KYSNY (SEQ ID NO: 139), KNSSYS/KHNNI (SEQ ID NO: 130), KNSGYS/KYSNV +45M (SEQ ID NO: 141), and KNSCYS/KYSNQ (SEQ ID NO: 144), respectively, according to the nomenclature of Table XVIII and Table XIX) were pooled, randomly mutagenized and transformed into yeast. 2304 transformed clones were then screened for cleavage against the HBV3.6 DNA target. 114 clones were found to cleave the HBV3.6 target more efficiently than the original variant. An example of positives is shown in Figure 26. Sequencing of these positive clones indicates that 65 distinct variants were identified (see examples Table XXII).
Table XXII: Functional variant combinations displaying strong cleavage activity for HBV3.6.

<table>
<thead>
<tr>
<th>Optimized Variants HBV3.4 (SEQ ID NO: 155 to 164)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D 33S 38Y 44K 68Y 70S 75N 77Y 140M</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 54L 68Y 70S 75N 77Y</td>
</tr>
<tr>
<td>33C 38Y 44K 64I 68Y 70S 75N 77Y 85R</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 68Y 70S 75N 77V 105A</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 59A 68Y 70S 75N 77V</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 68Y 70S 75N 77V 85R</td>
</tr>
<tr>
<td>33S 38Y 44K 45M 68Y 70S 75N 77V 86T</td>
</tr>
<tr>
<td>33S 38Y 44K 68Y 70S 75N 77L</td>
</tr>
<tr>
<td>33C 38Y 44K 57E 68Y 70S 75N 77V</td>
</tr>
<tr>
<td>32F 33C 38Y 44K 45M 68Y 70S 75N 77V</td>
</tr>
</tbody>
</table>

* Mutations resulting from random mutagenesis are in bold.

**Example 21: Making of meganucleases cleaving HBV3**

Optimized 1-OeI variants able to cleave each of the pseudo-palindromic targets HBV3.5 and HBV3.6 were identified in example 19 and example 20. Pairs of such optimized variants (one cutting HBV3.5 and one cutting HBV3.6) were co-expressed in yeast. Upon co-expression, there should be three active molecular species, two homodimers, and one heterodimer. It was determined whether the heterodimers that should be formed cut the non-palindromic HBV3 target.

**A) Materials and Methods**

a) Co-expression of variants

Yeast DNA was extracted from optimized HBV3.3 variants cleaving the HBV3.5 target (pCLS542 expression vector) as well as those optimized HBV3.4 variants cleaving HBV3.6 (pCLS1107 expression vector) using standard protocols and were used to transform *E. coli*. Plasmid DNA derived from an optimized HBV3.3 variant and an optimized HBV3.4 variant was then co-transformed into the yeast *Saccharomyces cerevisiae* strain FYC2-6A (MATα, trpl Δ 63, leu2Δ 1, his3Δ 200) using a high efficiency LiAc transformation protocol (Gietz and Woods, Methods Enzymol., 2002, 350, 87-96). Transformants were selected on synthetic medium lacking leucine and containing G418.
b) MatinR of meganucleases coexpressinR clones and screening in yeast

The experimental procedure is as described in example 18.

**B) Results**

Co-expression of an optimized HBV3.3 variants cleaving the HBV3.5 target (7 variants chosen among those described in example 19) and eleven optimized HBV3.4 variants cleaving the HBV3.6 target (described in example 20) resulted in efficient cleavage of the HBV3 target in some cases (Figure 27). Functional combinations cleaving HBV3 are summarized in Table XXIII.

