The invention relates to methods and compositions for treatment of human cytomegalovirus (HCMV) infection. The compositions include isolated nucleic acid sequences comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a particular target sequence in the HCMV genome.
METHODS AND COMPOSITIONS FOR RNA-GUIDED TREATMENT OF HUMAN CYTOMEGALOVIRUS (HCMV) INFECTION

FIELD OF THE INVENTION:
The invention relates to fields of virology and more particularly to methods and compositions for treatment of human cytomegalovirus (HCMV) infection.

BACKGROUND OF THE INVENTION:
HCMV, the prototype β-herpesvirus, is the cause of a "silent pandemic" that continuously inflicts suffering upon people including immunocompromised patients as well as pregnant mothers and their unborn or prematurely newborn babies. HCMV primary infection is followed by lifelong persistence in the host organism as a chronic and usually asymptomatic infection. However, in immunocompromised patients, HCMV primary infection and/or reactivation is still a significant cause of morbidity and mortality.

In the absence of an approved vaccine, viral DNA polymerase inhibitors, including the nucleoside analog ganciclovir, have provided major advances in HCMV disease management. However, use of these drugs is limited by significant toxicity and problems with viral drug resistance. Furthermore, no drug has an effect on the virus in a latent state, and recent trials with newly developed antivirals targeting HCMV have not lived up to their expectations. Thus, it is still imperative to develop new anti-HCMV strategies directed at appropriate viral molecular targets.

SUMMARY OF THE INVENTION:
In a first aspect, the invention relates to a method of inactivating a human cytomegalovirus (HCMV) in a cell, comprising exposing the cell to a composition comprising an isolated nucleic acid sequence encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV immediate-early (IE) gene.

In a second aspect, the invention relates to an isolated nucleic acid sequence encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide
RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene.

In a third aspect, the invention relates to an expression vector comprising the nucleic acid sequence of the invention.

In a fourth aspect, the invention relates to a composition comprising the nucleic acid sequence of the invention, or the expression vector comprising thereof.

In a fifth aspect, the invention relates to a method of reducing the risk of the transmission of a HCMV infection from an HCMV-infected donor to a recipient in need thereof, comprising treating the cell, tissue or organ obtained from said donor with a therapeutically effective amount of at least one isolated nucleic acid sequence of the invention.

**DETAILED DESCRIPTION OF THE INVENTION:**

The present invention is based on the discovery that we could eliminate the HCMV genome from HCMV infected cells by using the RNA-guided Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-Cas 9 nuclease system (Cas9/gRNA) in single and multiplex configurations. The inventors identified highly specific targets within the HCMV IE gene that were efficiently edited by Cas9/gRNA, inactivating viral gene expression and replication in all infected cells (i.e. replicating and latently-infected cells).

Thus, they cloned some gRNA targeting HCMV IE gene in plasmid allowing lentiviral production in order to validate their strategy on infected cells eventually combined with classical anti-HCMV agent (e.g. ganciclovir). The present results suggest that Cas9/gRNA can be engineered to provide a specific, efficacious prophylactic and therapeutic approach against HCMV. Interestingly, such an approach is not only useful for definitively treating HCMV-infected patients but also for reducing the risk of the transmission of a HCMV infection from an HCMV-infected donor to a recipient in need thereof by pretreating the graft.

Additionally, amongst the gRNA targeting HCMV IE gene, the inventors identified gRNA#2 targeting HCMV IE exon 2 as the most promising gRNA in therapy since gRNA#2 does not show an off-target on human genome, shows a high percentage of indels (80%) and induces an appropriate reduction in the amount of IE+ cells. Moreover, the multiplex strategy comprising the gRNA#2, gRNA#4 and gRNA#5 is even more efficient that the singleplex strategy comprising only the gRNA#2.
Accordingly, in an aspect, the invention relates to an isolated nucleic acid sequence encoding a gene editing complex comprising a sequence encoding a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene.

The nucleic acid sequences of the invention include nucleic acids encoding a CRISPR-associated endonuclease, e.g., Cas9, Cas9-HF, and a guide RNA that is complementary to a target nucleic acid sequence within the HCMV IE gene.

In bacteria the CRISPR/Cas loci encode RNA-guided adaptive immune systems against mobile genetic elements (viruses, transposable elements and conjugative plasmids). Three types (I-III) of CRISPR systems have been identified. CRISPR clusters contain spacers, the sequences complementary to antecedent mobile elements. CRISPR clusters are transcribed and processed into mature CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) RNA (crRNA). The CRISPR-associated endonuclease, Cas9, belongs to the type II CRISPR/Cas system and has strong endonuclease activity to cut target DNA. Cas9 is guided by a mature crRNA that contains about 20 base pairs (bp) of unique target sequence (called spacer) and a trans-activated small RNA (tracrRNA) that serves as a guide for ribonuclease III-aided processing of pre-crRNA. The crRNA: tracrRNA duplex directs Cas9 to target DNA via complementary base pairing between the spacer on the crRNA and the complementary sequence (called protospacer) on the target DNA. Cas9 recognizes a trinucleotide (NGG) protospacer adjacent motif (PAM) to specify the cut site (the 3rd nucleotide from PAM). The crRNA and tracrRNA can be expressed separately or engineered into an artificial fusion small guide RNA (sgRNA) via a synthetic stem loop to mimic the natural crRNA/tracrRNA duplex. Such sgRNA, like shRNA, can be synthesized or in vitro transcribed for direct RNA transfection or expressed from U6 or H1-promoted RNA expression vector, although cleavage efficiencies of the artificial sgRNA are lower than those for systems with the crRNA and tracrRNA expressed separately.

In some embodiments, the CRISPR-associated endonuclease can be a Cas9 nuclease. The Cas9 nuclease can have a nucleotide sequence identical to the wild type Streptococcus pyogenes sequence. In some embodiments, the CRISPR-associated endonuclease can be a sequence from other species, for example other Streptococcus species, such as thermophilus;
Pseudomonas aeruginosa, Escherichia coli, or other sequenced bacteria genomes and archaea, or other prokaryotic microorganisms. Alternatively, the wild type Streptococcus pyogenes Cas9 sequence can be modified. The nucleic acid sequence can be codon optimized for efficient expression in mammalian cells, i.e., "humanized." A humanized Cas9 nuclease sequence can be for example, the Cas9 nuclease sequence encoded by any of the expression vectors listed in Genbank accession numbers KM099231.1 GL669193757; KM099232.1 GL669193761; or KM099233.1 GL669193765. Alternatively, the Cas9 nuclease sequence can be for example, the sequence contained within a commercially available vector such as PX330 or PX260 or VP12 (Plasmid #72247) from Addgene (Cambridge, MA).

In some embodiments, the Cas9 endonuclease can have an amino acid sequence that is a variant or a fragment of any of the Cas9 endonuclease sequences of Genbank accession numbers KM099231.1 GL669193757; KM099232.1; GL669193761; or KM099233.1 GL669193765 or Cas9 amino acid sequence of PX330 or PX260 or VP12 (Plasmid #72247) (Addgene, Cambridge, MA). The Cas9 nuclease sequence can be modified to encode biologically active variants of Cas9, and these variants can have or can include, for example, an amino acid sequence that differs from a wild type Cas9 by virtue of containing one or more mutations (e.g., an addition, deletion, or substitution mutation or a combination of such mutations). One or more of the substitution mutations can be a substitution (e.g., a conservative amino acid substitution). For example, a biologically active variant of a Cas9 polypeptide can have an amino acid sequence with at least or about 50% sequence identity (e.g., at least or about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity) to a wild type Cas9 polypeptide. Conservative amino acid substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine, and leucine; aspartic acid and glutamic acid; asparagine, glutamine, serine and threonine; lysine, histidine and arginine; and phenylalanine and tyrosine.

