

- 1 -

TITLE OF THE INVENTION**DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 6A****CROSS-RELATED TO OTHER APPLICATIONS**

5 This is a continuation of U.S. Serial Number 08/310,468
filed September 22, 1994, now pending.

FIELD OF THE INVENTION

10 The present invention is directed to DNA molecules
encoding purified human papillomavirus type 6a and derivatives thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1. HPV6a nucleotide sequence.

15 **BACKGROUND OF THE INVENTION**

Papillomavirus (PV) infections occur in a variety of animals, including humans, sheep, dogs, cats, rabbits, monkeys, snakes and cows. Papillomaviruses infect epithelial cells, generally inducing benign epithelial or fibroepithelial tumors at the site of infection. PV are species specific infective agents; a human papillomavirus do not infect a nonhuman animal.

Papillomaviruses may be classified into distinct groups based on the host that they infect. Human papillomaviruses (HPV) are further classified into more than 70 types based on DNA sequence homology. PV types appear to be type-specific immunogens in that a neutralizing immunity to infection by one type of papillomavirus does not confer immunity against another type of papillomavirus.

In humans, different HPV types cause distinct diseases. HPV types 1, 2, 3, 4, 7, 10 and 26-29 cause benign warts in both normal and immunocompromised individuals. HPV types 5, 8, 9, 12, 14, 15, 17, 19-25, 36 and 46-50 cause flat lesions in immunocompromised individuals. HPV types 6, 11, 34, 39, 41-44 and 51-55 cause benign condylomata of the genital or respiratory mucosa. HPV types 16 and 18 cause epithelial dysplasia of the genital mucosa and are associated with

- 2 -

the majority of in situ and invasive carcinomas of the cervix, vagina, vulva and anal canal.

Papillomaviruses are small (50-60 nm), nonenveloped, icosahedral DNA viruses that encode for up to eight early and two late genes. The open reading frames (ORFs) of the virus genomes are designated E1 to E7 and L1 and L2, where "E" denotes early and "L" denotes late. L1 and L2 code for virus capsid proteins. The early (E) genes are associated with functions such as viral replication and cellular transformation.

The L1 protein is the major capsid protein and has a molecular weight of 55-60 kDa. L2 protein is a minor capsid protein which has a predicted molecular weight of 55-60 kDa and an apparent molecular weight of 75-100 kDa as determined by polyacrylamide gel electrophoresis. Immunological data suggest that most of the L2 protein is internal to the L1 protein. The L1 ORF is highly conserved among different papillomaviruses. The L2 proteins are less conserved among different papillomaviruses.

The L1 and L2 genes have been identified as good targets for immunoprophylactics. Some of the early genes have also been demonstrated to be potential targets of vaccine development. Studies in the cottontail rabbit papillomavirus (CRPV) and bovine papillomavirus (BPV) systems have shown that immunizations with these proteins expressed in bacteria or by using vaccinia vectors protected animals from viral infection. Expression of papillomavirus L1 genes in baculovirus expression systems or using vaccinia vectors resulted in the assembly of virus-like particles (VLP) which have been used to induce high-titer virus-neutralizing antibody responses that correlate with protection from viral challenge.

HPV6 and HPV 11, which are only rarely associated with malignancies, are the causative agents of ~90% of condyloma acuminata, benign lesions of the respiratory and genital mucosa. HPV6 is detected three times more often in these lesions than HPV11.

The complete nucleotide sequence of HPV6b, the original HPV6 isolate, has been determined (Schwarz, E., et al. 1983. EMBO J.

- 3 -

2:2341-8.). Other HPV6 subtypes have been identified on the basis of restriction enzyme digest patterns (Gissmann, L., et al. 1983. Proc. Natl. Acad. Sci. USA 80:560-3; Mounts, P., et al. 1982. Proc. Natl. Acad. Sci. USA 79:5425-9).

5 Several groups have demonstrated that HPV6a is the predominant subtype found in condyloma acuminatum biopsies from patients in the U.S. and Europe. A recent report suggests that HPV6a is the HPV6 prototype (Kitasato, H., et al. 1994. J. Gen. Virol. 75:1157-1162). It is estimated that in the U.S. alone approximately one percent of
10 all men and women in the 15 to 49 year age group present to physicians with condyloma acuminatum. Unfortunately, there is no effective treatment for HPV-related disease. Therefore, a vaccine would be highly desirable. For the development of a prophylactic or therapeutic vaccine however, the sequence determination of late and early genes of the most
15 common HPV subtypes is of critical importance.

 The limited sequence information about HPV6a concerns the long control region (LCR) and the E6 and E7 ORFs. The instant application describes the cloning of HPV6a from a condyloma acuminatum biopsy, the determination of its complete viral DNA
20 sequence and the corresponding amino acid sequences of the major HPV6a open reading frames (ORFs).

 The present invention is directed to DNA molecules encoding purified human papillomavirus type 6a (HPV type 6a; HPV6a) and uses of the DNA molecules.

25

SUMMARY OF THE INVENTION

 The present invention is directed to DNA molecules encoding purified human papillomavirus type 6a (HPV type 6a; HPV6a) and derivatives thereof.

30

- 4 -

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to DNA molecules encoding purified human papillomavirus type 6a (HPV type 6a; HPV6a) and derivatives thereof. Such derivatives include but are not limited to peptides and proteins encoded by the DNA, antibodies to the DNA or antibodies to the proteins encoded by the DNA, vaccines comprising the DNA or vaccines comprising proteins encoded by the DNA, immunological compositions comprising the DNA or the proteins encoded by the DNA, kits containing the DNA or RNA derived from the DNA or proteins encoded by the DNA.

HPV6 is the principal causative agent of condyloma acuminata (benign lesions of the respiratory and genital mucosa). The complete nucleotide sequence of HPV6b, the original HPV6 isolate, has been determined. Other HPV6 subtypes have been identified on the basis of restriction enzyme digest patterns.

Several groups have demonstrated that HPV6a is the predominant subtype found in condyloma acuminatum biopsies from patients in the U.S. and Europe. It is estimated that in the U. S. alone approximately one percent of all men and women in the 15 to 49 year age group present to physicians with condyloma acuminatum. Unfortunately, there is no effective treatment for HPV-related disease. Therefore, a vaccine would be highly desirable. For the development of a prophylactic or therapeutic vaccine however, the sequence determination of late and early genes of the most common HPV subtypes is of critical importance.

The limited sequence information about HPV6a concerns the long control region (LCR) and the E6 and E7 ORFs. The instant application describes the cloning of HPV6a from a condyloma acuminatum biopsy, the determination of its complete viral DNA sequence and the corresponding amino acid sequences of the major HPV6a open reading frames (ORFs). The present invention is directed to DNA molecules encoding purified human papillomavirus type 6 and derivatives of the DNA molecules.

- 5 -

Pharmaceutically useful compositions comprising the DNA or proteins encoded by the DNA may be formulated according to known methods such as by the admixture of a pharmaceutically acceptable carrier. Examples of such carriers and methods of formulation may be found in Remington's Pharmaceutical Sciences. To form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of the protein or VLP. Such compositions may contain proteins or VLP derived from more than one type of HPV.

Therapeutic or diagnostic compositions of the invention are administered to an individual in amounts sufficient to treat or diagnose PV infections. The effective amount may vary according to a variety of factors such as the individual's condition, weight, sex and age. Other factors include the mode of administration. Generally, the compositions will be administered in dosages ranging from about 1 μ g to about 1 mg.

The pharmaceutical compositions may be provided to the individual by a variety of routes such as subcutaneous, topical, oral, mucosal, intravenous and intramuscular.

The vaccines of the invention comprise DNA, RNA or proteins encoded by the DNA that contain the antigenic determinants necessary to induce the formation of neutralizing antibodies in the host. Such vaccines are also safe enough to be administered without danger of clinical infection; do not have toxic side effects; can be administered by an effective route; are stable; and are compatible with vaccine carriers.

The vaccines may be administered by a variety of routes, such as orally, parenterally, subcutaneously, mucosally, intravenously or intramuscularly. The dosage administered may vary with the condition, sex, weight, and age of the individual; the route of administration; and the type PV of the vaccine. The vaccine may be used in dosage forms such as capsules, suspensions, elixirs, or liquid solutions. The vaccine may be formulated with an immunologically acceptable carrier.

The vaccines are administered in therapeutically effective amounts, that is, in amounts sufficient to generate a immunologically protective response. The therapeutically effective amount may vary

- 6 -

according to the type of PV. The vaccine may be administered in single or multiple doses.

The DNA and DNA derivatives of the present invention may be used in the formulation of immunogenic compositions. Such
5 compositions, when introduced into a suitable host, are capable of inducing an immune response in the host.

The DNA or its derivatives may be used to generate antibodies. The term "antibody" as used herein includes both polyclonal and monoclonal antibodies, as well as fragments thereof, such as, Fv, Fab
10 and F(ab)2 fragments that are capable of binding antigen or hapten.

The DNA and DNA derivatives of the present invention may be used to serotype HPV infection and HPV screening. The DNA, recombinant proteins, VLP and antibodies lend themselves to the formulation of kits suitable for the detection and serotyping of HPV.
15 Such a kit would comprise a compartmentalized carrier suitable to hold in close confinement at least one container. The carrier would further comprise reagents such as HPV6a DNA, recombinant HPV protein or VLP or anti-HPV antibodies suitable for detecting a variety of HPV types. The carrier may also contain means for detection such as labeled
20 antigen or enzyme substrates or the like.

The DNA and derived proteins therefrom are also useful as molecular weight and molecular size markers.

Because the genetic code is degenerate, more than one codon may be used to encode a particular amino acid, and therefore, the
25 amino acid sequence can be encoded by any of a set of similar DNA oligonucleotides. Only one member of the set will be identical to the HPV6a sequence but will be capable of hybridizing to HPV6a DNA even in the presence of DNA oligonucleotides with mismatches under appropriate conditions. Under alternate conditions, the mismatched DNA
30 oligonucleotides may still hybridize to the HPV6a DNA to permit identification and isolation of HPV6a encoding DNA.

The purified HPV6a DNA of the invention or fragments thereof may be used to isolate and purify homologues and fragments of HPV6a from other sources. To accomplish this, the first HPV6a DNA

- 7 -

may be mixed with a sample containing DNA encoding homologues of HPV6a under appropriate hybridization conditions. The hybridized DNA complex may be isolated and the DNA encoding the homologous DNA may be purified therefrom.

