



<p>(51) Internationale Patentklassifikation <sup>6</sup> : C12N 15/37, C07K 14/025, 16/08, C12N 5/10, C12Q 1/68, G01N 33/50, A61K 39/12</p>	A2	<p>(11) Internationale Veröffentlichungsnummer: <b>WO 99/09177</b></p> <p>(43) Internationales Veröffentlichungsdatum: 25. Februar 1999 (25.02.99)</p>
<p>(21) Internationales Aktenzeichen: PCT/DE98/02379</p> <p>(22) Internationales Anmeldedatum: 12. August 1998 (12.08.98)</p> <p>(30) Prioritätsdaten: 197 35 118.2      13. August 1997 (13.08.97)      DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES ÖFFENTLICHEN RECHTS [DE/DE]; Im Neuenheimer Feld 280, D-69120 Heidelberg (DE).</p> <p>(72) Erfinder; und</p> <p>(75) Erfinder/Anmelder (nur für US): DE VILLIERS ZUR HAUSEN, Ethel-Michelle [DE/DE]; (DE). ZUR HAUSEN, Harald [DE/DE]; Eichenstrasse 1, D-69483 Waldmichelbach (DE).</p> <p>(74) Anwalt: SCHÜSSLER, Andrea; Huber &amp; Schüssler, Trud- eringer Strasse 246, D-81825 München (DE).</p>	<p>(81) Bestimmungsstaaten: JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p><b>Veröffentlicht</b> <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i></p>	
<p>(54) Title: PAPILOMA VIRUS, AGENTS FOR THE DETECTION THEREOF AND FOR THE THERAPY OF THE DISEASES CAUSED BY SAID VIRUS</p> <p>(54) Bezeichnung: PAPILOMVIREN, MITTEL ZU DEREN NACHWEIS SOWIE ZUR THERAPIE VON DURCH SIE VERURSACHTEN ERKRANKUNGEN</p> <p>(57) Abstract</p> <p>The invention relates to a DNA coding a peptide of a papilloma virus main capsid protein or a papilloma virus genome. The invention further relates to the proteins coded by the papilloma virus genome, to the antibodies directed against said virus and to the use of said proteins and antibodies for diagnostic, therapeutic and vaccination purposes.</p> <p>(57) Zusammenfassung</p> <p>Die Erfindung betrifft eine DNA, die für ein Peptid eines Papillomvirus-Hauptcapsid-Proteins bzw. ein Papillomvirus-Genom codiert. Des weiteren betrifft die Erfindung durch das Papillomvirus-Genom codierte Proteine und gegen sie gerichtete Antikörper sowie deren Verwendung in Diagnose, Therapie und Vakzinierung.</p>		

**LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

## **Papillomviren, Mittel zu deren Nachweis sowie zur Therapie von durch sie verursachten Erkrankungen**

Die Erfindung betrifft eine DNA, die für ein Peptid eines Papillomvirus-Hauptcapsid-Proteins bzw. ein Papillomvirus-Genom codiert. Desweiteren betrifft die Erfindung durch das Papillomvirus-Genom codierte Proteine und gegen sie gerichtete Antikörper sowie deren Verwendung in Diagnose, Therapie und Vakzinierung.

Es ist bekannt, daß Papillomviren das Epithelgewebe von Mensch und Tier infizieren. Human-Papillomviren (nachstehend mit HP-Viren bezeichnet) finden sich in benignen, z.B. Warzen, Kondylome im Genitalbereich, und malignen, z.B. Karzinome der Haut und der Gebärmutter, epithelialen Neoplasmen (vgl. zur Hausen, H., *Biochimica et Biophysica Acta (BBA)* 1288, (1996), Seiten 55-78). Auch werden HP-Viren für die Entwicklung maligner Tumoren im Oropharyngealbereich in Betracht gezogen (vgl. zur Hausen, H., *Curr. Top. Microbiol. Immunol.* 78, (1977), Seiten 1-30).

Papillomviren weisen ein ikosaedrisches Capsid ohne Hülle auf, in dem ein zirkuläres, doppelsträngiges DNA-Molekül von etwa 7900 bp vorliegt. Das Capsid umfaßt ein Hauptcapsid-Protein (L1) und ein Nebencapsid-Protein (L2). Beide Proteine, coexprimiert oder L1 alleine exprimiert, führen in vitro zur Ausbildung von Virus-ähnlichen Partikeln (vgl. Kirnbauer, R. et al., *Journal of Virology*, (1993), Seiten 6929-6936).

Papillomviren lassen sich nicht in Monolayer-Zellkultur vermehren. Ihre Charakterisierung ist daher äußerst schwierig, wobei bereits der Nachweis von Papillomviren erhebliche Probleme schafft. Dies trifft insbesondere für Papillomviren in Karzinomen der Haut zu.

Der vorliegenden Erfindung liegt somit die Aufgabe zugrunde, ein Mittel bereitzustellen, mit dem Papillomviren, insbesondere in Karzinomen der Haut, nachgewiesen werden können. Ferner sollte ein Mittel bereitgestellt werden, um gegen diese Papillomviren therapeutisch vorgehen zu können.

Erfindungsgemäß wird dies durch die Bereitstellung der Gegenstände in den Patentansprüchen erreicht.

Gegenstand der Erfindung ist somit eine für ein Peptid eines Papillomvirus-Hauptcapsid-Proteins (L1) codierende DNA, wobei das Peptid die Aminosäuresequenz von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7 oder eine davon durch ein oder mehrere Aminosäuren unterschiedliche Aminosäuresequenz umfaßt.