The Cas9 nuclease sequence can be a mutated sequence. For example the Cas9 nuclease can be mutated in the conserved FiNH and RuvC domains, which are involved in strand specific cleavage. For example, an aspartate-to-alanine (D10A) mutation in the RuvC catalytic domain allows the Cas9 nickase mutant (Cas9n) to nick rather than cleave DNA to yield single-stranded breaks, and the subsequent preferential repair through HDR can potentially decrease the frequency of unwanted indel mutations from off-target double-stranded breaks.
In some embodiments, the CRISPR-associated endonuclease can be a Cas9-HF nuclease. This kind of CRISPR Cas9 is a high fidelity (HF) variant harbouring alterations designed to reduce non-specific interactions with the target DNA (Kleinstiver et al 2016).

The polypeptides that are biologically active variants of a CRISPR-associated endonuclease can be characterized in terms of the extent to which their sequence is similar to or identical to the corresponding wild-type polypeptide. For example, the sequence of a biologically active variant can be at least or about 80% identical to corresponding residues in the wild-type polypeptide. For example, a biologically active variant of a CRISPR-associated endonuclease can have an amino acid sequence with at least or about 80% sequence identity (e.g., at least or about 85%, 90%, 95%, 97%, 98%, or 99% sequence identity) to a CRISPR-associated endonuclease or to a homolog or ortholog thereof.

A biologically active variant of a CRISPR-associated endonuclease polypeptide will retain sufficient biological activity to be useful in the present methods. The biologically active variants will retain sufficient activity to function in targeted DNA cleavage. The biological activity can be assessed in ways known to one of ordinary skill in the art and includes, without limitation, in vitro cleavage assays or functional assays.

Alternatively, the CRISPR-associated endonuclease polypeptide is a catalytically inactive CRISPR-associated endonuclease polypeptide such as a catalytically inactive Cas9 (dCas9) which can be used to knock-down gene expression by interfering with transcription of the gene of interest (i.e. IE HCMV). Thus, the guide RNA of interest is fused to the nucleic acid sequence of such a dCas9 in order to obtain dCas9-repressor peptide fusion. For instance, said dCAS9 is a Cas9-KRAB (e.g. pHR-SFFV-KRAB-dCas9-P2A-mCherry (Addgene)).

In some embodiments, the guide RNA can be a sequence complementary to a coding or a non-coding sequence. For example, the guide RNA can be sequence complimentary to an HCMV IE sequence, such as a protein coding sequence or a regulatory sequence.

As used herein, the term "cytomegalovirus" refers to a viral genus of the viral family known as Herpesviridae or herpesviruses. It is typically abbreviated as CMV. The species that infects humans is commonly known as human CMV (HCMV) or human herpesvirus-5.
(HHV-5), and is the most studied of all cytomegaloviruses. All herpesviruses share a characteristic ability to remain latent within the body over long periods.

As used herein, the term "HCMV immediate-early (IE) gene" refers to a gene that is believed to have a decisive role in acute infection and reactivation from viral latency. Although a variety of protein products expressed from these mRNAs have been identified, the 72-kDa IE1 and the 86-kDa IE2 nuclear phosphoproteins are the most abundant and important. They share 85 amino-terminal amino acids corresponding to major IE exons 2 and 3 but have distinct carboxy-terminal parts encoded by exon 4 (IE1; UL123) or exon 5 (IE2; UL122). IE2 is absolutely indispensable for productive viral replication, since IE2 is the principal transcriptional activator of the CMV early genes (Paulus and Nevers 2009). The naturally occurring HCMV immediate-early (IE) proteins have amino acid sequences as shown in UniProt Accession number P03169 (IE1, UL123, Towne strain) and Q6SWP7 (IE2, UL122, Towne strain) and is encoded by an alternative splicing of the Immediate Early gene provided in the GenBank database under accession number FJ616285 (coding sequence for IE1 (UL123) is the junction of the following exons 171498..172717,172888..173072, 173185..173255 ; coding sequence for IE2 (UL122) is the junction of the following exons 169859..171342,172888..173072,173185..173255).

The nucleic acid sequences of the invention include sequence encoding a guide RNA (gRNA) comprising a sequence that is complementary to a target nucleic acid sequence within the coding region of the HCMV immediate-early (IE) gene. In some embodiments, the guide RNA comprises a sequence that is complementary to major IE exon 2 and 3 but also to carboxy-terminal parts encoded by exon 4 (IE1) or exon 5 (IE2). Thus, the guide RNA sequence is complementary to a polynucleotide selected from the group consisting of a polynucleotide comprised in IE exon 2, in IE exon 3, in IE exon 4 and IE exon 5.

Exemplary guide RNA sequences that target the IE exon 2 and 5 of HCMV are shown in Table 1. A guide RNA sequence can comprise or consist of, for example, the sequences of:

<table>
<thead>
<tr>
<th>Name gRNA</th>
<th>Sequence 5'-3'</th>
<th>SED ID N°</th>
<th>Targeted Exon IE</th>
</tr>
</thead>
</table>

Table 1: Exemplary guide RNA sequences that target the IE exons 2 and 5 of HCMV
<table>
<thead>
<tr>
<th>gRNA#1</th>
<th>gtccatctttctttgccag</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>gRNA#2</td>
<td>ggactccatcgtcaagga</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>gRNA#3</td>
<td>gtcacctcgccacgacga</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>gRNA#4</td>
<td>GTCCTGGATGGCTGCCTCGA</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>gRNA#5</td>
<td>GGTGCTACTGGGAATCGATAC</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Accordingly, in one embodiment, the guide RNA sequence comprises a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5. In a particular embodiment, the guide RNA sequence consists of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5.

The invention is not so limiting however, and the guide RNA sequences can be selected to target any variant or mutant HCMV sequence. In some embodiments, the guide RNA can include a variant sequence or quasi-species sequence. In some embodiments, the guide RNA can be a sequence corresponding to a sequence in the genome of the virus harbored by the subject undergoing treatment. Thus for example, the sequence of the particular IE gene in the HCMV virus harbored by the subject can be obtained and guide RNAs complementary to the patient's particular sequences can be used.

The guide RNA sequence can be a sense or anti-sense sequence. The complementary target sequence to the guide RNA sequence includes a proto-spacer adjacent motif (PAM) which is sensed by the Cas9. The sequence of the PAM can vary depending upon the specificity requirements of the CRISPR endonuclease used. In the CRISPR-Cas system derived from *S. pyogenes*, the target DNA typically immediately precedes a 5'-NGG proto-spacer adjacent motif (PAM). Thus, for the *S. pyogenes* Cas9, the PAM sequence can be AGG, TGG, CGG or GGG. Other Cas9 orthologs may have different PAM specificities. The specific sequence of the guide RNA may vary, but, regardless of the sequence, useful guide RNA sequences will be those that minimize off-target effects while achieving high efficiency and complete ablation of the genomically integrated HCMV genome. The length of the guide RNA sequence can vary from about 19 to about 60 or more nucleotides, for example about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 45, about 50, about 55, about 60 or more nucleotides.
Useful selection methods identify regions having extremely low homology between the foreign viral genome and host cellular genome including endogenous retroviral DNA, include bioinformatic screening using 12-bp+NGG target-selection criteria to exclude off-target human transcriptome or (even rarely) untranslated-genomic sites, Sanger sequencing and SURVEYOR assay, to identify and exclude potential off-target effects.