5 It is known that there is a substantial amount of redundancy in the various codons which code for specific amino acids. Therefore, this invention is also directed to those DNA sequences which contain alternative codons which code for the eventual translation of the identical amino acid. For purposes of this specification, a sequence bearing one or
10 more replaced codons will be defined as a degenerate variation. Also included within the scope of this invention are mutations either in the DNA sequence or the translated protein which do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for
15 glutamine may not cause a change in functionality of the polypeptide.

It is known that DNA sequences coding for a peptide may be altered so as to code for a peptide having properties that are different than those of the naturally-occurring peptide. Methods of altering the DNA sequences include, but are not limited to site-directed mutagenesis.

20 As used herein, a "functional derivative" of HPV6a is a compound that possesses a biological activity (either functional or structural) that is substantially similar to the biological activity of HPV6a. The term "functional derivatives" is intended to include the "fragments," "variants," "degenerate variants," "analogs" and
25 "homologues" or to "chemical derivatives" of HPV6a. The term "fragment" is meant to refer to any polypeptide subset of HPV6a. The term "variant" is meant to refer to a molecule substantially similar in structure and function to either the entire HPV6a molecule or to a fragment thereof. A molecule is "substantially similar" to HPV6a if both
30 molecules have substantially similar structures or if both molecules possess similar biological activity. Therefore, if the two molecules possess substantially similar activity, they are considered to be variants even if the structure of one of the molecules is not found in the other or

- 8 -

even if the two amino acid sequences are not identical. The term "functional derivative" does not include HPV6b.

The term "analog" refers to a molecule substantially similar in function to either the entire HPV6a molecule or to a fragment thereof.

5 A variety of procedures may be used to molecularly clone HPV6a DNA. These methods include, but are not limited to, direct functional expression of the HPV6a genes following the construction of a HPV6a-containing cDNA or genomic DNA library in an appropriate expression vector system. Another method is to screen HPV6a-
10 containing cDNA or genomic DNA library constructed in a bacteriophage or plasmid shuttle vector with a labeled oligonucleotide probe designed from the amino acid sequence of the HPV6a. An additional method consists of screening a HPV6a-containing cDNA or genomic DNA library constructed in a bacteriophage or plasmid shuttle
15 vector with a partial DNA encoding the HPV6a. This partial DNA is obtained by the specific polymerase chain reaction (PCR) amplification of HPV6a DNA fragments through the design of degenerate oligonucleotide primers from the amino acid sequence of purified HPV6a. Another method is to isolate RNA from HPV6a-producing cells
20 and translate the RNA into protein via an *in vitro* or an *in vivo* translation system. The translation of the RNA into a peptide or a protein will result in the production of at least a portion of HPV6a protein which can be identified by, for example, the activity of HPV6a protein or by immunological reactivity with an anti-HPV6a antibody. In this method,
25 pools of RNA isolated from HPV6a-producing cells can be analyzed for the presence of an RNA which encodes at least a portion of the HPV6a. Further fractionation of the RNA pool can be done to purify the HPV6a RNA from non-HPV6a RNA. The peptide or protein produced by this method may be analyzed to provide amino acid sequences which in turn
30 are used to provide primers for production of HPV6a cDNA, or the RNA used for translation can be analyzed to provide nucleotide sequences encoding HPV6a and produce probes for the screening of a HPV6a cDNA library. These methods are known in the art and can be found in, for example, Sambrook, J., Fritsch, E.F., Maniatis, T. in Molecular

- 9 -

Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989.

It is apparent that other types of libraries, as well as libraries constructed from other cells or cell types, may be useful for isolating HPV6a-encoding DNA. Other types of libraries include, but are not limited to, cDNA libraries derived from other cells or cell lines containing HPV type 6a and genomic DNA libraries.

Preparation of cDNA libraries can be performed by a variety of techniques. cDNA library construction techniques can be found for example, Sambrook, J., et al., *supra*. It is apparent that DNA encoding HPV6a may also be isolated from a suitable genomic DNA library. Construction of genomic DNA libraries can be performed by a variety of techniques. Genomic DNA library construction techniques can be found in Sambrook, J., et al. *supra*.

The cloned HPV6a DNA or fragments thereof obtained through the methods described herein may be recombinantly expressed by molecular cloning into an expression vector containing a suitable promoter and other appropriate transcription regulatory elements, and transferred into prokaryotic or eukaryotic host cells to produce recombinant HPV6a. Techniques for such manipulations are fully described in Sambrook, J., *et al.*, *supra*, and are known in the art.

Expression vectors are defined herein as DNA sequences that are required for the transcription of cloned copies of genes and the translation of their mRNAs in an appropriate host. Such vectors can be used to express eukaryotic genes in a variety of hosts such as bacteria, bluegreen algae, plant cells, insect cells, fungal cells and animal cells. Specifically designed vectors allow the shuttling of DNA between hosts such as bacteria-yeast or bacteria-animal cells or bacteria-fungal cells or bacteria-invertebrate cells. An appropriately constructed expression vector should contain: an origin of replication for autonomous replication in host cells, selectable markers, a limited number of useful restriction enzyme sites, a potential for high copy number, and active promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and initiate RNA synthesis. A strong promoter is one

- 10 -

which causes mRNAs to be initiated at high frequency. Expression vectors may include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

5 A variety of mammalian expression vectors may be used to express HPV6a DNA or fragments thereof in mammalian cells.

Commercially available mammalian expression vectors which may be suitable for recombinant HPV6a expression, include but are not limited to, pcDNA3 (Invitrogen), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593) pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and λ ZD35 (ATCC 37565).

A variety of bacterial expression vectors may be used to express HPV6a DNA or fragments thereof in bacterial cells.

15 Commercially available bacterial expression vectors which may be suitable include, but are not limited to pET11a (Novagen), lambda gt11 (Invitrogen), pcDNAII (Invitrogen), pKK223-3 (Pharmacia).

A variety of fungal cell expression vectors may be used to express HPV6a or fragments thereof in fungal cells. Commercially
20 available fungal cell expression vectors which may be suitable include but are not limited to pYES2 (Invitrogen) and *Pichia* expression vector (Invitrogen).

A variety of insect cell expression vectors may be used to express HPV6a DNA or fragments thereof in insect cells. Commercially
25 available insect cell expression vectors which may be suitable include but are not limited to pBlue Bac III (Invitrogen).

An expression vector containing DNA encoding HPV6a or fragments thereof may be used for expression of HPV6a proteins or fragments of HPV6a proteins in a cell, tissues, organs, or animals
30 (including humans). Host cells may be prokaryotic or eukaryotic, including but not limited to bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including but not limited to cell lines of human, bovine, porcine, monkey and rodent origin, and insect cells including but not limited to *Drosophila* and silkworm derived cell lines. Cell lines

- 11 -

derived from mammalian species which may be suitable and which are commercially available, include but are not limited to, L cells L-M(TK-) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), 293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26) and MRC-5 (ATCC CCL 171).

The expression vector may be introduced into host cells via any one of a number of techniques including but not limited to transformation, transfection, lipofection, protoplast fusion, and electroporation. The expression vector-containing cells are clonally propagated and individually analyzed to determine whether they produce HPV6a protein. Identification of HPV6a expressing host cell clones may be done by several means, including but not limited to immunological reactivity with anti-HPV6a antibodies, and the presence of host cell-associated HPV6a activity, such as HPV6a-specific ligand binding or signal transduction defined as a response mediated by the interaction of HPV6a-specific ligands at the HPV6a.

Expression of HPV DNA fragments may also be performed using *in vitro* produced synthetic mRNA or native mRNA. Synthetic mRNA or mRNA isolated from HPV6a producing cells can be efficiently translated in various cell-free systems, including but not limited to wheat germ extracts and reticulocyte extracts, as well as efficiently translated in cell based systems, including but not limited to microinjection into frog oocytes, with microinjection into frog oocytes being preferred.

Following expression of HPV6a protein(s) in a host cell, HPV6a protein may be recovered to provide HPV6a in purified form. Several HPV6a purification procedures are available and suitable for use. As described herein, recombinant HPV6a protein may be purified from cell lysates and extracts by various combinations of, or individual application of salt fractionation, ion exchange chromatography, size exclusion chromatography, hydroxylapatite adsorption chromatography and hydrophobic interaction chromatography.

- 12 -

In addition, recombinant HPV6a may be separated from other cellular proteins by use of an immunoaffinity column made with monoclonal or polyclonal antibodies specific for full length nascent HPV6a, or polypeptide fragments of HPV6a. Monoclonal and polyclonal antibodies may be prepared according to a variety of methods known in the art. Monoclonal or monospecific antibody as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for HPV6a. Homogenous binding as used herein refers to the ability of the antibody species to bind to a specific antigen or epitope.

It is apparent that the methods for producing monospecific antibodies may be utilized to produce antibodies specific for HPV6a polypeptide fragments, or full-length nascent HPV6a polypeptide. Specifically, it is apparent that monospecific antibodies may be generated which are specific for the fully functional HPV6a or fragments thereof.

The present invention is also directed to methods for screening for compounds which modulate the expression of DNA or RNA encoding HPV6a as well as the function(s) of HPV6a protein(s) *in vivo*. Compounds which modulate these activities may be DNA, RNA, peptides, proteins, or non-proteinaceous organic molecules. Compounds may modulate by increasing or attenuating the expression of DNA or RNA encoding HPV6a, or the function of HPV6a protein. Compounds that modulate the expression of DNA or RNA encoding HPV6a or the function of HPV6a protein may be detected by a variety of assays. The assay may be a simple "yes/no" assay to determine whether there is a change in expression or function. The assay may be made quantitative by comparing the expression or function of a test sample with the levels of expression or function in a standard sample.

Kits containing HPV6a DNA, fragments of HPV6a DNA, antibodies to HPV6a DNA or HPV6a protein, HPV6a RNA or HPV6a protein may be prepared. Such kits are used to detect DNA which hybridizes to HPV6a DNA or to detect the presence of HPV6a protein(s) or peptide fragments in a sample. Such characterization is useful for a

- 13 -

variety of purposes including but not limited to forensic analyses and epidemiological studies.