**Table XXIII: Cleavage of the HBV3 target by the heterodimeric variants**

<table>
<thead>
<tr>
<th>Optimized HBV3.4 variants cleaving HBV3.6</th>
<th>Optimized HBV3.3 variant cleaving HBV3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>33S 38Y 44K 68Y 76S 75N 77Y (SEQ ID NO: 162)</td>
<td>+</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 68Y 76S 75N 77Y 121Y (SEQ ID NO: 168)</td>
<td>+</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 59A 66Y 76S 75N 77Y (SEQ ID NO: 159)</td>
<td>+</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 61V 68Y 76S 75N 77Y 95Y (SEQ ID NO: 169)</td>
<td>+</td>
</tr>
<tr>
<td>33C 38Y 44K 68Y 76S 75N 77Y 85R (SEQ ID NO: 157)</td>
<td>+</td>
</tr>
<tr>
<td>33S 38Y 44K 45M 68Y 76S 75N 77Y 61T (SEQ ID NO: 170)</td>
<td>+</td>
</tr>
<tr>
<td>31R 33C 38Y 44K 68Y 76S 73A 75N 77Y (SEQ ID NO: 171)</td>
<td>+</td>
</tr>
<tr>
<td>2S 33C 38Y 44K 68Y 76S 75N 77Y 68A (SEQ ID NO: 172)</td>
<td>+</td>
</tr>
<tr>
<td>33S 38Y 44K 45M 68Y 76S 75N 77Y 68T (SEQ ID NO: 161)</td>
<td>+</td>
</tr>
<tr>
<td>33C 38Y 44K 45M 68Y 76S 75N 77Y 105A (SEQ ID NO: 159)</td>
<td>+</td>
</tr>
<tr>
<td>2D 33S 38Y 44K 68Y 76S 75N 77Y 140M (SEQ ID NO: 159)</td>
<td>+</td>
</tr>
</tbody>
</table>

+ indicates a functional combination
Improved of meganucleases cleaving the HBV3 target site by random mutagenesis of I-Cre\ variants cleaving the HBV3.4 target and assembly with variants cleaving HBV3.3 in CHO cells.

I-Crel variants able to cleave the HBV3 target in yeast were previously identified in example 21 by assembly of optimized variants cleaving HBV3.3 and optimized variants cleaving HBV3.4.

In this example, it was determined if the activity of the meganucleases could be increased and at the same time establish if the meganucleases are active in CHO cells. The variants cleaving HBV3.4 described in example 20 (Table XXII) were subjected to random mutagenesis and more efficient variants cleaving HBV3 in combination with variants cleaving HBV3.3 (identified in example 19) were identified 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 HBV3 target in a vector for CHO screen

The target was cloned as follow: oligonucleotide corresponding to the HBV3 target sequence flanked by gateway cloning sequence was ordered from PROLIGO: 5' tggcatataagtttcctgccttacttttggaagagaaacaatcgtctgtca 3' (SEQ ID NO: 136). 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 18). Cloned target was verified by sequencing (MILLEGEN).

b) Construction of libraries by random mutagenesis

I-Crel variants cleaving HBV3.4 were pooled and randomly mutagenized by PCR in the presence of Mn2+. Primers used are attBl-ICreIFor (5'-ggggacagtttgtcataaaaaagcaggttctgaaggagatagacatggccaatatacaaatataaagagttcG-3'; SEQ ID NO: 110) and attB2-ICrelRev (5'-ggggaccactttgtcataaaaaagcaggttctgaaggagatagacatggccaatatacaaatataaagagttcG-3'; SEQ ID NO: 111). PCR products obtained were cloned in pCDNA6.2 from INVITROGEN
c) Re-cloning of meganucleases

The ORF of and 1-OeI variants cleaving the HBV3.3 target were re-cloned in pCLS1768 (Figure 19). ORFs were amplified by PCR on yeast DNA using the above described attBl-ICreIFor and attB2-ICreIRev primers. PCR products were cloned in CHO expression vector pCDNA6.2 from INVITROGEN (pCLS1768, Figure 19) using the Gateway protocol (INVITROGEN). Resulting clones were verified by sequencing (MILLEGEN).

d) Extrachromosomal assay in mammalian cells

CHO 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 IOOX buffer (MgCl₂ 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 Velocity 11 BioCel platform. Positives clones resulting of the screen of libraries were secondary screened and verified by sequencing (MILLEGEN).