The guide RNA sequence can be configured as a single sequence or as a combination of one or more different sequences, e.g., a multiplex configuration. Multiplex configurations can include combinations of two, three, four, five, six, seven, eight, nine, ten, or more different guide RNAs, for example any combination of sequences in IE exon 2 and exon 5.

In some embodiments, the RNA molecules e.g. crRNA, tracrRNA, gRNA are engineered to comprise one or more modified nucleobases. For example, known modifications of RNA molecules can be found, for example, in Genes VI, Chapter 9 ("Interpreting the Genetic Code"), Lewis, ed. (1997, Oxford University Press, New York), and Modification and Editing of RNA, Grosjean and Benne, eds. (1998, ASM Press, Washington DC). Modified RNA components include the following: 2'-0-methylcytidine; 5-methylcytidine; 5,2'-0-dimethylcytidine; 5-hydroxymethylcytidine; 5-formylcytidine; 3'-methylcytidine; 2-thiocytidine; lysidine; 2'-0-methyluridine; 2-thiouridine; 2-thio-2'-0-methyluridine; 3,2'-0-dimethyluridine; 4-thiouridine; 5-methyl-2-thiouridine; 5-hydroxyuridine; 5-methoxyuridine; dihydrouridine; inosine; 20'-methyl inosine; 1-methyl inosine; and 1-methyl guanosine.

An "isolated" nucleic acid can be, for example, a naturally-occurring DNA molecule or a fragment thereof, provided that at least one of the nucleic acid sequences normally found immediately flanking that DNA molecule in a naturally-occurring genome is removed or absent. Thus, an isolated nucleic acid includes, without limitation, a DNA molecule that exists as a separate molecule, independent of other sequences (e.g., a chemically synthesized nucleic acid, or a cDNA or genomic DNA fragment produced by the polymerase chain reaction (PCR) or restriction endonuclease treatment). An isolated nucleic acid also refers to a DNA molecule that is incorporated into a vector, an autonomously replicating plasmid, a virus, or into the genomic DNA of a prokaryote or eukaryote. In addition, an isolated nucleic acid can include an engineered nucleic acid such as a DNA molecule that is part of a hybrid or fusion nucleic acid. A nucleic acid existing among many (e.g., dozens, or hundreds to millions) of
other nucleic acids within, for example, cDNA libraries or genomic libraries, or gel slices containing a genomic DNA restriction digest, is not an isolated nucleic acid. Isolated nucleic acid molecules can be produced by standard techniques. For example, polymerase chain reaction (PCR) techniques can be used to obtain an isolated nucleic acid containing a nucleotide sequence described herein, including nucleotide sequences encoding a polypeptide described herein. PCR can be used to amplify specific sequences from DNA as well as RNA, including sequences from total genomic DNA or total cellular RNA. Various PCR methods are described in, for example, PCR Primer: A Laboratory Manual, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Generally, sequence information from the ends of the region of interest or beyond is employed to design oligonucleotide primers that are identical or similar in sequence to opposite strands of the template to be amplified. Various PCR strategies also are available by which site-specific nucleotide sequence modifications can be introduced into a template nucleic acid.

Isolated nucleic acids also can be chemically synthesized, either as a single nucleic acid molecule (e.g., using automated DNA synthesis in the 3’ to 5’ direction using phosphoramidite technology) or as a series of oligonucleotides. For example, one or more pairs of long oligonucleotides (e.g., >50-100 nucleotides) can be synthesized that contain the desired sequence, with each pair containing a short segment of complementarity (e.g., about 15 nucleotides) such that a duplex is formed when the oligonucleotide pair is annealed. DNA polymerase is used to extend the oligonucleotides, resulting in a single, double-stranded nucleic acid molecule per oligonucleotide pair, which then can be ligated into a vector.

The invention also relates to an isolated nucleic acid sequence encoding a guide RNA which is complementary to a target nucleic acid sequence within the HCMV IE gene as above-described. For instance, the guide RNA sequence is complementary to a polynucleotide selected from the group consisting of a polynucleotide comprised in IE exon 2, in IE exon 3, in IE exon 4 and IE exon 5. In one embodiment, the guide RNA sequence comprises a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5. In a particular embodiment, the guide RNA sequence consists of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5.
In another aspect, the invention relates to an expression vector comprising an isolated nucleic acid sequence of the invention as above-described.

When the nucleic acid sequences are administered in an expression vector, the guide RNAs can be encoded by a single vector. Alternatively, multiple vectors can be engineered to each include two or more different guide RNAs. Useful configurations will result in the excision of viral sequences between cleavage sites resulting in the ablation of HCMV genome or HCMV protein expression. Thus, the use of two or more different guide RNAs promotes excision of the viral sequences between the cleavage sites recognized by the CRISPR endonuclease. The excised region can vary in size from a single nucleotide to several thousand nucleotides.

When the nucleic acid sequences are contained within an expression vector, the CRISPR endonuclease can be encoded by the same nucleic acid or vector as the guide RNA sequences. Alternatively or in addition, the CRISPR endonuclease can be encoded in a physically separate nucleic acid from the guide RNA sequences or in a separate vector.

Recombinant constructs are also provided herein and can be used to transform cells in order to express Cas9 and/or a guide RNA complementary to a target sequence in HCMV IE gene. A recombinant nucleic acid construct comprises a nucleic acid encoding a Cas9 and/or a guide RNA complementary to a target sequence in HCMV IE as described herein, operably linked to a regulatory region suitable for expressing the Cas9 and/or a guide RNA complementary to a target sequence in HCMV IE gene in the cell. It will be appreciated that a number of nucleic acids can encode a polypeptide having a particular amino acid sequence. The degeneracy of the genetic code is well known in the art. For many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. For example, codons in the coding sequence for Cas9 can be modified such that optimal expression in a particular organism is obtained, using appropriate codon bias tables for that organism.

Vectors containing nucleic acids such as those described herein also are provided. A "vector" is a replicon, such as a plasmid, phage, or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. Generally, a vector is capable of replication when associated with the proper control elements. Suitable vector backbones include, for example, those routinely used in the art such as plasmids,
viruses, artificial chromosomes, BACs, YACs, or PACs. The term "vector" includes cloning and expression vectors, as well as viral vectors and integrating vectors. An "expression vector" is a vector that includes a regulatory region. A wide variety of host/expression vector combinations may be used to express the nucleic acid sequences described herein. Suitable expression vectors include, without limitation, plasmids and viral vectors derived from, for example, bacteriophage, baculo viruses, and retroviruses. Numerous vectors and expression systems are commercially available from such corporations as Novagen (Madison, WI), Clontech (Palo Alto, CA), Stratagene (La Jolla, CA), and Invitrogen/Life Technologies (Carlsbad, CA).