Ein weiterer Gegenstand der Erfindung ist eine für ein Peptid eines Papillomvirus-Hauptcapsid-Proteins codierende DNA, wobei die DNA die Basensequenz von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7 oder eine davon durch ein oder mehrere Basenpaare unterschiedliche Basensequenz umfaßt.

Fig. 1 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid DL314 bei der DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen) unter DSM 11604 am 12. Juni 1997 hinterlegt.

Fig. 2 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid DL347 bei der DSMZ unter DSM 11605 am 12. Juni 1997 hinterlegt.

Fig. 3 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid DL369 bei der DSMZ unter DSM 11606 am 12. Juni 1997

hinterlegt.

Fig. 4 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid GA1-3 bei der DSMZ unter DSM 11607 am 12. Juni 1997 hinterlegt.

Fig. 5 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid GA3-1 bei der DSMZ unter DSM 11608 am 12. Juni 1997 hinterlegt.

Fig. 6 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid GA6-2 bei der DSMZ unter DSM 11609 am 12. Juni 1997 hinterlegt.

Fig. 7 zeigt die Basensequenz und die davon abgeleitete Aminosäuresequenz einer für ein Peptid von L1 eines Papillomvirus codierenden DNA. Diese DNA wurde als Plasmid GA9-4 bei der DSMZ unter DSM 11610 am 12. Juni 1997 hinterlegt.

Vorstehende DNA wurde mit der DNA bekannter Papillomviren verglichen. Es wurden Sequenzhomologie-Studien durchgeführt. Eine Homologie, die weniger als 90 % beträgt, weist eine erfindungsgemäße DNA als neues HP-Virus aus. Die erfindungsgemäßen DNAs weisen zu bekannten Papillomviren folgende Sequenzhomologien auf:

DNA von Fig. 1: 78 % zu HP-Virus 15

DNA von Fig. 2: 80 % zu HP-Virus 5b

DNA von Fig. 3: 76 % zu HP-Virus 15

DNA von Fig. 4: 80 % zu HP-Virus 24

- 4 -

DNA von Fig. 5: 79 % zu HP-Virus 8  
DNA von Fig. 6: 81 % zu HP-Virus 12  
DNA von Fig. 7: 84 % zu HP-Virus 15

Erfindungsgemäß kann vorstehende DNA in einem Vektor bzw. Expressionsvektor vorliegen. Beispiele solcher sind dem Fachmann bekannt. Im Falle eines Expressionsvektors für *E. coli* sind dies z.B. pGEMEX, pUC-Derivate, pGEM-T und pGEX-2T. Für die Expression in Hefe sind z.B. pY100 und Ycpad1 zu nennen, während für die Expression in tierischen Zellen z.B. pKCR, pEF-BOS, cDM8 und pCEV4, anzugeben sind.

Der Fachmann kennt geeignete Zellen, um vorstehende, in einem Expressionsvektor vorliegende DNA zu exprimieren. Beispiele solcher Zellen umfassen die *E. coli*-Stämme HB101, DH1, x1776, JM101, JM 109 und XL1-Blue, den Hefe-Stamm *Saccharomyces cerevisiae* und die tierischen Zellen L, NH-3T3, FM3A, CHO, COS, Vero, und Hela.

Der Fachmann weiß, in welcher Weise vorstehende DNA in einen Expressionsvektor inseriert werden muß. Ihm ist auch bekannt, daß vorstehende DNA in Verbindung mit einer für ein anderes Protein bzw. Peptid codierenden DNA inseriert werden kann, so daß vorstehende DNA in Form eines Fusionsproteins exprimiert werden kann.

Ein weiterer Gegenstand der Erfindung ist ein Papillomvirus-Genom, das vorstehende DNA umfaßt. Der Ausdruck "Papillomvirus-Genom" umfaßt auch ein unvollständiges Genom, d.h. Fragmente eines Papillomvirus-Genoms, die vorstehende DNA umfassen. Dies kann z.B. eine für L1 codierende DNA oder ein Teil davon sein.

Zur Bereitstellung vorstehenden Papillomvirus-Genoms kann ein übliches Verfahren verwendet werden. Günstig ist ein Verfahren, das folgende Verfahrensschritte umfaßt:

- 5 -

- (a) Isolierung der Gesamt-DNA aus einer Biopsie epithelialen Neoplasmas,
- (b) Hybridisierung der Gesamt-DNA von (a) mit vorstehender DNA, wodurch ein in der Gesamt-DNA von (a) enthaltenes Papillomvirus-Genom nachgewiesen wird, und
- (c) Klonierung der das Papillomvirus-Genom enthaltenden Gesamt-DNA von (a) in einem Vektor, und gegebenenfalls Subklonierung des erhaltenen Klons, wobei sämtliche Verfahrensschritte üblicher DNA-Rekombinationstechnik entstammen.

Hinsichtlich der Isolierung, Hybridisierung und Klonierung von Zell-DNA wird ergänzend auf Sambrook et al., Molecular Cloning, A Laboratory Manual, zweite Ausgabe, Cold Spring Harbor Laboratory (1989) verwiesen.

Der Ausdruck "epitheliales Neoplasma" umfaßt jegliche Neoplasmen des Epithelgewebes bei Mensch und Tier. Beispiele solcher Neoplasmen sind Warzen, Kondylome im Genitalbereich und Karzinome der Haut. Letztere werden vorliegend bevorzugt verwendet, um vorstehendes Papillomvirus-Genom zu isolieren.

