Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2686772 C 2015/03/24

(11)(21) 2 686 772

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2008/01/29
(87) Date publication PCT/PCT Publication Date: 2008/11/20
(45) Date de délivrance/Issue Date: 2015/03/24
(85) Entrée phase nationale/National Entry: 2009/11/06
(86) N° demande PCT/PCT Application No.: EP 2008/051031
(87) N° publication PCT/PCT Publication No.: 2008/138648
(30) Priorité/Priority: 2007/05/15 (EP07360019.9)

(51) Cl.Int./Int.Cl. *C12N 15/86* (2006.01)

(72) Inventeurs/Inventors:
SILVESTRE, NATHALIE, FR;
SCHMITT, DORIS, FR

(73) Propriétaire/Owner:
TRANSGENE S.A., FR

(74) Agent: NORTON ROSE FULBRIGHT CANADA LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : VECTEURS D'EXPRESSION GENIQUE DE PLUSIEURS SEQUENCES
(54) Title: VECTORS FOR MULTIPLE GENE EXPRESSION

(57) Abrégé/Abstract:
The present invention provides a vector for expressing at least a first and a second nucleic acid molecules which exhibit a percentage of homology of approximately 80% or greater than 80% over a portion of 40 or more continuous nucleotides and wherein said first nucleic acid molecule and/or said second nucleic acid molecule is modified so as to reduce said percentage of homology to less than 75%. The present invention also relates to substantially isolated nucleic acid molecules comprising a nucleotide sequence as defined in any of SEQ ID NO: 9-15 and 66-69. It also provides a host cell and a pharmaceutical composition comprising such a nucleic acid molecule or vector as well as their use for therapeutic or preventive purposes.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 November 2008 (20.11.2008)

PCT

(10) International Publication Number
WO 2008/138648 A1

(51) International Patent Classification:
C12N 15/86 (2006.01)

(21) International Application Number:
PCT/EP2008/051031

(22) International Filing Date: 29 January 2008 (29.01.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07360019.9 15 May 2007 (15.05.2007) EP

(71) Applicant *(for all designated States except US)*: **TRANSGENE S.A.** [FR/FR]; 11, rue de Molsheim, F-67000 Strasbourg (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SILVESTRE, Nathalie** [FR/FR]; 6, rue des Iris, F-67120 Ergersheim (FR). **SCHMITT, Doris** [FR/FR]; 6, rue Aristide Briand, F-67115 Plobsheim (FR).

(74) Agent: **WARCOIN, AHNER, TEXIER, LE FORESTIER, CALLON DE LAMARCK, COLLIN, TETAZ**; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

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

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

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

WO 2008/138648 A1

(54) Title: VECTORS FOR MULTIPLE GENE EXPRESSION

(57) Abstract: The present invention provides a vector for expressing at least a first and a second nucleic acid molecules which exhibit a percentage of homology of approximately 80% or greater than 80% over a portion of 40 or more continuous nucleotides and wherein said first nucleic acid molecule and/or said second nucleic acid molecule is modified so as to reduce said percentage of homology to less than 75%. The present invention also relates to substantially isolated nucleic acid molecules comprising a nucleotide sequence as defined in any of SEQ ID NO: 9-15 and 66-69. It also provides a host cell and a pharmaceutical composition comprising such a nucleic acid molecule or vector as well as their use for therapeutic or preventive purposes.

Vectors for multiple gene expression

This invention relates to a recombinant vector engineered for independently expressing multiple nucleotide sequences of interest which are obtained from the same organism or from closely related organisms. This invention relates to the field of recombinant nucleic acid technology for expressing multiple nucleotide sequences exhibiting homology with each other, in various prokaryotic as well as eukaryotic *in vitro* systems or in an animal or human subject for therapeutic or prophylactic purposes. The present invention is particularly useful in the field of immunotherapy especially for treating or preventing pathological conditions caused by infectious organisms such as papillomavirus and hepatitis virus.

Recombinant DNA technology has made it possible to express nucleotide sequences in cultured host cells or in living organisms. Several plasmid DNA and viral vectors have been generated and employed for a variety of purposes, including vaccination, gene therapy, immunotherapy and expression in cultured cells. Vectors such as adenoviral and poxviral vectors have the advantage to accommodate a large cloning capacity, with potential of expressing multiple nucleotide sequences in a wide range of host cells. Expression of multiple nucleotide sequences may be advantageous in order to improve the therapeutic efficacy provided by the encoded polypeptides (e.g. combining humoral and cellular immunity). Rather than producing a plurality of recombinant vectors engineered separately to express each of the desired nucleotide sequences, it would be advantageous to produce a single recombinant vector, at least to facilitate production steps and regulatory approval.

For example, with respect to papillomavirus infections, it would be of interest of expressing immunogenic polypeptides from several papillomavirus genotypes, in order to broaden or reinforce the host's immune response especially in subject at risk of multiple infections, e.g. HPV-16 and HPV-18. However, the nucleotide sequences encoding such immunogenic polypeptides are highly homologous between related HPV genotypes. For example, the HPV-16 E6 and HPV-18 E6 sequences which show an overall homology of 63% at the nucleotide level, nevertheless comprise particular regions of very high

homology beyond 75% which may jeopardize expression of HPV-16 and HPV-18 genes from a single vector.

Moreover, when expressing polypeptides of viral origin, homologous nucleotide sequences may also arise from the overall organization of virus genome. It is rather frequent that a virus use the same nucleotide sequence to encode two different proteins through biological mechanisms such as internal translation initiation or reading frame shifting, i.e. the same sequence of DNA is translated in more than one reading frame. For example, in the HPV-16 genome, the adjacent E1 and E2 genes overlap over 59 nucleotides which are translated in different reading frames. In other words, the last 59 nucleotides of the E1 gene overlap with the first 59 nucleotides of the E2 gene.

However, the presence of homologous sequences in a vector is expected to negatively influence its stability especially during the vector production steps, leading to loss of gene sequences due to recombination events that occur between the homologous sequences. Thus, expressing HPV-16 E1 and E2 genes in a single vector involves the presence of a common portion of 59 nucleotides which could potentially lead to homologous recombination events and ultimately to loss of the sequences comprised between the E1 and E2 homologous sequences. Such undesired homologous recombination events may also occur when expressing HPV-16 and HPV-18 gene sequences in the same vector. This instability problem can render vector stock unusable, especially for human clinical trial.

In this respect, WO92/16636 propose to insert in the recombinant vector the homologous nucleotide sequences in opposite orientation with respect of each other so as to reduce the likelihood of recombination events. However, this strategy was described in connection with vaccinia virus vector and not for other recombinant vectors such as adenoviruses. Moreover, the arrangement in opposite orientation is not always possible due to possible promoter interference and construction constraint.

There is a need in the art for generating recombinant vectors capable of expressing in a host cell or subject nucleotide sequences obtained from the same or from closely related organisms, which, in the native context, contains highly homologous portions. The present invention addresses this need in providing a novel strategy designed to minimise the likelihood of the recombination events, by altering either or both of the homologous nucleotide sequences using the degeneracy of the genetic code to make them less

homologous than before modification while not altering or not altering significantly the encoded amino acid sequence. The present invention permits to circumvent the deleterious effect of homologous recombination that may occur between the homologous sequences, especially during vector production steps and lead to the loss of nucleotide sequences contained in between. It has been found that the vector of the present invention is surprisingly effective in expressing E1 and E2 papillomavirus genes which in the native context share a 100% homologous portion of 59 nucleotides and surprisingly stable during the vector production steps. It has also been found that the vector of the present invention is surprisingly effective in expressing E6 and E7 genes obtained from the closely related HPV-16 and HPV-18 genotypes.

This technical problem is solved by the provision of the embodiments as defined in the claims.

Other and further aspects, features and advantages of the present invention will be apparent from the following description of the presently preferred embodiments of the invention. These embodiments are given for the purpose of disclosure.

Accordingly, in a first aspect, the present invention provides a vector comprising at least a first nucleic acid molecule encoding a first polypeptide and a second nucleic acid molecule encoding a second polypeptide wherein:

- said first and second nucleic acid molecules are obtained respectively from a first and second native nucleic acid sequences which exhibit a percentage of homology of approximately 80% or greater than 80% over a portion of 40 or more continuous nucleotides, and
- said first nucleic acid molecule and/or said second nucleic acid molecule comprised in the vector is modified so as to reduce said percentage of homology to less than 75%.

As used herein throughout the entire application, the terms “a” and “an” are used in the sense that they mean “at least one”, “at least a first”, “one or more” or “a plurality”

of the referenced compounds or steps, unless the context dictates otherwise. For example, the term “a cell” includes a plurality of cells including a mixture thereof.

The term “and/or” wherever used herein includes the meaning of “and”, “or” and “all or any other combination of the elements connected by said term”. For example, “the first nucleic acid molecule and/or the second nucleic acid molecule” means the first nucleic acid molecule, or the second nucleic acid molecule or both the first and the second nucleic acid molecules.

The term “about” or “approximately” as used herein means within 5%, preferably within 4%, and more preferably within 2% of a given value or range.

As used herein, when used to define products, compositions and methods, the term "comprising" is intended to mean that the products, compositions and methods include the referenced components or steps, but not excluding others. "Consisting essentially of” shall mean excluding other components or steps of any essential significance. Thus, a composition consisting essentially of the recited components would not exclude trace contaminants and pharmaceutically acceptable carriers. "Consisting of” shall mean excluding more than trace elements of other components or steps. For example, a polypeptide “consists of” an amino acid sequence when the polypeptide does not contain any amino acids but the recited amino acid sequence. A polypeptide “consists essentially of” an amino acid sequence when such an amino acid sequence is present together with only a few additional amino acid residues, typically from about 1 to about 50 or so additional residues. A polypeptide “comprises” an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the polypeptide. Such a polypeptide can have a few up to several hundred additional amino acids residues. Such additional amino acid residues may play a role in polypeptide trafficking, facilitate polypeptide production or purification; prolong half-life, among other things. The same can be applied for nucleotide sequences.

As used herein, a "vector" may be any agent capable of delivering and expressing at least the first and second nucleic acid molecules in a host cell or subject. The vector may be extrachromosomal (e.g. episome) or integrating (for being incorporated into the host chromosomes), autonomously replicating or not, multi or low copy, double-stranded or single-stranded, naked or complexed with other molecules (e.g. vectors complexed with lipids or polymers to form particulate structures such as liposomes, lipoplexes or

nanoparticles, vectors packaged in a viral capsid, and vectors immobilised onto solid phase particles, etc.). The definition of the term "vector" also encompasses vectors that have been modified to allow preferential targeting to a particular host cell. A characteristic feature of targeted vectors is the presence at their surface of a ligand capable of recognizing and binding to a cellular and surface-exposed component such as a cell-specific marker (e.g. an HPV-infected cell), a tissue-specific marker or a tumor-specific marker. The ligand can be genetically inserted into a polypeptide present on the surface of the vector (e.g. adenoviral fiber, penton, pIX as described in WO94/10323 and WO02/96939 or vaccinia p14 gene product as described in EP 1 146 125).

Within the context of the present invention, the terms “nucleic acid”, “nucleic acid molecule”, “polynucleotide” and “nucleotide sequence” are used interchangeably and define a polymer of any length of either polydeoxyribonucleotides (DNA) or polyribonucleotides (RNA) molecules or any combination thereof. The definition encompasses single or double-stranded, linear or circular, naturally-occurring or synthetic polynucleotides. Moreover, such polynucleotides may comprise non-naturally occurring nucleotides (e.g. methylated nucleotides and nucleotide analogs such as those described in US 5,525,711, US 4,711,955 or EPA 302 175) as well as chemical modifications (e.g. see WO 92/03568; US 5,118,672) in order to increase the in vivo stability of the nucleic acid, enhance the delivery thereof, or reduce the clearance rate from the host subject. If present, modifications may be imparted before or after polymerization.

The terms “polypeptide”, “peptide” and “protein” are used herein interchangeably to refer to polymers of amino acid residues which comprise 9 or more amino acids bonded via peptide bonds. The polymer can be linear, branched or cyclic. In the context of this invention, a "polypeptide" may include amino acids that are L stereoisomers (the naturally occurring form) or D stereoisomers and may include amino acids other than the 20 common naturally occurring amino acids, such as [beta]-alanine, ornithine, or methionine sulfoxide, or amino acids modified on one or more alpha-amino, alpha-carboxyl, or side-chain, e.g., by appendage of a methyl, formyl, acetyl, glycosyl, phosphoryl, and the like. As a general indication, if the amino acid polymer is long (e.g. more than 50 amino acid residues), it is preferably referred to as a polypeptide or a protein. By way of consequence, a "peptide" refers to a fragment of about 9 to about 50 amino acids in length. In the context of the invention, a peptide preferably comprises a

selected region of a naturally-occurring (or native) protein, e.g. an immunogenic fragment thereof containing an epitope.

The term "polypeptide" as defined herein encompasses native as well as modified polypeptides. The term "native" as used herein refers to a material recovered from a source in nature as distinct from material artificially modified or altered by man in the laboratory. For example, a native polypeptide is encoded by a gene that is present in the genome of a wild-type organism or cell. By contrast, a modified polypeptide is encoded by a nucleic acid molecule that has been modified in the laboratory so as to differ from the native polypeptide, e.g. by insertion, deletion or substitution of one or more amino acid(s) or any combination of these possibilities. When several modifications are contemplated, they can concern consecutive residues and/or non consecutive residues. Examples of modification(s) contemplated by the present invention may result in alteration of the biological activity exhibited by the native polypeptide. Amino acids that are critical for a given biological activity can be identified by routine methods, such as by structural and functional analysis and one skilled in the art can readily determine the type of mutation(s) that is able to reduce or abolish such a biological activity. Such modifications can be performed by routine techniques such as site-directed mutagenesis. Alternatively, one may generate a synthetic nucleic acid molecule encoding the modified polypeptide by chemical synthesis in automatised process (e.g. assembled from overlapping synthetic oligonucleotides as described in the appended example section).

The term "obtained" as used herein refers to material that is found, isolated, purified, or derived from a source in nature. "Isolated" means removed from its natural environment. "Purified" denotes that it is substantially free from at least one other component(s) with which it is naturally associated. "Derived" denotes one or more modification(s) as compared to the native material (in particular mutations such as substitutions, deletions and/or insertions). Techniques of isolation, purification and modification are routine in the art and depend on the material to be obtained (e.g. cloning of a nucleic acid molecule can be performed from a source in nature by using restriction enzyme, by PCR or by chemical synthesis).