Per assay, 150 ng of target vector was cotransfected with 12.5 ng of each one of the variants (12.5 ng of variant cleaving palindromic HBV3.3 target and 12.5 ng of variant cleaving palindromic HBV3.4 target).

2) Results

Four optimized variants cleaving HBV3.4

(33C,38Y,44K,45M,54L,68Y,70S,75N,77Y) (SEQ ID NO: 156),
33S,38Y,44K,68Y,70S,75N,77L (SEQ ID NO: 162),
33S,38Y,44K,45M,68Y 70S,75N,77V,86T (SEQ ID NO: 161), and
2D,33S,38Y,44K,68Y,70S,75N,77Y,140M (SEQ ID NO: 155), according to the nomenclature of Table XXII in example 20) were subjected to another round of optimization. They were pooled, randomly mutagenized and a library of new 1-OeI variants was cloned in the pCLS1768 vector allowing expression of the variant in
CHO cells (Figure 19). 3456 clones were screened using the extrachromosomal assay in CHO cells. The screen is carried out by co-transfection of 3 plasmids in CHO cells: one expressing a variant resulting from random mutagenesis of the variant cleaving HBV3.4, a second expressing a chosen variant cleaving HBV3.3 re-cloned in pCLS1768 (Figure 19) and a third one containing the HBV3 target cloned in pCLS1058 (Figure 18). The I-Crel variant cleaving HBV3.3 used for the screen of the library: $\text{I-Crel } 26R,33C,38S,44N,68Y,70S,75R,77Y,81T$, SEQ ID NO: 146, according to Table XXI in example 19.

Six clones were found to trigger cleavage of the HBV3 target in the CHO assay when forming heterodimers with the optimized HBV3.3 variant (1-OeI 26R,33C,38S,44N,68Y,70S,75R,77Y,81T SEQ ID NO: 146) in a primary screen. The 6 clones (SEQ ID NO: 173 to 178) were validated in a secondary screen (Figure 28) and sequenced (Table XXIV). In the secondary screen, the efficiency of the 6 clones was compared to one of the initial HBV3.4 variants (I-Crel 33S,38Y,44K,68Y,70S,75N,77L, SEQ ID NO: 162 according to Table XXII in example 20) co-expressed with the optimized HBV3.3 variant (I-Crel 26R,33C,38S,44N,68Y,70S,75R,77Y,81T SEQ ID NO: 146). All six new refined HBV3.4 variants were able to cleave the HBV3 target with an efficacy superior to that observed with the heterodimer formed by the initial HBV3.4 variant (1-OeI 33S,38Y,44K,68Y,70S,75N,77L).


<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4_A1</td>
<td>19S 33C 38Y 44K 68Y 70S 75N 77Q 96E 140M</td>
<td>173</td>
</tr>
<tr>
<td>3.4_A5</td>
<td>19S 33C 38Y 44K 68Y 70S 75N 77Y</td>
<td>174</td>
</tr>
<tr>
<td>3.4_A7</td>
<td>19S 33C 38Y 44K 68Y 70S 75N 77Q 140M</td>
<td>175</td>
</tr>
<tr>
<td>3.4_E10</td>
<td>19S 33C 38R 44K 68Y 70S 75N 77Q 132M</td>
<td>176</td>
</tr>
<tr>
<td>3.4_B4</td>
<td>19S 33C 38Y 44K 68Y 70S 75N 77Q</td>
<td>177</td>
</tr>
<tr>
<td>3.4_D3</td>
<td>19S 33C 38Y 44K 61G 68Y 70S 75N 77Q 128C</td>
<td>178</td>
</tr>
</tbody>
</table>
Example 23: Improvement of meganucleases cleaving the HBV3 target site by random mutagenesis of I-Crel variants cleaving the HBV3.3 target and assembly with variants cleaving HBV3.4 in CHO cells.