Nucleotide sequences that are complementary to the HPV6a encoding DNA sequence may be synthesized for antisense therapy.

- 5 These antisense molecules may be DNA, stable derivatives of DNA such as phosphorothioates or methylphosphonates, RNA, stable derivatives of RNA such as 2'-O-alkylRNA, or other HPV6a antisense oligonucleotide mimetics. HPV6a antisense molecules may be introduced into cells by microinjection, liposome encapsulation or by expression from vectors
10 harboring the antisense sequence. HPV6a antisense therapy may be particularly useful for the treatment of diseases where it is beneficial to reduce HPV6a activity.

- The term "chemical derivative" describes a molecule that contains additional chemical moieties which are not normally a part of
15 the base molecule. Such moieties may improve the solubility, half-life, absorption, etc. of the base molecule. Alternatively the moieties may attenuate undesirable side effects of the base molecule or decrease the toxicity of the base molecule. Examples of such moieties are described in a variety of texts, such as Remington's Pharmaceutical Sciences.

- 20 Compounds identified according to the methods disclosed herein may be used alone at appropriate dosages defined by routine testing in order to obtain optimal inhibition of the HPV6a or its activity while minimizing any potential toxicity. In addition, co-administration or sequential administration of other agents may be desirable.

- 25 Advantageously, compounds of the present invention may be administered in a single daily dose, or the total daily dosage may be administered in several divided doses. Furthermore, compounds for the present invention may be administered via a variety of routes including but not limited to intranasally, transdermally, by suppository, orally, and
30 the like

For combination treatment with more than one active agent, where the active agents are in separate dosage formulations, the active agents can be administered concurrently, or they each can be administered at separately staggered times.

- 14 -

The dosage regimen utilizing the compounds of the present invention is selected in accordance with a variety of factors including type, species, age, weight, sex and medical condition of the patient; the severity of the condition to be treated; the route of administration; the renal and hepatic function of the patient; and the particular compound thereof employed. A physician of ordinary skill can readily determine and prescribe the effective amount of the drug required to prevent, counter or arrest the progress of the condition. Optimal precision in achieving concentrations of drug within the range that yields efficacy without toxicity requires a regimen based on the kinetics of the drug's availability to target sites. This involves a consideration of the distribution, equilibrium, and elimination of a drug.

In the methods of the present invention, the compounds herein described in detail can form the active ingredient, and are typically administered in admixture with suitable pharmaceutical diluents, excipients or carriers (collectively referred to herein as "carrier" materials) suitably selected with respect to the intended form of administration, that is, oral tablets, capsules, elixirs, syrup, suppositories, gels and the like, and consistent with conventional pharmaceutical practices.

For instance, for oral administration in the form of a tablet or capsule, the active drug component can be combined with an oral, non-toxic pharmaceutically acceptable inert carrier such as ethanol, glycerol, water and the like. Moreover, when desired or necessary, suitable binders, lubricants, disintegrating agents and coloring agents can also be incorporated into the mixture. Suitable binders include without limitation, starch, gelatin, natural sugars such as glucose or beta-lactose, corn sweeteners, natural and synthetic gums such as acacia, tragacanth or sodium alginate, carboxymethylcellulose, polyethylene glycol, waxes and the like. Lubricants used in these dosage forms include, without limitation, sodium oleate, sodium stearate, magnesium stearate, sodium benzoate, sodium acetate, sodium chloride and the like. Disintegrators include, without limitation, starch, methyl cellulose, agar, bentonite, xanthan gum and the like.

- 15 -

For liquid forms the active drug component can be combined in suitably flavored suspending or dispersing agents such as the synthetic and natural gums, for example, tragacanth, acacia, methyl-cellulose and the like. Other dispersing agents which may be employed
5 include glycerin and the like. For parenteral administration, sterile suspensions and solutions are desired. Isotonic preparations which generally contain suitable preservatives are employed when intravenous administration is desired.

Topical preparations containing the active drug component
10 can be admixed with a variety of carrier materials well known in the art, such as, e.g., alcohols, aloe vera gel, allantoin, glycerine, vitamin A and E oils, mineral oil, PPG2 myristyl propionate, and the like, to form, e.g., alcoholic solutions, topical cleansers, cleansing creams, skin gels, skin lotions, and shampoos in cream or gel formulations.

15 The compounds of the present invention can also be administered in the form of liposome delivery systems, such as small unilamellar vesicles, large unilamellar vesicles and multilamellar vesicles. Liposomes can be formed from a variety of phospholipids, such as cholesterol, stearylamine or phosphatidylcholines.

20 Compounds of the present invention may also be delivered by the use of monoclonal antibodies as individual carriers to which the compound molecules are coupled. The compounds of the present invention may also be coupled with soluble polymers as targetable drug carriers. Such polymers can include polyvinyl-pyrrolidone, pyran
25 copolymer, polyhydroxypropylmethacryl-amidephenol, polyhydroxy-ethylaspartamidephenol, or polyethyl-eneoxidepolylysine substituted with palmitoyl residues. Furthermore, the compounds of the present invention may be coupled to a class of biodegradable polymers useful in achieving controlled release of a drug, for example, polylactic acid,
30 polyepsilon caprolactone, polyhydroxy butyric acid, polyorthoesters, polyacetals, polydihydro-pyrans, polycyanoacrylates and cross-linked or amphipathic block copolymers of hydrogels.

The following examples illustrate the present invention without, however, limiting the same thereto.

- 16 -

EXAMPLE 1

Extraction of nucleic acid from biopsy.

5 A large vulvar condyloma acuminatum lesion was obtained from a 25 year old, post partum female patient. A fragment of the lesion was frozen in liquid nitrogen, then processed with a Braun mikro-dismembrator II (B. Braun Instruments, Melsungen, Germany). The resulting material was solubilized with 0.6% (w/v) sodium dodecyl sulfate (SDS), treated with proteinase K (50 µg/ml), and extracted with
10 phenol/chloroform/isoamyl alcohol. DNA was ethanol-precipitated and quantified by UV spectrophotometry. The presence of high-molecular-weight DNA was established by agarose gel electrophoresis followed by staining with ethidium bromide.

EXAMPLE 2

Typing of HPV DNA.

15 The HPV DNA type was determined using the hybrid capture assay marketed as ViraType Plus (Digene Diagnostics, Beltsville, MD). The HPV probes used were divided into two pools whose
20 composition is based on the association of each type with genital tract malignancies. Probe group A contained the "low-risk" types HPV6, 11, 42, 43, and 44 while probe B contained the "high-risk" types 16, 18, 31, 33, 35, 45, 51, 52, and 56. Total DNA was digested with *Pst*I, *Bam*HI, and *Hind*III and Southern blots were performed under high stringency
25 conditions (T_m -15°C) to determine the HPV subtype.

EXAMPLE 3

Cloning of HPV6a genome.

30 Total DNA extracted from the HPV6a-positive biopsy sample was digested with *Hind*III endonuclease. Following size-fractionation through a 0.8% low-melting-temperature agarose preparative gel, a region corresponding to DNA of ~ 8 kilobase pairs (kbp) was excised from the gel and the agarose was digested with GelaseTM enzyme (Epicentre Technologies, Inc., Madison, WI). The

- 17 -

sample was ligated with pUC18 (Pharmacia, Inc., Piscataway, NJ) which had been digested with *Hind*III and dephosphorylated. Following transformation of competent *E. coli* DH5 cells (Gibco, BRL, Gaithersburg, MD), the plasmid library was screened for HPV6a-positive clones by colony-hybridization using an antisense ³²P-labeled oligonucleotide that was complementary to the 3'-end of the HPV6b L1 gene (5'-GAG AGA TCT TAC CTT TTA GTT TTG GCG CGC TTA C-3'; SEQ ID NO:1). A pUC18 plasmid containing the 8.1-kbp HPV6a genome was isolated and characterized by restriction enzyme and Southern blot analyses. This plasmid was designated pUC18-HPV6a. Plasmid DNA was prepared using the QiagenTM Plasmid Maxi kit (Qiagen Inc., Chatsworth, CA).

EXAMPLE 4

Sequence analysis of pUC18-HPV6a.

To determine the complete HPV6a sequence, sequencing primers were synthesized based on the published HPV6b sequence. Both strands of the complete 8.1-kbp HPV6a genome were sequenced by the dideoxy chain termination method using the PRISMTM kit and an Applied Biosystems (ABI) automated sequencer (#373A) according to the manufacturers' instructions (ABI, Inc., Foster City, CA). In cases where the sense and antisense sequence did not match, additional HPV6a specific primers were synthesized to resequence in both directions over the area in question to obtain a consensus.

The complete HPV6a sequence in comparison to the published HPV6b sequence is shown in Figure 1. Bases shown below the HPV6a sequence correspond to the HPV6b sequence. The DNA sequences of HPV6a and HPV6b exhibited over 97% identity with a total of 229 bp changes identified out of 8010 bp. The most significant differences compared to the HPV6b sequence were found in the LCR (nt 7205 - nt 106). Apart from several single nucleotide (nt) changes in the HPV6a LCR, a 94-bp insertion at nt 7350 and another 19-bp insertion at nt 7804 were found. At nt 7615, six base pairs were deleted from the HPV6a genome.

- 18 -

EXAMPLE 5

HPV6a sequence variation of ORFs compared to HPV6b.

5 Open reading frames were determined in the HPV6a
sequence and the major ORFs translated into amino acid sequences and
compared to the respective HPV6b sequences.

10 The major capsid protein L1 was the only ORF identical to
the HPV6b sequence. All other ORFs showed amino acid changes which
are summarized in Table 1. The minor capsid protein L2 showed five
amino acid changes; the E6 and E7 ORFs showed one amino acid change
each. In the E1 protein six amino acids and in the E2 protein 11 amino
acids were different. In the E4 protein four amino acid changes were
detected. The E5a ORF had changes in four positions, the ORF E5b in
seven positions.