As used herein the term "homology" is generally expressed as a percentage and denotes nucleotide sequences that retain a given degree of identity each other over a portion of at least 40 consecutive nucleotides (e.g. approximately 40, 45, 50, 55, 57, 58, 59, 60, 70 or even more consecutive nucleotides). "At least 80%" refers to approximately

80% or greater than 80% (e.g. any value beyond 80%, advantageously at least 85%, desirably at least 87%, preferably at least 90%, more preferably at least 95%, still more preferably at least 97% up to 100% of sequence homology). "Less than 75%" refers to any value below 75, e.g. approximately 74, 72, 70, 68, 65, 62, 60% or even less. The percent homology between two nucleotide sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps which need to be introduced for optimal alignment and the length of each gap. Various computer programs and mathematical algorithms are available in the art to determine percentage identities between nucleotide sequences such as GCG Wisconsin package and the Basic Local alignment Search Tool (BLAST) program which is publicly available at National Center for Biotechnology Information (NCBI) and described in printed publications (e.g. Altschul et al., 1990, J. Mol. Biol. 215, 403-410).

As a starting point, a sequence alignment between the first and second nucleic acid molecules before modification may be used in order to reveal the one or more portions of 40 or more continuous nucleotides that share a percentage of homology of 80% or greater than 80%, i.e. the "homologous" portion(s). In a particular embodiment, the codon usage pattern of the first nucleic acid molecule or the second nucleic acid molecule or both the first and second nucleic acid molecules is modified (e.g. by degenerescence of the codon usage pattern) at least in said homologous portion(s) of 40 or more (e.g. approximately 40, 45, 50, 55, 57, 58, 59, 60, 70 or even more) continuous nucleotides so as to reduce the percentage of homology to less than 75% (e.g. approximately 74, 72, 70, 68, 65, 62, 60% or even less).

Whereas methionine and tryptophane residues are each encoded by a unique nucleic acid triplet (i.e. codon), different codons can be used to code for the 18 other amino acids (degeneracy of the genetic code). For example, amino acids are encoded by codons as follows: Alanine (Ala or A) is encoded by codons GCA, GCC, GCG, and GCU; cysteine (C or Cys) by codons UGC, and UGU; aspartic acid (D or Asp) by codons GAC, and GAU; glutamic acid (E or Glu) by codons GAA, and GAG; phenylalanine (F or Phe) by codons UUC, and UUU; glycine (G or Gly) by codons GGA, GGC, GGG, and GGU; histidine (H or His) by codons CAC, and CAU; isoleucine (I or Ile) by codons AUA, AUC, and AUU; lysine (K or Lys) by codons AAA, and AAG; leucine (L or Leu) by codons UUA, UUG, CUA, CUC, CUG, and CUU; methionine (M or Met) by codon AUG; asparagine (N or Asn) by codons AAC, and AAU; proline (P or

Pro) by codons CCA, CCC, CCG, and CCU; glutamine (Q or Gln) by codons CAA, and CAG; arginine (R or Arg) by codons AGA, AGG, CGA, CGC, CGG, and CGU; serine (S or Ser) by codons AGC, AGU, UCA, UCC, UCG, and UCU; threonine (T or Thr) by codons ACA, ACC, ACG, and ACU; valine (V or Val) by codons GUA, GUC, GUG, and GUU; tryptophan (W or Trp) by codon UGG and tyrosine (Y or Tyr) by codons UAC, and UAU.

Reduction of the percentage of homology in the one or more homologous portion(s) present in said first and second nucleic acid molecules can be achieved by taking advantage of the degeneracy of the genetic code and modifying the codon usage pattern in the first nucleic acid molecule and/or the second nucleic acid molecule. Modification of the codon usage pattern is typically performed by replacing one or more "native" codon(s) with another codon(s). For example, the replacement of the Arg-encoding AGA codon with the Arg-encoding CGC codon will reduce homology in 2 of 3 positions of the codon. It is not necessary to degenerate all native codons since homology can be sufficiently reduced with partial replacement. Moreover, modification of the codon usage pattern can be performed over the entire nucleic acid molecule or can be restricted to the homologous portion(s) present before modification. Desirably, in the context of the invention, degenerescence is performed in the first nucleic acid molecule and is restricted to the homologous portion(s). Preferably, the codon usage pattern is modified at the nucleotide level and the modifications are silent at the amino acid level, i.e. when it is possible, each "native" codon is replaced with a codon encoding the same amino acid so that such modifications do not translate in the encoded polypeptide. More preferably, when it is possible, the codon usage pattern is modified in such a way that homologous portions between the first and second nucleic acid molecules are restricted to less than 9 or 8 consecutive nucleotides, advantageously to less than 7 consecutive nucleotides, preferably to less than 6 consecutive nucleotides and, more preferably, to less than 5 consecutive nucleotides. Modification of the codon usage pattern can be generated by a number of ways known to those skilled in the art, such as site-directed mutagenesis (e.g. using the Sculptor™ *in vitro* mutagenesis system of Amersham, Les Ullis, France), PCR mutagenesis, DNA shuffling and by chemical synthetic techniques (e.g. resulting in a synthetic nucleic acid molecule).

When the vector according to the invention comprises more than two nucleic acid molecules, then any nucleic acid molecule comprised in the vector and obtained from a

native nucleic acid sequence which exhibit a percentage of homology of approximately 80% or greater than 80% over a portion of 40 or more continuous nucleotides with at least one other native nucleic acid sequence from which another nucleic acid molecule is obtained, can be modified so as to reduce the percentage of homology to less than 75%, i.e. so that no pair of nucleic acid molecules comprised in the vector may comprise a portion of 40 or more consecutive nucleotides exhibiting a percentage of identity greater than 75%.

A sequence alignment between each (pair of) native sequences from which the nucleic acid molecules are obtained may be used in order to reveal the one or more portions exhibiting a percentage of homology of 80% or greater than 80%. Then, the sequence of one or more of the native sequences is modified, in particular by degenerating the codon usage, so as to reduce the percentage of homology at least in the homologous portions to less than 75%. In the end, no nucleic acid molecule comprised in the vector should comprise a portion of 40 or more (e.g., 45, 50, 55, 57, 58, 59, 60, 70 or even more) consecutive nucleotides exhibiting a percentage of identity greater than 75% with any other nucleic acid molecule comprised in said vector.

As mentioned above, the polypeptide encoded by the nucleic acid molecules comprised in the vector may or not have the same amino acid sequence as the native polypeptide. In particular, in addition to mutations for degenerating the codon usage so as to reduce homology at least in the homologous portions of nucleic acid molecules comprised in the vector, said nucleic acid molecules comprised in the vector may also comprise additional mutations resulting or not in a modification of the amino acid sequence of the encoded polypeptide.

The vector of the invention encompasses viral as well as non-viral (e.g. plasmid DNA) vectors. Suitable non viral vectors include plasmids such as pREP4, pCEP4 (Invitrogene), pCI (Promega), pCDM8 (Seed, 1987, Nature 329, 840), pVAX and pgWiz (Gene Therapy System Inc; Himoudi et al., 2002, J. Virol. 76, 12735-12746). A “viral vector” is used herein according to its art-recognized meaning. It refers to any vector that comprises at least one element of viral origin, including a complete viral genome, a portion thereof or a modified viral genome as described below as well as viral particles generated thereof (e.g. viral vector packaged into a viral capsid to produce infectious viral

particles). Viral vectors of the invention can be replication-competent, or can be genetically disabled so as to be replication-defective or replication-impaired. The term "replication-competent" as used herein encompasses replication-selective and conditionally-replicative viral vectors which are engineered to replicate better or selectively in specific host cells (e.g. tumoral cells). Viral vectors may be obtained from a variety of different viruses, and especially from a virus selected from the group consisting of retrovirus, adenovirus, adeno-associated virus (AAV), poxvirus, herpes virus, measle virus and foamy virus.

In one embodiment, the vector of the invention is an adenoviral vector (for a review, see "Adenoviral vectors for gene therapy", 2002, Ed D. Curiel and J. Douglas, Academic Press). It can be derived from any human or animal adenovirus. Any serotype and subgroup can be employed in the context of the invention. One may cite more particularly subgroup A (e.g. serotypes 12, 18, and 31), subgroup B (e.g. serotypes 3, 7, 11, 14, 16, 21, 34, and 35), subgroup C (e.g. serotypes 1, 2, 5, and 6), subgroup D (e.g. serotypes 8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, and 42-47), subgroup E (serotype 4), and subgroup F (serotypes 40 and 41). Particularly preferred are human adenoviruses 2 (Ad2), 5 (Ad5), 6 (Ad6), 11 (Ad11), 24 (Ad24) and 35 (Ad35). Such adenovirus are available from the American Type Culture Collection (ATCC, Rockville, Md.) and have been the subject of numerous publications describing their sequence, organization and methods of producing, allowing the artisan to apply them (see for example US 6,133,028; US 6,110,735; WO 02/40665; WO 00/50573; EP 1016711; Vogels et al., 2003, J. Virol. 77, 8263-8271).

The adenoviral vector of the present invention can be replication-competent. Numerous examples of replication-competent adenoviral vectors are readily available to those skilled in the art (see for example Hernandez-Alcoceba et al., 2000, Human Gene Ther. 11, 2009-2024; Nemunaitis et al., 2001, Gene Ther. 8, 746-759; Alemany et al., 2000, Nature Biotechnology 18, 723-727; WO00/24408; US5,998,205, WO99/25860, US5,698,443, WO00/46355, WO00/15820 and WO01/36650).

Alternatively, the adenoviral vector of the invention can be replication-defective (see for example WO94/28152). Preferred replication-defective adenoviral vectors are E1-defective (e.g. US 6,136,594 and US 6,013,638), with an E1 deletion extending from approximately positions 459 to 3328 or from approximately positions 459 to 3510 (by reference to the sequence of the human adenovirus type 5 disclosed in the GeneBank

under the accession number M 73260 and in Chroboczek et al., 1992, Virol. 186, 280-285). The cloning capacity and safety can further be improved by deleting additional portion(s) of the adenoviral genome (e.g. in the non essential E3 region or in other essential E2, E4 regions as described in Lusky et al., 1998, J. Virol 72, 2022-2032).

The first and second nucleic acid molecules can be independently inserted in any location of the adenoviral vector of the invention, as described in Chartier et al. (1996, J. Virol. 70, 4805-4810) and independently positioned in sense and/or antisense orientation relative to the natural transcriptional direction of the region of insertion. For example, they can be both inserted in replacement of the E1 region or alternatively, the one is inserted in replacement of the E1 region and the other in replacement of the E3 region.

In another embodiment, the vector of the invention is a poxviral vector (see for example Cox et al. in "Viruses in Human Gene Therapy" Ed J. M. Hos, Carolina Academic Press). It may be obtained from any member of the poxviridae, in particular canarypox (e.g. ALVAC as described in WO95/2778O), fowlpox (e.g. TROVAC as described in Paoletti et al., 1995, Dev. Biol. Stand. 84, 159-163) or vaccinia virus, the latter being preferred. A suitable vaccinia virus can be selected from the group consisting of the Copenhagen strain (Goebel et al., 1990, Virol. 179, 247-266 and 517-563; Johnson et al., 1993, Virol. 196, 381-401), the Wyeth strain, NYVAC (see WO92/15672 and Tartaglia et al., 1992, Virology 188, 217-232) and the highly attenuated modified Ankara (MVA) strain (Mayr et al., 1975, Infection 3, 6-16). Such vectors and methods of producing are described in numerous documents accessible to the man skilled in the art (e.g. Paul et al., 2002, Cancer gene Ther. 9, 470-477; Piccini et al., 1987, Methods of Enzymology 153, 545-563; US 4,769,330 ; US 4,772,848 ; US 4,603,112 ; US 5,100,587 and US 5,179,993). The first and second nucleic acid molecules in use in the present invention are preferably inserted in a nonessential locus of the poxviral genome, in order that the recombinant poxvirus remains viable and infectious. Nonessential regions are non-coding intergenic regions or any gene for which inactivation or deletion does not significantly impair viral growth, replication or infection. One may also envisage insertion in an essential viral locus provided that the defective function is supplied *in trans* during production of viral particles, for example by using an helper cell line carrying the complementing sequences corresponding to those deleted in the poxviral genome.

When using the Copenhagen vaccinia virus, the at least first and second nucleic acid molecules are preferably inserted in the thymidine kinase gene (tk) (Hruby et al., 1983, Proc. Natl. Acad. Sci USA 80, 3411-3415; Weir et al., 1983, J. Virol. 46, 530-537). However, other insertion sites are also appropriate, e.g. in the hemagglutinin gene (Guo et al., 1989, J. Virol. 63, 4189-4198), in the K1L locus, in the u gene (Zhou et al., 1990, J. Gen. Virol. 71, 2185-2190) or at the left end of the vaccinia virus genome where a variety of spontaneous or engineered deletions have been reported in the literature (Altenburger et al., 1989, Archives Virol. 105, 15-27 ; Moss et al. 1981, J. Virol. 40, 387-395 ; Panicali et al., 1981, J. Virol. 37, 1000-1010 ; Perkus et al, 1989, J. Virol. 63, 3829-3836 ; Perkus et al, 1990, Virol. 179, 276-286 ; Perkus et al, 1991, Virol. 180, 406-410).

When using MVA, the at least first and second nucleic acid molecules can be independently inserted in anyone of the identified deletions I to VII which occurred in the MVA genome (Antoine et al., 1998, Virology 244, 365-396) as well as in the D4R locus, but insertion in deletion II and/or III is preferred (Meyer et al., 1991, J. Gen. Virol. 72, 1031-1038 ; Sutter et al., 1994, Vaccine 12, 1032-1040).

When using fowlpox virus, although insertion within the thymidine kinase gene may be considered, the at least first and second nucleic acid molecules are preferably introduced in the intergenic region situated between ORFs 7 and 9 (see for example EP 314 569 and US 5,180,675).

In another embodiment of the invention, the at least first and second nucleic acid molecules independently encode a polypeptide capable of providing a therapeutic or protective activity in a subject exhibiting or susceptible to exhibit a pathological condition. The term “subject” as used herein refers to a vertebrate, particularly a member of the mammalian species and especially domestic animals, farm animals, sport animals, and primates including humans. Such a polypeptide is preferably selected from the group consisting of immunogenic polypeptides and anti-tumor polypeptides.

An “immunogenic” polypeptide refers to a polypeptide able to induce, stimulate, develop or boost an immune system in a subject into which it is expressed. Such immune response can be humoral or cellular or both humoral and cellular. Humoral response elicits antibody production against the polypeptide in question whereas cellular response elicits T-helper cell and/or CTL response and/or stimulation of cytokine production.

Typically, the immunogenic property of a polypeptide can be evaluated either *in vitro* or *in vivo* by a variety of assays which are standard in the art (for a general description of techniques available to evaluate the onset and activation of an immune response, see for example the latest edition of Coligan et al., Current Protocols in Immunology; ed J Wiley & Sons Inc, National Institute of Health). For example, detection can be colorimetric, fluorometric or radioactive and suitable techniques include ELISA, Western Blot, radioimmunoassays and immunoprecipitation assays. Measurement of cellular immunity can be performed by measurement of cytokine profiles secreted by activated effector cells including those derived from CD4+ and CD8+ T-cells (e.g. quantification IFNg-producing cells by ELIspot), by determination of the activation status of immune effector cells (e.g. T cell proliferation assays by a classical [^{3}H] thymidine uptake), by assaying for antigen-specific T lymphocytes in a sensitized subject (e.g. peptide-specific lysis in a cytotoxicity assay). The immunogenic property of a polypeptide could also be evaluated in suitable animal models by ELIspot, tetramer-based analytical techniques or other standard techniques for analysis T cell-mediated immunity. Suitable immunogenic polypeptides can be obtained from hepatitis B virus (HBV) (e.g. S, preS2 or preS1-polypeptide as described in EP 414 374; EP 304 578 or EP 198 474); hepatitis C virus (HCV) (e.g. Core (C), the envelop glycoprotein E1, E2, the non-structural polypeptide NS2, NS3, NS4, or NS5 or any combination thereof); human immunodeficiency virus (HIV) (e.g. gp120 or gp160), and papillomavirus (as illustrated hereinafter).