The vectors provided herein also can include, for example, origins of replication, scaffold attachment regions (SARs), and/or markers. A marker gene can confer a selectable phenotype on a host cell. For example, a marker can confer biocide resistance, such as resistance to an antibiotic (e.g., kanamycin, G418, bleomycin, or hygromycin). As noted above, an expression vector can include a tag sequence designed to facilitate manipulation or detection (e.g., purification or localization) of the expressed polypeptide. Tag sequences, such as green fluorescent protein (GFP), glutathione S-transferase (GST), polyhistidine, c-myc, hemagglutinin, or Flag™ tag (Kodak, New Haven, CT) sequences typically are expressed as a fusion with the encoded polypeptide. Such tags can be inserted anywhere within the polypeptide, including at either the carboxyl or amino terminus.

Additional expression vectors also can include, for example, segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, e.g., E. coli plasmids col El, pCRI, pBR322, pMal-C2, pET, pGEX, pMB9 and their derivatives, plasmids such as RP4; phage DNAs, e.g., the numerous derivatives of phage 1, e.g., NM989, and other phage DNA, e.g., M1 3 and filamentous single stranded phage DNA; yeast plasmids such as the 2µ plasmid or derivatives thereof, vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences. Yeast expression systems can also be used. For example, the non-fusion pYES2 vector (Xbal, SplI, Shol, NotI, GslXI, EcoRI, BstXI, BamHI, Sad, Kpnl, and HindIII cloning sites; Invitrogen) or the fusion pYESHisA, B, C (Xbal, SplI, Shol, NotI, BstXI, EcoRI, BamHI, Sad, Kpnl, and HindIII cloning sites, N-terminal peptide purified with ProBond resin.
and cleaved with enterokinase; Invitrogen), to mention just two, can be employed according to the invention. A yeast two-hybrid expression system can also be prepared in accordance with the invention.

The vector can also include a regulatory region. The term "regulatory region" refers to nucleotide sequences that influence transcription or translation initiation and rate, and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5' and 3' untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, nuclear localization signals, and introns.

As used herein, the term "operably linked" refers to positioning of a regulatory region and a sequence to be transcribed in a nucleic acid so as to influence transcription or translation of such a sequence. For example, to bring a coding sequence under the control of a promoter, the translation initiation site of the translational reading frame of the polypeptide is typically positioned between one and about fifty nucleotides downstream of the promoter. A promoter can, however, be positioned as much as about 5,000 nucleotides upstream of the translation initiation site or about 2,000 nucleotides upstream of the transcription start site. A promoter typically comprises at least a core (basal) promoter. A promoter also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). The choice of promoters to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level, and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning promoters and other regulatory regions relative to the coding sequence.

Vectors include, for example, viral vectors (such as adenoviruses ("Ad"), adeno-associated viruses (AAV), and vesicular stomatitis virus (VSV) and retroviruses), liposomes and other lipid-containing complexes, and other macromolecular complexes capable of mediating delivery of a polynucleotide to a host cell. Vectors can also comprise other components or functionalities that further modulate gene delivery and/or gene expression, or that otherwise provide beneficial properties to the targeted cells.
A "recombinant viral vector" refers to a viral vector comprising one or more heterologous gene products or sequences. Since many viral vectors exhibit size-constraints associated with packaging, the heterologous gene products or sequences are typically introduced by replacing one or more portions of the viral genome. Such viruses may become replication-defective, requiring the deleted function(s) to be provided in trans during viral replication and encapsidation (by using, e.g., a helper virus or a packaging cell line carrying gene products necessary for replication and/or encapsidation). Modified viral vectors in which a polynucleotide to be delivered is carried on the outside of the viral particle have also been described (see, e.g., Curiel, D T, et al. PNAS 88: 8850-8854, 1991).

Suitable nucleic acid delivery systems include recombinant viral vector, typically sequence from at least one of an adenovirus, adenovirus-associated virus (AAV), helper-dependent adenovirus, retrovirus, or hemagglutinating virus of Japan-liposome (HVJ) complex. In such cases, the viral vector comprises a strong eukaryotic promoter operably linked to the polynucleotide. The recombinant viral vector can include one or more of the polynucleotides therein, preferably about one polynucleotide. In some embodiments, the viral vector used in the invention methods has a pfu (plaque forming units) of from about 10^8 to about 5x10^10 pfu. In embodiments in which the polynucleotide is to be administered with a non-viral vector, use of between from about 0.1 nanograms to about 4000 micrograms will often be useful e.g., about 1 nanogram to about 100 micrograms.

Additional vectors include viral vectors, fusion proteins and chemical conjugates.

In another aspect, the invention relates to a composition comprising an isolated nucleic acid sequence encoding a gene editing complex comprising a sequence encoding a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene as above-described or an expression vector comprising thereof.

Alternatively, the invention relates to a composition comprising a CRISPR-associated endonuclease and an isolated nucleic acid sequence encoding one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene as above-described.
When the CRISPR-associated endonuclease such as Cas9 or Cas9-HF is administered as a polypeptide, the bonds between the amino acid residues can be conventional peptide bonds or another covalent bond (such as an ester or ether bond), and the polypeptides can be modified by amidation, phosphorylation or glycosylation. A modification can affect the polypeptide backbone and/or one or more side chains. Chemical modifications can be naturally occurring modifications made in vivo following translation of an mRNA encoding the polypeptide (e.g., glycosylation in a bacterial host) or synthetic modifications made in vitro. A biologically active variant of a CRISPR-associated endonuclease can include one or more structural modifications resulting from any combination of naturally occurring (i.e., made naturally in vivo) and synthetic modifications (i.e., naturally occurring or non-naturally occurring modifications made in vitro). Examples of modifications include, but are not limited to, amidation (e.g., replacement of the free carboxyl group at the C-terminus by an amino group); biotinylation (e.g., acylation of lysine or other reactive amino acid residues with a biotin molecule); glycosylation (e.g., addition of a glycosyl group to either asparagines, hydroxylysine, serine or threonine residues to generate a glycoprotein or glycopeptide); acetylation (e.g., the addition of an acetyl group, typically at the N-terminus of a polypeptide); alkylation (e.g., the addition of an alkyl group); isoprenylation (e.g., the addition of an isoprenoid group); lipoylation (e.g. attachment of a lipoate moiety); and phosphorylation (e.g., addition of a phosphate group to serine, tyrosine, threonine or histidine).

The amino acid residues in the CRISPR-associated endonuclease sequence such as in the Cas9 or Cas9-HF amino acid sequence can be non-naturally occurring amino acid residues. Naturally occurring amino acid residues include those naturally encoded by the genetic code as well as non-standard amino acids (e.g., amino acids having the D-configuration instead of the L-configuration). The polypeptides can also include amino acid residues that are modified versions of standard residues (e.g. pyrrolysine can be used in place of lysine and selenocysteine can be used in place of cysteine). Non-naturally occurring amino acid residues are those that have not been found in nature, but that conform to the basic formula of an amino acid and can be incorporated into a polypeptide.