Der Ausdruck "Vektor" umfaßt jegliche zur Klonierung von chromosomaler bzw. extrachromosomaler DNA geeignete Vektoren. Beispiele solcher Vektoren sind Cosmide, wie pWE15 und Super Cos1, und Phagen, wie  $\lambda$ -Phagen, z.B.  $\lambda$ ZAP Expressvector,  $\lambda$ ZAPII Vector und  $\lambda$ gt10 Vektor. Vorliegend werden  $\lambda$ -Phagen bevorzugt verwendet. Vorstehende Vektoren sind bekannt und bei der Firma Stratagene erhältlich.

Erfindungsgemäße Papillomvirus-Genome können integriert in chromosomaler DNA oder extrachromosomal vorliegen. Dem Fachmann sind Verfahren bekannt, dies abzuklären. Auch weiß er um Verfahren, die zur Klonierung der Papillomvirus-Genome optimalen Restriktionsenzyme herauszufinden. Er wird sich an

Genomen bekannter Papillomviren orientieren. Insbesondere wird der Fachmann die vorstehend genannten HP-Viren entsprechend beachten.

Beispielhaft wird die Bereitstellung eines mit DL314-G bezeichneten Papillomvirus-Genoms beschrieben. Hierzu wird die Gesamt-DNA aus einer Biopsie eines Basalzellkarzinoms isoliert, mit BamHI gespalten und in einem Agarosegel elektrophoretisch aufgetrennt. Das Agarosegel wird danach einem Blotting-Verfahren unterzogen, wodurch die DNA auf eine Nitrozellulosemembran übertragen wird. Diese wird in ein Hybridisierungsverfahren eingesetzt, in dem die DNA von Fig. 1, ggfs. in Kombination mit einer DNA von HP-Virus 15 als markierte Probe verwendet wird. Es wird eine Hybridisierung mit der in der Gesamt-DNA vorliegenden Papillomvirus-DNA erhalten.

Im weiteren wird vorstehende mit BamHI gesplattene Gesamt-DNA in einem  $\lambda$ -Phagen kloniert. Die entsprechenden Klone, d.h. die die Papillomvirus-DNA enthaltenden Klone, werden durch Hybridisierung mit der DNA von Fig. 1, ggfs. in Kombination mit einer DNA des HP-Virus 15 identifiziert. Das Insert dieser Klone wird dann einer weiteren Klonierung in einem Plasmid-Vektor unterzogen, wodurch ein Klon erhalten wird, der das Papillomvirus-Genom DL314-G enthält. Das Genom wird durch Sequenzierung bestätigt.

In analoger Weise werden weitere Papillomvirus-Genome bereitgestellt. Sie werden entsprechend der zu ihrer Bereitstellung verwendeten DNAs bezeichnet, mit: DL347-G, DL369-G, GA1-3-G, GA3-1-G, GA6-2-G bzw. GA9-4-G.

Ein weiterer Gegenstand der Erfindung ist ein Protein, das durch vorstehendes Papillomvirus-Genom codiert wird. Ein solches Protein ist z.B. ein Hauptcapsid-Protein (L1) oder ein Nebencapsidprotein (L2). Die Herstellung eines vorstehenden Proteins erfolgt in üblicher Weise. Beispielhaft wird die Herstellung von L1 bzw. L2 des Papillomvirus-Genoms DL314-G beschrieben. Hierzu wird das zu der DNA von Fig. 1 verwandte HP-Virus 15 herangezogen. Von diesem ist die vollständige Sequenz und die Lage einzelner für Proteine codierender DNA-

Bereiche bekannt. Durch parallele Restriktionsspaltungen beider Genome und anschließender Hybridisierung mit verschiedenen, die L1 bzw. L2 codierende DNA betreffenden Fragmenten werden diese DNAs auf dem Papillomvirus-Genom DL314-G identifiziert. Sie werden durch Sequenzierung bestätigt. Die für L1 codierende DNA wird mit DL314-G-L1-DNA und die für L2 codierende DNA mit DL314-G-L2-DNA bezeichnet.

Im weiteren wird die für L1 bzw. L2 codierende DNA in einen Expressionsvektor inseriert. Beispiele eines solchen für E. coli, Hefe und tierische Zellen sind vorstehend genannt. Insbesondere wird für die Expression in E. coli auf den Vektor pGEX-2T verwiesen (vgl. Kirnbauer, R. et al., supra). Nach Insertion der DL314-G-L1- bzw. DL314-G-L2-DNA wird pGEX-2T-DL314-G-L1 bzw. pGEX-2T-DL314-G-L2 erhalten. Diese Expressionsvektoren exprimieren nach Transformation von E. coli ein Glutathion S-Transferase-L1- bzw. Glutathion S-Transferase-L2-Fusionsprotein. Die Reinigung dieser Proteine erfolgt in üblicher Weise.

Für eine weitere Expression vorstehender L1 bzw. L2 codierender DNA wird das Baculovirus- bzw. Vacciniavirus-System genannt. Hierfür verwendbare Expressionsvektoren sind z.B. pEV mod. und pSynwtVI für das Baculovirus-System (vgl. Kirnbauer, R. et al., supra). Für das Vacciniavirus-System sind insbesondere Vektoren mit dem Vacciniavirus "early" (p7.5k)- bzw. "late" (pSynth, p11K)-Promotor zu nennen (vgl. Hagensee, M., E. et al., Journal of Virology (1993), Seiten 315-322). Vorliegend wird das Baculovirus-System bevorzugt. Nach Insertion vorstehender L1 bzw. L2 codierender DNA in pEV mod. wird pEVmod.-DL314-G-L1 bzw. pEVmod.-DL314-G-L2 erhalten.