An "anti-tumor" polypeptide refers to a polypeptide able to provide suppression or a net reduction in the expansion of tumor cells. The antitumor property of a polypeptide can be determined in appropriate animal models or in the treated subject by a decrease of the actual tumor size over a period of time. A variety of methods may be used to estimate tumor size including radiologic methods (e.g., single photon and positron emission computerized tomography; see generally, "Nuclear Medicine in Clinical Oncology," Winkler, C. (ed.) Springer-Verlag, New York, 1986), methods employing conventional imaging reagents (e.g., Gallium-67 citrate), immunologic methods (e.g., radiolabeled monoclonal antibody directed to specific tumor markers) as well as ultrasound methods (see, "Ultrasonic Differential Diagnosis of Tumors", Kossoff and Fukuda, (eds.), Igaku-Shoin, New York, 1984). Alternatively, the anti-tumor property of a polypeptide may be determined based upon a decrease in the presence of a tumor marker. Examples include PSA for the detection of prostate cancer and CEA for the detection of colorectal and

certain breast cancers. Further, the anti-tumor property of a polypeptide can be determined in a suitable animal model, e.g. using mice injected with a representative human cancer cell line. After palpable tumors have developed, the mice are injected with the vector of the invention, and then monitored for reduced tumor growth rate and increased survival. In addition, a variety of in vitro methods can be used to predict in vivo tumor inhibition. Suitable antitumor polypeptides include tumour-associated antigens (TAAs) such as MUC-1 (WO92/07000; Acres et al., 2005, Exp. Rev. Vaccines 4(4)), BRCA-1, BRCA-2 (Palma et al., 2006, Critical Reviews in Oncology/haematology 27, 1-23), Carcinoembryonic antigen CEA (Conroy et al., 1995, Gene Ther; 2, 59-65), MAGE (WO99/40188; De Plaen et al., 1994, Immunogenetics 40, 360-369), MART-1, gp 100 (Bakker et al., 1994, J. Exp. Med. 179, 1005-9), PRAME, BAGE, Lage (also known as NY Eos 1) SAGE, HAGE (WO99/53061), GAGE (Robbins and Kawakami, 1996. Current Opinions in Immunol. 8, 628-36) and Prostate specific antigen (PSA) (Ferguson, et al., 1999, Proc. Natl. Acad. Sci. USA. 96, 3114-9; WO98/12302, WO98/20117 and WO00/04149) as well as viral polypeptides from viruses having tumor-inducing potential (e.g. papillomavirus).

In another embodiment of the invention, the at least first and second nucleic acid molecules are obtained from the same organism or from closely related organisms.

As used herein, the term “organism” encompasses microorganisms preferably having pathogenic potential and well as higher eukaryotes. The term “microorganism” denotes fungi, bacteria, protozoa and viruses. Representative examples of viruses include without limitation HIV (HIV-1 or HIV-2), human herpes viruses (e.g. HSV1 or HSV2), cytomegalovirus (CMV), Epstein Barr virus (EBV), hepatitis viruses (e.g. hepatitis A virus (HAV), HBV, HCV and hepatitis E virus), flaviviruses (e.g. Yellow Fever Virus), varicella-zoster virus (VZV), paramyxoviruses, respiratory syncytial viruses, parainfluenza viruses, measles virus, influenza viruses, and papillomaviruses (as defined above). Representative examples of suitable bacteria include without limitation Neisseria (e.g. N. gonorrhea and N. meningitidis); Bordetella (e.g. B. pertussis, B. parapertussis and B. bronchiseptica), Mycobacteria (e.g. M. tuberculosis, M. bovis, M. leprae, M. avium, M. paratuberculosis, M. smegmatis); Legionella (e.g. L. pneumophila); Escherichia (e.g. enterotoxic E. coli, enterohemorragic E. coli, enteropathogenic E. coli); Shigella (e.g. S. sonnei, S. dysenteriae, S. flexnerii); Salmonella (e.g. S. typhi, S. paratyphi, S.

choleraesuis, S. enteritidis); Listeria (e.g. L. monocytogenes); Helicobacter (e.g. H. pylori); Pseudomonas (e.g. P. aeruginosa); Staphylococcus (e.g. S. aureus, S. epidermidis); Enterococcus (e.g. E. faecalis, E. faecium); Bacillus (e.g. B. anthracis); Corynebacterium (e.g. C. diphtheriae), and Chlamydia (e.g. C. trachomatis, C. pneumoniae, C. psittaci). Representative examples of parasites include without limitation Plasmodium (e.g. P. falciparum); Toxoplasma (e.g. T. gondii); Leshmania (e.g. L. major); Pneumocystis (e.g. P. carinii); and Schisostoma (e.g. S. mansoni). Representative examples of fungi include without limitation Candida (e.g. C. albicans) and Aspergillus. The higher eukaryotes are preferably mammals including humans.

The "same organism" defines organisms which originate from a common ancestor and have followed the same evolution pathway. Representative examples include various isolates of viruses having the same serotype or genotype. For example two isolates of HPV-16 are classified in this category. "Closely related organisms" define organisms which originate from a common ancestor but have diverged during evolution. Representative examples include viruses having different serotypes or genotypes. For example HPV-16 and HPV-18 are classified in this category.

In a preferred embodiment, the organism for which the at least first and second nucleic acid molecules are obtained is a papillomavirus and each encodes a papillomavirus polypeptide. A "Papillomavirus" can be defined as a virus that belongs to the papillomavirinae subfamily and this term encompasses animal papillomavirus of non-human species origin including but not limited to cattle, horses, rabbits, sheep, dogs, non-human primate, and rodents as well as human papillomavirus (HPV). More than 100 HPV genotypes have been identified at present time (Van Ranst et al., 1992, J. Gen. Virol. 73, 2653; De Villiers et al., 2004, Virology 324, 17-27) which have been classified in "low" (LR) and "high risk" (HR) serotypes depending on their oncogenic potential. LR HPV causes benign tumors in infected subjects whereas HR bears a high risk for malignant progression.

For general guidance, papillomaviruses possess a double-stranded circular DNA of about 7900 base pairs which is surrounded by a protein capsid (see for example Pfister, 1987, in *The papovaviridae: The Papillomaviruses*, Salzman and Howley edition, Plenum Press, New York, p 1-38). Their genome consists of three functional regions, the early

(E), the late (L), and the long control (LCR) regions. The LCR contains transcriptional regulatory sequences such as enhancers and promoters. The late region encodes the structural L1 and L2 proteins, respectively the major and minor capsid proteins, whereas the early region encodes regulatory proteins (E1-E7) found predominantly in the nucleus that control viral replication, transcription and cellular transformation. More specifically, the E1 protein is a DNA binding phosphoprotein with ATP-dependent helicase activity (Desaintes and Demeret, 1996, Semin. Cancer Biol. 7, 339-347; Wilson *et al*, 2002, Virus Gene 24, 275-290). The E2 protein is a multifunctional DNA binding phosphoprotein that regulates viral gene transcription and controls DNA replication (Bechtold et al., 2003, J. Virol. 77, 2021-8). The E4-encoded protein binds and disrupts the cytoplasmic keratin network and plays a role in viral maturation. The function for E5 protein is still controversial and its expression is often lost during viral integration in the host chromosomes. The E6 and E7-encoded gene products of HR HPV genotypes are involved in the oncogenic transformation of infected cells (Kanda et al., 1988, J. Virol. 62, 610-3; Vousden et al., 1988, Oncogene Res. 3, 1-9; Bedell et al., 1987, J. Virol. 61, 3635-40), presumably through binding of these viral proteins to cellular tumor suppressor gene products p53 and retinoblastoma (Rb), respectively (reviewed in Howley, 1996, Papillomaviruses and their replication, p 2045-2076. In B.N. Fields, D.M. Knipe and P.M. Howley (ed), Virology, 3rd ed. Lippincott-Raven Press, New York, N.Y.). The amino acid residues involved in the binding of the native HPV-16 E6 polypeptide to p53 have been clearly defined from residues 118 to 122 (+1 being the first Met residue or from residues 111 to 115 starting from the preferably used second Met residue) (Crook et al., 1991, Cell 67, 547-556) and those involved in the binding of the native HPV-16 E7 polypeptide to Rb are located from residues 21 to 26 (Munger et al., 1989, EMBO J. 8, 4099-4105; Heck et al., 1992, Proc. Natl. Acad. Sci. USA 89, 4442-4446).

Preferably, the at least first and second nucleic acid molecules are independently obtained from a high risk papillomavirus selected from the group consisting of HPV-16, HPV-18, HPV-30, HPV-31, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-56, HPV-58, HPV-59, HPV-66, HPV-68, HPV-70 and HPV-85.

A "papillomavirus polypeptide" as used herein refers to an art-recognized polypeptide encoded by a nucleic acid molecule obtained from a papillomavirus genome/source. As defined above in connection with the term "polypeptide", a

"papillomavirus polypeptide" encompasses native, modified papillomavirus polypeptides and peptides thereof. Sources of papillomavirus include without limitation biological samples (e.g. biological samples, tissue sections, biopsy specimen and tissue cultures collected from a subject that has been exposed to a papillomavirus), cultured cells (e.g. CaSki cells available at ATCC), as well as recombinant materials available in depositary institutions, in commercial catalogues or described in the literature. The nucleotide sequences of a number of papillomavirus genomes and the amino acid sequences of the encoded polypeptides have been described in the literature and are available in specialized data banks, e.g. Genbank. For general information, HPV-16 genome is described in Genbank under accession numbers NC_01526 and K02718; HPV-18 under NC_001357 and X05015; HPV-31 under J04353; HPV-33 under M12732; HPV-35 under NC_001529; HPV-39 under NC_001535; HPV-45 under X74479; HPV-51 under NC_001533; HPV-52 under NC_001592; HPV-56 under X74483; HPV-58 under D90400; HPV-59 under NC_001635; HPV-68 under X67160 and M73258; HPV-70 under U21941; and HPV-85 under AF131950.

The papillomavirus polypeptide(s) encoded by the first and/or the second nucleic acid molecule(s) can be an early, a late or any combination thereof. Early papillomavirus polypeptides include E1, E2, E4, E5, E6 and E7 whereas late polypeptides can be L1 or L2. The nucleotide and amino acid sequences of the early and late polypeptides of a vast number of papillomavirus serotypes are described in the literature available to the skilled person.

Desirably, the at least first and second nucleic acid molecules encode independently an early polypeptide selected from the group consisting of E1, E2, E6 and E7. For purpose of illustration, the amino acid sequences of native HPV-16 E1, E2, E6 and E7 polypeptides are given respectively in SEQ ID NO: 1-4. However, the present invention is not limited to these exemplary sequences. Indeed the nucleotide and amino acid sequences can vary between different papillomavirus isolates and this natural genetic variation is included within the scope of the invention as well as non-natural modification(s) such as those described below. Exemplary illustration of suitable modified papillomavirus polypeptides are given hereinafter (e.g. in SEQ ID NO: 5-8 and 64-65), however, it is within the reach of the skilled person to adapt the modifications described herein (e.g. to polypeptides originating from other papillomavirus genotypes by sequence comparison).

Suitable papillomavirus E1 polypeptides for use in the present invention encompass mutants that are defective for stimulating viral replication., i.e. their ability to stimulate viral replication is statistically significantly lower than that of the corresponding native E1 polypeptide (e.g. less than 75%, advantageously less than 50%, preferably less than 10%, and more preferably less than 5%). For general guidance, the domain responsible for stimulating viral replication is located in the central portion of E1 (e.g. Hugues and Romanos, 1993, Nucleic Acids Res.21, 5817-23). Representative examples of replication-defective E1 polypeptides are described in the literature available to the man skilled in the art, e.g. in Yasugi et al. (1997, J. Virol 71, 5942-51). A preferred modification in the context of the invention includes the substitution of the Gly residue at position 482 of the HPV-16 E1 polypeptide with another residue (preferably with an Asp residue) (e.g. see SEQ ID NO: 5) or the substitution of the Gly residue at position 489 of HPV-18 E1 polypeptide with another residue (preferably with an Asp residue (e.g., see SEQ ID NO: 6)

Suitable E2 polypeptides for use in the invention encompass mutants that are defective in transcriptional activation and/or replication activities as compared to the native E2 polypeptide (e.g. less than 75%, advantageously less than 50%, preferably less than 10%, and more preferably less than 5%). For general guidance, the domain responsible for transcriptional activation and stimulation of replication is located in the N-terminal portion of E2 (Seedorf et al., 1985, Virology, 145,181-185; Kennedy et al., 1991, J. Virol. 65, 2093-2097; Cole et al., 1987, J. Mol.Biol. 193, 599-608 ; McBride et al., 1989, Proc. Natl. Acad. Sci. USA, 86, 510-514) and the reduction or lack of replication and transcriptional E2 activities can be easily determined using standard methods (see for example Sakai et al., 1996, J. Virol. 70, 1602-1611). Suitable defective E2 mutants for use in the present invention are described in the literature available to the man skilled in the art, e.g. in Demeret et al. (1995, Nucleic Acids Res. 23, 4777-4784), Sakai et al. (1996, J. Virol. 70, 1602-1611), Brokaw et al. (1996, J. Virology 70, 23-29) and Ferguson et al. (1996, J. Virology 70, 4193-4199). Preferred modifications in the context of the invention include the substitution of the Glu residue at position 39 of HPV-16 E2 preferably with an Ala residue (E39A) and/or the substitution of the Ile residue at position 73 preferably with an Ala residue (I73A) (e.g. see SEQ ID NO: 7). For purposes of illustration, the Glu and Ile residues at positions 39 and 73 of HPV-16 E2 correspond

respectively to the Glu and the Ile residues at positions 43 and 77 of HPV-18 E2 (e.g. see SEQ ID NO: 8).

Suitable E6 polypeptides for use in the invention encompass non-oncogenic mutants that are defective in binding to the cellular tumor suppressor gene product p53. Representative examples of non-oncogenic E6 polypeptides are described e.g. in Pim et al. (1994, Oncogene 9, 1869-1876), and Crook et al. (1991, Cell 67, 547-556). Preferred modifications in this context include the deletion in HPV-16 E6 of residues 118 to 122 (CPEEK) (e.g. see SEQ ID NO: 64) or the deletion in HPV-18 E6 of residues 113 to 117 (NPAEK).