Thus, the compositions of the invention can be prepared in a variety of ways known to one of ordinary skill in the art. Regardless of their original source or the manner in which they are obtained, the compositions of the invention can be formulated in accordance with their use. For example, the nucleic acids and vectors described above can be formulated within
compositions for application to cells in tissue culture or for administration to a patient or
subject. Any of the pharmaceutical compositions of the invention can be formulated for use in
the preparation of a medicament, and particular uses are indicated below in the context of
treatment, e.g., the treatment of a subject having an HCMV infection or at risk for contracting
and HCMV infection. When employed as pharmaceuticals, any of the nucleic acids and
vectors can be administered in the form of pharmaceutical compositions. These compositions
can be prepared in a manner well known in the pharmaceutical art, and can be administered
by a variety of routes, depending upon whether local or systemic treatment is desired and
upon the area to be treated. Administration may be topical (including ophthalmic and to
mucous membranes including intranasal, vaginal and rectal delivery), pulmonary (e.g., by
inhalation or insufflation of powders or aerosols, including by nebulizer; intratracheal,
intranasal, epidermal and transdermal), ocular, oral or parenteral. Methods for ocular delivery
can include topical administration (eye drops), subconjunctival, periocular or intravitreal
injection or introduction by balloon catheter or ophthalmic inserts surgically placed in the
conjunctival sac. Parenteral administration includes intravenous, intraarterial, subcutaneous,
intraperitoneal or intramuscular injection or infusion; or intracranial, e.g., intrathecal or
intraventricular administration.

Parenteral administration can be in the form of a single bolus dose, or may be, for
example, by a continuous perfusion pump. Pharmaceutical compositions and formulations for
topical administration may include transdermal patches, ointments, lotions, creams, gels,
drops, suppositories, sprays, liquids, powders, and the like. Conventional pharmaceutical
carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.
This invention also includes pharmaceutical compositions which contain, as the active
ingredient, nucleic acids and vectors described herein in combination with one or more
pharmaceutically acceptable carriers. As used herein, the terms "pharmaceutically acceptable"
(or "pharmacologically acceptable") refer to molecular entities and compositions that do not
produce an adverse, allergic or other untoward reaction when administered to an animal or a
human, as appropriate. The term "pharmaceutically acceptable carrier," as used herein,
includes any and all solvents, dispersion media, coatings, antibacterial, isotonic and
absorption delaying agents, buffers, excipients, binders, lubricants, gels, surfactants and the
like, that may be used as media for a pharmaceutically acceptable substance. In making the
compositions of the invention, the active ingredient is typically mixed with an excipient,
diluted by an excipient or enclosed within such a carrier in the form of, for example, a
capsule, tablet, sachet, paper, or other container. When the excipient serves as a diluent, it can be a solid, semisolid, or liquid material (e.g., normal saline), which acts as a vehicle, carrier or medium for the active ingredient. As is known in the art, the type of diluent can vary depending upon the intended route of administration.

The nucleic acid sequences of the invention can be delivered to an appropriate cell of a subject. This can be achieved by, for example, the use of a polymeric, biodegradable microparticle or microcapsule delivery vehicle, sized to optimize phagocytosis by phagocytic cells such as macrophages. For example, PLGA (poly-lacto-co-glycolide) microparticles approximately 1-10 µm in diameter can be used. The polynucleotide is encapsulated in these microparticles, which are taken up by macrophages and gradually biodegraded within the cell, thereby releasing the polynucleotide. Once released, the DNA is expressed within the cell. A second type of microparticle is intended not to be taken up directly by cells, but rather to serve primarily as a slow-release reservoir of nucleic acid that is taken up by cells only upon release from the micro-particle through biodegradation. These polymeric particles should therefore be large enough to preclude phagocytosis (i.e., larger than 5µm and preferably larger than 20µm). Another way to achieve uptake of the nucleic acid is using liposomes, prepared by standard methods. The nucleic acids can be incorporated alone into these delivery vehicles or co-incorporated with tissue-specific antibodies, for example antibodies that target cell types that are commonly latently infected reservoirs of HCMV infection, for example, CD34+ hematopoietic cells. Alternatively, one can prepare a molecular complex composed of a plasmid or other vector attached to poly-L-lysine by electrostatic or covalent forces. Poly-L-lysine binds to a ligand that can bind to a receptor on target cells. Delivery of "naked DNA" (i.e., without a delivery vehicle) to an intramuscular, intradermal, or subcutaneous site, is another means to achieve in vivo expression. In the relevant polynucleotides (e.g., expression vectors) the nucleic acid sequence encoding an isolated nucleic acid sequence comprising a sequence encoding a CRISPR-associated endonuclease and a guide RNA is operatively linked to a promoter or enhancer-promoter combination.

In another aspect, the invention relates to a method of inactivating a human cytomegalovirus HCMV in a cell, comprising exposing the cell to a composition comprising an isolated nucleic acid encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene.
In still another aspect, the invention relates to a method of inactivating a human cytomegalovirus HCMV in a cell, comprising exposing the cell to a composition comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene.

The methods can include exposing the cell to a composition comprising an isolated nucleic acid encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNAs wherein the guide RNA is complementary to a target nucleic acid sequence in the HCMV IE gene. The contacting step can take place in vivo, that is, the compositions can be administered directly to a subject having HCMV infection. The methods are not so limited however, and the contacting step can take place ex vivo. For example, a cell or plurality of cells, or a tissue explant or a graft, can be removed from a donor having an HCMV infection and placed in culture, and then contacted with a composition comprising a CRISPR-associated endonuclease and a guide RNA wherein the guide RNA is complementary to the nucleic acid sequence in the HCMV IE gene. As described above, composition can be a nucleic acid encoding a CRISPR-associated endonuclease and a guide RNA wherein the guide RNA is complementary to the nucleic acid sequence in the HCMV IE gene; an expression vector comprising the nucleic acid sequence; or a pharmaceutical composition comprising a nucleic acid encoding a CRISPR-associated endonuclease and a guide RNA wherein the guide RNA is complementary to the nucleic acid sequence in the HCMV IE gene; or an expression vector comprising the nucleic acid sequence. In some embodiments, the gene editing complex can comprise a CRISPR-associated endonuclease polypeptide and a guide RNA wherein the guide RNA is complementary to the nucleic acid sequence in the HCMV IE gene.

In one embodiment, the method further comprises treating said cell with a therapeutically effective amount of an anti-HCMV agent.

In a particular embodiment, the anti-HCMV agent is selected from the group consisting of acyclovir, cidofovir, fomivirsen, foscarnet and ganciclovir.

Regardless of whether compositions are administered as nucleic acids or polypeptides, they are formulated in such a way as to promote uptake by the human cell. Useful vector systems and formulations are described above. In some embodiments the vector can deliver
the compositions to a specific cell type. The invention is not so limited however, and other methods of DNA delivery such as chemical transfection, using, for example calcium phosphate, DEAE dextran, liposomes, lipoplexes, surfactants, and perfluoro chemical liquids are also contemplated, as are physical delivery methods, such as electroporation, micro injection, ballistic particles, and "gene gun" systems.

Standard methods, for example, immunoassays to detect the CRISPR-associated endonuclease, or nucleic acid-based assays such as PCR to detect the gRNA, can be used to confirm that the complex has been taken up and expressed by the cell into which it has been introduced. The engineered cells can then be reintroduced into the subject from whom they were derived as described below.