Der erstere Expressionsvektor alleine bzw. beide Expressionsvektoren zusammen führen nach Infektion von SF-9 Insektenzellen zur Ausbildung von Virus-ähnlichen Partikeln. Im ersteren Fall umfaßt ein solches Partikel ein L1-Protein, während es im letzteren Fall neben einem L1- auch ein L2-Protein enthält.

Ein Virus-ähnliches Partikel letzteren Falls wird auch erhalten, indem die vor-

stehenden DL314-G-L1- und DL314-G-L2-DNAs gemeinsam in den Expressionsvektor pSynwtVI inseriert werden und das erhaltene pSynwtVI:DL314-G-L1/L2 zur Infektion von SF-9 Insektenzellen verwendet wird. Die Reinigung vorstehender Virus-ähnlicher Partikel erfolgt in üblicher Weise. Sie stellen auch einen Gegenstand der Erfindung dar.

Ein weiterer Gegenstand der Erfindung ist ein gegen ein vorstehendes Protein bzw. Virus-ähnliches Partikel gerichteter Antikörper. Die Herstellung eines solchen erfolgt in üblicher Weise. Beispielhaft wird es für die Herstellung eines Antikörpers beschrieben, der gegen ein L1 von DL314-G umfassendes Virus-ähnliches Partikel gerichtet ist. Hierzu wird das Virus-ähnliche Partikel BALB/c-Mäusen subcutan injiziert. Diese Injektion wird im Abstand von jeweils 3 Wochen wiederholt. Etwa 2 Wochen nach der letzten Injektion wird das den Antikörper enthaltende Serum isoliert und in üblicher Weise getestet.

In bevorzugter Ausführungsform ist der Antikörper ein monoklonaler Antikörper. Zu seiner Herstellung werden nach vorstehender vierten Injektion den Mäusen Milzzellen entnommen und diese in üblicher Weise mit Myelomzellen fusioniert. Die weitere Klonierung erfolgt ebenso nach bekannten Verfahren.

Mit der vorliegenden Erfindung wird es ermöglicht, Papillomviren, insbesondere in Karzinomen der Haut, nachzuweisen. Hierzu kann die erfindungsgemäße DNA als solche oder von einer weiteren DNA umfaßt eingesetzt werden. Letztere kann auch ein Papillomvirus-Gom oder ein Teil davon sein.

Die vorliegende Erfindung ermöglicht ferner die Bereitstellung von bisher nicht gekannten Papillomviren. Diese finden sich insbesondere in Karzinomen der Haut. Desweiteren liefert die Erfindung Proteine und Virus-ähnliche Partikel, die auf diese Papillomviren zurückgehen. Darüberhinaus werden Antikörper bereitgestellt, die gegen diese Proteine bzw. Partikel gerichtet sind.

Die vorliegende Erfindung ermöglicht es also, diagnostische und therapeutische

Maßnahmen bei Papillomvirus-Erkrankungen zu ergreifen. Darüberhinaus liefert sie die Möglichkeit, eine Vakzine gegen Papillomvirus-Infektionen aufzubauen. Die vorliegende Erfindung stellt somit einen Durchbruch auf dem Gebiet der Papillomvirus-Forschung dar.

Die Erfindung wird durch die Beispiele erläutert.

**Beispiel 1: Identifizierung des Papillomvirus-Genoms DL314-G**

Aus einer Biopsie eines Basalzellkarzinoms wird die Gesamt-DNA isoliert. 10 $\mu$ g dieser DNA werden mit dem Restriktionsenzym Bam-HI gespalten und in einem 0,5 % Agarosegel elektrophoretisch aufgetrennt. Gleichzeitig werden auch 10 $\mu$ g vorstehender DNA aufgetrennt, die nicht gespalten worden ist. Das Agarosegel wird einem Blotting-Verfahren unterzogen, wodurch die DNA aus dem Agarosegel auf eine Nitrozellulosemembran übertragen wird. Diese wird in ein Hybridisierungsverfahren eingesetzt, in dem die vorstehende DNA von Fig. 1 in Kombination mit HP-Virus-15 DNA als p<sup>32</sup>-markierte Probe verwendet wird. Es wird eine Hybridisierung mit der geblotteten DNA erhalten.

Vorstehende Verfahren sind dem Fachmann auf dem Gebiet der DNA-Rekombinationstechnik bekannt. Ergänzend wird auf Sambrook et al., supra verwiesen.

**Beispiel 2: Klonierung des Papillomvirus-Genoms DL314-G**

Die aus Beispiel 1 erhaltene Biopsie-DNA wird mit dem Restriktionsenzym BamHI gespalten. Die erhaltenen Fragmente werden in eine Ligasereaktion eingesetzt, in der ebenfalls der mit BamHI gesplattene und dephosphorylierte Vektor  $\lambda$ ZAP Express vorliegt. Die hierbei erhaltenen rekombinanten DNA-Moleküle werden in

- 10 -

Bakteriophagen verpackt und diese zur Infektion von Bakterien verwendet. Für diese Verfahrensschritte wird der von der Firma Stratagene angebotene ZAP Express Vektor Kit verwendet. Die erhaltenen Phagenplaques werden dann einem Hybridisierungsverfahren unterzogen, in dem die in Beispiel 1 verwendete p<sup>32</sup>-markierte DNA von Fig. 1 in Kombination mit p<sup>32</sup>-markierter HP-Virus-15-DNA verwendet wird. Es wird eine Hybridisierung mit entsprechenden Phagenplaques erhalten. Aus diesen werden die BamHI-Fragmente von DL314-G isoliert und zusammen mit einem BamHI-gespaltenen, dephosphorylierten Plasmid-Vektor, pBluescript, in eine weitere Ligasereaktion eingesetzt. Die erhaltenen rekombinanten DNA-Moleküle werden zur Transformation von Bakterien, E. coli XL1-Blue, verwendet. Durch Restriktionsspaltungen bzw. Hybridisierung mit vorstehenden DNA-Proben wird ein das Papillomvirus-Genom DL314-G enthaltender Bakterienklon identifiziert. Das Plasmid dieses Bakterienklons wird mit pBlue-DL314-G bezeichnet.