Suitable E7 polypeptides for use in the invention encompass non-oncogenic mutants that are defective in binding to the cellular tumor suppressor gene product Rb. Representative examples of non-oncogenic E7 polypeptides are described, e.g. in Munger et al. (1989, EMBO J. 8, 4099-4105), Heck et al. (1992, Proc. Natl. Acad. Sci. USA 89, 4442-4446) and Phelps et al. (1992, J. Virol. 66, 2148-2427). Preferred modifications in this context include the deletion in HPV-16 E7 of residues 21 to 26 (DLYCYE) (e.g. see SEQ ID NO: 65) or the deletion in HPV-18 E7 of residues 24 to 28 (DLLCH).

Moreover, the polypeptides (e.g. papillomavirus polypeptides) encoded by the at least first and/or second nucleic acid molecules may further comprise additional modifications which are beneficial to the processing, stability and/or solubility of the encoded polypeptides, e.g. suppression of potential cleavage site(s), suppression of potential glycosylation site(s) and/or presentation at the surface of the expressing cells. For example, the encoded polypeptide(s) can include suitable signals for being anchored within the plasma membrane of the expressing cells. Indeed, it has been previously shown that membrane presentation permits to improve MHC class I and/or MHC class II presentation resulting in an enhancement of recognition by the host's immune system (see for example WO99/0388). As native early papillomavirus polypeptides (E1, E2, E6 and E7) are nuclear proteins (although no typical nuclear localization signal could be clearly identified), it could be beneficial to address them at the plasma membrane, in order to improve their immunogenic potential and thus their therapeutic efficacy in the host subject.

Efficient membrane presentation of a polypeptide at the surface of the expressing host cell can be achieved by fusing the polypeptide to a signal peptide and a membrane-

anchoring peptide. Such peptides are known in the art. Briefly, signal peptides are generally present at the N-terminus of membrane-presented or secreted polypeptides and initiate their passage into the endoplasmic reticulum (ER). They comprise 15 to 35 essentially hydrophobic amino acids which are then removed by a specific ER-located endopeptidase to give the mature polypeptide. Membrane-anchoring peptides are usually highly hydrophobic in nature and serve to anchor the polypeptides in the cell membrane (see for example Branden and Tooze, 1991, in Introduction to Protein Structure p. 202-214, NY Garland). The choice of the signal and membrane-anchoring peptides which can be used in the context of the present invention is vast. They may be independently obtained from any secreted or membrane-anchored polypeptide (e.g. cellular or viral polypeptides) such as the rabies glycoprotein, the HIV virus envelope glycoprotein or the measles virus F protein or may be synthetic. The preferred site of insertion of the signal peptide is the N-terminus downstream of the codon for initiation of translation and that of the membrane-anchoring peptide is the C-terminus, for example immediately upstream of the stop codon. If necessary, a linker peptide can be used to connect the signal peptide and/or the membrane anchoring peptide to the encoded polypeptide.

Representative examples of membrane-anchored and defective E1 polypeptides suitable for use in the invention are given in SEQ ID NO: 5 (defining the HPV-16 SS-E1*-TMR polypeptide illustrated in the example section) and in SEQ ID NO: 6 (defining the HPV-18 SS-E1*-TMF polypeptide illustrated in the example section). Representative examples of membrane-anchored and defective E2 polypeptides suitable for use in the invention are given in SEQ ID NO: 7 (defining the HPV-16 SS-E2*-TMR polypeptide illustrated in the example section) and in SEQ ID NO: 8 (defining the HPV-18 SS-E2*-TMR polypeptide illustrated in the example section). Representative examples of membrane-anchored and non-oncogenic E6 and E7 polypeptides suitable for use in the invention are given respectively in SEQ ID NO: 64 (defining the HPV-16 SS-E6*-TMF polypeptide illustrated in the example section) and in SEQ ID NO: 65 (defining the HPV-16 SS-E7*-TMR polypeptide illustrated in the example section.

In a particularly preferred embodiment, the at least first nucleic acid molecule and the second nucleic acid molecule encode two different papillomavirus polypeptides obtained from the same HPV serotype.

In a preferred aspect of this embodiment, the first nucleic acid molecule encodes an E1 polypeptide and the second nucleic acid molecule encodes an E2 polypeptide or vice versa. Desirably, the E1 and E2-encoding nucleic acid molecules are obtained from HPV-16 or from HPV-18. Preferably, the first nucleic acid molecule encodes a polypeptide comprising or essentially consisting of, or consisting of the amino acid sequence shown in SEQ ID NO: 5 and the second nucleic acid molecule encodes a polypeptide comprising or essentially consisting of, or consisting of the amino acid sequence shown in SEQ ID NO: 7. Alternatively, the first nucleic acid molecule encodes a polypeptide comprising or essentially consisting of, or consisting of the amino acid sequence shown in SEQ ID NO: 6 and the second nucleic acid molecule encodes a polypeptide comprising or essentially consisting of, or consisting of the amino acid sequence shown in SEQ ID NO: 8.

In the native context (e.g. the HPV-16 or HPV-18 genome), the 3' portion of the E1- encoding sequence overlaps the 5' portion of the E2-encoding sequence over 59 nucleotides. The presence of these 100% homologous 59 nucleotides is expected to negatively influence the stability of a vector expressing both E1 and E2-encoding nucleic acid molecules. Homologous recombination can occur between these common portions and lead to the loss of the nucleotide sequences comprised between them.

In accordance with the present invention, the 100% homology between the overlapping portion of 59 nucleotides present before modification in E1 and E2-encoding nucleic acid molecules can be reduced to less than 75% by degenerating the codon usage pattern in one of the nucleic acid molecules. A representative example of degenerated sequences is given in SEQ ID NO: 9 in which homology in the E1/E2 overlapping 59 nucleotides is reduced to 69% (as illustrated in Figure 1) and a preferred vector of the invention encoding HPV-16 E1 and E2 polypeptides comprises the nucleotide sequence shown in SEQ ID NO: 9. The same strategy can be applied to the overlapping portion present in HPV-18 E1 and E2-encoding sequences. Such degenerated sequences can be introduced in the E1-encoding first nucleic acid molecule in replacement of the native overlapping 59 nucleotides (e.g. SEQ ID NO: 10 and 11, respectively).

Accordingly, a preferred vector of the invention comprises a first nucleic acid molecule comprising or essentially consisting of, or consisting of the nucleotide sequence shown in SEQ ID NO: 10 (which encodes the HPV-16 E1 polypeptide of SEQ ID NO: 5) and a second nucleic acid molecule comprising, or essentially consisting of, or consisting

of the nucleotide sequence shown in SEQ ID NO: 12 (which encodes the HPV-16 E2 polypeptide of SEQ ID NO: 7). Another preferred vector of the invention comprises a first nucleic acid molecule comprising or essentially consisting of, or consisting of the nucleotide sequence shown in SEQ ID NO: 11 (which encodes the HPV-18 E1 polypeptide of SEQ ID NO: 6) and a second nucleic acid molecule comprising, or essentially consisting of, or consisting of the nucleotide sequence shown in SEQ ID NO: 13 (which encodes the HPV-18 E2 polypeptide of SEQ ID NO: 8). More preferably, the vector of the invention is a MVA vector, the first (E1-encoding) nucleic acid molecule is placed under the control of the vaccinia 7.5K promoter and the second (E2-encoding) nucleic acid molecule under the control of the vaccinia H5R promoter and the first and second nucleic acid molecules are both inserted in deletion III of said MVA vector.

The invention also pertains to a vector comprising a first nucleic acid molecule encoding an HPV-16 E1 polypeptide, a second nucleic acid molecule encoding an HPV-16 E2 polypeptide, a third nucleic acid molecule encoding an HPV-18 E1 polypeptide and a fourth nucleic acid molecule encoding an HPV-18 E2 polypeptide, wherein said first, second, third and fourth nucleic acid molecules do not comprise a portion of 40 or more continuous nucleotides exhibiting a percentage of homology of 75% or greater than 75%. Preferably, said HPV-16 E1 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 5, said HPV-16 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 7, said HPV-18 E1 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 6 and/or said HPV-18 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 8.

In the native context, HPV-16 and HPV-18 E1-encoding sequences comprise several portions of 40 or more continuous nucleotides that exhibit a percentage of homology of 80% or greater than 80%. The same is true with respect to HPV-16 and HPV-18 E2-encoding sequences. Moreover, the adjacent E1 and E2-encoding sequences overlap over a portion of approximately 59 nucleotides in HPV-16 and HPV-18 genomes. In this context, it is advisable to modify the HPV-18 E1 and E2-encoding nucleic acid molecules sequences so as to reduce homology with their HPV-16 counterparts to less than 75% especially in the homologous portions shared by the both serotypes. For this purpose, nucleotide sequences of HPV-16 and HPV-18 E1 and E2 genes can be aligned and modifications can be designed at the nucleotide level so as to reduce homology to less than 8, 7, 6 or preferably 5 consecutive nucleotides. Moreover, HPV-18 E1 sequence can

be further modified to reduce homology to less than 75% with the portion of 59 nucleotides overlapping the 5' end of the HPV-18 E2 sequence. Preferably, the codon usage is modified but modifications do not translate at the amino acid level, except for generating modifications as defined herein, e.g. resulting in defective enzymatic functions. Representative examples of "degenerated" HPV-18 E1- and HPV-18 E2-encoding nucleotide sequences that can be suitably used as third and fourth nucleic acid molecules are given in SEQ ID NO: 11 and SEQ ID NO: 13, respectively. A preferred vector of the invention comprises a first nucleic acid molecule comprising, or essentially consisting of or consisting of the nucleotide sequence shown in SEQ ID NO: 10 (encoding the HPV-16 E1 polypeptide shown in SEQ ID NO: 5), a second nucleic acid molecule comprising, or essentially consisting of or consisting of the nucleotide sequence shown in SEQ ID NO: 12 (encoding the HPV-16 E2 polypeptide shown in SEQ ID NO: 7), a third nucleic acid molecule comprising, or essentially consisting of or consisting of the nucleotide sequence shown in SEQ ID NO: 11 (encoding the HPV-18 E1 polypeptide shown in SEQ ID NO: 6) and a fourth nucleic acid molecule comprising, or essentially consisting of or consisting of the nucleotide sequence shown in SEQ ID NO: 13 (encoding the HPV-18 E2 polypeptide shown in SEQ ID NO: 8). Preferably, the vector is a MVA vector, the first, second, third and fourth nucleic acid molecules are introduced in deletion III of the MVA vector, the first and third (E1-encoding) nucleic acid molecules are placed in opposite orientation, each under the control of the vaccinia p7.5K promoter and the second and fourth (E2-encoding) nucleic acid molecules are placed in opposite orientation, each under the control of the vaccinia pH5R promoter.

In another particularly preferred embodiment, the at least first nucleic acid molecule and the second nucleic acid molecule encode the same polypeptide obtained from closely related organisms, e.g. closely related HPV serotypes such as HPV-16, HPV-18, HPV-33 and/or HPV-52.

In a first aspect of this embodiment, the same polypeptide obtained from closely related organisms is preferably an E2 polypeptide. The encoded E2 polypeptides are preferably modified so as to be membrane-anchored and defective for viral replication, as defined herein. In the native context, E2-encoding sequences of various genotypes exhibit a high degree of homology at the nucleotide level, especially in the most conserved portions. The presence of these homologous sequences is expected to negatively influence

the stability of a vector co-expressing two or more (e.g. 3, 4 or even more) E2 genes, for example E2 genes from HR HPV such as HPV-16, HPV-18, HPV-33 and HPV-52. Homologous recombination can occur between these homologous gene sequences and lead to the loss of the nucleotide sequences comprised between them, and thus to gene silencing.

In accordance with the present invention, the nucleic acid molecules encoding E2 polypeptides comprised in the vector of the invention can be modified by degenerating the codon usage pattern so as to reduce homology to less than 75% especially in the highly homologous portions. Representative examples of degenerated nucleic acid molecules encoding E2 polypeptides are given in SEQ ID NO: 13, 66, 67, 68 and 69. More specifically, SEQ ID NO: 13 encodes a membrane-presented and replication defective HPV-18 E2 polypeptide which nucleotide sequence has been designed so as to reduce homology with its E2-encoding counterparts to less than 8 or 7 consecutive nucleotides. SEQ ID NO: 66 and 67 both encode a replication-defective HPV-33 E2 polypeptide (it is further membrane-presented in SEQ ID NO: 67) which nucleotide sequences have been designed so as to reduce homology with the other E2-encoding couterparts to less than 8 or 7 consecutive nucleotides. SEQ ID NO: 68 and 69 both encode a replication-defective HPV-52 E2 polypeptide (it is further membrane-presented in SEQ ID NO: 69) which nucleotide sequences have been designed so as to reduce homology with the other E2-encoding counterparts to less than 8 or 7 consecutive nucleotides. However, the present invention is not limited to these exemplary sequences and alternative versions of degenerated nucleic acid molecules encoding E2 papillomavirus polypeptides as defined above can be designed on this principle.

A preferred vector of the invention comprises a first nucleic acid molecule encoding an HPV-16 E2 polypeptide as defined herein (e.g. the membrane-presented and replication-defective E2 polypeptide comprising the amino acid sequence shown in SEQ ID NO: 7), a second nucleic acid molecule encoding an HPV-18 E2 polypeptide as defined herein (e.g. the membrane-presented and replication-defective E2 polypeptide comprising the amino acid sequence shown in SEQ ID NO: 8), a third nucleic acid molecule encoding an HPV-33 E2 polypeptide as defined herein (e.g. the membrane-presented and replication-defective E2 polypeptide comprising the amino acid sequence shown in SEQ ID NO: 70), and a fourth nucleic acid molecule encoding an HPV-52 E2 polypeptide as defined herein (e.g. the membrane-presented and replication-defective E2

polypeptide comprising the amino acid sequence shown in SEQ ID NO: 71). More preferably, the first nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 12; the second nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 13;. the third nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 67 and/or the fourth nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 69. More preferably, the vector of the invention is a MVA vector and the four E2-encoding nucleic acid molecules are inserted in deletion III. Even more preferably, the first and the second nucleic acid molecules are under the control of the vaccinia H5R promoter and placed in inverted orientation each other whereas the third and fourth nucleic acid molecules are under the control of the vaccinia p7.5K promoter and placed in inverted orientation each other.

In another aspect of this embodiment, the same polypeptide obtained from closely related organisms is preferably an E6 polypeptide, an E7 polypeptide or both E6 and E7 polypeptides. E6 and E7 can be expressed independently or as a fusion polypeptide. The encoded E6 and/or E7 polypeptides are preferably modified so as to be membrane-anchored and non-oncogenic as defined herein.

In the native context HPV-16 and HPV-18 native E6 sequences have 63% of homology at the nucleotide level whereas HPV-16 and HPV-18 native E7 sequences are 57% homologous each other. However, in both cases, the HPV-16 and HPV-18 native sequences share several portions of 40 nucleotides or more that exhibit 80% or greater than 80% of homology (see Figure 2). The presence of these homologous portions is expected to negatively influence the stability of a vector co-expressing HPV-16 and HPV-18 E6 and/or E7 genes. Homologous recombination can occur between these homologous portions and lead to the loss of the nucleotide sequences comprised between them, and thus to gene silencing.