The gene editing complex comprises a CRISPR-associated nuclease, e.g., Cas9 or Cas9-HF, and a guide RNA complementary to the HCMV IE target sequence, for example, within the coding region of the HCMV IE gene. The gene editing complex can introduce various mutations into the viral DNA. The mechanism by which such mutations inactivate the virus can vary, for example the mutation can affect viral replication, viral gene expression. The mutations may be located in regulatory sequences or structural gene sequences and result in defective production of HCMV. The mutation can comprise a deletion. The size of the deletion can vary from a single nucleotide base pair to about 10,000 base pairs. In some embodiments, the deletion can include all or substantially all of the viral sequence. In some embodiments the deletion can include the entire viral sequence. The mutation can comprise an insertion that is the addition of one or more nucleotide base pairs to the viral sequence. The size of the inserted sequence also may vary, for example from about one base pair to about 300 nucleotide base pairs. The mutation can comprise a point mutation, that is, the replacement of a single nucleotide with another nucleotide. Useful point mutations are those that have functional consequences, for example, mutations that result in the conversion of an amino acid codon into a termination codon or that result in the production of a nonfunctional protein.

In other embodiments, the compositions comprise a cell which has been transformed or transfected with one or more Cas/gRNA vectors. In some embodiments, the methods of the invention can be applied ex vivo. That is, the subject's cells or donor's cells can be removed from the body and treated with the compositions in culture to excise HCMV sequences and
the treated cells are administrated to the recipient's body. Accordingly, such donor's cells may be CD34+ cells used as allogeneic hematopoietic stem cell transplant (HSCT). The cell can be the subject's cells or they can be haplotype matched (donor's cells) or a cell line. In some embodiments, the cells are human leukocyte antigen (HLA)-matched, autologous, cell lines, or combinations thereof. In other embodiments the cells can be a stem cell. For example, an embryonic stem cell or an artificial pluripotent stem cell (induced pluripotent stem cell (iPS cell)). Embryonic stem cells (ES cells) and artificial pluripotent stem cells (induced pluripotent stem cell, iPS cells) have been established from many animal species, including humans. These types of pluripotent stem cells would be the most useful source of cells for regenerative medicine because these cells are capable of differentiation into almost all of the organs by appropriate induction of their differentiation, with retaining their ability of actively dividing while maintaining their pluripotency. iPS cells, in particular, can be established from self-derived somatic cells, and therefore are not likely to cause ethical and social issues, in comparison with ES cells which are produced by destruction of embryos. Further, iPS cells, which are self-derived cell, make it possible to avoid rejection reactions, which are the biggest obstacle to regenerative medicine or transplantation therapy.

Transduced cells are prepared for reinfusion according to established methods. After a period of about 2-4 weeks in culture, the cells may number between 1x10⁶ and 1x10⁹. In this regard, the growth characteristics of cells vary from patient to patient and from cell type to cell type. About 72 hours prior to reinfusion of the transduced cells, an aliquot is taken for analysis of phenotype, and percentage of cells expressing the therapeutic agent. For administration, cells of the invention can be administered at a rate determined by the LD50 of the cell type, and the side effects of the cell type at various concentrations, as applied to the mass and overall health of the patient. Administration can be accomplished via single or divided doses. Adult stem cells may also be mobilized using exogenously administered factors that stimulate their production and egress from tissues or spaces that may include, but are not restricted to, bone marrow or adipose tissues.

In another aspect, the invention relates to an isolated nucleic acid encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene for use in a method for preventing and/or treating HCMV infection.
In another aspect, the invention relates to a composition comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene for use in a method for preventing and/or treating HCMV infection.

In another aspect, the invention relates to a method of reducing the risk of the transmission of a HCMV infection from an HCMV-infected donor to a recipient in need thereof, comprising treating the cell, tissue or organ obtained from said donor with a therapeutically effective amount of at least one isolated nucleic acid sequence of the invention.

In one embodiment, the cell is selected from the group consisting of multipotent hematopoietic stem cells derived from bone marrow, peripheral blood, or umbilical cord blood; or pluripotent (i.e. ES or iPS) or multipotent stem cell-derived differentiated cells of different cell lineages such as cardiomyocytes, beta-pancreatic cells, hepatocytes, neurons.

As used herein, the term "organ" refers to a solid vascularized organ that performs a specific function or group of functions within an organism. The term organ includes, but is not limited to, heart, lung, kidney, liver, pancreas, skin, uterus, bone, cartilage, small or large bowel, bladder, brain, breast, blood vessels, esophagus, fallopian tube, gallbladder, ovaries, pancreas, prostate, placenta, spinal cord, limb including upper and lower, spleen, stomach, testes, thymus, thyroid, trachea, ureter, urethra, uterus.

In one embodiment, the organ is a kidney transplant/graft.

As used herein, the term "tissue" refers to any type of tissue in human, and includes, but is not limited to, vascular tissue, skin tissue, hepatic tissue, pancreatic tissue, neural tissue, urogenital tissue, gastrointestinal tissue, skeletal tissue including bone and cartilage, adipose tissue, connective tissue including tendons and ligaments, amniotic tissue, chorionic tissue, dura, pericardia, muscle tissue, glandular tissue, facial tissue, ophthalmic tissue.

In one embodiment, the methods further comprise treating said cell, tissue or organ with a therapeutically effective amount of an anti-HCMV agent.
In a particular embodiment, the anti-HCMV agent is selected from the group consisting of acyclovir, cidofovir, fomivirsen, foscarnet and ganciclovir.

The invention will be further illustrated by the following figures and examples. However, these examples and figures should not be interpreted in any way as limiting the scope of the present invention.

FIGURES:

Figure 1: U373-MG cell lines constitutively expressing gRNA anti-CMV/Cas9 GFP. FACS analysis of IE expression 4 and 8 days after infection with HCMV strain TB40-GFP (MOI 1, MOI 0.1 and MOI 0.01).

Figure 2: Viral genome extracted from U373-MG cell lines constitutively expressing gRNA anti-CMV/Cas9-GFP, are harboring indels. Nested PCR at the target site followed by a T7 assay stress the effects of each of the three gRNA anti-HCMV on the viral genome.

Figure 3: Analysis of the indels consequences on the nucleotide sequence. The percentage of indels regarding the total amount of clone sequence is presented for each gRNA (top panel). Indels were analyzed for their efficiency on the change of the open-reading frame (bottom panel).

Figure 4: Multiplex strategy is more efficient than the singleplex to reduced IE expression after HCMV infection. U373-MG stably transduced to express gRNA#2, gRNA#4, gNRA#5 and CAS9-GFP or gRNA#2 CAS9-GFP or U373-MG were infected with TOLEDO (MOI 1), TB40 (MOII) or VR1814 (MOI 0.1). 8 days later, cells were labeled for IE expression in the nucleus.

Figure 5: Multiplex strategy induces deletion of the targeted exon in viral genome. U373-MG stably transduced to express gRNA#2, gRNA#4, gNRA#5 and CAS9-GFP or gRNA#2 CAS9-GFP or U373-MG were infected with TOLEDO (MOI 1), TB40 (MOII). 8 days later the viral genome was extracted and a PCR was done using primers allowing the amplification of the all IE gene. Cells harboring the multiplex strategy were deleted of the target exon resulting in a PCR product of 500bp as oppose to the 4Kbp of the WT IE gene.
Figure 6 Multiplex strategy impair the expression of late viral protein. U373-MG stably transduced to express gRNA#2, gRNA#4, gRNA#5 and CAS9-GFP or gRNA#2 CAS9-GFP or U373-MG were infected with TOLEDO (MOI 1). 8 days later, cells were labeled for gb (late viral protein) expression.