15

- 19 -

Table 1. Sequence variation in the HPV6a ORFs E6, E7, E1, E2, E4, E5, and L2 compared to HPV6b

Open reading frame	Position (nucleotide)	Position (amino acid)	Amino acid change
E6	252	50	His→Gln
E7	792	88	Asp→Asn
E1	1535, 1536	235	Leu→Ala
	1670	280	Leu→Val
	1741	303	Glu→Asp
	2208	459	Thr→Ser
	2557	575	Asp→Glu
	2654	608	Thr→Ala
E2	2802	27	His→Asp
	2974	94	Arg→Lys
	3148	142	Asn→Thr
	3153	144	Thr→Ser
	3272	193	His→Gln
	3388	222	Leu→Pro
	3405	227	Lys→Gln
	3643	307	Arg→Lys
	3693	324	Ser→Pro
	3735	338	Asp→His
	3765	348	Asp→Asn
	3794	357	Ser→Arg
E4	3272	6	Ile→Asn
	3388	60	Gly→Glu
	3461	69	Pro→His
	3552	99	Asp→Glu

- 20 -

E5a	3935	16	Phe→Leu
	4004	40	Glu→Asp
	4137	84	Tyr→His
	4150	88	Thr→Asn
E5b	4235	25	Met→Val
	4297	45	Lys→Asn
	4314	51	Asn→Thr
	4323	54	Asp→Ala
	4343	61	Tyr→His
	4346,4347	62	Thr→Asp
	4353	64	Asp→Ala
L2	4646,4647	75	Gln→Gly
	4976	185	Val→Ile
	5021	200	Val→Ile
	5490	356	Gly→Asp
	5597	392	Leu→Ile

EXAMPLE 6**Sub-cloning of the HPV6a cDNA into expression vectors**

- 5 The cDNA encoding HPV6a is sub-cloned into several vectors for expression of the HPV6a protein in transfected host cells and for *in vitro* transcription/translation. These vectors include pBluescript II SK+ (where expression is driven by T7 or T3 promoters) pcDNA I/Amp (where expression is driven by the cytomegalovirus (CMV) promoter),
- 10 pSZ9016-1 (where expression is driven by the HIV long terminal repeat (LTR) promoter) and the baculovirus transfer vector pVL1393 (where expression is driven by the polyhedrin (PH) promoter) for producing recombinant baculovirus containing the HPV6a encoding DNA sequence.

- a) pBluescript II SK+:HPV6a. The full length HPV6a
- 15 cDNA clone is retrieved from lambda bacteriophage by limited *Eco* RI

- 21 -

digestion and ligated into *Eco* RI-cut, CIP-treated pBluescript II SK+. Separate subclones are recovered in which the sense orientation of HPV6a followed either the T7 or T3 promoters.

b)pcDNA I/Amp:HPV6a. To facilitate directional cloning,
5 HPV6a is excised from a purified plasmid preparation of pBluescript II SK+:HPV6a in which the HPV6a DNA sequence is downstream of the T7 promoter using *Eco* RV and *Xba* I. The resulting *Eco* RV, *Xba* I HPV6a fragment is purified and ligated into *Eco* RV-cut, *Xba* I-cut, CIP-treated pcDNA I/Amp such that the HPV6a encoding DNA is
10 downstream of the CMV promoter.

c)pSZ9016-1:HPV6a. HPV6a is excised from pBluescript II SK+:HPV6a by limited *Eco* RI digestion and subsequent purification of the 1.3 Kb fragment from agarose gels. The resulting *Eco* RI HPV6a fragment is ligated into *Eco* RI-cut, CIP-treated pSZ9016-1. Subclones
15 are selected in which the sense orientation of HPV6a is downstream of the HIV LTR promoter.

d)pVL1393:HPV6a and pVL1393:T7 HPV6a HA
Directional cloning of the HPV6a encoding DNA into the baculovirus transfer vector pVL1393 is mediated by excising HPV6a from pcDNA
20 I/Amp:HPV6a with *Bam* HI and *Xba* I then ligating the resulting 1.3 Kb fragment into *Bam* HI-cut, *Xba* I-cut, CIP-treated pVL1393 producing pVL1393:HPV6a. Similarly, HPV6a is epitope tagged by engineering a T7 tag at the 5' amino terminus of the HPV6a open reading frame and a FluHA epitope at the 3' carboxy terminus. The HPV6a DNA modified in
25 this manner is ligated into the *Bam* HI/*Xba* I sites of pVL1393 to produce pVL1393:T7 HPV6a HA.

- 22 -

EXAMPLE 7

Expression Of The HPV6a Polypeptide By *In Vitro* Transcription/ Translation And By Transfection Into Host Cells

Vectors containing HPV DNA sequences are used to drive
5 the translation of the HPV6a polypeptide in rabbit reticulocyte lysates,
mammalian host cells, and in baculovirus infected insect cells. The
experimental procedures are essentially those outlined in the
manufacturers' instructions.

- a) *In vitro* Transcription/Translation. pBluescript III
10 SK+:HPV6a plasmid DNA (with HPV6a in the T7 orientation) is
linearized by *Bam* HI digestion downstream of the HPV6a insert. The
linearized plasmid is purified and used as a template for run-off
transcription using T7 RNA polymerase in the presence of
m7G(5')ppp(5')G. The resulting capped HPV6a transcripts are purified
15 by LiCl precipitation and used to drive the translation of HPV6a in
nuclease-pretreated rabbit reticulocyte lysate in the presence of L-
[³⁵S]methionine.

- b) Expression in Mammalian Cells. The HPV6a protein is
expressed in mammalian host cells following transfection with either
20 pcDNA I/Amp:HPV6a (under control of the CMV promoter) or
pSZ9016-1:HPV6a (under control of the HIV LTR promoter). In the
latter case (pSZ9016-1:HPV6a), cells are co-transfected with the TAT
expressing plasmid pSZ90161:TAT. For both HPV6a expression
plasmids, COS-7 cells are transfected using either DEAE-dextran or
25 lipofection with Lipofectamine (BRL).

- c) Expression in Insect Cells. The HPV6a - containing
baculovirus transfer vector pVL1393:T7 HPV6a HA is used to produce
recombinant baculovirus (*Autographa californica*) by *in vivo*
homologous recombination. Epitope tagged HPV6a is then expressed in
30 Sf9 (*Spodoptera frugiperda*) insect cells grown in suspension culture
following infection with the HPV6a - containing recombinant
baculovirus.

- 23 -

EXAMPLE 8

Compounds that affect HPV6a activity may be detected by a variety of methods. A method of identifying compounds that affect HPV6a comprises:

- 5 (a) mixing a test compound with a solution containing HPV6a to form a mixture;
- (b) measuring HPV6a activity in the mixture; and
- (c) comparing the HPV6a in the mixture to a standard.

- 10 Compounds that affect HPV6a activity may be formulated into pharmaceutical compositions. Such pharmaceutical compositions may be useful for treating diseases or conditions that are characterized by HPV6a infection.

EXAMPLE 9

- 15 DNA which is structurally related to DNA encoding HPV6a is detected with a probe. A suitable probe may be derived from DNA having all or a portion of the nucleotide sequence of Figure 1, RNA encoded by DNA having all or a portion of the nucleotide sequence of Figure 1 or degenerate oligonucleotides derived from a portion of the
- 20 sequence of Figure 1.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: JANSEN, KATHRIN U.
HOFMANN, KATHRYN J.
- (ii) TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMA VIRUS TYPE 6A
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: CHRISTINE E. CARTY
 - (B) STREET: 126 E. LINCOLN AVENUE, P.O. BOX 2000
 - (C) CITY: RAHWAY
 - (D) STATE: NJ
 - (E) COUNTRY: USA
 - (F) ZIP: 07065-0907
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/310,468
 - (B) FILING DATE: 22-SEP-1994
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: CARTY, CHRISTINE E.
 - (B) REGISTRATION NUMBER: 36,099
 - (C) REFERENCE/DOCKET NUMBER: 19307 PCT
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (908) 594-6734
 - (B) TELEFAX: (908) 594-4720

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAGAGATCTT ACCTTTTAGT TTTGGCGCGC TTAC

34

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8010 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTTAATAACA ATCTTGTTTT TAAAAAATAG GAGGGACCGA AAACGGTTCA ACCGAAAACG	60
GTTGTATATA AACCAGCCCT AAAATTTAGC AAACGAGGCA TTATGGAAAG TGCAAATGCC	120
TCCACGTCTG CAACGACCAT AGACCAGTTG TGCAAGACGT TTAATCTATC TATGCATACG	180
TTGCAAATTA ATTGTGTGTT TTGCAAGAAT GCACTGACCA CTGCAGAGAT TTATTCATAT	240
GCATATAAAC AGCTAAAGGT CCTGTTTCGA GCGGCTATC CATATGCAGC CTGCGCGTGC	300
TGCCTAGAAT TTCATGGAAA AATCAACCAA TATAGACACT TTGATTATGC TGGATATGCA	360
ACAACGTGTG AAGAAGAAAC TAAACAAGAC ATTTTAGACG TGCTAATTCG GTGCTACCTG	420
TGTCACAAAC CGCTGTGTGA AGTAGAAAAG GTAAACATA TACTAACCAA GGCGCGGTTT	480
ATAAAGCTAA ATTGTACGTG GAAGGGTCGC TGCCTACACT GCTGGACAAC ATGCATGGAA	540
GACATGTTAC CCTAAAGGAT ATTGTATTAG ACCTGCAACC TCCAGACCCT GTAGGGTTAC	600
ATTGCTATGA GCAATTAGTA GACAGCTCAG AAGATGAGGT GGACGAAGTG GACGGACAAG	660
ATTCACAACC TTAAAAACAA CATTTCCAAA TAGTGACCTG TTGCTGTGGA TGTGACAGCA	720
ACGTTTCACT GGTGTGTCAG TGTACAGAAA CAGACATCAG AGAAGTGCAA CAGCTTCTGT	780
TGGGAACACT AGACATAGTG TGTCCCATCT GCGCACCGAA GACATAACAA CGATGGCGGA	840
CGATTTCAGT ACAGAAAATG AGGGGTCTGG GTGTACAGGA TGGTTTATGG TAGAAGCTAT	900
AGTGCAACAC CCAACAGGTA CACAAATATC AGACGATGAG GATGAGGAGG TGGAGGACAG	960
TGGGTATGAC ATGGTGGACT TTATTGATGA CAGCAATATT ACACACAATT CCTTGAAGC	1020
ACAGGCATTG TTAAACAGGC AGGAGGCGGA CACCCATTAT GCGACTGTGC AGGACCTAAA	1080
ACGAAAGTAT TTAGGTAGTC CATATGTTAG TCCTATAAAC ACTATAGCCG AGGCAGTGGA	1140
AAGTGAAATA AGTCCACGAT TGGACGCCAT TAACTTACA AGACAGCCAA AAAAGGTAAA	1200