In accordance with the present invention, the nucleic acid molecules encoding HPV-16 and/or HPV-18 E6 and E7 polypeptides can be modified by degenerating the codon usage pattern so as to reduce homology to less than 75% especially in the homologous portions. A representative example of a degenerated nucleic acid molecule

encoding an HPV-18 E6 polypeptide is given in SEQ ID NO: 14 and a representative example of a degenerated modified nucleic acid molecule encoding an HPV-18 E7 polypeptide is given in SEQ ID NO:15. More specifically, SEQ ID NO: 14 and SEQ ID NO: 15 have been designed so as to reduce homology with the HPV-16 counterparts to less than 8, 7, 6 or preferably 5 consecutive nucleotides while encoding HPV-18 membrane-anchored and non-oncogenic E6 and E7 polypeptides. However, alternative versions of degenerated nucleic acid molecules encoding E6 and/or E7 papillomavirus polypeptides as defined above can be designed on this principle.

A preferred vector of the invention comprises a first nucleic acid molecule encoding an HPV-16 E6 polypeptide as defined herein (e.g. membrane-anchored and non oncogenic) and a second nucleic acid molecule encoding an HPV-18 E6 polypeptide as defined herein (e.g. a membrane-anchored non oncogenic), wherein the second nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 14. Another preferred vector of the invention comprises a first nucleic acid molecule encoding an HPV-16 E7 polypeptide as defined herein (e.g. membrane-anchored and non oncogenic) and a second nucleic acid molecule encoding an HPV-18 E7 polypeptide as defined herein (e.g. membrane-anchored and non oncogenic), wherein the second nucleic acid molecule comprises or essentially consists of the nucleotide sequence shown in SEQ ID NO: 15. More preferably, the vector of the invention is a MVA vector, the first nucleic acid molecule is placed under the control of the vaccinia 7.5K promoter and the second nucleic acid molecule under the control of the vaccinia H5R promoter and the first and second nucleic acid molecules are both inserted in deletion III of said MVA vector.

The invention also pertains to a vector comprising a first nucleic acid molecule encoding a fusion of an HPV-16 E6 polypeptide with an HPV-16 E7 polypeptide and a second nucleic acid molecule encoding a fusion of an HPV-18 E6 polypeptide with an HPV-18 E7 polypeptide wherein said first and second nucleic acid molecules do not comprise a portion of 40 or more continuous nucleotides exhibiting a percentage of homology of approximately 75% or greater than 75%.

The invention also pertains to a vector comprising a first nucleic acid molecule encoding an HPV-16 E6 polypeptide, a second nucleic acid molecule encoding an HPV-18 E6 polypeptide, a third nucleic acid molecule encoding an HPV-16 E7 polypeptide and a fourth nucleic acid molecule encoding an HPV-18 E7 polypeptide wherein said first,

second, third and fourth nucleic acid molecules do not comprise a portion of 40 or more continuous nucleotides exhibiting a percentage of homology of 75% or greater than 75%. Preferably, the second nucleic acid molecule comprises, essentially consists in or consists in SEQ ID NO: 14 and/or the fourth nucleic acid molecule comprises, essentially consists in or consists in SEQ ID NO: 15. More preferably, the vector of the invention is a MVA vector, the first and second nucleic acid molecules are placed in inverted orientation each under the control of the vaccinia 7.5K promoter and the third and fourth nucleic acid molecules in inverted orientation each under the control of the vaccinia H5R promoter and the first, second, third and fourth nucleic acid molecules are inserted in deletion III of said MVA vector.

In another aspect, the present invention also provides a substantially isolated nucleic acid molecule comprising, essentially consisting of or consisting of the nucleotide sequence shown in any SEQ ID NO: 9, 10, 11, 12, 13, 14, 15, 66, 67, 68 or 69.

In another embodiment of the invention, the first and second and if present third and fourth nucleic acid molecules comprised in the vector of the invention are in a form suitable for expression of the encoded polypeptides in a host cell or subject, which means that they are placed under the control of the regulatory sequences necessary to their expression.

As used herein, the term "regulatory sequences" refers to any sequence that allows, contributes or modulates the expression of a nucleic acid molecule in a given host cell, including replication, duplication, transcription, splicing, translation, stability and/or transport of the nucleic acid or one of its derivative (i.e. mRNA) into the host cell. In the context of the present invention, the regulatory sequences are "operably linked" to the nucleic acid molecule to be expressed, i.e. they are placed in a functional relationship which allows for expression in a host cell or subject. Such regulatory sequences are well known in the art (see for example Goeddel, 1990, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego). It will be appreciated by those skilled in the art that the choice of the regulatory sequences can depend on factors such as the vector type, the host cell, the level of expression desired, etc.

The promoter is of special importance and the present invention encompasses the use of constitutive promoters which direct expression of the nucleic acid molecules in many types of host cells and those which direct expression only in certain host cells (e.g., tissue-specific regulatory sequences) or in response to specific events or exogenous factors (e.g. by temperature, nutrient additive, hormone or other ligand). Suitable promoters for constitutive expression in eukaryotic systems include viral promoters, such as SV40 promoter, the cytomegalovirus (CMV) immediate early promoter or enhancer (Boshart et al., 1985, Cell 41, 521-530), the adenovirus early and late promoters, the thymidine kinase (TK) promoter of herpes simplex virus (HSV)-1 and retroviral long-terminal repeats (e.g. MoMuLV and Rous sarcoma virus (RSV) LTRs) as well as cellular promoters such as the phosphoglycero kinase (PGK) promoter (Hitzeman et al., 1983, Science 219, 620-625 ; Adra et al., 1987, Gene 60, 65-74). Suitable promoters useful to drive expression of the nucleic acid molecules in a poxviral vector include the 7.5K, H5R, TK, p28, p11 or K1L promoters of vaccinia virus. Alternatively, one may use a synthetic promoter such as those described in Chakrabarti et al. (1997, Biotechniques 23, 1094-1097), Hammond et al. (1997, J. Virological Methods 66, 135-138) and Kumar and Boyle (1990, Virology 179, 151-158) as well as chimeric promoters between early and late poxviral promoters.

Inducible promoters are regulated by exogenously supplied compounds, and include, without limitation, the zinc-inducible metallothionein (MT) promoter (Mc Ivor et al., 1987, Mol. Cell Biol. 7, 838-848), the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088), the ecdysone insect promoter (No et al., 1996, Proc. Natl. Acad. Sci. USA 93, 3346-3351), the tetracycline-repressible promoter (Gossen et al., 1992, Proc. Natl. Acad. Sci. USA 89, 5547-5551), the tetracycline-inducible promoter (Kim et al., 1995, J. Virol. 69, 2565-2573), the RU486-inducible promoter (Wang et al., 1997, Nat. Biotech. 15, 239-243 and Wang et al., 1997, Gene Ther. 4, 432-441), the rapamycin-inducible promoter (Magari et al., 1997, J. Clin. Invest. 100, 2865-2872) and the lac, TRP, and TAC promoters from E. coli.

The regulatory sequences in use in the context of the present invention can also be tissue-specific to drive expression of the nucleic acid molecules in specific tissues where therapeutic benefit is desired. Suitable promoters can be taken from genes that are preferentially expressed in tumor cells. Such genes can be identified for example by

display and comparative genomic hybridization (see for example US 5,759,776 and 5,776,683).

Those skilled in the art will appreciate that the regulatory elements controlling the expression of the nucleic acid molecules comprised in the vector of the invention may further comprise additional elements for proper initiation, regulation and/or termination of transcription (e.g. polyA transcription termination sequences), mRNA transport (e.g. nuclear localization signal sequences), processing (e.g. splicing signals), stability (e.g. introns and non-coding 5' and 3' sequences), and translation (e.g. tripartite leader sequences, ribosome binding sites, Shine-Dalgamo sequences, etc.) into the host cell or subject.

In another aspect, the present invention provides infectious viral particles comprising the above-described vector. No attempts to describe in detail the various methods known for the production of infectious viral particles will be made here. Typically, such viral particles are produced by a process comprising the steps of (a) introducing the viral vector in an appropriate cell line, (b) culturing the cell line under suitable conditions so as to allow the production of said infectious viral particle, recovering the produced infectious viral particle from the culture of said cell line, and optionally purifying said recovered infectious viral particle.

When the viral vector is defective, the infectious particles are usually produced in a complementation cell line or via the use of a helper virus, which supplies *in trans* the non functional viral genes. For example, suitable cell lines for complementing E1-deleted adenoviral vectors include the 293 cells (Graham et al., 1997, J. Gen. Virol. 36, 59-72), the PER-C6 cells (Fallaux et al., 1998, Human Gene Ther. 9, 1909-1917) and the HER96 cells. Cells appropriate for propagating poxvirus vectors are avian cells, and most preferably primary chicken embryo fibroblasts (CEF) prepared from chicken embryos obtained from fertilized eggs. The producer cells can be cultured in conventional fermentation bioreactors, flasks and Petri plates under appropriate temperature, pH and oxygen content conditions.

The infectious viral particles may be recovered from the culture supernatant or from the cells after lysis. They can be further purified according to standard techniques (chromatography, ultracentrifugation as described for example in WO96/27677,

WO98/00524, WO98/22588, WO98/26048, WO00/40702, EP1016700 and WO00/50573).

In another aspect, the present invention provides host cells comprising the above-described nucleic acid molecules, vectors or infectious viral particles. The term "host cell" as used herein defines any cell which can be or has been the recipient of the vector or the infectious viral particle of this invention and progeny of such cells. This term should be understood broadly so as to encompass isolated cells, a group of cells, as well as particular organization of cells, e.g. in tissue or organ. Such cells can be primary, transformed or cultured cells.

Host cells in the context of the invention include prokaryotic cells (e.g. *Escherichia coli*, Bacillus, Listeria), lower eukaryotic cells such as yeast (e.g. *Saccharomyces cerevisiae*, *Saccharomyces pombe* or *Pichia pastoris*), and other eukaryotic cells such as insect cells, plant and higher eukaryotic cells, with a special preference for mammalian cells (e.g. human or non-human cells). Representative examples of suitable host cells include but are not limited to BHK (baby hamster kidney) cells, MDCK cells (Madin-Darby canine kidney cell line), CRFK cells (Crandell feline kidney cell line), CV-1 cells (African monkey kidney cell line), COS (e.g., COS-7) cells, chinese hamster ovary (CHO) cells, mouse NIH/3T3 cells, HeLa cells and Vero cells. The term “host cell” also encompasses complementing cells capable of complementing at least one defective function of a replication-defective vector of the invention (e.g. adenoviral vector) such as those cited above.

Host cells can be used for producing by recombinant means the polypeptides encoded by the nucleic acid molecules comprised in the vector or infectious particles of the invention. Such techniques are well known in the art (see for example Ausubel, Current Protocols in Molecular Biology, John Wiley, 1987-2002; and the latest edition of Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press).

In another aspect, the present invention provides a composition comprising the above-described nucleic acid molecules, vector, infectious viral particle, or host cell (also referred herein to the "active agent") or any combination thereof. Advantageously, the

composition is a pharmaceutical composition which comprises a therapeutically effective amount of the active agent(s) and a pharmaceutically acceptable vehicle.

The term "pharmaceutically acceptable vehicle" as used herein is intended to include any and all carriers, solvents, diluents, excipients, adjuvants, dispersion media, coatings, antibacterial and antifungal agents, and absorption delaying agents, and the like, compatible with pharmaceutical administration. As used herein a “therapeutically effective amount” is a dose sufficient for the alleviation of one or more symptoms normally associated with the pathological condition desired to be treated or prevented in a subject. When prophylactic use is concerned, this term means a dose sufficient to prevent or to delay the establishment of a pathological condition in a subject. For example, a therapeutically effective amount could be that amount that is sufficient to induce or enhance an immune response in the treated subject, or that amount that is sufficient to palliate, ameliorate, stabilize, reverse, slow or delay the progression of the pathological condition (e.g. for instance size reduction or regression of a lesion or a tumor in a subject, reversion of a viral infection in an infected subject).

Desirably, the composition of the invention comprises one or more carrier and/or diluent non-toxic at the dosage and concentration employed. Such carrier and/or diluent are preferably selected from those usually employed to formulate compositions in either unit dosage or multi-dose form for systemic or mucosal administration. A suitable carrier can be a solvent, a dispersing medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol, and the like), a vegetable oil or suitable mixtures thereof. The diluent is preferably isotonic, hypotonic or weakly hypertonic and has a relatively low ionic strength. Representative examples of suitable diluents include sterile water, physiological saline (e.g. sodium chloride), Ringer’s solution, glucose, trehalose or saccharose solutions, Hank’s solution, and other aqueous physiologically balanced salt solutions (see for example the most current edition of Remington : The Science and Practice of Pharmacy, A. Gennaro, Lippincott, Williams&Wilkins). The pH of the composition of the invention is suitably adjusted and buffered in order to be appropriate for use in humans or animals, preferably at a physiological or slightly basic pH (between approximately pH 7.5 to approximately pH 9, with a special preference for a pH of approximately 8 or 8.5). Suitable buffers include phosphate buffer (e.g. PBS), bicarbonate buffer and/or Tris buffer.

The composition of the invention can be in various forms, e.g. frozen, solid (e.g. dry powdered or lyophilized form), or liquid (e.g. aqueous). A solid composition of the active agent plus any additional desired ingredient(s) can be obtained from a previously sterile-filtered solution thereof submitted to vacuum drying and freeze-drying. It can, if desired, be stored in a sterile ampoule ready for reconstitution by the addition of a suitable vehicle before use.

A particularly preferred composition (especially when the active agent is an adenoviral vector) is formulated in 1M saccharose, 150 mM NaCl, 1mM $MgCl_2$, 54 mg/l Tween 80, 10 mM Tris pH 8.5. Another preferred composition is formulated in 10 mg/ml mannitol, 1 mg/ml HSA, 20 mM Tris, pH 7.2, and 150 mM NaCl. Such formulations are particularly suited for preserving stability of the composition of the invention over a period of at least two months at either freezing (e.g. -70°C, -40°C, -20°C), or refrigerated (e.g. 4°C) temperature.

The composition may also contain one or more pharmaceutically acceptable excipients for providing desirable pharmaceutical or pharmacodynamic properties, including for example modifying or maintaining the pH, osmolarity, viscosity, clarity, colour, sterility, stability, release or absorption into an the human or animal subject. Representative examples of stabilizing components include polysorbate 80, L-arginine, polyvinylpyrrolidone, trehalose, and polymers such as polyethylene glycol which may be used to obtain desirable properties of solubility, stability, and half-life (Davis et al., 1978, Enzyme Eng. 4, 169-173; Burnham et al., 1994, Am. J. Hosp. Pharm. 51, 210-218). Viscosity enhancing agents include sodium carboxymethylcellulose, sorbitol, and dextran. The composition can also contain substances known in the art to promote penetration or transport across a mucosal barrier or in a particular organ. For example, a composition suited for vaginal administration can eventually include one or more absorption enhancers useful to increase the pore size of the mucosal membranes.