Figure 7 Multiplex strategy mostly abrogate the production of new virion. U373-MG stably transduced to express gRNA#2, gRNA#4, gRNA#5 and CAS9-GFP or gRNA#2 CAS9-GFP or U373-MG were infected with TOLEDO (MOI 1). 8 days later, cells were wash, counted and culture over monolayer of MRC5 in semi-solid media. Seven to 14 days later, plaque unit formation were counted and multiplex show very little if not transinfection as compared to the non-treated or singleplex strategy.

EXAMPLE: Combining gRNA/Cas9 targeting the HCMV genome with approved antiviral drug to pretreat a cell suspension before their transplantation.

Material & Methods

Design and cloning of gRNA: gRNA targeting the IE gene of HCMV were designed following Mali et al. 2013, GN(i>2i)GG. The first set of gRNA was designed to be as close as possible to the ATG of the IE gene (gRNA#1, gRNA#2 and gRNA#3). The second set of gRNA was designed to be included in a multiplex strategy and target both side of the 5th exon of IE gene (gRNA#4 and gRNA#5). gRNA#1, gRNA#2, gRNA#3 were cloned in the pX330 from Addgene (CRISPR pX330-U6-(gRNA)-FLAGXnlsCas9nls) as previously described.

Selected gRNA are working with the spCas9 (Cas9 Nuclease, S. pyogenes).
Table 1: Information regarding the selected gRNA anti-HCMV:

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5'-3' (short trinucleotide protospacer adjacent motif (PAM))</th>
<th>Strand</th>
</tr>
</thead>
<tbody>
<tr>
<td>gRNA#1</td>
<td>gtcacatcttctcttggeagg(agg)</td>
<td>+</td>
</tr>
<tr>
<td>gRNA#2</td>
<td>gactcactgtgaagaga(cgg)</td>
<td>+</td>
</tr>
<tr>
<td>gRNA#3</td>
<td>gtcacgctctgtgaacaga(tgg)</td>
<td>-</td>
</tr>
<tr>
<td>gRNA#4</td>
<td>GTCCTGGATGGCTGCCCTGCA(TGG)</td>
<td>+</td>
</tr>
<tr>
<td>gRNA#5</td>
<td>GGTGCTACTGGAATCGATAC(CGG)</td>
<td>+</td>
</tr>
</tbody>
</table>

Validation of the gRNA: Each gRNA cloned in the pX330 were retrotranscribed thanks to floating primers with T7 promoters. PCR amplification with T7 floating primers and lower primer were done using the MEGA Short Script (Life technology) according to the manufacture protocol. gRNA were then purify using the kit MegaClear (Ambion) according to the manufacture protocol, and were validated using the Cas9 Nuclease #M0386S (NEB) on a PCR product harboring the target site.

gRNA were then cloned into a plasmid allowing the production of lentivirus.

Multiplex strategy: In order to increase the efficacy of the gRNA strategy for the destruction of the IE gene, gRNA#2, gRNA#4 and gRNA#5 were cloned together in the same plasmid allowing their expression together with the human-codon-optimized spCas9-GFP in a lentiviral backboned. The construct was synthetized by Genscript.

Cells and virus: U373-MG and MRC-5 cells were transduced with lentiviral vector encoding gRNA#1/Cas9-GFP, gRNA#2/Cas9-GFP or gRNA#3/Cas9-GFP, or the multiplex strategy were cultured during 3 weeks before being FACS-sorted based on their GFP expression. Cells were then expanded prior to be infected with HCMV (TB40-GFP) at a MOI of 1 or 0.1 or 0.01. Two, 4 and 8 days later, cells were fixed with PFA, permeabiliized with triton X-100 0.1% and then stained for IE expression. Two, 4 and 8 days later, cells were centrifuged and HMCV genome was extracted according to the Hirt viral genome extraction protocol.

Analyzing indels and off-targets: For the singleplex strategy: Nested PCR was processed on HCMV genome extracted from the different U373-MG cells lines. The presence...
of mutations in PCR products was examined by a T7 assay and PCR products were then sequenced with white/blue topo PCR cloning kit (Life technology). For the multiplex strategy: a PCR from the whole IE gene was performed. No T7 assay was needed to appreciate the effect of the multiplex strategy.

**Virion released analysis by plaque assay:** U373-MG cell line and U373-MG cell line of the singleplex or the multiplex strategy were infected with HCMV at a MOI of 1 and 8 days later, cells were trypsined, washed and counted. Equivalent number of serially diluted U373-MG cells were added over a monolayer of MRC5 in duplicate. The next day, the media was carefully removed and replaced by a semisolid media. Seven to 14 days later, plaques were counted under the microscope to assess the of U373-MG able to release new virion.

**Ganciclovir experiments:** Ganciclovir (Cymevan®) blocks DNA synthesis and therefore CMV replication cycle by inhibiting the viral DNA. MRC-5 cells infected for 2 hours with HCMV (MOI 0,1) were treated with different doses of Ganciclovir (from 0,25 to 20M) for 4 days. Cells were fixed and permeabilized prior to be stained with anti-IE and anti-Gb antibodies.

**gRNA targeting latently infected cells:** CD34+ haematopoietic cells were transduced with lentiviral vectors suitable for the gRNA simplex or multiplex strategy in order to efficiently target latently infected.

**Results**

**First set of selected gRNA are efficient:** gRNA#1, gRNA#2 and gRNA#3 were retrotranscribed and mixed with Cas9 and a PCR product harboring the target site for 2 hours at RT to validate their efficiency. Each gRNA was functional as shown by the digestion of the PCR product.

**Selection of the best gRNA targeting the ATG of Exon2 of IE gene:** U373-MG transduced with lentiviral vector encoding each gRNA and FACS-sorted based on the GFP expression were then HCMV infected. As shown in Figure 1, the lower the MOI, the better is the effect of the gRNA. When analyzing the IE expression gRNA#1 and #2 seems better than gRNA#3 (Figure 1). Four days after infection, viral genome was purified and amplified by nested PCR to analyze the nucleotide sequence at the target site for each condition. PCR
products were then subject to a T7 assay in order to highlight the *indels* at the target site. As shown in Figure 2 the three gRNA tested were efficiently cutting the HCMV genome in infected cells. PCR products were sequenced and *indels* were further characterized as deletion or insertion and regarding their effectiveness on the protein synthesis disruption. Because gRNA#1 was the only one cutting after the ATG, it was expected to be the most efficient in terms of change in the open reading frame. Indeed, gRNA#1 is not only the gRNA harboring an efficacy of 100% on the sequenced cloned but also it is also the one impacting the most the genome (efficient *indels*) (Figure 3). Prediction of off-targets based on the use of Zanhg’s lab software ([http://crispr.mit.edu/](http://crispr.mit.edu/)) prompt us to analyze possible off-targets on human genome. The three off-targets harboring the highest score were analyzed on the genome of each U373-MG gRNA/Cas9 cell lines. By T7 assay, gRNA#1 shows an off-target on human genome and therefore was not considered anymore an interesting gRNA.

When thinking of a multiplex strategy to completely destroy the IE gene, we consider gRNA#2 and gRNA#3 as the most appropriate gRNA to use in order to completely delete the ATG regardless of the percentage of efficient *indels* (Figure 3). Yet the gRNA#2 shows 80% of *indels* on the sequenced cloned and induces also a reduction in the amount of IE*+* cells after HCMV infection (Figure 1). Therefore gRNA#2 was selected for the multiplex strategy.