GCGACGGCTG	TTTCAAACCA	GGGAACTAAC	GGACAGTGGA	TATGGCTATT	CTGAAGTGGA	1260
AGCTGGAACG	GGAACGCAGG	TAGAGAAACA	TGGCGTCCCG	GAAAATGGGG	GAGATGGTCA	1320
GGAAAAGGAC	ACAGGAAGGG	ACATAGAGGG	GGAGGAACAT	ACAGAGGCGG	AAGCGCCCAC	1380
AAACAGTGTA	CGGGAGCATG	CAGGCACAGC	AGGAATATTG	GAATTGCTAA	AATGTAAAGA	1440
TTTACGGGCA	GCATTACTTG	GTAAGTTTAA	AGAATGCTTT	GGGCTGTCTT	TTATTGATTT	1500
AATTAGGCCA	TTTAAAAGTG	ATAAAACAAC	ATGTGCAGAC	TGGGTGGTAG	CAGGATTTGG	1560
TATACATCAT	AGCATATCAG	AGGCATTTCA	AAAATTAATT	GAGCCATTAA	GTTTATATGC	1620
ACATATACAA	TGGCTAACAA	ATGCATGGGG	AATGGTATTG	TTAGTATTAG	TAAGATTTAA	1680
AGTAAATAAA	AGTAGAAGTA	CCGTTGCACG	TACACTTGCA	ACGCTATTAA	ATATACCTGA	1740
CAATCAAATG	TTAATAGAGC	CACCAAAAAT	ACAAAGTGGT	GTTGCAGCCC	TGTATTGGTT	1800
TCGTACAGGT	ATATCAAATG	CCAGTACAGT	TATAGGGGAA	GCACCAGAAT	GGATAACACG	1860
CCAAACTGTT	ATTGAACATG	GGTTGGCAGA	CAGTCAGTTT	AAATTAACAG	AAATGGTGCA	1920
GTGGGCATAT	GATAATGACA	TATGCGAGGA	GAGTGAAATT	GCATTTGAAT	ATGCACAAAG	1980
GGGAGATTTT	GATTCTAATG	CACGAGCATT	TTTAAATAGC	AATATGCAGG	CAAAATATGT	2040
GAAAGATTGT	GCAACTATGT	GTAGACATTA	TAAACATGCA	GAAATGAGGA	AGATGTCTAT	2100
AAAACAATGG	ATAAAACATA	GGGGTTCTAA	AATAGAAGGC	ACAGGAAATT	GGAAACCAAT	2160
TGTACAATTC	CTACGACATC	AAAATATAGA	ATTTATTCCA	TTTTTAAGTA	AATTTAAATT	2220
ATGGCTGCAC	GGTACGCCAA	AAAAAACTG	CATAGCCATA	GTAGGCCCTC	CAGATACTGG	2280
GAAATCGTAC	TTTTGTATGA	GTTTAATAAG	CTTTTATAGGA	GGTACAGTTA	TTAGTCATGT	2340
AAATTCCAGC	AGCCATTTTT	GGTTGCAACC	GTTAGTAGAT	GCTAAGGTAG	CATTGTTAGA	2400
TGATGCAACA	CAGCCATGTT	GGATATATAT	GGATACATAT	ATGAGAAATT	TGTTAGATGG	2460
TAATCCTATG	AGTATTGACA	GAAAGCATAA	AGCATTGACA	TTAATTAAAT	GTCCACCTCT	2520
GCTAGTAACG	TCCAACATAG	ATATTACTAA	AGAAGAGAAA	TATAAGTATT	TACATACTAG	2580
AGTAACAACA	TTTACATTTT	CAAATCCATT	CCCTTTTGAC	AGAAATGGGA	ATGCAGTGTA	2640
TGAACTGTCA	AATGCAAACT	GGAAATGTTT	TTTTGAAAGA	CTGTCGTCAA	GCCTAGACAT	2700
TCAGGATTCA	GAGGACGAGG	AAGATGGAAG	CAATAGCCAA	GCGTTTAGAT	GCGTGCCAGG	2760
AACAGTTGTT	AGAACTTTAT	GAAGAAAACA	GTA CTGACCT	AAACAAACAT	GTATTGCATT	2820
GGAAATGCAT	GAGACATGAA	AGTGTATTAT	TATATAAAGC	AAAACAAATG	GGCCTAAGCC	2880

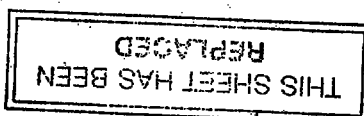
ACATAGGAAT	GCAAGTAGTG	CCACCATTAA	AGGTGTCCGA	AGCAAAAGGA	CATAATGCCA	2940
TTGAAATGCA	AATGCATTTA	GAATCATTAT	TAAAGACTGA	GTATAGTATG	GAACCGTGGA	3000
CATTACAAGA	AACAAGTTAT	GAAATGTGGC	AAACACCACC	TAAACGCTGT	TTTAAAAAAC	3060
GGGGCAAAAC	TGTAGAAGTT	AAATTTGATG	GCTGTGCAAA	CAATACAATG	GATTATGTGG	3120
TATGGACAGA	TGTGTATGTG	CAGGACACTG	ACTCCTGGGT	AAAGGTGCAT	AGTATGGTAG	3180
ATGCTAAGGG	TATATATTAC	ACATGTGGAC	AATTTAAAC	ATATTATGTA	AACTTTGTAA	3240
AAGAGGCAGA	AAAGTATGGG	AGCACCAAAC	AATGGGAAGT	ATGTTATGGC	AGCACAGTTA	3300
TATGTTCTCC	TGCATCTGTA	TCTAGCACTA	CACAAGAAGT	ATCCATTCTT	GAATCTACTA	3360
CATACACCCC	CGCACAGACC	TCCACCCCTG	TGTCCTCAAG	CACCCAGGAA	GACGCAGTGC	3420
AAACGCCGCC	TAGAAAACGA	GCACGAGGAG	TCCAACAGTC	ACCTTGCAAC	GCCTTGTGTG	3480
TGGCCCATAT	TGGACCCGTG	GACAGTGGAA	ACCACAACCT	CATCACTAAC	AATCACGACC	3540
AGCACCAAAG	AAGGAACAAC	AGTAACAGTT	CAGCTACGCC	TATAGTGCAA	TTTCAAGGTG	3600
AATCTAATTG	TTTAAAGTGT	TTTAGATATA	GGCTAAATGA	CAAACACAGA	CATTTATTTG	3660
ATTTAATATC	ATCAACGTGG	CACTGGGCCT	CCCCAAAGGC	ACCACATAAA	CATGCCATTG	3720
TAACTGTAAC	ATATCATAGT	GAGGAACAAA	GGCAACAGTT	TTTAAATGTT	GTAAAAATAC	3780
CACCTACTAT	TAGGCACAAA	CTGGGGTTTA	TGTCACTGCA	CCTATTGTAA	TTTGTATATA	3840
TGTAAATGTG	TAAATATATG	GTATTGGTGT	AATACAACCT	TACATGTATG	GAAGTGGTAC	3900
CTGTACAAAT	AGCTGCAGGA	ACAACCAGCA	CATTAATACT	GCCTGTTATA	ATTGCATTTG	3960
TTGTATGTTT	TGTTAGCATC	ATACTTATTG	TATGGATATC	TGACTTTATT	GTGTACACAT	4020
CTGTGCTAGT	ACTAACACTG	CTTTTATACT	TACTATTGTG	GCTGCTATTA	ACAACCCCTT	4080
TGCAATTTTT	CCTACTAACT	CTACTTGTGT	GTTACTGTCC	CGCATTGTAT	ATACACCACT	4140
ACATTGTTAA	CACACAGCAA	TGATGCTAAC	ATGTCAATTT	AATGATGGAG	ATACATGGCT	4200
GGGTTTGTGG	TTGTTATGTG	CCTTTATTGT	AGGGGTGTTG	GGGTTATTAT	TAATGCACTA	4260
TAGAGCTGTA	CAAGGCGATA	AACACACCAA	ATGTAACAAG	TGTAACAAAC	ACACCTGTAA	4320
TGCTGATTAT	GTAACATATG	ATCATGATAC	TGCTGGTGAT	TATATATATA	TGAATTAGAG	4380
TAAAACTTTT	TTTATATTTG	TAACAGTGTA	TGTTTTGTAT	ACCATGGCAC	ATAGTAGGGC	4440
CCGACGACGC	AAGCGTGCCT	CAGCTACACA	GCTATATCAA	ACATGTAAAC	TTACTGGAAC	4500
ATGCCCCCCA	GATGTAATTC	CTAAGGTGGA	GCACAACACC	ATTGCAGATC	AAATATTAAA	4560