In addition, the composition of the invention may comprise one or more adjuvant(s) suitable for systemic or mucosal administration in humans. The term "adjuvant" denotes a compound having the ability to enhance the immune response to a particular antigen. The adjuvant can be delivered at or near the site of antigen. Enhancement of humoral immunity is typically manifested by a significant increase (usually greater than 10 fold) in the titer of antibody raised to the antigen. Enhancement of cellular immunity can be measured for example by a positive skin test, cytotoxic T-cell

assay, ELIspot assay for IFNg or IL-2. Preferably, the adjuvant in use in the invention is capable of stimulating immunity to the active agent, especially through toll-like receptors (TLR), such as TLR-7, TLR-8 and TLR-9. Representative examples of useful adjuvants include without limitation alum, mineral oil emulsion such as Freunds complete and incomplete (IFA), lipopolysaccharide or a derivative thereof (Ribi et al., 1986, Immunology and Immunopharmacology of Bacterial Endotoxins, Plenum Publ. Corp., NY, p407-19), saponins such as QS21 (Sumino et al., 1998, J.Virol. 72, 4931-9; WO 98/56415), imidazoquinoline compounds such as Imiquimod (Suader, 2000, J. Am Acad Dermatol. 43, S6-S11), 1H-imidazo (4, 5-c) quinolon-4-amine derivative (Aldara^{TM}) and related compound (Smorlesi, 2005, Gene Ther. 12, 1324-32), cytosine phosphate guanosine oligodeoxynucleotides such as CpG (Chu et al., 1997, J. Exp. Med. 186: 1623; Tritel et al., 2003, J. Immunol. 171: 2358-2547) and cationic peptides such as IC-31 (Kritsch et al., 2005, J. Chromatogr Anal. Technol Biomed Life Sci 822, 263-70).

The nucleic acid molecule, vector, infectious particle or composition of the invention can be administered by a variety of modes of administration, including systemic, topical and mucosal administration. Systemic administration can be performed by any means, e.g. by subcutaneous, intradermal, intramuscular, intravenous, intraperitoneal, intravascular, intraarterial injection. Injections can be made with conventional syringes and needles, or any other appropriate devices available in the art. Mucosal administration can be performed by oral, nasal, intratracheal, intrapulmonary, intravaginal or intra-rectal route. Topical administration can be performed using transdermal means (e.g. patch and the like). Intramuscular or subcutaneous administration is particularly preferred with viral vectors and infectious particles as active agent.

The appropriate dosage may vary depending upon known factors such as the pharmacodynamic characteristics of the particular active agent, age, health, and weight of the subject, the pathological condition(s) to be treated, nature and extent of symptoms, kind of concurrent treatment, frequency of treatment, the need for prevention or therapy and/or the effect desired. The dosage will also be calculated dependent upon the particular route of administration selected. Further refinement of the calculations necessary to determine the appropriate dosage for treatment is routinely made by a practitioner, in the light of the relevant circumstances. For general guidance, suitable dosage for adenovirus particles varies from about 10^5 to about 10^{13} iu (infectious units), desirably from about 10^7 to about 10^{12} iu and preferably from about 10^8 to about 10^{11} iu. Suitable dosage for

vaccinia virus particles varies from about 10^4 to about 10^{10} pfu (plaque-forming particle), desirably from about 10^5 to about 10^9 pfu and preferably from about 10^6 to about $5x10^8$ pfu. Vector plasmids can be administered in doses of between 10 μg and 20 mg, and preferably between 100 μg and 2 mg.

Further, the administration may take place in a single dose or, alternatively, in multiple doses according to standard protocols, dosages and regimens over several hours, days and/or weeks. Moreover, the administration can be by bolus injection or continuous infusion. For example, the subject may be treated with at least two (e.g. from 2 to 10) administrations of the above-described nucleic acid molecule, vector, infectious particle or composition. Preferably, a first series of administrations is carried out sequentially within a period of time varying from few days to 4 weeks followed by a second series of administrations (e.g. one or two administrations) carried out within one to 6 months following the latest administration of the first series. The period of time between each of the administrations of the second series can be from few days to 4 weeks. In a preferred embodiment, the first series of administrations comprises three sequential administrations at week interval and the second series comprises one administration within 4 to 6 months following the first series. As a general guidance, with MVA vector, preferred administration route is subcutaneous with a dose of MVA particles comprised between 10^6 to $5x10^8$ pfu.

The nucleic acid molecule, vector, infectious particle, host cell or composition of the invention may be introduced in a subject for treating or preventing a variety of pathological conditions, including genetic diseases, congenital diseases and acquired diseases. The present invention also pertains to the use of the nucleic acid molecule, vector, infectious particle, host cell or composition of the invention for the preparation of a drug intended for treating or preventing such pathological conditions. It is particularly appropriate for treating or preventing infectious diseases (e.g. viral and/or bacterial infections), cancers and immune deficiency diseases. The term “cancer” as used herein encompasses any cancerous condition which results from unwanted cell proliferation including diffuse or localized tumors, metastasis, cancerous polyps and preneoplastic lesions (e.g. neoplasia).

Infectious diseases which are contemplated in the context of the invention encompass any condition associated with infection by a pathogenic microorganism as described above. Cancers which are contemplated in the context of the invention include without limitation glioblastoma, sarcoma, melanoma, mastocytoma, carcinoma as well as breast cancer, prostate cancer, testicular cancer, ovarian cancer, endometrial cancer, cervical cancer (in particular, those associated with a papillomavirus infection), lung cancer (e.g. including large cell, small cell, squamous and adeno-carcinomas), renal cancer, bladder cancer, liver cancer, colon cancer, anal cancer, pancreatic cancer, stomach cancer, gastrointestinal cancer, cancer of the oral cavity, larynx cancer, brain and CNS cancer, skin cancer (e.g. melanoma and non-melanoma), blood cancer (lymphomas, leukemia, especially if they have developed in solid mass), bone cancer, retinoblastoma and thyroid cancer.

In a preferred embodiment, the invention is used for the preventive or curative treatment of a condition associated with infection by a papillomavirus (especially a HR HPV) such as persistent infection, pre-malignant and malignant lesions. "Persistent infection" refers to the asymptomatic phase of the papillomavirus infection in an infected subject that has not achieved viral eradication. Typically no clinical signs are observed. Examples of pre-malignant lesions include without limitation intraepithelial neoplasia of low, moderate or high grade that can be detected in various tissues such as CIN (cervical intraepithelial neoplasia), vulvar intraepithelial neoplasia (VIN), anal intraepithelial neoplasia (AIN), penile intraepithelial neoplasia (PIN), and vaginal intraepithelial neoplasia (VaIN). Examples of malignant lesions include without limitation cervical carcinoma, anal carcinoma, vaginal cancer, penile cancer and oral cancer. The nucleic acid molecule, vector, infectious particle, host cell or composition of the invention encoding papillomavirus polypeptides is particularly destined for treating pre-malignant, especially CIN2/3 lesions, or malignant lesions, especially cervical carcinoma. In another embodiment, the invention can also be used for the preventive or curative treatment of a condition associated with infection by a hepatitis virus (e.g. HBV or HCV) such as persistent infection, chronic or fulgurant hepatitis and liver cancer.

The active agent can be used alone or, if desired, in conjunction with conventional therapeutic modalities (e.g. radiation, chemotherapy and/or surgery). The conventional therapeutic modalities are delivered in the animal or human subject according to standard protocols using standard agents, dosages and regimens and such modalities may be

performed before during and/or after the administration of the active agent(s) of the invention. For example, for treating conditions associated with HCV infection, the method or use of the invention is preferably associated with e.g. protease inhibitors (e.g. serine protease inhibitors such as VX950 of Vertex), polymerase inhibitors, helicase inhibitors, antifibrotics, nucleoside analogs, TLR agonists, siRNA, antisense oligonucleotides, anti-HCV antibodies, immune modulators, therapeutic vaccines and antitumor agents conventionally used in the treatment of HCV-associated hepatocarcinomas (e.g. adriamycin or a mixture of adriamycin lipiodol and spongel usually administered by chimioembolisation in the hepatic artery). A particularly suitable combination includes treatment with pegylated IFN-α (IFN-α2a or IFN-α2b) and/or ribavirin, preferably for 24 to 48 weeks before administration of the active agent(s) of the invention. For treating conditions associated with papillomavirus infection, the method or use of the invention can be associated with ablative procedures, such as loop electrosurgical excision. The method or use according to the invention can also be carried out in conjunction with immunostimulator(s) such as cytokines (e.g. IL-2, IL-7, IL-15, IL-18, IL-21, IFNg) or suicide gene products (e.g. the thymidine kinase of HSV-1 described in Caruso et al., 1993, Proc. Natl. Acad. Sci. USA 90, 7024-28; *FCU-1* described in WO 99/54481) or vector(s) expressing such polypeptide(s).

In another embodiment, the method or use of the invention is carried out according to a prime boost therapeutic modality which comprises sequential administrations of primer composition(s) and booster composition(s). Typically, the priming and the boosting compositions use different vehicles which comprise or encode at least an antigenic domain in common. The method or use of the invention may comprise one to ten administrations of the priming composition followed by one to ten administrations of the boosting composition. Desirably, injection intervals are a matter of one day to twelve months. A preferred modality includes three or four sequential administrations of the primer independently separated by a period of time varying from 3 to 10 days (e.g. a week) followed by one or two administration(s) of the booster one to several weeks after the latest primer administration. Moreover, the priming and boosting compositions can be administered at the same site or at alternative sites by the same route or by different routes of administration. For example, compositions based on polypeptide can be administered by a mucosal route whereas compositions based on vectors are preferably injected, e.g. subcutaneous injection for a MVA vector, intramuscular injection

for a DNA plasmid and for an adenoviral vector. The vector, infectious particle or composition of the invention can be used to either prime or boost or both prime and boost an immune response in the treated subject. In one embodiment, priming is performed with a plasmid vector of the invention and boosting with a vaccinia virus infectious particle of the invention. In another embodiment, priming is performed with an adenovirus infectious particle of the invention and boosting with a vaccinia virus infectious particle of the invention. In still another embodiment, priming is performed with a vaccinia virus infectious particle of the invention and boosting with an adenovirus infectious particle of the invention.

While the invention has been described in connection with specific embodiments thereof, it will be understood that the scope of the claims should not be limited by the preferred embodiments set forth in the examples, but should be given the broadest interpretation consistent with the description as a whole.

Legends of Figures

Figure 1 illustrates the sequence alignment (A) between the 59 nucleotides present (a) at the end of the native HPV-16 E1 sequences and (b) at the beginning of the native HPV-16 E2 sequences and (B) between the 59 nucleotides present (a) at the end of the native HPV-16 E1 sequences or at the beginning of the native HPV-16 E2 sequences and (b) SEQ ID NO: 9.

Figure 2 illustrates the sequence alignment (A) between HPV-16 and HPV-18 E6-encoding sequences and (B) between HPV-16 E6-encoding sequences and SEQ ID NO: 14.

The following examples serve to illustrate the present invention.

EXAMPLES

The constructions described below are carried out according to the general genetic engineered and molecular cloning techniques detailed in Maniatis et al. (1989, Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY) or according to the manufacturer's recommendations when a commercial kit is used. PCR amplification techniques are known to the person skilled in the art (see for example PCR protocols –A guide to methods and applications, 1990, published by Innis, Gelfand, Sninsky andWhite, Academic Press). The recombinant plasmids carrying the ampicillin resistance gene are replicated in the *E. coli* C600 (Stratagene), BJ5183 (Hanahan, 1983, J. Mol. Biol. 166, 557-580) and NM522 on agar or liquid medium supplemented with 100µg/ml of antibiotic. The constructions of the recombinant vaccinia viruses are performed according to the conventional technology in the field in the documents above cited and in Mackett et al. (1982, Proc. Natl. Acad. Sci. USA 79, 7415-7419) and Mackett et al. (1984, J. Virol. 49, 857-864). The selection gene *gpt* (xanthine guanine phosphoribosyltransferase) of *E. coli* (Falkner and Moss, 1988, J. Virol. 62, 1849-1854) is used to facilitate the selection of the recombinant vaccinia viruses.

Example 1: Construction of a MVA vector expressing HPV-16 E1 and E2 genes (MVATG17410)

a) Construction of a recombinant MVA vector encoding HPV-16 E2 gene (MVATG17408)

Cloning of HPV16 E2 gene

The nucleotide sequences encoding HPV-16 E2 were cloned from the genomic DNA isolated from CaSki cells (ATCC CRL-1550). E2 gene was amplified using primers OTG16809 (SEQ ID NO: 16) and OTG16810 (SEQ ID NO: 17). The resulting fragment was digested by *Bam*HI and *Eco*RI and inserted in pGEX2T (Amersham Biosciences) restricted by the same enzymes, giving rise to pTG17239. Sequencing of the cloned E2 gene showed five mutations comparing to HPV16 E2 prototype sequence (described in Genbank NC-01526). Two mutations were silent and the three non-silent

mutations (T210I, S219P, K310T) were corrected using the QuickChange Site Directed Mutagenesis kit (Stratagene), giving rise to pTG17268.

Modification of the HPV-16 E2 polypeptide

The E2 nucleotide sequences incorporated in pTG17268 were modified by site directed mutagenesis, in order to generate an HPV-16 E2 variant (E39A and I73A), designated E2*. More specifically, the E2 replication function was abolished by substituting the Glu residue in position 39 with an Ala and the transactivation function by substituting the Ile residue in position 73 with an Ala. The resulting plasmid pTG17318 comprises the modified sequences encoding HPV-16 E2*.

HPV-16 E2* was further modified by fusion at its N-terminus to a peptide signal and at its C-terminus to a membrane-anchoring sequences derived from the glycoprotein of the rabies virus (PG strain; Genbank ay009097) so as to direct presentation of HPV-16 E2* in the expressing host cells at the plasma membrane surface. The nucleotide sequences (SEQ ID NO: 12) encoding the membrane-presented E2 defective variant, designated SS-E2*-TMR, were reassembled by triple PCR using the following primers: OTG17500 (SEQ ID NO: 18), OTG17501 (SEQ ID NO: 19), OTG17502 (SEQ ID NO: 20), OTG17503 (SEQ ID NO: 21), OTG17504 (SEQ ID NO: 22) and OTG17505 (SEQ ID NO: 23). The reassembled sequence was inserted in a pBS-derived vector (Stratagene), to give pTG17360, and then cloned in a vaccinia transfer plasmid downstream the pH5R promoter (Rosel *et al*, 1986, J Virol. 60, 436-449) resulting in pTG17408.