As a complement to example 22 we also decided to perform random mutagenesis with variants that cleave HBV3.3. The variants cleaving HBV3.3 described in example 19 (Table XXI) were subjected to random mutagenesis and more efficient variants cleaving HBV3 in combination with a variant cleaving HBV3.4 (identified in example 22) were identified in CHO cells.

1) Materials and Methods

a) Construction of libraries by random mutagenesis

I-Crel variants cleaving HBV3.3 were pooled and randomly mutagenized by PCR in the presence of Mn²⁺. Primers used are attBl-I-CrelFor (5'-ggggacaagtgttacaaaaagcaggtctgaaggagatagaaccatggccataacaaataacaagagttcc-3'; SEQ ID NO: 110) and attB2-I-CrelRev (5'-ggggaccactttgatagaaagcaggtctgaaggagatagaaccatggccataacaaataacaagagttcc-3'; SEQ ID NO: 111). PCR products obtained were cloned in pCDNA6.2 from INVITROGEN (pCLS1768, Figure 19), a vector for expression in CHO cells, using the Gateway protocol (INVITROGEN).

b) Extrachromosomal assay in mammalian cells

Extrachromosomal assay in mammalian cells was performed as described in example 22.

2) Results

Four optimized variants cleaving HBV3.3 (26R,33C,38S,44N,68Y,70S,75R,77Y,81T (SEQ ID NO: 146), 33C,38R,44A,68Y,70S,75R,77T,132V (SEQ ID NO: 147), 33C,38R,44A,68Y,70S,75R,77T,83S (SEQ ID NO: 150), and 33C,38R,44A,57N,68Y,70S,75R,77Y,80G (SEQ ID NO: 149), according to the nomenclature of Table XXI in example 19) were subjected to a round of optimization. They were pooled, randomly mutagenized and a library of new I-Crel variants was cloned in the pCLS1768 vector allowing expression of the variant in CHO cells (Figure 19). 2976 clones were screened using the extrachromosomal assay in CHO cells. The screen is carried out by co-transfection of 3 plasmids in CHO cells:
one expressing a variant resulting from random mutagenesis of the variant cleaving HBV3.3, a second expressing a chosen variant cleaving HBV3.4 re-cloned in pCLS1768 (Figure 19) and a third one containing the HBV3 target cloned in pCLS1058 (Figure 18). The optimized 1-OeI variant cleaving HBV3.4 (3.4JB4) was used for the screen of the library: 19S,33C,38Y,44K,68Y,70S,75N,77Q (SEQ ID NO: 177) according to Table XXIV in example 22.

Six clones were found to trigger cleavage of the HBV3 target in the CHO assay when forming heterodimers with the HBV3.4 variant (I-Crel 19S,33C,38Y,44K,68Y,70S,75N,77Q, Seq ID NO: 177) in a primary screen. The 6 clones (Seq ID NO: 179 to 184) were validated in a secondary screen (Figure 29) and sequenced (Table XXV). In the secondary screen, the efficiency of the 6 clones was compared to one of the initial HBV3.3 variants (I-Crel 26R,33C,38S,44N,68Y,70S,75R,77Y,81T, Seq ID NO: 146, according to Table XXI in example 19) co-expressed with the HBV3.4 variant (1-OeI 19S,33C,38Y,44K,68Y,70S,75N,77Q, Seq ID NO: 177). One of the new HBV3.3 variants (I-Crel 26R,33C,38S,44N,68Y,70S,75R,77Y,81T,139R, Seq ID NO:183) was able to cleave the HBV3 target with an efficacy superior to that observed with the heterodimer formed by the initial HBV3.3 variant (I-Crel 26R,33C,38S,44N,68Y,70S,75R,77Y,81T, Seq ID NO: 146).

Table XXV: I-Crel variants displaying improved cleavage activity for HBV3 DNA target when forming heterodimers with HBV3.4 (I-Crel 19S, 33C, 38Y, 44K, 68Y, 70S, 77Q).