ATGGGGAAGT TTGGGGGTTT TTTTGGAGG GTTGGGTATA GGCACCGGTT CCGGCACTGG 4620
GGGTCGTA CT GGCTATGTTT CCTTAGGAAC TTCTGCAAAA CCTTCTATTA CTAGTGGGCC 4680
TATGGCTCGT CCTCCTGTGG TGGTGGAGCC TGTGGCCCCT TCGGATCCAT CCATTGTGTC 4740
TTTAATTGAA GAATCAGCAA TCATTACGC AGGGGCGCCT GAAATTGTGC CCCCTGCACA 4800
CGGTGGGTTT ACAATTACAT CCTCTGAAAC AACTACCCCT GCAATATTGG ATGTATCAGT 4860
TACTAGTCAT ACTACTACTA GTATATTTAG AAATCCTGTC TTTACAGAAC CTTCTGTAAC 4920
ACAACCCCAA CCACCCGTGG AGGCTAATGG ACATATATTA ATTTCTGCAC CCACTATAAC 4980
GTCACACCCT ATAGAGGAAA TTCCTTTAGA TACTTTTGTG ATATCCTCTA GTGATAGCGG 5040
TCCTACATCC AGTACCCCTG TTCCTGGTAC TGCACCTAGG CCTCGTGTGG GCCTATATAG 5100
TCGTGCATTG CACCAGGTGC AGGTTACAGA CCCTGCATTT CTTTCCACTC CTCAACGCTT 5160
AATTACATAT GATAACCCTG TATATGAAGG GGAGGATGTT AGTGTA CAAT TTAGTCATGA 5220
TTCTATACAC AATGCACCTG ATGAGGCTTT TATGGACATA ATTCGTTTGC ACAGACCTGC 5280
TATTGCGTCC CGACGTGGCC TTGTGCGGTA CAGTCGCATT GGACAACGGG GGTCTATGCA 5340
CACTCGCAGC GGAAAGCACA TAGGGGCCCCG CATTCAATAT TTTTATGATA TTTCACCTAT 5400
TGCACAAGCT GCAGAAGAAA TAGAAATGCA CCCTCTTGTG GCTGCACAGG ATGATACATT 5460
TGATATTTAT GCTGAATCTT TTGAACCTGA CATTAAACCT ACCCAACACC CTGTTACAAA 5520
TATATCAGAT ACATATTTAA CTTCCACACC TAATACAGTT ACACAACCGT GGGGTAACAC 5580
CACAGTTCCA TTGTCAATTC CTAATGACCT GTTTTTACAG TCTGGCCCTG ATATAACTTT 5640
TCCTACTGCA CCTATGGGAA CACCCTTTAG TCCTGTAACT CCTGCTTTAC CTACAGGCCC 5700
TGTTTTTATT ACAGGTTCTG GATTTTATTT GCATCCTGCA TGGTATTTTG CACGTAAACG 5760
CCGTAAACGT ATTCCCTTAT TTTTTTCAGA TGTGGCGGCC TAGCGACAGC ACAGTATATG 5820
TGCCTCCTCC TAACCCTGTA TCCAAAGTTG TTGCCACGGA TGCTTATGTT ACTCGCACCA 5880
ACATATTTTA TCATGCCAGC AGTTCTAGAC TTCTTGCACT GGGTCATCCT TATTTTTCCA 5940
TAAACGGGC TAACAAACT GTTGTGCCAA AGGTGTCAGG ATATCAATAC AGAGTATTTA 6000
AGGTGGTGTT ACCAGATCCT AACAAATTTG CATTGCCTGA CTCGTCTCTT TTTGATCCCA 6060
CAACACAACG TTTGGTATGG GCATGCACAG GCCTAGAGGT GGGCCGGGGA CAGCCATTAG 6120
GTGTGGGTGT AAGTGGACAT CCTTTCCTAA ATAAATATGA TGATGTTGAA AATTCAGGGA 6180
GTGGTGGTAA CCCTGGACAG GATAACAGGG TTAATGTAGG TATGGATTAT AAACAAACAC 6240

AATTATGCAT	GGTTGGATGT	GCCCCCCTT	TGGGCGAGCA	TTGGGGTAAA	GGTAAACAGT	6300
GTACTAATAC	ACCTGTACAG	GCTGGTGACT	GCCCGCCCTT	AGAACTTATT	ACCAGTGTTA	6360
TACAGGATGG	CGATATGGTT	GACACAGGCT	TTGGTGCTAT	GAATTTTGCT	GATTTGCAGA	6420
CCAATAAATC	AGATGTTCCCT	ATTGACATAT	GTGGCACTAC	ATGTAAATAT	CCAGATTATT	6480
TACAAATGGC	TGCAGACCCA	TATGGTGATA	GATTATTTTT	TTTTCTACGG	AAGGAACAAA	6540
TGTTTGCCAG	ACATTTTTTT	AACAGGGCTG	GCGAGGTGGG	GGAACCTGTG	CCTGATACTC	6600
TTATAATTAA	GGGTAGTGGA	AATCGCACGT	CTGTAGGGAG	TAGTATATAT	GTTAACACCC	6660
CAAGCGGCTC	TTGGTGTC	TCTGAGGCAC	AATTGTTTAA	TAAGCCATAT	TGGCTACAAA	6720
AAGCCCAGGG	ACATAACAAT	GGTATTTGTT	GGGGTAATCA	ACTGTTTGTT	ACTGTGGTAG	6780
ATACCACACG	CAGTACCAAC	ATGACATTAT	GTGCATCCGT	AACTACATCT	TCCACATACA	6840
CCAATTCTGA	TTATAAAGAG	TACATGCGTC	ATGTGGAAGA	GTATGATTTA	CAATTTATTT	6900
TTCAATTATG	TAGCATTACA	TTGTCTGCTG	AAGTAATGGC	CTATATTCAC	ACAATGAATC	6960
CCTCTGTTTT	GGAAGACTGG	AACTTTGGGT	TATCGCCTCC	CCCAAATGGT	ACATTAGAAG	7020
ATACCTATAG	GTATGTGCAG	TCACAGGCCA	TTACCTGTCA	AAAGCCCACT	CCTGAAAAGG	7080
AAAAGCCAGA	TCCCTATAAG	AACCTTAGTT	TTTGGGAGGT	TAATTTAAAA	GAAAAGTTTT	7140
CTAGTGAATT	GGATCAGTAT	CCTTTGGGAC	GCAAGTTTTT	GTTACAAAGT	GGATATAGGG	7200
GACGGTCCTC	TATTCGTACC	GGTGTTAAGC	GCCCTGCTGT	TTCCAAAGCC	TCTGCTGCCC	7260
CTAAACGTAA	GCGCGCCAAA	ACTAAAAGGT	AATATATGTG	TATATGTACT	GTTATATATA	7320
TGTGTGTATG	TACTGTTATG	TATATGTGTG	TATGTACTGT	TATATGTATG	TGTGTTGTAT	7380
ATATGTGTGT	ATATATGTGT	ATGTGTGTAT	ATGTATATGT	ATGTGTTGTG	TATATATATG	7440
TGTGTGTGTG	TTATGTGTGT	AATGTAATTT	ATTTGTGTAA	TGTGTATGTG	TGTTTATGTG	7500
CAATAAACAA	TAACTACAC	CCTGTGACTC	AGTGGCTGTT	GCACGCGTTT	TGGTTTGCAC	7560
GCGCCTTACA	CACATAAGTA	ATATACATGC	ACAATATATA	TATTTTTGTT	ACAATAATAT	7620
ATTTTTATAT	TTGCAACCGT	TTTCGGTTGC	CCTTGGCATA	CACTTTCCAC	CAATTTGTTA	7680
CAACGTGTTG	CCTGTTAATC	CTATATATTT	TGTGCCAGGT	ACACATTGCC	CTGCCAAGTT	7740
CATTGCCAAG	TGCATCATAT	CCTGCCAACC	ACACACCTGG	CGCCAGGGTG	CGGTATTGCC	7800
TTACTCATAT	GTTTATTGCC	ACTGCAATAA	ACCTGTCTTT	GTGTTATACT	TTTCTGCACT	7860
GTAGCCAAC	CTTAAAAGCA	TTTTTGGCTT	GTAGCAGAAC	ATTTTTTTGC	TCTTACTGTT	7920

TGGTATACAA TAACATAAAA ATGAGTAACC TAAGGTCACA CACCTGCAAC CGGTTTCGGT 7980

TATCCACACC CTACATATTT CCTTCTTATA 8010



- 31 -

**WHAT IS CLAIMED IS:**

1. An isolated and purified DNA molecule which encodes human papillomavirus type 6a or a functional derivative thereof.
2. The isolated and purified DNA molecule of Claim 1, having a nucleotide sequence shown in Figure 1 or a functional derivative thereof.
3. An expression vector for expression of the DNA molecule of Claim 1 in a host.
4. An essentially purified protein encoded by the DNA molecule of Claim 1.
5. The protein of Claim 4 selected from the group HPV6a L2, HPV6a E1, HPV6a E2, HPV6a E4, HPV6a E5a, HPV6a E5b, HPV6a E6 and HPV6a E7 or a functional derivative thereof.
6. A monospecific antibody immunologically reactive with the a compound selected from the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1.
7. A process for expression of a human papillomavirus type 6a protein in a host, comprising:
 - (a) transferring the expression vector of Claim 4 into a suitable host; and
 - (b) culturing the host of step (a) under conditions which allow expression of the human papillomavirus type 6a protein from the expression vector.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN
REPLACED



- 32 -

8. A method of identifying compounds that modulate human papillomavirus type 6a activity, comprising:

- 5 (a) combining a suspected modulator of human papillomavirus type 6a activity with the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1a DNA; and
- 10 (b) measuring an effect of the modulator on the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1.

15 9. A compound active in the method of Claim 8.

10. A pharmaceutical composition comprising a compound active in the method of Claim 8.

20 11. A method of treating a patient in need of such treatment for a condition which is mediated by a human papillomavirus type 6a, comprising administration of the compound of Claim 9 to the patient.

25 12. A composition capable of inducing an immune response in a subject treated with the composition, the composition containing a compound selected from the DNA molecule of Claim 1, peptides encoded by the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1, or combinations thereof.

30 13. A vaccine for the prevention or treatment of human papillomavirus infection, the vaccine containing a compound selected from the group consisting of the DNA molecule of Claim 1, peptides

This Sheet Has Been Replaced



THIS SHEET HAS BEEN
REPLACED



- 33 -

encoded by the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1, or combinations thereof.

14. A method for inducing immune responses against
5 infection or disease caused by human papillomavirus which comprises introducing into an animal the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1, protein encoded by the DNA molecule of Claim 1, or combinations thereof.