The transfer plasmid is designed to permit insertion of the nucleotide sequence to be transferred by for homologous recombination in deletion III of the MVA genome. It originates from plasmid pTG1E (described in Braun *et al.,* 2000, Gene Ther. 7, 1447-57) into which were cloned the flanking sequences (BRG3 and BRD3) surrounding the MVA deletion III, which sequences were obtained by PCR from MVATGN33.1 DNA (Sutter and Moss, 1992, Proc. Natl. Acad. Sci. USA 89, 10847-51). The transfer plasmid also contains a fusion between the *Aequorea victoria* enhanced Green Fluorescent protein (*eGFP* gene, isolated from pEGP-C1, Clontech) and the *Escherichia coli* xanthine-guanine phosphoribosyltransferase gene (*gpt* gene) under the control of the early late vaccinia virus synthetic promoter p11K7.5 (kindly provided by R. Wittek, University of Lausanne). Synthesis of xanthine-guanine phosphoribosyltransferase enables GPT^+ recombinant MVA to form plaques in a selective medium containing mycophenolic acid, xanthine, and hypoxanthine (Falkner *et al*, 1988, J. Virol. 62, 1849-

54) and *eGFP* enables the visualisation of recombinant MVA plaques. The selection marker *eGPP-GPT* is placed between two homologous sequences in the same orientation. When the clonal selection is achieved, the selection marker is easily eliminated by several passages without selection allowing the growth of eGPP-*GPT* recombinant MVA.

Construction of a recombinant MVA expressing the HPV-16 SS-E2-TMR gene*

Generation of MVATG17408 virus was performed by homologous recombination in primary chicken embryos fibroblasts (CEF) infected with MVATGN33.1 (at a MOI of 0.1 pfu/cell) and transfected with pTG17408 (according to the standard calcium phosphate DNA precipitation). Viral selection was performed by three round of plaque purification in the presence of a selective medium containing mycophenolic acid, xanthine and hypoxanthine. As mentioned above, the selection marker was then eliminated by passage in a non-selective medium. Absence of contamination by parental MVA was verified by PCR.

Analysis of E2 expression was performed by Western-blot. CEF were infected at MOI 0.2 with MVATG17408 and after 24 hours, cells were harvested. Western-blot analysis was performed using commercial monoclonal anti-E2 antibody TVG271 (Abcam). Expression of a protein with an apparent molecular weight of 55 kDA was detected, while theoretical molecular weight of E2*-TMR is 48.9 kDa. After treatment of cell extracts with endoglycosydase F, a reduction of the size of the recombinant protein was observed, suggesting that the E2* TMR polypeptide expressed from MVATG17408 is N-glycosyled.

<u>b) Construction of a recombinant MVA encoding an HPV-16 E1 gene degenerated in the portion overlapping with HPV-16 E2 gene (MVATG17409)</u>

The nucleotide sequences encoding HPV-16 E1 polypeptide were cloned from CaSki cell DNA (ATCC CRL-1550). More specifically, the E1 gene was amplified in two parts E1a (nt 1 – 1102) and E1b (nt 1001 to 1950). Primers OTG16811 (SEQ ID NO: 24) and OTG 16814 (SEQ ID NO: 25) were used to amplify E1a fragment, which was digested by *Bam*HI and *Eco*RI and inserted in pGEX2T restricted by the same enzymes, giving rise to pTG17240. E1b fragment was generated using OTG16813 (SEQ ID NO: 26) and OTG16812 (SEQ ID NO: 27) and digested by *Bam*HI and *Eco*RI before being inserted in pGEX2T, resulting in pTG17241. Sequencing showed 4 mutations comparing to HPV-16 E1 prototype sequence (described in Genbank NC-01526). One

mutation was silent and the three non-silent mutations present in E1a (K130Q, N185T and T220S) were corrected by site-directed mutagenesis. The complete E1 gene was then reassembled by cloning the corrected E1a fragment in pTG17241 digested by *BsrG*I and *Eco*RI. The resulting plasmid was named pTG17289.

In the HPV-16 genome, the 59 last nucleotides of the E1 gene are identical to the 59 first nucleotides of the E2 gene. In order to avoid problem of instability during production steps of an E1 and E2-encoding MVA vector, this portion of E1-encoding sequences was modified by codon usage modifications so as to decrease the sequence homology with the E2-encoding sequence. The degenerated sequence (SEQ ID NO: 9) was obtained by amplification of the 3' end of E1 gene using degenerated primers OTG17408 (SEQ ID NO: 28) and OTG17409 (SEQ ID NO: 29). The amplified fragment was digested by *Nsi*I and *Bgl*II and inserted in pTG17289 restricted by the same enzymes, giving rise to pTG17340.

The HPV-16 degenerated E1 sequences were also mutated by site-directed mutagenesis in order to abolish the replication function of the encoded E1 polypeptide, by substituting the Gly residue in position 482 of HPV-16 E1 with an Asp residue (G482D; also designated herein E1*), resulting in pTG17373.

The HPV-16 E1deg* sequences were also modified so as to direct expression of the encoded polypeptide at the plasma cell surface, by fusion with the signal and the membrane-anchoring peptides derived from the glycoprotein of the rabies virus (ERA isolate; described in Genbank N° M38452). The SS-E1deg*-TMR sequence (SEQ ID NO: 10) was reconstituted by triple PCR using the following primers OTG17560 (SEQ ID NO: 30), OTG17561 (SEQ ID NO: 31), OTG17562 (SEQ ID NO: 32), OTG17563 (SEQ ID NO: 33), OTG17564 (SEQ ID NO: 34) and OTG17565 (SEQ ID NO: 35). The resulting fragment was inserted in a pBS-derived vector (Stratagene), giving pTG17404. The SS-E1deg*-TMR sequence was then cloned in the vaccinia transfer plasmid downstream of the p7.5K promoter (Cochran *et al,* 1985, J. Virol. 54, 30-7) giving rise to pTG17409.

Generation of MVATG17409 viruses was performed in CEF by homologous recombination as described above.

<u>c) Construction of a recombinant MVA encoding HPV-16 E1 and E2 genes (MVATG17410)</u>

The SS-E1deg*-TMR sequenced controlled by the p7.5K promoter was isolated from pTG17409 and inserted in pTG17408, giving rise to pTG17410.

Generation of MVATG17410 viruses was performed in CEF by homologous recombination as described above.

Example 2: Construction of a MVA vector encoding HPV-18 E1 and E2 genes (MVATG17582)

HPV-18 E1 and E2 genes were reconstituted as synthetic genes and the oligonucleotides were designed so as (i) to reduce the percentage of homology between the homologous portions shared by the native HPV-16 and HPV-18 sequences to less than 75% (Sequences of HPV-16 and HPV-18 E1 and E2 genes were aligned and oligonucleotides were designed so as to reduce homology to less than 5 consecutive nucleotides) (ii) to reduce homology to less than 75% between the portion of 59 nucleotides present both in the 3' end of the native HPV-18 E1 sequence and in the 5' end of the HPV-18 E2 sequence and (ii) to introduce the mutations abolishing the enzymatic functions of the HPV-18 E1 and E2 gene product (E1: G489D, E2 : E43A and I77A).

HPV-18 degE1* sequence was reconstituted by assembling 50 oligonucleotides and cloned in a pBS vector giving rise to pTG17473. The E1 sequence was then fused to the signalling sequences clones from measles virus F protein (SS-18E1deg*-TMF) by a triple PCR using primers OTG15315 (SEQ ID NO: 36), OTG17881 (SEQ ID NO: 37), OTG17882 (SEQ ID NO: 38), OTG17883 (SEQ ID NO: 39), OTG17884 (SEQ ID NO: 40) and OTG17885 (SEQ ID NO: 41). The resulting fragment (SEQ ID NO: 11) was cloned in a MVA transfer vector under the control of p7.5K promoter, to generate pTG17521.

HPV-18 degE2* sequence was reconstituted by assembling 26 oligonucleotides and cloned in a pBS vector, giving rise to pTG17498. The fusion with the signal and the membrane-anchoring peptides of the glycoprotein of the rabies virus (ERA strain; Genbank n° M38452) was performed by triple PCR using primers OTG17875 (SEQ ID NO: 42), OTG17876 (SEQ ID NO: 43), OTG17877 (SEQ ID NO: 44), OTG17878 (SEQ ID NO: 45), OTG17879 (SEQ ID NO: 46) and OTG17880 (SEQ ID NO: 47). The SS-18E2*-TMR cassette was inserted in the MVA transfer plasmid downstream the pH5R promoter, giving rise to pTG17552. Finally, the p7.5K-SS-E1deg*-TMF cassette was isolated from pTG17521 and inserted in pTG17552, giving rise to pTG17582.

Generation of recombinant MVATG17521, MVATG17552 and MVATG17582 was performed as described above.

Example 3: Construction of a multivalent MVA vector expressing HPV-16 and HPV-18 E1 and E2 genes (MVATG17583)

The p7.5K-SS-18E1deg*-TMF cassette and the pH5R-SS-18E2*-TMR cassette were introduced in pTG17410 (containing the p7.5K-SS-16E1deg*-TMR cassette and the pH5R -SS-16E2*-TMR) and the resulting transfer plasmid was named pTG17583. Generation of MVATG17583 was performed as described above.

Example 4: Construction of a multivalent recombinant virus expressing HPV16 and HPV18 E6 and E7 genes

MVATG16327 is a recombinant MVA virus expressing membrane anchored and non-oncogenic variants of HPV-16 and HPV-18 E6 and E7 polypeptides. The E6 and E7 nucleotide sequences were mutated in order to eliminate their oncogenic properties (E6* and E7*) and were fused to sequences encoding appropriate signal and membrane anchoring peptides (E6*tm, E7*tm). More specifically, HPV-18 E7* was fused respectively at its N- and C-termini with the signal and membrane-anchoring peptides of the F glycoprotein of the measle virus whereas HPV-16 E6*, HPV-16 E7* and HPV-18 E6* with signal and membrane-anchoring peptides derived from those of the rabies virus glycoprotein. Moreover, HPV-18 E6 and E7 nucleotide sequences were further modified by codon usage modification so as to decrease homology with their HPV16 counterparts. For this purpose, sequences of native HPV16 and HPV18 genes were aligned and codon degeneration was performed to reduce homology to less than 5 consecutive nucleotides. In the vector, the HPV-16 and HPV-18 E6 sequences are both placed under the control of the p7.5K promoter in opposite orientation each other whereas the HPV-16 and HPV-18 E7 sequences are driven by the H5R promoter and all expression cassettes are inserted into the region of excision III of the MVA genome.

a) Construction of the HPV-16 E7*tm expression cassette

The HPV-16 E7 gene was isolated from Caski cells and modified so as to encode a non-oncogenic and membrane-presented E7 polypeptide (16E7*tmR) as described in WO99/03885. Non-oncogenic mutations were performed by deletion of amino acid

residues 21-26 (ΔDLYCYE) and membrane presentation by fusion of the E7* mutated sequence respectively at its 5' end 3' ends to sequences encoding the signal and membrane-anchoring peptides cloned from the rabies virus glycoprotein (ERA Strain; Genbank accession number M38452). The resulting sequence was cloned under the control of the early-late pH5R promoter (Rosel *et al*, 1986. J. Virol. 60, 436-9) and the cassette was introduced in a pBS derived vector, giving rise to pTG16161.

b) Cloning of HPV-16 E6*tm and HPV-18 E6*tm expression cassettes

The HPV-16 E6 gene was isolated and modified so as to encode a non-oncogenic and membrane-presented E6 polypeptide as described in WO99/03885. Non-oncogenic mutations were performed by deletion of amino acid residues 118-122 (ΔCPEEK) and membrane presentation by fusion of the E6*-mutated sequence respectively at its 5' end 3' ends to sequences encoding the signal and membrane-anchoring peptides derived from the rabies virus glycoprotein (PG strain; Genbank accession number ay009097). This was performed by inserting the E6*-mutated sequence in a vector containing the signal peptide and the membrane-anchoring peptide sequence separated by a BamHI site, leading to pTG16097.

A synthetic HPV-18 E6 sequence was generated by assembling oligonucleotides OTG15174 (SEQ ID NO: 48), OTG15175 (SEQ ID NO: 49), OTG15176 (SEQ ID NO: 50), OTG15177 (SEQ ID NO: 51), OTG15178 (SEQ ID NO: 52), OTG15179 (SEQ ID NO: 53), OTG15180 (SEQ ID NO: 54) and OTG15181 (SEQ ID NO: 55). The oligonucleotides were designed so as to introduce deletion of codons encoding amino acid residues 113-117 (non-oncogenic mutation ΔNPAEK) and to degenerate codon usage in order to reduce homology with the HPV-16 E6 gene (degenerated sequence). The resulting synthetic sequence was then fused respectively at its 5' and 3' end with the sequences encoding signal and membrane-anchoring peptides derived from the rabies virus glycoprotein gene, to provide the sequence shown in SEQ ID NO: 14, leading to pTG16160. The HPV-16 E6*tmR and HPV-18 degE6*tmR sequences were inserted in opposite orientation, each under the control of the p7.5K promoter. The cassettes were then introduced in pTG16161 to generate pTG16215.

c) Cloning of HPV-18 E7*tmF expression cassette

A synthetic HPV-18 E7 sequence was generated by assembling oligonucleotides OTG14773 (SEQ ID NO: 56), OTG14774 (SEQ ID NO: 57), OTG14775 (SEQ ID NO:

58), OTG14776 (SEQ ID NO: 59), OTG14777 (SEQ ID NO: 60) and OTG14778 (SEQ ID NO: 61). The oligonucleotides were designed so as to introduce deletion of codons encoding amino acid residues 24-28 (non-oncogenic mutation ΔDLLCH) and to degenerate codon usage in order to reduce homology with the HPV-16 E7 gene (degenerated sequence). The resulting synthetic sequence was then fused at its 5' and 3' ends respectively with the coding sequences for signal and membrane-anchoring peptides cloned from F protein gene of measles virus (described in EP 0305229). The resulting sequence (SEQ ID NO: 15) was cloned under the control of the pH5R promoter and the cassette was introduced in a pBS derived-vector to generate pTG16015.

d) Construction of transfer plasmid pTG16327

The transfer plasmid pTG6019 (described in Example 2 of WO99/03885) contains homologous sequences flanking MVA deletion III. It was modified as follow. A synthetic polylinker, obtained by hybridation of primers OTG15040 (SEQ ID NO: 62) and OTG15041 (SEQ ID NO: 63), was introduced in pTG6019 digested by *Bam*HI and *Sac*I, giving rise to pTG16007. A *Sac*I-*Sac*I fragment containing the expression cassette coding for *E.coli gpt* placed under the control of the early-late pH5R promoter was isolated from pTG14033 (described in Example 2 of EP 1 146 125) and introduced in pTG16007 digested by *Sac*I, giving rise to pTG16093. Synthesis of xanthine-guanine phosphoribosyltransferase enables GPT^+ recombinant MVA to form plaques in a selective medium containing mycophenolic acid, xanthine, and hypoxanthine (Falkner *et al*, 1988. J.Virol. 62, 1849-54). The selection marker *GPT* is placed between two homologous sequences in the same orientation. When the clonal selection is achieved, the selection marker is easily eliminated by several passages without selection allowing the growth of GPT^- recombinant MVA.