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3_B10</td>
<td>23V 26R 33C 38S 44N 68Y 70S 75R 77Y 81T</td>
<td>179</td>
</tr>
<tr>
<td>3.3_C9</td>
<td>33C 34R 38R 44A 68Y 70S 75R 77T 81T 83S</td>
<td>180</td>
</tr>
<tr>
<td>3.3_D3</td>
<td>26R 33C 38S 44N 68Y 70S 75R 77Y 81T 130G</td>
<td>181</td>
</tr>
<tr>
<td>3.3_E9</td>
<td>26R 33C 38S 44N 53R 68Y 70S 75R 77Y 81T 89A</td>
<td>182</td>
</tr>
<tr>
<td>3.3_F1</td>
<td>26R 33C 38S 44N 68Y 70S 75R 77Y 81T 139R</td>
<td>183</td>
</tr>
<tr>
<td>3.3_G6</td>
<td>26R 33C 38S 44N 56G 68Y 70S 75R 77Y 81T</td>
<td>184</td>
</tr>
</tbody>
</table>
**Example 24:** Improvement of meganucleases cleaving the HBV3 DNA target by multiple site-directed mutagenesis of HBV3.3 and HBV3.4 variants

Optimized HBV3.3 and HBV3.4 variants able to cleave the HBV3 target in CHO cells when forming heterodimers were identified in examples 22 and 23. However, these variants displayed cleavage activity for the HBV3 target that was inferior to that of the I-Scel meganuclease for its target. To try and further improve that activity of the HBV3 meganuclease, HBV3.3 and HBV3.4 variants were subjected to an additional step of optimization by introducing selected amino-acid substitutions.

Three 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), Valine 105 with Alanine (V105A) and Isoleucine 132 with Valine (II 32V). One, two or all three of these mutations were introduced into the coding sequence of proteins cleaving HBV3.3 and HBV3.4; and the resulting heterodimers were tested for their ability to induce cleavage of the HBV3 target in an extrachromosomal assay in CHO cells.

1) **Materials and methods**

a) Construction of site directed variants (single mutations)

Site-directed variants containing a single mutation were created by PCR. For example, to introduce the F54L substitution into the coding sequence of the variant, two separate overlapping PCR reactions were carried out that amplify the 5' end (amino acid residues 1-59) or the 3' end (amino acid residues 49-167) of the I-CreI coding sequence. For both the 5' and 3' end, PCR amplification is carried out using a primer with homology to the vector attBl-ICreIFor (5'-ggggacaagttgtgacaaaaaacggctcagaagagatgaaccatgccccatcacaatatataacaggttcc-3'; SEQ ID NO: 110) and attB2-ICreIRev (5'-ggggaccacttttgacaaaaagctggttagagccccgccgagattctctectge-3'; SEQ ID NO: 111) and a primer specific to the I-CreI coding sequence for amino acids 49-59 that contain the substitution mutation F54L (F54LF: 5'-accacgagccgtttgtcgtggccatgagttg-3' (SEQ ID NO: 51) or F54LR: 5'-cactagttttgcagcaccagggccagtggtgt-3' (SEQ ID NO: 52)). The resulting PCR products
contain 33 bp of homology with each other. An intact I-CreI coding sequence is obtained by assembly PCR and subsequently cloned in pCDNA6.2 from INVITROGEN (pCLS1768, Figure 19), a vector for expression in CHO cells, using the Gateway protocol (INVITROGEN). Resulting clones were verified by sequencing (MILLEGEN).

The same strategy is used with the following pair of oligonucleotides to create variants containing the V105A and 1132V substitutions, respectively:

* V105AF: 5'-aaacgcaaaccctggcttgaaattatcga-3' and V105AR: 5'-ttgatatttctagcagctttgtcgcctgttt-3' (SEQ ID NO: 57 and 58);

* 1132VF: 5'-acctgggtagttgtggctgtgctgaacagat-3' and 1132VR: 5'-atcgctgcaacactgtatacaccaggt-3' (SEQ ID NO: 59 and 60).

b) Construction of site directed variants (multiple mutations)