This Sheet Has Been Replaced



1/7

1 [E6→
GTTAATAACAATCTTGGTTTTAAAAATAGGAGGGACCGAAAACGGTTCAACCGAAAACGGTTGTATATAAACCAGCCCTAAAATTTAGCAAACGAGGCA
101
TTATGGAAGTGCAAATGCCTCCACGTCTGCAACGACCATAGACCAGTTGTGCAAGACGTTAATCTATCTATGCATACGTTGCAAATTAATTGTGTGTT
201
TTGCAAGAATGCACTGACCACTGCAGAGATTTATTCATATGCATATAAACAGCTAAAGGTCCTGTTTCGAGGCGGCTATCCATATGCAGCCTGCGCGTGC
A C
301
TGCCTAGAATTTTCATGGAAAAATCAACCAATATAGACACTTTGATTATGCTGGATATGCAACAACCTGTTGAAGAAGAACTAAACAAGACATTTTAGACG
A A C
401 [E7→
TGCTAATTCGGTGCTACCTGTGTACAAAACCGCTGTGTGAAGTAGAAAAGGTAAACATATACTAACCAAGGCGCGTTTATAAAGCTAAATTGTACGTG
C
501 ←E6]
GAAGGGTCGCTGCCTACACTGCTGGACAACATGCAATGGAAGACATGTTACCTAAAGGATATTGTATTAGACCTGCAACCTCCAGACCCTGTAGGGTTAC
601
ATTGCTATGAGCAATTAGTAGACAGCTCAGAAGATGAGGTGGACGAAGTGGACGGACAAGATTCAACCTTTAAACAACATTTCCAAATAGTGACCTG
701 [E1→
TTGCTGTGGATGTGACAGCAACGTTGCACTGGTTGTGCAGTGTACAGAAACAGACATCAGAGAAGTGAACAGCTTCTGTTGGGAACACTAGACATAGTG
A
801 ←E7]
TGTCCCATCTGGCACCAGAGACATAACAACGATGGGGACGATTACGTTACAGAAAATGAGGGCTCTGGGTGTACAGGATGGTTTATGCTAGAAGCTAT
C
901
AGTGCAACACCCAACAGGTACACAAATATCAGACGATGAGGATGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATGACAGCAATATT
1001
ACACACAATTCCTTGAAGCACAGGCATTGTTTAACAGGCAGGAGGCGACACCCATTATCGCACTGTGCAGGACCTAAAACGAAAGTATTTAGGTAGTC
AC
1101
CATATGTTAGTCTATAAACAATATAGCCGAGGCAGTGGAAAGTGAAATAAGTCCAGGATTGGACGCCATTAACTTACAAGACAGCCAAAAAGGTAAA
1201
GCGACGGCTGTTTCAAACCAGGGAACCTAACGACAGTGGATATGGCTATTCTGAAGTGAAGCTGGAACGGGAACGAGGTAGAGAAACATGGCGTCCCG
A

FIG.1A

2/7

1301

GAAAATGGGGAGATGGTCAGGAAAAGGACACAGGAAGGGACATAGAGGGGAGGAACATACAGAGCGGAAGCGCCACAAACAGTGTACGGGAGCATG

1401

CAGGCACAGCAGGAATATTGGAATTGCTAAAATGTAAGATTTACGGGCAGCATTACTTGGTAAGTTTAAAGAATGCTTTGGGCTGTCTTTTATTGATTT
T A

1501

AATTAGGCCATTTAAAAGTGATAAAACAACATGTGCAGACTGGGTGGTAGCAGGATTTGGTATACATCATAGCATATCAGAGGCATTTCAAAAATTAATT
TT T G

1601

GAGCCATTAAGTTTATATGCACATATACAATGGCTAACAAATGCATGGGAATGGTATTGTTAGTATTAGTAAGATTTAAAGTAAATAAAAGTAGAAGTA
T

1701

CCGTTGCAGGTACACTTGCAACGCTATTAATATACCTGACAATCAAATGTTAATAGAGCCACCAAAATACAAAGTGGTGTGCAGCCCTGTATTGGTT
A C

1801

TCGTACAGGTATATCAAATGCCAGTACAGTTATAGGGAAGCACCAGAATGGATAACAGCCAACTGTTATTGAACATGGGTGGCAGACAGTCAGTTT
A C

1901

AAATTAACAGAAATGGTGCAGTGGGCATATGATAATGACATATCGGAGGAGAGTGAAATTGCATTTGAATATGCACAAAGGGGAGATTTTGATTCTAATG
G

2001

CAGCAGCATTTTTAAATAGCAATATGCAGGCAAAATATGTGAAGATTGTGCAACTATGTGTAGACATTATAACATGCAGAAATGAGGAAGATGTCTAT

2101

AAAACAATGCATAAAACATAGGGTTCTAAAATAGAAGGCACAGGAAATTGGAACCAATTGTACAATTCCTACGACATCAAAATATAGAATTTATTCCA
C T

2201

TTTTAAGTAAATTTAAATTATGGCTGCACGTACGCCAAAAAAAACATGATAGCCATAGTAGCCCTCCAGATACTGGGAAATCGTACTTTTGTATGA
C

2301

GTTTAATAAGCTTTTTAGGAGGTACAGTTATTAGTCATGTAAATTCAGCAGCCATTTTGGTTGCAACCGTTAGTAGATGCTAAGGTAGCATTGTTAGA
C

2401

TGATGCAACACAGCCATGTTGGATATATATGGATACATATATGAGAAATTTGTTAGATGGTAATCCTATGAGTATTGACAGAAAGCATAAAGCATTGACA

2501

TTAATTAATGTCCACCTCTGCTAGTAACGTCCAACATAGATATTACTAAGAAGAGAAATATAAGTATTTACATACTAGAGTAACAACATTTACATTTT
T

FIG.1B

3/7

2601 [E2→
CAAATCCATTCCCTTTTGACAGAAATGGGAATGCAGTGTATGAACTGTCAAATGCAAATGGAATGTTTTTTGAAAGACTGTCGTCAAGCCTAGACAT
A

2701 ←E1]
TCAGGATTCAGAGGACCAGGAAGATGGAAGCAATAGCCAAGCGTTTAGATCGGTGCCAGGAACAGTTGTTAGAAGTTTATGAAGAAAACAGTACTGACCT
T

2801
AAACAAACATGTATTGCATTGGAATGCATGAGACATGAAAGTGTATTATTATATAAAGCAAAACAAATGGCCCTAAGCCACATAGGAATGCAAGTAGTG
C

2901
CCACCATTAAAGGTGTCCGAAGCAAAGGACATAATGCCATTGAAATGCAAATGCATTTAGAATCATTATTAAGACTGAGTATAGTATGGAACCGTGA
G

3001
CATTACAAGAAACAAGTTATGAAATGTGGCAAACACCACCTAAACGCTGTTTTAAAAACGGGGCAAACTGTAGAAGTTAAATTTGATGGCTGTGCAAA

3101
CAATACAATGGATTATGTGGTATGGACAGATGTGTATGTGCAGGACACTGACTCCTGGGTAAAGGTGCATAGTATGGTAGATGCTAAGGTATATATTAC
A A

3201 [E4→
ACATGTGGACAATTTAAACATATTATGTAACCTTTGTAAGAGGGCAGAAAAGTATGGGAGCACCAACAATGGGAAGTATGTTATGGCAGCACAGTTA
T

3301
TATGTTCTCCTGCATCTGTATCTAGCACTACACAAGAAGTATCCATTCTGAATCTACTACATACACCCCGCACAGACCTCCACCCCTGTGTCTCAAG
T

3401
CACCAGGAAGACGCAGTGCAAACGCCCTAGAAAACGAGCAGGAGTCCAACAGTCACCTTGCAACGCCCTTGTGTGTGGCCACATTGGACCGTG
A G C

3501 ←E4]
GACAGTGGAAACCACAACCTCATCACTAACAATCAGGACCAGCACCAAGAAGGAACAACAGTAACAGTTCAGCTACGCCTATAGTGCAATTTCAAGGTG
C

3601
AATCTAATTGTTTAAAGTGTTTAGATATAGGCTAAATGACAAACACAGACATTTATTTGATTTAATATCATCAAGTGGCACTGGGCCTCCCAAAGGC
C G T

3701
ACCACATAAACATGCCATTGTAACGTGAACATATCATAGTGAGGAACAAAGGCAACAGTTTTTAAATGTTGTAATAATACCACCTACTATTAGGCACAAA
G G C C C

3801 ←E2] [E5→
CTGGGGTTTATGTCAGTGCACCTATTGTAATTTGTATATATGTAATGTGTAATATATGCTATTGGTGAATACAAGTGTACATGATGGAAGTGGTAC
A G

FIG.1C

4/7

3901
CTGTACAAATAGCTGCAGGAACAACCAGCACATTAATACTGCCTGTTATAATTGCATTGTTGTATGTTTTGTTAGCATCATACTTATTGTATGGATATC
C

4001
TGACTTTATTGTGTACACATCTGTGCTAGTACTAAGCTGCTTTTATACTTACTATTGTGGCTGCTATTAAACAACCCCTTGCAATTTTCCTACTAACT
G T

4101
CTACTTGTGTGTTACTGTCCCGCATTGTATATACCACTACATTGTTAAACACACAGCAATGATGCTAACATGTCAATTAATGATGGAGATACATGGCT
T T C -E5]

4201
GGGTTTGTGGTTGTATGTGCCTTTATTGTAGGGGTGTTGGGTTATTATTAATGCACTATAGAGCTGTACAAGGGGATAAACACACCAAATGTAACAAG
A G G G

4301
TGTAACAAACACACCTGTAATGCTGATTATGTAAGTATGCATCATGATACTGCTGGTGATTATATATATGAATTAGAGTAAACCTTTTTTATATTTG
A A T AC A [L2- C G