*Ganciclovir inhibits the replication cycle of the HCMV*: Cymevan® is actually used in clinic to block CMV replication in AIDS and transplanted patients. This drugs blocks viral DNA replication. In order to potentiate the effect of the gRNA, we are thinking of combining both therapies together. Initiation of CMV replication cycle might be very fast after viral entry depending on the target cell. The gRNA/Cas9 strategy alone might therefore not be fast enough to completely cut all the viral genome. Hence, by blocking the DNA replication with anti-viral drugs already approved, gRNA and Cas9 enable to cut as much HCMV genome as possible. When using Cymevan® on MRC-5 previously infected with HCMV for 2 hours, we observed a reduction of IE*+* cells and gB*+* cells 4 days after infection.

*Multiplex is more efficient than singleplex strategy*: U373-MG cell line stably transduced to express gRNA#2, gRNA#4, gRNA#5, Cas-9 and GFP were FACS-sorted based on the GFP expression. These multiplex cell line was then compared to the singleplex cell line harboring gRNA#2 Cas-9 and GFP and the U373-MG. All those cell line were HCMV infected with different strain (TB40; Toledo or VR1814) at a MOI of 1 or 0.1 for
the VR1814. 8 days after infection cells were FACS-labeled for intranuclear IE molecules. As shown in Figure 4, multiplex strategy was more efficient on the three virus strain than singleplex strategy. When focusing on the viral genome extracting from those infected cell lines, again, multiplex strategy shows impressive efficiency with deletion of the targeted exon as shown by the PCR in Figure 5. Both strategy did affect the expression of viral late protein as shown in Figure 6 by the reduction of gb expression. Again the multiplex strategy is more efficient than the singleplex. Finally, we address the viral release potential of HCMV infected U373-MG. While we were not able to assess some viral production in the supernatant (since U373-MG are known to be low HCMV producing cells), we were able to show trans-infection when HCMV infected U373-MG were culture in semisolid media over a monolayer of highly permissive MRC5 cells. The read-out of this analysis was done by counting the lysed plaques within the monolayer cells. Again, while HCMV infected U373-MG were able to transinfect MRC5 cells, the U373-MG harboring our multiplex strategy were producing very low level (if not) of virion (Figure 7).

Altogether those data clearly show that Crispr/CAS9 is effiently targeting the CMV genome, but that the multiplex strategy is even more efficient and almost abrogate new virion release.

REFERENCES:

Throughout this application, various references describe the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present disclosure.


Christina Paulus, Michael Nevels; The Human Cytomegalovirus Major Immediate-Early Proteins as Antagonists of Intrinsic and Innate Antiviral Host Responses; Viruses. 2009 December; 1(3): 760-779.
CLAIMS:

1. A method of inactivating a human cytomegalovirus (HCMV) in a cell, comprising exposing the cell to a composition comprising an isolated nucleic acid encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV immediate-early (IE) gene.

2. The method of claim 1, wherein the guide RNA has the sequence comprising or consisting of SEQ ID NO: 2.

3. The method of claim 1 or 2, wherein the cell is a latently infected cell.

4. The method of any of claims 1-3, wherein the inactivating is in vivo.

5. The method of any of claims 1-3, wherein the inactivating is ex vivo.

6. The method of claim 5, wherein the cell comprises a cultured cell, tissue or organ obtained from a donor having a HCMV infection.

7. The method of claim 6, wherein the cultured cell, tissue or organ obtained from the donor is transplanted into a recipient following the exposing step.

8. The method of any of claims 1-7, wherein the gene editing complex introduces one or more mutations in the viral DNA and wherein the mutation inactivates viral replication or viral gene expression.

9. The method of any of claims 1-8, wherein the CRISPR-associated endonuclease is Cas9.

10. An isolated nucleic acid sequence encoding a gene editing complex comprising a CRISPR-associated endonuclease and one or more guide RNA, wherein the guide RNA is complementary to a target nucleic acid sequence within the HCMV IE gene.

11. The isolated nucleic acid sequence of claim 10, wherein the guide RNA has the sequence comprising or consisting of SEQ ID NO: 2.
12. An expression vector comprising the nucleic acid sequence of claim 10 or 11.

13. A composition comprising the nucleic acid sequence of claim 10 or 11, or the expression vector of claim 12.

14. A method of reducing the risk of the transmission of a HCMV infection from an HCMV-infected donor to a recipient in need thereof, comprising treating the cell, tissue or organ obtained from said donor with a therapeutically effective amount of at least one isolated nucleic acid sequence of claim 10.

15. The method of claim 14, further comprising treating said cell, tissue or organ with a therapeutically effective amount of an anti-HCMV agent selected from the group consisting of acyclovir, cidofovir, fomivirsen, foscarnet and ganciclovir.
Figure 1
Figure 2
Figure 3
Facs analysis 8 days post infection

% of IE positive cells

- Toledo
- TB40GFP
- VR1814

Figure 4
Figure 5
Figure 6
Figure 7
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/113 A61K 31/7088 C12N 9/22

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N A61K

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

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

EPO-Internal, BIOSIS, EMBASE, WPI Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>TADAHI no SUENAGA ET AL: &quot;Engineering large viral DNA genomes using the CRISPR-Cas9 system&quot;, MICROBIOLOGY AND IMMUNOLOGY, vol. 58, no. 9, 1 September 2014 (2014-09-01), pages 513-522, XP055205700, ISSN: 0385-5600, DOI: 10.1111/1348-0421.12180 the whole document -----</td>
<td>1-15</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. See patent family annex.

Date of the actual completion of the international search

31 August 2016

Name and mailing address of the ISA

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

Date of mailing of the international search report

14/09/2016

Authorized officer

Andres, Serge

Form PCT/ISA/210 (second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>GERGEN JANINA ET AL: &quot;The CRISPR/Cas9 System as an Anti-Viral Treatment to Prevent Primary Infection by HCMV Positive Hematopoietic Stem Cells&quot;, MOLECULAR THERAPY, vol. 24, no. Suppl. 1, May 2016 (2016-05), page S304, XP002761289, ISSN: 1525-0016 &amp; 19TH ANNUAL MEETING OF THE AMERICAN-SOCIETY-OF-GENE-AND-CELL-THERAPY (ASGCT); WASHINGTON, DC, USA; MAY 04 -07, 2016</td>
<td>1-15</td>
</tr>
<tr>
<td>Category</td>
<td>Citation of document, with indication, where appropriate, of the relevant passages</td>
<td>Relevant to claim No.</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>T</td>
<td>FERDY R. VAN DIEMEN ET AL: &quot;CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections&quot;, PLOS PATHOGENS, vol. 12, no. 6, 30 June 2016 (2016-06-30), page el005701, XP055298459, DOI: 10.1371/journal.ppat.1005701</td>
<td>-</td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 7103398 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2285921 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0975654 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2000514093 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20010006089 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 6153595 A</td>
</tr>
<tr>
<td>WO 2004071430 A2</td>
<td>26-08-2004</td>
<td>AU 2004211949 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2514912 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1589931 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004248839 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2004071430 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2894668 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 105658796 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HK 1206388 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HK 1208056 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2016505256 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20150105635 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2014179006 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2014189896 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2014273231 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2015031134 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2015079681 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2016115488 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2016115489 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2014093595 A1</td>
</tr>
</tbody>
</table>

Form PCT/ISA/210 (patent family annex) (April 2005)