To obtain multiple insertions of the F54L, V105A and I132V substitutions into the coding sequence of HBV3.3 and HBV3.4 variants, 4 groups of separate overlapping PCR reactions were carried out that amplify internal fragments of the I-CreI N75 coding sequence containing the sequences between the different mutations. As an example, for the multiple site directed mutagenesis for the insertion of the mutations F54L and V105A, PCR amplification is carried out using a forward primer specific to the I-CreI coding sequence for amino acids 49-59 either with or without the substitution F54L (F54LF: 5'-accagcggctttgcgtctggacaactagtg-3' (SEQ ID NO: 51) and F54wtF: 5'-accagcggctttgcgtctggacaactagtg-3' (SEQ ID NO: 194)) and a reverse primer specific to the I-CreI coding sequence for amino acids 100-110 either with or without V105A (V105AR 5'-ttgatattttcagcaccaggttggcctgttt-3' (SEQ ID NO: 58) and V105wtR 5'-ttgatattttcagcaccaggttggcctgttt-3' (SEQ ID NO: 195)), leading to the generation of four different PCR fragments.

The same strategy was used with the following groups of oligonucleotides to create the other internal fragments:

* attBl-ICreIFor (5'-gggacaagttgttacaaaaaaagcagcttcggaagagataaaccatggaactggccaatataaataataaaagagttcG-3'; SEQ ID NO: 110) with F54LR 5'-cactagttttgcctcgcagcacaaggcggtgtcgtggggtcgtggt-3' (SEQ ID NO: 52) and F54wtR 5'-cactagttttgcctcgcagcacaaggcggtgtcgtggggtcgtggt-3' (SEQ ID NO: 196)
The resulting overlapping PCR products contain 15 bp of homology with each other. The PCR fragments corresponding to each internal region were then purified and I-Crel coding sequences containing the mutations F54L, V105A and I132V were generated by PCR assembly. The PCR products are subsequently cloned in pCDNA6.2 from INVITROGEN (pCLS1768, Figure 19), a vector for expression in CHO cells, using the Gateway protocol (INVITROGEN). Resulting clones were verified by sequencing (MILLEGEN).

c) Extrachromosomal assay in mammalian cells

Extrachromosomal assay in mammalian cells was performed as described in example 22.

2) Results

All possible combinations of three site directed mutations (F54L, V105A and I132V) were inserted into a HBV3.3 variant, HBV3.3_F1 (1-OeI 26R,33C,38S,44N 568Y,70S,75R/77Y,81T,139R, SEQ ID NO: 183) and an HBV3.4 variant, HBV3.4_A7 (I-Crel 19S,33C,38Y,44K,68Y,70S,75N,77Q,140M, SEQ ID NO: 175). Thus, seven site-directed variants were generated for each variant (+54L5 +54L 105A5 +54L 132V, +105A, +105A 132V, +132V, +54L 105A 132V) and were cloned in the pCLS1768 vector allowing expression of the variant in CHO cells (Figure 19).

AU pair-wise combinations of the HBV3.3 and HBV3.4 variants containing site-directed mutations were screened using the extrachromosomal assay in CHO cells. The screen is carried out by co-transfection of 3 plasmids in CHO cells:
one expressing a variant cleaving HBV3.3, a second expressing a variant cleaving HBV3.4 and a third one containing the HBV3 target cloned in pCLS1058 (Figure 18).

Five different heterodimers were found to trigger improved cleavage of the HBV3 target in a primary screen in CHO cells. These heterodimers consisted of a HBV3.3 variant containing site-directed mutations co-expressed with either the initial HBV3.4 variant or one of four HBV3.4 variants containing site-directed mutations (Table XXVI). These five heterodimers were validated in a secondary screen (Figure 30). In the secondary screen, the efficiency of the five heterodimers was compared to the initial heterodimer (HBV3.3_F1 variant I-OeI 26R,33C,38S,44N,68Y,70S,75R,77Y,81T,139R, SEQ ID NO: 183 co-expressed with the HBV3.4_A7 variant l-Crel 19S,33C,38Y,44K,68Y,70S,75N,77Q,140M, SEQ ID NO: 175). All five optimized heterodimers were able to cleave the HBV3 target with an efficacy superior to that observed with the initial heterodimer.