4401
TAACAGTGTATGTTTTGTATACCATGGCACATAGTAGGGCCGACGACGCAAGCGTGGCTCAGCTACACAGCTATATCAACATGTAACCTTACTGGAAC
C C

4501
ATGCCCCCAGATGTAATTCCTAAGGTGGAGCACAACACCATTGCAGATCAAATATTAAATGGGAAGTTTGGGGTTTTTTTGGAGGGTTGGGTATA
G

4601
GGCACCGGTTCCGGCACTGGGGTCTACTGGCTATGTTCCCTTAGGAAGTTCTGCAAAACCTTCTATTACTAGTGGCCTATGCCTCGTCCCTCTGTGG
G CA

4701
TGGTGGAGCCTGTGGCCCTTGGATCCATCCATTGTGTCTTTAATTGAAGAATCAGCAATCATTAAAGCAGGGGGCCTGAAATTGTCCCCCTGCACA
T G

4801
CGGTGGGTTTACAATTACATCCTCTGAACAACCTACCCCTGCAATATTGGATGTATCAGTTACTAGTCATACTACTAGTATATTAGAAATCCTGTC
C

4901
TTTACAGAACCTTCTGTAACACAACCCCAACCACCGTGGAGGCTAATGGACATATATTAATTTCTGCACCCACTATAAGCTCACACCCTATAGAGGAAA
G

5001
TTCCTTTAGATACTTTTGTGATATCCTCTAGTGATAGCGTCTACATCCAGTACCCCTGTTCTCTGGTACTGCACCTAGGCCTCGTGTGGGCCTATATAG
G A C

5101
TCGTGCATTGCACCAGGTGCAGGTTACAGACCCCTGCATTCTTTCCACTCCTCAACGCTTAATTACATATGATAACCCCTGTATATGAAGGGGAGGATGTT

FIG.1D

5/7

5201

AGTGTAACAATTTAGTCATGATTCTATACACAATGCACCTGATGAGGCTTTATGGACATAATTCGTTTGCACAGACCTGCTATTGCGTCCCGACGTGGCC
C

5301

TTGTGGGTACAGTCCATTGGACAACGGGGTCTATGCACACTGCCAGCGAAAGCACATAGGGCCCCGATTCATTATTTTTATGATATTTACCTAT

5401

TGCACAAGCTGCAGAAGAAATAGAAATGCACCTCTTGTGGCTGCACAGGATGATACATTTGATATTTATGCTGAATCTTTGAACCTGACATTAAACCT
G G

5501

ACCCAACACCCCTGTTACAAATATATCAGATACATATTTAACTTCCACACCTAATACAGTTACACAACCGTGGGTAACACCACAGTTCATTGTCAATTC
C

5601

CTAATGACCTGTTTTACAGTCTGGCCCTGATATAACTTTTCTACTGCACCTATGGGAACACCTTTAGTCCTGTAACCTCTGCTTTACCTACAGGCC
A [L1→

5701

TGTTTTATTACAGGTTCTGGATTTATTTGCATCCTGCATGGTATTTTGCACGTAACGGCGTAAACGTATTCCTTATTTTTTTCAGATGTCGGGGC

→L2]

TAGCGACAGCAGTATATGTCCTCCTCTAACCTGTATCCAAAGTTGTGCCACGGATGCTTATGTTACTCGCACCAACATATTTTATCATGCCAGC

5901

AGTTCTAGACTTCTTGCAGTGGGTCATCCTTATTTTTCCATAAAACGGGCTAACAAACTGTTGTGCCAAAGGTGCAGGATATCAATACAGATATTTA
A G

6001

AGGTGGTGTTACCAGATCCTAACAAATTTGCATTGCCTGACTGTCTCTTTTGTATCCACAACACACGTTTGGTATGGGCATGCACAGCCTAGAGGT
C A

6101

GGCCCGGGACAGCCATTAGGTGTGGGTGTAAGTGGACATCCTTTCTAAATAAATATGATGATGTTGAAAATTCAGGGAGTGGTGGTAACCTGGACAG
A

6201

GATAACAGGTTAATGTAGGTATGCATTATAACAAACACAATTATGCATGGTTGGATGTGCCCCCTTTGGGCGAGCATTGGGGTAAAGGTAACAGT

6301

GTACTAATACACCTGTACAGGCTGGTACTGCCCGCCCTTAGAACTTATTACCAGTGTTATACAGGATGGCGATATGGTTGACACAGCCTTTGGTGCTAT

6401

GAATTTTGCTGATTTGCAGACCAATAAATCAGATGTTCTATTGACATATGTGGCACTACATGTAATATCCAGATTATTTACAAATGGCTGCAGACCCA

FIG.1E

6/7

6501

TATGGTGATAGATTATTTTTTTTCTACGAAGGAACAAATGTTGCCAGACATTTTTTTAACAGGGCTGGCGAGGTGGGGAACCTGTCCTGATACTC
A

6601

TTATAATTAAGGGTAGTGGAATCGCACGCTGTAGGGAGTAGTATATATGTTAACACCCCAAGCGGCTCTTTGGTGTCTCTGAGGCACAATTGTTAA
G

6701

TAGCCATATTGGCTACAAAAGCCAGGGACATAACAATGGTATTTGTTGGGTAATCAACTGTTTGTACTGTGGTAGATACCACAGCAGTACCAAC

6801

ATGACATTATGTGCATCCGTAACATACATCTCCACATACACCAATTCTGATTATAAAGAGTACATGGTCATGTGGAAGAGTATGATTACAATTTATT

6901

TTCAATTATGTAGCATTACATTGTCTGCTGAAGTAATGGCCTATATTCACACAATGAATCCCTCTGTTTGAAGACTGGAACCTTGGGTTATCGCCTCC

7001

CCCAATGCTACATTAGAAGATACCTATAGGTATGTGCAGTCACAGCCATTACCTGTCAAAGCCCACTCCTGAAAAGGAAAAGCCAGATCCCTATAAG

7101

AACCTTAGTTTTTGGGAGGTTAATTTAAAGAAAAGTTTTCTAGTGAATTGGATCAGTATCCTTGGGACGCAAGTTTTGTACAAAGTGGATATAGGG

7201

GACCGTCTCTATTGCTACCGGTGTTAAGCGCCCTGCTGTTTCCAAAGCCTCTGCTGCCCTAAACGTAAGCGCGCCAAACTAAAAGSTAATATATGTG
A

7301

TATATGTAAGTATATATATATGTGTATGTACTGTTATGTATATGTGTATGTACTGTTATATGTATGTGTGTATATATGTGTATATATGTGT

7401

ATGTGTATATGTATATGTATGTGTGTGTATATATGTGTGTGTGTATGTGTATGTATGTATGTGTGTATATGTGTGTATATGTGT
C G

7501

CAATAACAATTAAC—ACACCTGTGACTCAGTGGCTGTTGCACGCGTTTGGTTTGCACGGCCTTACACACATAAGTAATATACATGCACAA
C CTGTT

7595

TATATATATTTTGT—ACAATAATATATTTTATATTTGCAACCGTTTTCGGTTGCCCTTGCATACACTTTCCACCAATTTGTTACAACGTGTTGCC
T A C C A T

FIG.1F

7/7

7694

GTTAATCCTATATATTTTGTGCCAGGTACACATTGCCCTGCCAAGTTCATTGCCAAGTGCATCATCCTGCCAACCACACACCTGGCGCCAGGGTGCGG

C

GC

7794

TATTGCCTTACTCATATGTTTATTGCCACTGCAATAAACCTGTCTTTGTGTTATACTTTTCTGCACTGTAGCCAACCTCTAAAAGCATTTTTGGCTTGTA

A

7894

GCAGAACATTTTTTCTCTTACTGTTTGGTATACAATAACATAAAAAATGAGTAACCTAAGGTCACACACCTGCAACCGGTTTCGGTTATCCACACCTA

C

G

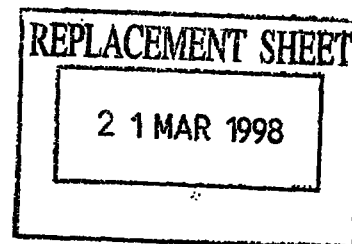
7994

CATATTCCTTCTTATA

FIG.1G



-31-

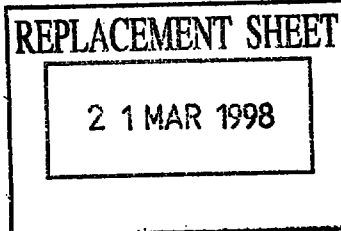


WHAT IS CLAIMED IS :

1. An isolated and purified DNA molecule which encodes human papillomavirus type 6a which is shown in SEQ.ID.NO.2 (or Figure 1).
2. An essentially purified protein encoded by the DNA molecule of Claim 1.
3. The protein of Claim 2 selected from the group HPV6a L2, HPV6a E1, HPV6a E2, HPV6a E4, HPV6a E5a, HPV6a E5b, HPV6a E6 and HPV6a E7 or a functional derivative thereof.
4. A monospecific antibody immunologically reactive with a compound selected from the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1.
5. A method of identifying compounds that modulate human papillomavirus type 6a activity, comprising :
 - (a) combining a suspected modulator of human papillomavirus type 6a activity with the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1a DNA; and
 - (b) measuring an effect of the modulator on the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1 or a protein encoded by the DNA molecule of Claim 1.



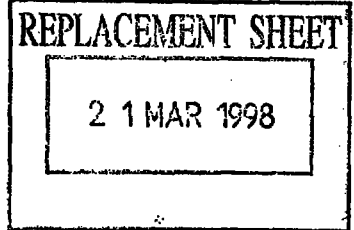
-32-



6. A compound active in the method of Claim 5.
7. A pharmaceutical composition comprising a compound active in the method of Claim 5.
8. A composition capable of inducing an immune response in a subject treated with the composition, the composition containing a compound selected from the DNA molecule of Claim 1, peptides encoded by the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1, or combinations thereof.
9. A vaccine for the prevention or treatment of human papillomavirus infection, the vaccine containing a compound selected from the group consisting of the DNA molecule of Claim 1, peptides encoded by the DNA molecule of Claim 1, RNA complementary to the DNA molecule of Claim 1, or combinations thereof.
10. An open reading frame encoding a human papillomavirus type 6a protein, the open reading frame being derived from the DNA molecule of Claim 1.
11. An expression vector comprising the open reading frame of Claim 10.
12. A process for expression of a human papillomavirus type 6a protein in a host, comprising :
 - (a) transferring the expression vector of Claim 11 into a suitable host; and



-33-



(b) culturing the host of step(a) under conditions which allow expression of the human papillomavirus type 6a protein from the expression vector.

13. A host cell transformed with the expression vector of Claim 11.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

86 9701136-5.



RTMP
COPY

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 7/00, 15/37, 15/63, C12P 21/02		A1	(11) International Publication Number: WO 96/09375 ✓
			(43) International Publication Date: 28 March 1996 (28.03.96)
(21) International Application Number: PCT/US95/11859 ✓		(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KR, KZ, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, (SG) SI, SK, TJ, TM, TT, UA, US, UZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).	
(22) International Filing Date: 18 September 1995 (18.09.95) ✓			
(30) Priority Data: 310,468 22 September 1994 (22.09.94) ✓ US			
(60) Parent Application or Grant (63) Related by Continuation US 310,468 (CON) Filed on 22 September 1994 (22.09.94)		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). ✓			
(72) Inventors; and (75) Inventors/Applicants (for US only): JANSEN, Kathrin, U. [DE/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). HOFMANN, Kathryn, J. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).			
(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).			
(54) Title: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 6A			
(57) Abstract The present invention is directed to DNA molecules encoding purified human papillomavirus type 6a and compounds derived therefrom.			