



(12) 发明专利

(10) 授权公告号 CN 101688223 B

(45) 授权公告日 2013. 07. 31

(21) 申请号 200880016061. X

(22) 申请日 2008. 01. 29

(30) 优先权数据

07360019. 9 2007. 05. 15 EP

(85) PCT申请进入国家阶段日

2009. 11. 13

(86) PCT申请的申请数据

PCT/EP2008/051031 2008. 01. 29

(87) PCT申请的公布数据

W02008/138648 EN 2008. 11. 20

(73) 专利权人 特兰斯吉恩股份有限公司

地址 法国伊尔基希 - 格拉苏斯塔登

(72) 发明人 N · 西尔韦斯特雷 D · 施密特

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 左路 林晓红

(51) Int. Cl.

C12N 15/86 (2006. 01)

(56) 对比文件

US 6410298 B1, 2002. 06. 25, 全文.

WO 92/16636 A1, 1992. 10. 01, 全文.

US 2005/075303 A1, 2005. 04. 07, 全文.

WO 03/018055 A, 2003. 03. 06, 全文.

审查员 劳芳

权利要求书3页 说明书27页

序列表32页 附图3页

(54) 发明名称

用于多基因表达的载体

(57) 摘要

本发明提供用于表达至少第一和第二核酸分子的载体,其中所述第一和第二核酸分子在 40 个或更多个连续核苷酸的部分显示大约 80%或大于 80%的同源性百分比,并且中所述第一核酸分子和 / 或所述第二核酸分子被修饰,以使得所述的同源性百分比减低至小于 75%。本发明也涉及基本上分离的核酸分子,其包含如 SEQ ID NO: 9-15 和 66-69 任一所定义的核苷酸序列。还提供包含此种核酸分子或载体的宿主细胞和药物组合物以及它们用于治疗或预防目的用途。

1. 一种改良的 Ankara (MVA) 痘病毒载体, 其至少包含编码第一多肽的第一核酸分子和编码第二多肽的第二核酸分子, 其中:

- 所述第一核酸分子和所述第二核酸分子编码从密切相关的 HPV 血清型获得的同一种多肽,

- 所述第一和第二核酸分子分别从第一和第二天然核酸序列获得, 其中所述第一和第二天然核酸序列在 40 个或更多个连续核苷酸的部分显示 80% 或大于 80% 的同源性百分比, 和

- 所述第一和 / 或第二天然核酸序列被修饰, 以使得所获得的第一核酸分子和第二核酸分子的同源性百分比减低至小于 75%,

其中所述修饰是: 改变所述第一和 / 或第二天然核酸序列中的多个密码子, 以使得所获得的第一与第二核酸分子之间的同源部分限制在小于 8 个连续核苷酸。

2. 根据权利要求 1 的载体, 其中所述修饰是: 改变所述第一和 / 或第二天然核酸序列中的多个密码子, 以使得在所述 40 个或更多个连续核苷酸的部分中的同源性百分比减低至小于 75%。

3. 根据权利要求 2 的载体, 其中所述修饰是: 改变所述第一和 / 或第二天然核酸序列中的多个密码子, 以使得第一和 / 或第二天然核酸序列中的所述修饰在氨基酸水平是沉默的。

4. 根据权利要求 1 的载体, 其中所述修饰是: 改变所述第一和 / 或第二天然核酸序列中的多个密码子, 以使得所获得的第一与第二核酸分子之间的同源部分限制在小于 5 个连续核苷酸。

5. 根据权利要求 1 的载体, 其中第一和第二核酸分子独立地从选自 HPV-16、HPV-18、HPV-30、HPV-31、HPV-33、HPV-35、HPV-39、HPV-45、HPV-51、HPV-52、HPV-56、HPV-58、HPV-59、HPV-66、HPV-68、HPV-70 和 HPV-85 组成的组中的高危型乳头状瘤病毒获得。

6. 根据权利要求 5 的载体, 其中第一和第二核酸分子编码选自 E1、E2、E6 和 E7 组成的组中的早期乳头状瘤病毒多肽。

7. 根据权利要求 1 的载体, 其中从密切相关的 HPV 血清型获得的所述同一种多肽是 E2 多肽。

8. 根据权利要求 7 的载体, 其中密切相关的 HPV 血清型是 HPV-16、HPV-18、HPV-33 和 / 或 HPV-52。

9. 根据权利要求 8 的载体, 其中所述载体包含编码 HPV-16E2 多肽的第一核酸分子、编码 HPV-18E2 多肽的第二核酸分子、编码 HPV-33E2 多肽的第三核酸分子和编码 HPV-52E2 多肽的第四核酸分子。

10. 根据权利要求 9 的载体, 其中所述 HPV-16E2 多肽由 SEQ ID NO:7 所示的氨基酸序列所组成, 所述 HPV-18E2 多肽由 SEQ ID NO:8 所示的氨基酸序列所组成, 所述 HPV-33E2 多肽由 SEQ ID NO:70 所示的氨基酸序列所组成, 所述 HPV-52E2 多肽由 SEQ ID NO:71 所示的氨基酸序列所组成。