### Patentansprüche

1. DNA, codierend für ein Peptid eines Papillomvirus-Hauptcapsid-Proteins, wobei das Peptid die Aminosäuresequenz von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7 oder eine davon durch ein oder mehrere Aminosäuren unterschiedliche Aminosäuresequenz umfaßt, und wobei die DNA durch folgende Verfahrensschritte erhältlich ist:
  - (a) Isolierung der Gesamt-DNA aus einer Biopsie epithelialen Neoplasmas,
  - (b) Hybridisierung der Gesamt-DNA von (a) mit einer DNA von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7, wodurch ein in der Gesamt-DNA von (a) enthaltenes Papillomvirus-Genom nachgewiesen wird, und
  - (c) Klonierung der das Papillomvirus-Genom enthaltenden Gesamt-DNA von (a) in einem Vektor sowie Sequenzierung des Klons.
  
2. DNA nach Anspruch 1, dadurch gekennzeichnet, daß die für das Peptid des Papillomvirus-Hauptcapsid-Proteins codierende DNA die Basensequenz von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7, oder eine davon durch ein oder mehrere Basenpaare unterschiedliche Basensequenz umfaßt, und wobei die DNA durch folgende Verfahrensschritte erhältlich ist:
  - (a) Isolierung der Gesamt-DNA aus einer Biopsie epithelialen Neoplasmas,
  - (b) Hybridisierung der Gesamt-DNA von (a) mit einer DNA von Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6 oder Fig. 7, wodurch ein in der Gesamt-DNA von (a) enthaltenes Papillomvirus-Genom nachgewiesen wird, und

- 12 -

- (c) Klonierung der das Papillomvirus-Genom enthaltenden Gesamt-DNA von (a) in einem Vektor sowie Sequenzierung des Klons.
3. DNA nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die DNA ein Papillomvirus-Genom umfaßt.
  4. Protein, codiert durch das Papillomvirus-Genom nach Anspruch 3.
  5. Protein nach Anspruch 4, dadurch gekennzeichnet, daß das Papillomvirus-Hauptcapsid-Protein als Virus-ähnliches Partikel vorliegt.
  6. Protein nach Anspruch 5, dadurch gekennzeichnet, daß das Virus-ähnliche Partikel auch ein Papillomvirus-Nebencapsid-Protein enthält.
  7. Expressionsvektor, umfassend eine für das Protein nach Anspruch 4 codierende DNA.
  8. Transformante, enthaltend den Expressionsvektor nach Anspruch 7.
  9. Verfahren zur Herstellung des Proteins nach Anspruch 4, umfassend die Kultivierung der Transformante nach Anspruch 8 unter geeigneten Bedingungen.
  10. Antikörper, gerichtet gegen das Protein nach einem der Ansprüche 4-6.
  11. Verwendung der DNA nach einem der Ansprüche 1-3 als Reagens zur Diagnose.
  12. Verwendung des Proteins nach einem der Ansprüche 4-6 als Reagens zur Diagnose, Therapie und/oder Vakzinierung.
  13. Verwendung nach Anspruch 11 oder 12, wobei die Diagnose Papillom-

virus-Infektionen bzw. -Erkrankungen betrifft.

14. Verwendung nach Anspruch 12, wobei die Therapie und/oder Vakzinierung Papillomvirus-Infektionen bzw. -Erkrankungen betrifft.

DL314.seq from: 1 to: 380

```

AATCAGCTGTTTATTACAGTGGCTGATAACACTCGAAACACAAATTTCACTATTAGTGTT
1 -----+-----+-----+-----+-----+-----+ 60
TTAGTCGACAAATAATGTCACCGACTATTGTGAGCTTTGTGTTTAAAGTGATAATCACAA

N Q L F I T V A D N T R N T N F T I S V -

ACTACAGATGCTGGGGATATAAATGAATATACAGCCACAAATGTTAGAGAATTTTAAAGA
61 -----+-----+-----+-----+-----+-----+ 120
TGATGTCACGACCCCTATATTTACTTATATGTCGGTGTTTACAATCTCTTAAAAATTCT

T T D A G D I N E Y T A T N V R E F L R -

CATGTAGAAGAATTTCAAATATCAATTATTTACAATTATGTAAGGTCCTTTAGTTCCA
121 -----+-----+-----+-----+-----+-----+ 180
GTACATCTTCTTAAAGTTTATAGTTAATAAAATGTTAATACATTCCAAGGAAATCAAGGT

H V E E F Q I S I I L Q L C K V P L V P -

GAAGTTTATCACAGATAAATGCTATGAATTCAGGAATATTGGAAGAGTGGCAATTAGGA
181 -----+-----+-----+-----+-----+-----+ 240
CTTCAAAATAGTGTCTATTTACGATACTTAAGTCCTTATAACCTTCTCACCGTTAATCCT

E V L S Q I N A M N S G I L E E W Q L G -

TTTGTACCCACGCCAGACAATGCTGTACATGATACATATAGATACATTAAC TCAAAGCA
241 -----+-----+-----+-----+-----+-----+ 300
AAACATGGGTGCGGTCTGTTACGACATGTACTATGTATATCTATGTAATTGAGTTTTCGT

F V P T P D N A V H D T Y R Y I N S K A -

ACAAAATGTCCAGATGCTGCAACAGCTGAACAAAAGGAAGATCCTTTTGGTAAATTTACA
301 -----+-----+-----+-----+-----+-----+ 360
TGTTTTACAGGTCTACGACGTTGTGCGACTTGTTTTCTTCTAGGAAAACCATTTAAATGT

T K C P D A A T A E Q K E D P F G K F T -

TTTTGGAATGTAGATCTATC
361 -----+-----+ 380
AAAACCTTACATCTAGATAG

F W N V D L -