A *Hin*dIII-*Sma*I fragment containing the HPV-18 degE7*TMF expression cassette was isolated from pTG16015 and introduced in pTG16093 digested by the same enzymes, giving rise to pTG16105. On the other hand, pTG16215 was digested by *Sal*I and *Eco*RI, treated by T4 DNa polymerase, and the resulting fragment containing HPV-16 E7*tmR, HPV-16 E6*tmR and HPV-18 degE6*TMR expression cassettes was introduced in pTG16105 digested by *Sma*I, leading to pTG16327 (Figure 2).

e) Generation of MVATG16327

Generation of MVATG16327 was performed by homologous recombination in primary chicken embryos fibroblasts (CEF). For this purpose, pTG16327 was transfected according to the standard calcium phosphate DNA precipitation onto CEF previously infected with MVATGN33.1 at a MOI of 0.1 pfu/cell. Viral selection was performed by three round of plaque purification on CEF in the presence of a selective medium containing mycophenolic acid, xanthine and hypoxanthine. The selection marker was then eliminated by passage in non-selective medium. Absence of contamination by parental MVA was verified by PCR.

Analysis of gene expression was performed by Western-blot. CEF were infected at MOI 0.2 with MVATG16327 and after 24 hours, cells were harvested. Western-blot analysis was performed using rabbit polyclonal antibodies against HPV-16 and HPV-18 E6 and E7 proteins, respectively. The results show that all HPV polypeptides were correctly expressed from MVATG16327.

f) Study of genetic stability of MVATG16327

Five passages of MVATG16327 were done on CEF infected at an MOI of 10^{-2} pfu/cell and 10^{-4}pfu/cell. Genetic stability was evaluated on 100 viral clones isolated from the 5^{th} passage of the research stock. Two methods were used: PCR amplification to determine the structure of the expression cassettes, and antigens detection by Western Blot. Results of the PCR analysis showed that 99 % of the clones contained the expression cassettes of interest. Immuno-detection showed that 97% of the clones expressed the four antigens: HPV-16 and HPV-18 E6*tm and E7*tm polypeptides.

These analyses showed that 97% of clones derived from MVATG16327 were conformed after five passages, indicating a good genetic stability of this construct.

Example 5: Construction of a multivalent MVA vector expressing HPV-16, HPV-18, HPV-33 and HPV-52 E2 genes.

A synthetic gene encoding HPV-33 E2 polypeptide was synthetized by Geneart (Regensburg, Germany). The synthetic sequence was designed so as (i) to reduce the percentage of homology to less than 75 % with E2 genes from HPV-16, HPV-18 and HPV-52 (if possible homologous portions are reduced to less than 6 consecutive nucleotides) and (ii) to introduce the mutations abolishing the enzymatic functions of the HPV-33 gene product (E39A and I73A).

The HPV-33 degE2* sequence was then fused with nucleotide sequence encoding the signal and the membrane-anchoring peptides of the glycoprotein of the the rabies virus (ERA strain, Genbank n° M38452). This was performed by triple PCR using primers OTG18962 (SEQ ID NO: 72), OTG18963 (SEQ ID NO: 73), OTG18964 (SEQ ID NO: 74), OTG18965 (SEQ ID NO: 75), OTG18966 (SEQ ID NO: 76) and OTG18967 (SEQ ID NO: 77). The resulting fragment (SEQ ID NO: 67) encoding the SS-33degE2*-TMR polypeptide was cloned in a MVA transfer vector under the control of p7.5K promoter, and virus particles were generated as described above.

A synthetic gene encoding HPV-52 E2 polypeptide was synthetized by Geneart (Regensburg, Germany). The synthetic sequence was designed so as (i) to reduce the percentage of homology to less than 75 % with E2 genes from HPV-16, HPV-18 and HPV-33 (homologous portions are preferably reduced to less than 6 consecutive nucleotides) and (ii) to introduce the mutations abolishing the enzymatic functions of the HPV-52 gene product (E39A and I73A).

The synthetic HPV-52 E2*deg sequence was then fused with nucleotide sequences encoding the signal and the membrane-anchoring peptides of the measles virus F protein (giving SS-52E2*deg-TMF) by a triple PCR using primers OTG18968 (SEQ ID NO: 78), OTG18969 (SEQ ID NO: 79), OTG18970 (SEQ ID NO: 80), OTG18971 (SEQ ID NO: 81), OTG18972 (SEQ ID NO: 82) and OTG18973 (SEQ ID NO: 83).

The resulting fragment (SEQ ID NO: 69) encoding the SS-52E2*deg-TMF polypeptide was inserted in a MVA transfer plasmid downstream the p7.5K promoter, and virus particles were generated as described above.

The pH5R-SS-18E2*-TMR cassette encoding the membrane-presented and enzymatically defective HPV-18 E2 polypeptide (isolated from pTG17552), the p7.5K-SS-33degE2*-TMR cassette encoding the membrane-presented and enzymatically defective HPV-33 E2 polypeptide and the p7.5K-SS-52degE2*-TMF cassette encoding the membrane-presented and enzymatically defective HPV-52 E2 polypeptide were introduced in pTG17408 (containing the pH5R-SS-16E2*-TMR cassette), and virus particles were generated as described above.

Claims

1. A vector comprising at least a first nucleic acid molecule encoding a first polypeptide and a second nucleic acid molecule encoding a second polypeptide wherein:

   - said first and second nucleic acid molecules are obtained respectively from a first and second native nucleic acid sequences which exhibit a percentage of homology of approximately 80% or greater than 80% over a portion of 40 or more continuous nucleotides, and
   - said first nucleic acid molecule and/or said second nucleic acid molecule comprised in the vector is modified so as to reduce said percentage of homology to less than 75%; and

wherein said first nucleic acid molecule and said second nucleic acid molecule encode at least the same polypeptide obtained from closely related HPV serotypes.

2. The vector according to claim 1, wherein the codon usage pattern of the first nucleic acid molecule or the second nucleic acid molecule or both the first and second nucleic acid molecules is modified at least in said portion of 40 or more continuous nucleotides sharing 80% homology or more so as to reduce the percentage of identity to less than 75%.

3. The vector according to claim 2, wherein the codon usage pattern is modified at the nucleotide level and said modifications are silent at the amino acid level.

4. The vector according to claim 2 or 3, wherein the codon usage pattern is modified in such a way that homologous portions between the first and second nucleic acid molecules are restricted to less than 8 consecutive nucleotides.

5. The vector according to claim 4, wherein the codon usage pattern is modified in such a way that homologous portions between the first and second nucleic acid molecules are restricted to less than 5 consecutive nucleotides.

6. The vector according to any one of claims 1 to 5, wherein said vector is an adenoviral vector

7. The vector according to claim 6, wherein said adenoviral vector is replication-defective.

8. The vector according to any one of claims 1 to 5, wherein said vector is a poxviral vector.

9. The vector according to claim 8, wherein said poxviral vector is obtained from a vaccinia virus selected from the group consisting of the Copenhagen strain, the Wyeth strain, NYVAC and the highly attenuated modified Ankara (MVA) strain.

10. The vector according to any one of claims 1 to 9, wherein the first and second nucleic acid molecules are independently obtained from a high risk papillomavirus selected from the group consisting of HPV-16, HPV-18, HPV-30, HPV-31, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-56, HPV-58, HPV-59, HPV-66, HPV-68, HPV-70 and HPV-85.

11. The vector according to claim 10, wherein the first and second nucleic acid molecules encode an early papillomavirus polypeptide selected from the group consisting of E1, E2, E6 and E7.

12. The vector according to any one of claims 1 to 11, wherein said same polypeptide obtained from closely related organisms is an E2 polypeptide.

13. The vector according to claim 12, wherein closely related HPV serotypes are HPV-16, HPV-18, HPV-33 and HPV-52.

14. The vector according to claim 13, wherein said vector comprises a first nucleic acid molecule encoding an HPV-16 E2 polypeptide, a second nucleic acid

molecule encoding an HPV-18 E2 polypeptide, a third nucleic acid molecule encoding an HPV-33 E2 polypeptide, and a fourth nucleic acid molecule encoding an HPV-52 E2 polypeptide.

15. The vector according to claim 14, wherein said HPV-16 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 7, said HPV-18 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 8, said HPV-33 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 70 and said HPV-52 E2 polypeptide comprises the amino acid sequence shown in SEQ ID NO: 71.

16. The vector according to claim 14 or 15, wherein said first nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 12; said second nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 13; said third nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 67 and said fourth nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 69.

17. The vector according to any one of claims 1 to 11 and 13, wherein said same polypeptide obtained from closely related organisms is an E6 polypeptide, an E7 polypeptide or both E6 and E7 polypeptides.

18. The vector according to claim 17, wherein the first nucleic acid molecule encodes an HPV-16 E6 polypeptide and the second nucleic acid molecule encodes an HPV-18 E6 polypeptide, wherein the second nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 14.

19. The vector according to claim 17, wherein the first nucleic acid molecule encodes an HPV-16 E7 polypeptide and the second nucleic acid molecule encodes an HPV-18 E7 polypeptide, wherein the second nucleic acid molecule comprises the nucleotide sequence shown in SEQ ID NO: 15.

20. The vector according to claim 18 or 19, wherein said vector is a MVA vector, the first nucleic acid molecule is placed under the control of the vaccinia 7.5K promoter and the second nucleic acid molecule under the control of the vaccinia H5R promoter and the first and second nucleic acid molecules are both inserted in deletion III of said MVA vector.

21. The vector according to claim 17, wherein said vector comprises a first nucleic acid molecule encoding an HPV-16 E6 polypeptide, a second nucleic acid molecule encoding an HPV-18 E6 polypeptide, a third nucleic acid molecule encoding an HPV-16 E7 polypeptide and a fourth nucleic acid molecule encoding an HPV-18 E7 polypeptide wherein said first, second, third and fourth nucleic acid molecules do not comprise a portion of 40 or more continuous nucleotides exhibiting a percentage of homology of 75% or greater than 75%.

22. The vector of to claim 1, wherein at least one of said first and second nucleic acid molecules is selected from SEQ ID NO: 10, 11, 12, 13, 14, 15, 66, 67, 68 and 69.

23. A host cell comprising the vector according to any one of claims 1 to 22.

24. A pharmaceutical composition comprising a therapeutically effective amount of the vector according to any one of claims 1 to 22 or the host cell according to claim 23 and a pharmaceutically acceptable vehicle.

25. The pharmaceutical composition of claim 24, wherein said composition comprises one or more adjuvant(s) suitable for systemic or mucosal administration in humans.

26. The pharmaceutical composition of claim 25, wherein said adjuvant is an imidazoquinoline compound.

27. Use of the vector according to any one of claims 1 to 22, the host cell according to claim 23 or the composition according to anyone of claims 24 to 26, for the preparation of a drug intended for treating or preventing infectious diseases, cancers or immune deficiency diseases.

28. The use according to claim 27 for the preventive or curative treatment of a condition associated with infection by a papillomavirus.

29. The use according to claim 28, wherein the condition associated with infection by a papillomavirus is selected from a persistent infection, pre-malignant lesions and malignant lesions.

30. The use according to any one of claims 27 to 29, wherein said use is carried out according to a prime boost therapeutic modality and wherein said vector or composition is used to either prime or boost or both prime and boost an immune response in a subject.

Figure 1

Figure 1A

a)
…GCACGAGGACGAGGACAAGGAAAACGATGGTGATTCATTACCTACATTCAAGTGCGTATCTGGTCA
GAACACAAATACTTTG**TGA**

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b)
ATGGTGATTCATTACCTACATTCAAGTGCGTATCTGGTCAGAACACAAATACTTTGTGAAAATGATA
GT…

Percent identity: 100

Figure 1B

a)
…GCACGAGGACGAGGACAAGGAAAACGATGGTGATTCATTACCTACATTCAAGTGCGTATCTGGTCA
GAACACAAATACTTTG**TGA**
|||| || || || || || || || || || || || || || || || || || || || || || || || ||
b)
ATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTAACACATTATGAAAATGATA
GT…

Percent identity: 69.492

Figure 2

Figure 2A: E6 HPV16 (a) versus E6 HPV18 (b)

```
a  ATGCACCAAAAGAGAACTGCAATGTTTCAGGACCCACAGGAGCGACCCAG 50
                |  ||    ||| |||| |||     |||||||
b  ..............ATGGCGCGCTTTGAGGATCCAACACGGCGACCCTA 35
           .        .        .        .        .
a  AAAGTTACCACAGTTATGCACAGAGCTGCAAACAACTATACATGATATAA 100
    ||| ||||  |  | ||||| || ||| | ||  |  | || || |||
b  CAAGCTACCTGATCTGTGCACGGAACTGAACACTTCACTGCAAGACATAG 85
           .        .        .        .        .
a  TATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATAT 150
    | ||   ||||| || ||||||  |   ||    |     ||||||| |
b  AAATAACCTGTGTATATTGCAAGACAGTATTGGAACTTACAGAGGTATTT 135
           .        .        .        .        .
a  GACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATA 200
   || ||||| |||   |||||||   | || ||||||||  | |  ||  |
b  GAATTTGCATTTAAAGATTTATTTGTGGTGTATAGAGACAGTATACCCCA 185
           .        .        .        .        .
a  TGCTGTATGTGATAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATA 250
   |||||  |||  |||||||| || | |||||||||| |||||| ||  |  |
b  TGCTGCATGCCATAAATGTATAGATTTTTATTCTAGAATTAGAGAATTAA 235
           .        .        .        .        .
a  GACATTATTGTTATAGTTTGTATGGAACAACATTAGAACAGCAATACAAC 300
   |||||||||   |   | ||||||||   ||||| ||| | | |   |||
b  GACATTATTCAGACTCTGTGTATGGAGACACATTGGAAAAACTAACTAAC 285
           .        .        .        .        .
a  AAACCGTTGTGTGATTTGTTAATTAGGTGTATTAACTGTCAAAAGCCACT 350
   |    ||| |   |||| ||||| |||||  |    || || || ||  |
b  ACTGGGTTATACAATTTATTAATAAGGTGCCTGCGGTGCCAGAAACCGTT 335
           .        .        .        .        .
a  GTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAAAGATTCCATA 400
   |  ||| | |||||| |  ||||| ||  |  |||| | | |||| || |
b  GAATCCAGCAGAAAAACTTAGACACCTTAATGAAAAACGACGATTTCACA 385
           .        .        .        .        .
a  ATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCA 450
   | |||   || |  |  |  || |  ||    || || ||||    | ||
b  ACATAGCTGGGCACTATAGAGGCCAGTGCCATTCGTGCTGCAACCGAGCA 435
           .        .        .        .        .
a  AG..............AACACGTAGAGAAACCCAGCTGTAA 477
    |               |  ||| |||||||| ||  | |||
b  CGACAGGAACGACTCCAACGACGCAGAGAAACACAAGTATAA 477
```

Figure 2B: HPV16 E6 (a) versus SEQ ID n°14 (b)