**Table XXVI: l-Crel variant combinations displaying improved cleavage efficiency of the HBV3 target in CHO cells.**

<table>
<thead>
<tr>
<th>Optimized variant cleaving HBV3.3</th>
<th>Optimized variant derived from variants cleaving the HBV3.4 target</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3_R5</td>
<td>3.4_A7: 19S 33C 38Y 44K 68Y 70S 75N 77Q 140M</td>
<td>186</td>
</tr>
<tr>
<td>26R 33C 38S 44N 68Y 70S 75R 77Y 81T 105A 132V 139R (SEQ ID NO: 185)</td>
<td>3.4_R2: 19S 33C 38Y 44K 54L 68Y 70S 75N 77Q 105A 140M</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>3.4_R4: 19S 33C 38Y 44K 68Y 70S 75N 77Q 105A 140M</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>3.4_R5: 19S 33C 38Y 44K 68Y 70S 75N 77Q 105A 132V 140M</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>3.4_R6: 19S 33C 38Y 44K 68Y 70S 75N 77Q 132V 140M</td>
<td>190</td>
</tr>
</tbody>
</table>

* Mutations resulting from site-directed mutagenesis are in bold.

**Example 25: Single-chain**

The optimized HBV3 heterodimer obtained by co-expression of the two variants HBV3.3_R5 and HBV3.4_R4 efficiently cleaves the HBV3 target but will also cleave the HBV3.3 and HBV3.4 targets because of the presence of the two homodimers. To avoid this unwanted cleavage activity, a single chain molecule composed of the two l-Crel derived variants 3.3_R5 and 3.4_R4 was generated. The single chain construct was engineered using the linker RM2...
(AAGGSDKYNQALSKYNQALSKYNQALSGGGGS), resulting in the production of the single chain molecule 3.3_R5-RM2-3.4_R4, also called SC_34. In a second step, mutations K7E, K96E were introduced into the 3.3JR.5 variant and mutations E8K, E61R into the 3.4_R4 variant of 3.3_R5-RM2-3.4_R4 to create the single chain molecule: 3.3_R5(K7E K96E)-RM2-3.4_R4(E8K E61R) that is also called SC_OH_34. The resulting single chain constructs were then tested in an extrachromosomal assay in CHO for their ability to cleave the HBV3 target.

1) Materials and methods
   a) Cloning of the single chain molecules

The two single chain molecules 3.3_R5-RM2-3.4_R4 (SEQ ID NO: 191) and 3.3_R5(K7E K96E)-RM2-3.4_R4(E8K E61R) (SEQ ID NO: 192) were synthesized by MWG and cloned into pCLS1768 (Figure 19).

b) Extrachromosomal assay in mammalian cells

Extrachromosomal assay in mammalian cells was performed as described in example 22.

2) Results

The activity of the two HBV3 single chain molecules SC_34 and SC_OH_34 was monitored against the HBV3 target using the previously described extrachromosomal assay in CHO cells. The ability of these single-chain molecules to cleave the HBV3 target was compared to the heterodimeric meganuclease (3.3_R5 / 3.4_R4) as well as the I-Scel meganuclease against its proper target (tagggataacaggttaat: SEQ ID NO: 112).