```

Fig. 1

MAP of: DL347.seq from: 1 to: 389

```

AATCAAATGTTTATTACTGTGGTAGACAACACACGAAACACTAATTCAGTATTCAGTC
1  -----+-----+-----+-----+-----+-----+ 60
TTAGTTTACAAATAATGACACCATCTGTTGTGTGCTTTGTGATTAAAGTCATAAAGTCAG
N Q M F I T V V D N T R N T N F S I S V -

TATACTGAAGGTGGACAAATAAAAGATATCAGGGACTATACATCTACACAGTTCAGGGAA
61 -----+-----+-----+-----+-----+ 120
ATATGACTTCCACCTGTTATTTTCTATAGTCCCTGATATGTAGATGTGTCAAGTCCCTT
Y T E G G Q I K D I R D Y T S T Q F R E -

TATTTAAGACATGTGGAAGAATATGAAATATCAGTCATATTGCAGTTATGTAAAATACCT
121 -----+-----+-----+-----+-----+ 180
ATAAATTCTGTACACCTTCTTATACTTTATAGTCAGTATAACGTCAATACATTTTATGGA
Y L R H V E E Y E I S V I L Q L C K I P -

TTGAAGGCTGAAGTTTTAGCCCAGATAAATGCAATGAACTCCTCGTTATTGGAAGACTGG
181 -----+-----+-----+-----+-----+ 240
AACTTCCGACTTCAAAAATCGGGTCTATTTACGTTACTTGAGGAGCAATAACCTTCTGACC
L K A E V L A Q I N A M N S S L L E D W -

CAATTAGGATTTGTGCCTACACCTGATAATCCCATTTCATGATACCTACAGATTTATTGAT
241 -----+-----+-----+-----+-----+ 300
GTTAATCCTAAACACGGATGTGGACTATTAGGGTAAGTACTATGGATGTCTAAATAACTA
Q L G F V P T P D N P I H D T Y R F I D -

TCCTTGGAACCCGATGCCCTGACAAAAATCCCCAAAAGAAAAACCTGACCCTTATGAA
301 -----+-----+-----+-----+-----+ 360
AGGAACCGTTGGGCTACGGGACTGTTTTTAGGGGTTTTCTTTTTGGACTGGGAATACTT
S L A T R C P D K N P P K E K P D P Y E -

GGCTTAAACTTTTGGGAATGTTGATCTAAC
361 -----+-----+-----+-----+ 389
CCGAATTTGAAAACCTTACAACCTAGATTG
G L N F W N V D L -
    
```

Fig. 2

DL369.seq from: 1 to: 380

```

AATCAGATGTTTGTACTGTAGCAGATAATACAAGAAATACTAATTTTAGTATTAGTGTA
1 -----+-----+-----+-----+-----+-----+-----+ 60
TTAGTCTACAAACAATGACATCGTCTATTATGTTCTTTATGATTAATAATCATAATCACAT

N Q M F V T V A D N T R N T N F S I S V -

TCTACAGATGGCAATATACCACAGGAATATGATTCTTCAAATATTAGAGAATTTTAAAGA
61 -----+-----+-----+-----+-----+-----+-----+ 120
AGATGTCTACCGTTATATGGTGTCTTATACTAAGAAGTTTATAATCTCTTAAAAATTCT

S T D G N I P Q E Y D S S N I R E F L R -

CACGTGGAAGAATATCAAATTTTCAGTAATTTTGCAGCTATGTAAAGTATCATTGGATCCA
121 -----+-----+-----+-----+-----+-----+-----+ 180
GTGCACCTTCTTATAGTTTAAAGTCATTAACCGTCGATACATTTTCATAGTAACCTAGGT

H V E E Y Q I S V I L Q L C K V S L D P -

GATATTTTAGCTCAAATCAATGCTATGAATTCTGGAATCTTAGAAGACTGGCAATTAGGG
181 -----+-----+-----+-----+-----+-----+-----+ 240
CTATAAAATCGAGTTTAGTTACGATACTTAAGACCTTAGAATCTTCTGACCGTTAATCCC

D I L A Q I N A M N S G I L E D W Q L G -

TTTATTCCTGTCCCAGATAACTCAGTTCATGACACATACAGATATATTAATTCATTAGCT
241 -----+-----+-----+-----+-----+-----+-----+ 300
AAATAAGGACAGGGTCTATTGAGTCAAGTACTGTGTATGTCTATATAATTAAGTAATCGA

F I P V P D N S V H D T Y R Y I N S L A -

ACTAAATGTCCAGCTAAAGTACCACCGACAGAAAGAGAAGATCCTTTTGCTAAGTATGTG
301 -----+-----+-----+-----+-----+-----+-----+ 360
TGATTTACAGGTCGATTTTCATGGTGGCTGTCTTTCTCTTCTAGGAAAACGATTCATACAC

T K C P A K V P P T E R E D P F A K Y V -

TTCTGGAATGTAGATCTAAC
361 -----+-----+ 380
AAGACCTTACATCTAGATTG

F W N V D L -
    
```

Fig. 3

GA1-3 from: 1 to: 437

```

AATCAACTGTTTATTACAGTGGTGGACAACACAAGAAACACAACTTCAGTATTAGTGTG
1  -----+-----+-----+-----+-----+-----+-----+ 60
TTAGTTGACAAATAATGTCACCACCTGTTGTGTTCTTTGTGTTTGAAGTCATAATCACAC

N Q L F I T V V D N T R N T N F S I S V -

TATAGTGAAGCAGGTAAAGTAAAGGATATTTTCAGATTATGATGCAAACAAATTTAGGGAA
61  -----+-----+-----+-----+-----+-----+-----+ 120
ATATCACTTCGTCCATTTTCATTTCCCTATAAAGTCTAATACTACGTTTGTTTAAATCCCTT

Y S E A G K V K D I S D Y D A N K F R E -

TATCAAAAACATGTAGAAGAATATGAAATTTCTTTAATATTACAACATATGTAAGATACCT
121 -----+-----+-----+-----+-----+-----+-----+ 180
ATAGTTTTTGTACATCTTCTTATACTTTAAAGAAATATAATGTTGATACATTCTATGGA

Y Q K H V E E Y E I S L I L Q L C K I P -

TTAAAAGCCGACGTGTTGGCACAATTAATGCAATGAATCCATCGTTATTAGAAGAGTGG
181 -----+-----+-----+-----+-----+-----+-----+ 240
AATTTTCGGCTGCACAACCGTGTTTAATTACGTTACTTAGGTAGCAATAATCTTCTCACC

L K A D V L A Q I N A M N P S L L E E W -

CAACTGGGGTTTGTACCTGCACCAGACAATCCATTGCAAAGTACCTATAGGTATATCGAT
241 -----+-----+-----+-----+-----+-----+-----+ 300
GTTGACCCCAAACATGGACGTGGTCTGTTAGGTAACGTTTCATGGATATCCATATAGCTA

Q L G F V P A P D N P L Q S T Y R Y I D -

AGCTTGGCCACACCATGTCCTGATAAAGTGCCTACCAAAGAAAAGGAAGATCCATATGCT
301 -----+-----+-----+-----+-----+-----+-----+ 360
TCGAACCGGTGTGGTACAGGACTATTTACGGATGGTTTCTTTTCTTCTAGGTATACGA

S L A T P C P D K V P T K E K E D P Y A -

CCGTTTACATTTTGAACGTTGATTTGACAGAAAGACTTTCCTTGGAACTGGATCAATAT
361 -----+-----+-----+-----+-----+-----+-----+ 420
GGCAAATGTA AACCTTGCAACTAACTGTCTTTCTGAAAGGAACCTTGACCTAGTTATA

P F T F W N V D L T E R L S L E L D Q Y -

TCTCTGGGACGAAAGTT
421 -----+----- 437
AGAGACCCTGCTTTCAA

S L G R K -
    
```

Fig. 4

GA3-1 from: 1 to: 437

```

AATCAGATGTTTATTACTGTTGTAGACAACACACGCAGCACAAATTTTAGTATATCAGTT
1 -----+-----+-----+-----+-----+-----+-----+ 60
TTAGTCTACAAATAATGACAACATCTGTTGTGTGCGTGTGTTAAAATCATATAGTCAA

N Q M F I T V V D N T R S T N F S I S V -

CACACAGAAAATCAAGATATATCTAAAATTGACAGTTTTGATGCAACTCAGTTTAGGGAA
61 -----+-----+-----+-----+-----+-----+ 120
GTGTGTCTTTTAGTTCTATATAGATTTTAACTGTCAAAACTACGTTGAGTCAAATCCCTT

H T E N Q D I S K I D S F D A T Q F R E -

TACTTAAGACATGTAGAGGAATATGAGATTTCTATAATATTACAGTTATGTAAGATTCCT
121 -----+-----+-----+-----+-----+-----+ 180
ATGAATTCTGTACATCTCCTTATACTCTAAGATATTATAATGTCAATACATTCTAAGGA

Y L R H V E E Y E I S I I L Q L C K I P -

CTGAAAGCAGAAGTCTTAGCACAAATTAATGCAATGAATTCTTCATTACTTGAAGACTGG
181 -----+-----+-----+-----+-----+-----+ 240
GACTTTCGTCTTCAGAATCGTGTTTAATTACGTTACTTAAGAAGTAATGAAGTTCTGACC

L K A E V L A Q I N A M N S S L L E D W -

CAACTTGGCTTTGTGCCGACGCCTGATAATCCAATTCATGATACGTACAGATATATTGAT
241 -----+-----+-----+-----+-----+-----+ 300
GTTGAACCGAAACACGGCTGCGGACTATTAGGTTAAGTACTATGCATGTCTATATAACTA

Q L G F V P T P D N P I H D T Y R Y I D -

TCTTTGGCAACACGGTGCCCTGATAAGACGCCTCCAAAGGAAAAACCTGATCCATATGAA
301 -----+-----+-----+-----+-----+-----+ 360
AGAAACCGTTGTGCCACGGGACTATTCTGCGGAGGTTTCCTTTTGGACTAGGTATACTT

S L A T R C P D K T P P K E K P D P Y E -

AAGTTACATTTTTGGAATGTGGACCTTACCGAACGTCTGTCTTTAGATTTAGATCAATAT
361 -----+-----+-----+-----+-----+-----+ 420
TTCAATGTAAAAACCTTACACCTGGAATGGCTTGCAGACAGAAATCTAAATCTAGTTATA

K L H F W N V D L T E R L S L D L D Q Y -

CCTCTGGGACGAAAGTT
421 -----+----- 437
GGAGACCCTGCTTTCAA

P L G R K -
    
```