The results of this screen (Figure 31), indicate that both single chain molecules, SC_34 and SC_OH_34, display a cleavage activity for the HBV3 target that is similar if not greater than the heterodimeric meganuclease (3.3_R5 / 3.4_R4). In addition the cleavage activity, observed with the single-chain molecule is as active as I-Scel against its proper target. These results demonstrate that it is possible to improve the specificity of the HBV3 meganuclease by generating a single-chain molecule without affecting its activity toward the DNA target of interest.
CLAIMS

1. An I-Crel variant, characterized in that 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 non-integrating virus (NIV) genome, and being obtainable by a method comprising at least the steps of:

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

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

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

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

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

(f) selecting and/or screening the variants from the second series of step (b) which are able to cleave a mutant 1-Oei site wherein at least one of (i) the nucleotide triplet in positions +3 to +5 of the l-Crel site has been replaced with the nucleotide triplet which is present in positions +3 to +5 of said DNA target sequence from the NIV 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 +3 to +5 of said DNA target sequence from the NIV 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 l-Crel 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 NIV 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 NIV 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 NIV 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 NIV 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 1-Oei variant which cleaves a sequence wherein (i) the nucleotide triplet in positions +8 to +10 of the l-Crel site has been replaced with the nucleotide triplet which is present in positions +8 to +10 of said DNA target sequence from the NIV 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 NIV 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 NIV 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 NIV 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 NIV 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 NIV 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 NIV 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-Crel are in positions 44, 68, 70, 75 and/or 77.

6. The variant of any one of claims 1 to 5, wherein said substitutions) in the subdomain situated from positions 26 to 40 of I-Crel 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-Crel 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 NIV genome.

8. The variant of any one of claims 1 to 7, which comprises one or more substitutions on the entire \( l\)-Crel sequence that improve the binding and/or the cleavage properties of the variant towards said DNA target sequence from the NIV genome.

9. The variant according to any one of claims 1 to 8, wherein said DNA target is located in an element of the NIV 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 \( l\)-Crel, said heterodimer being able to cleave a non-palindromic DNA target sequence from the HBV.

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.


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 NIV is from a family selected from the group *Herpesviridae, Adenoviridae, Papovaviridae, Poxviridae, Parvoviridae, Hepadnaviridae.*
17. The variant according to any one of claims 1 to 16, wherein said NIV is hepatitis B and in particular hepatitis B strain A938,

18. The variant of claim 17 wherein said DNA target is selected from the group consisting of the SEQ ID NO: 6 to 9, 79 to 82, 117 to 122.

19. The variant according to claim 17 or 18, wherein said variant is selected from the group consisting of SEQ ID NO: 15 to 20, 22 to 27, 29, 32 to 48, 61 to 75, 84 to 89, 91, 93 to 96, 98 to 109, 124 to 128, 130 to 134, 137 to 144, 146 to 153, 155 to 192.

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

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

22. The vector of claim 21, 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 NIV genome.

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

24. The vector of claim 23, wherein the sequence which inactivates the NIV 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.

25. The vector of any one of claims 21 to 24, wherein said sequence sharing homologies with the regions surrounding DNA target sequence from the NIV genome is a fragment of the NIV genome comprising sequences upstream and downstream of the cleavage site, so as to allow the deletion of coding sequences flanking the cleavage site.

26. A host cell which is modified by a polynucleotide of claim 20 or a vector of anyone of claims 21 to 25.

27. A non-human transgenic animal which is modified by a polynucleotide of claim 20 or a vector of anyone of claims 21 to 25.
28. A transgenic plant which is modified by a polynucleotide of claim 20 or a vector of anyone of claims 21 to 25.

29. Use of at least one variant of anyone of claims 1 to 19, or at least one vector according to anyone of claims 21 to 25, for genome engineering, for non-therapeutic purposes.

30. Use of a variant according to any one of claims 1 to 19, a nucleic acid molecule according to claim 20 or a vector according to any one of claims 21 to 25 to prepare a medicament to treat an NIV infection.
Figure 2
Figure 11
Figure 25
Figure 31

OD420

3.3 R5
3.4 R4

SC_34
SC_OH_34

I-Scel target
Empty vector
Figure 35

HBV3: CCGCTTATTTTTTTGGAAGAA (core protein), SEQ ID NO: 117
HBV8: ATTGACCTTTAAGATGTTGGA (core protein), SEQ ID NO: 79
HBV12: ATATTTTTTGGAAAAGATGCTACA (polymerase gene), SEQ ID NO: 6