Fig. 5

GA6-2 from: 1 to: 389

```

AATCAAATGTTTATTACAGTTGTAGACAACACGCGAAACACCAATTTTCAGTATATCTATA
1  -----+-----+-----+-----+-----+-----+ 60
TTAGTTTACAAATAATGTCAACATCTGTTGTGCGCTTTGTGGTTAAAGTCATATAGATAT

N Q M F I T V V D N T R N T N F S I S I -

TCTAGTAAAAATCAAGATATACAGCAAATACAATCATATGACTCACAAAAGTTTAGGGAA
61  -----+-----+-----+-----+-----+-----+ 120
AGATCACTTTTAGTTCTATATGTCGTTTATGTTAGTATACTGAGTGTTTTCAAATCCCTT

S S E N Q D I Q Q I Q S Y D S Q K F R E -

TATTTAAGGCACGTAGAAGAATATGAAATTTCTATTATTTTACAGTTGTGTAAGATTCCA
121 -----+-----+-----+-----+-----+-----+ 180
ATAAATTCCGTGCATCTTCTTATACTTTAAAGATAATAAAATGTCAACACATTCTAAGGT

Y L R H V E E Y E I S I I L Q L C K I P -

CTACAAGCAGAAGTTTTAGCACAAATAAATGCAATGAACCCCTCCTTACTAGAGGATTGG
181 -----+-----+-----+-----+-----+-----+ 240
GATGTTTCGTCTTCAAATCGTGTTTATTTACGTTACTTGGGGAGGAATGATCTCCTAACC

L Q A E V L A Q I N A M N P S L L E D W -

CAGTTAGGATTTGTGCCAACTCCCGATAATCCTATCCAGGACACATACAGATTTATTGAT -
241 -----+-----+-----+-----+-----+-----+ 300
GTCAATCCTAAACACGGTTGAGGGCTATTAGGATAGGTCCTGTGTATGTCTAAATAACTA

Q L G F V P T P D N P I Q D T Y R F I D -

TCCTTAGCTACCAGGTGTCCCGATAAAAAATCCACCAAAGGAAAAACCTGATCCTTATGAA
301 -----+-----+-----+-----+-----+-----+ 360
AGGAATCGATGGTCCACAGGGCTATTTTTAGGTGGTTTCCTTTTTGGACTAGGAATACTT

S L A T R C P D K N P P K E K P D P Y E -

AAATTAACATTCTGGAATGTAGATCTAAC
361 -----+-----+-----+-----+-----+ 389
TTTAATTGTAAGACCTTACATCTAGATTG

K L T F W N V D L -
    
```

Fig. 6

GA9-4 from: 1 to: 380

```

AATCAACTGTTTGTACAGTTGCAGATAATACAAGGAATACCAATTTTACTATAAGTGTA
1 -----+-----+-----+-----+-----+-----+-----+ 60
TTAGTTGACAAACAATGTCAACGTCTATTATGTTTCCTTATGGTTAAAATGATATTCACAT

N Q L F V T V A D N T R N T N F T I S V -

ACATCTAATGGTACCCCATAGCAGAATATGATTCCAAAACCTATTAGAGAATTTTAAAGG
61 -----+-----+-----+-----+-----+-----+-----+ 120
TGTTAGATTACCATGGGGGTATCGTCTTATACTAAGGTTTTGATAATCTCTTAAAAATTCC

T S N G T P I A E Y D S K T I R E F L R -

CACGTAGAAGAATATCAGTTGTCCATGATATTGCAATTATGTAAAGTACCTTTAAAAGCA
121 -----+-----+-----+-----+-----+-----+-----+ 180
GTGCATCTTCTTATAGTCAACAGGTACTATAACGTTAATACATTTTCATGGAAATTTTCGT

H V E E Y Q L S M I L Q L C K V P L K A -

GAAGTTTTATCCCAGATTAATGCTATGAATTCAGGTATTTTGGAGGAGTGGCAGTTAGGT
181 -----+-----+-----+-----+-----+-----+-----+ 240
CTTCAAAAATAGGGTCTAATTACGATACTTAAGTCCATAAAACCTCCTCACCGTCAATCCA

E V L S Q I N A M N S G I L E E W Q L G -

TTTGTGCCTACACCAGACAACCTCTGTACATGATATTTATAGATATATTGACTCAAAGCA
241 -----+-----+-----+-----+-----+-----+-----+ 300
AAACACGGATGTGGTCTGTTGAGACATGTACTATAAATATCTATATAACTGAGTTTTTCGT

F V P T P D N S V H D I Y R Y I D S K A -

ACAAAATGTCCCGATGCAGTGCCTGCAAAAGAAAAGAAGATCCATTTGACAAATATACA
301 -----+-----+-----+-----+-----+-----+-----+ 360
TGTTTTTACAGGGCTACGTCACGGACGTTTTTCTTTTCTTCTAGGTAAACTGTTTATATGT

T K C P D A V P A K E K E D P F D K Y T -

TTTTGGAATGTAGATCTAAC
361 -----+-----+ 380
AAAACCTTACATCTAGATTG

F W N V D L -
    
```

Fig. 7