11. 根据权利要求 10 的载体, 其中所述第一核酸分子由 SEQ ID NO:12 所示的核苷酸序列所组成、所述第二核酸分子由 SEQ ID NO:13 所示的核苷酸序列所组成、所述第三核酸分子由 SEQ ID NO:67 所示的核苷酸序列所组成并且所述第四核酸分子由 SEQ ID NO:69 所示

的核苷酸序列所组成。

12. 根据权利要求 1 的载体,其中从密切相关的 HPV 血清型获得的所述同一种多肽是 E6 多肽、E7 多肽或 E6 和 E7 多肽。

13. 根据权利要求 12 的载体,其中从密切相关的 HPV 血清型获得的同一种多肽独立地从选自 HPV-16、HPV-18、HPV-30、HPV-31、HPV-33、HPV-35、HPV-39、HPV-45、HPV-51、HPV-52、HPV-56、HPV-58、HPV-59、HPV-66、HPV-68、HPV-70 和 HPV-85 组成的组中的高危型乳头状瘤病毒获得。

14. 根据权利要求 13 的载体,其中第一核酸分子编码 HPV-16E6 多肽并且第二核酸分子编码 HPV-18E6 多肽,其中第二核酸分子由 SEQ ID NO:14 所示的核苷酸序列所组成。

15. 根据权利要求 13 的载体,其中第一核酸分子编码 HPV-16E7 多肽并且第二核酸分子编码 HPV-18E7 多肽,其中第二核酸分子由 SEQ ID NO:15 所示的核苷酸序列所组成。

16. 根据权利要求 14 的载体,其中所述载体是 MVA 载体,第一核酸分子置于痘苗病毒 7.5K 启动子的控制下并且第二核酸分子在痘苗病毒 H5R 启动子的控制下,并且第一和第二核酸分子均插入所述 MVA 载体的缺失 III 中。

17. 根据权利要求 15 的载体,其中所述载体是 MVA 载体,所述第一核酸分子置于痘苗病毒 7.5K 启动子的控制下并且第二核酸分子在痘苗病毒 H5R 启动子的控制下,并且第一和第二核酸分子均插入所述 MVA 载体的缺失 III 中。

18. 根据权利要求 13 的载体,其中所述第一核酸分子编码 HPV-16E6 多肽与 HPV-16E7 多肽的融合物并且所述第二核酸分子编码 HPV-18E6 多肽与 HPV-18E7 多肽的融合物

19. 根据权利要求 13 的载体,其中所述载体包含编码 HPV-16E6 多肽的第一核酸分子、编码 HPV-18E6 多肽的第二核酸分子、编码 HPV-16E7 多肽的第三核酸分子和编码 HPV-18E7 多肽的第四核酸分子,其中所述第一、第二、第三和第四核酸分子不包含显示 75% 或大于 75% 同源性百分比的 40 个或更多个连续核苷酸的部分。

20. 根据权利要求 19 的载体,其中所述编码 HPV-18E6 多肽的第二核酸分子由 SEQ ID NO:14 组成,以及所述编码 HPV-18E7 多肽的第四核酸分子由 SEQ ID NO:15 组成。

21. 根据权利要求 19 的载体,其中所述编码 HPV-16E6 多肽的第一核酸分子和所述编码 HPV-18E6 多肽的第二核酸分子是相反方向。

22. 根据权利要求 1 的载体,其中所述第一和第二核酸分子的至少一个是选自 SEQ ID NO:10、11、12、13、14、15、66、67、68 和 69。

23. 一种宿主细胞,其包含根据权利要求 1 至 22 任一项的载体。

24. 一种药物组合物,其包含治疗有效量的根据权利要求 1 至 22 任一项的载体或根据权利要求 23 的宿主细胞和可药用介质。

25. 权利要求 24 的药物组合物,其中所述组合物包含适合给人类全身性或粘膜性施用的一种或多种佐剂。

26. 权利要求 25 的药物组合物,其中所述佐剂是咪唑喹啉化合物。

27. 根据权利要求 1 至 22 任一项的载体、根据权利要求 23 的宿主细胞或根据权利要求 24 至 26 任一项的组合物在制备用于治疗或预防与乳头状瘤病毒感染相关的病况的药物中的用途。

28. 根据权利要求 27 的用途,用于制备用于治疗或预防乳头状瘤病毒持续性感染、恶

变前病变和恶性病变的药物。

29. 根据权利要求 27 的用途,其中按照激发-增强治疗模式实施所述用途,并且其中所述载体或组合物用来激发或增强或激发并增强对象的免疫应答。

用于多基因表达的载体

[0001] 本发明涉及改造用于独立表达多个目的核苷酸序列的重组载体,其中所述目的核苷酸序列获得自同一种生物体或从密切相关的生物体。本发明涉及在多种原核及真核体外系统或在动物或人类对象中表达彼此显示同源性的多个核苷酸序列用于治疗或预防性目的重组核酸技术领域。本发明特别用于免疫治疗领域,尤其用于治疗或预防因感染性生物如乳头状瘤病毒和肝炎病毒所致的病理状况。

[0002] 重组 DNA 技术已经使得在培养的宿主细胞或在活生物中表达核苷酸序列成为可能。已经产生几种质粒 DNA 和病毒载体并用于多种目的,包括免疫接种、基因治疗、免疫治疗和在培养的细胞中表达。载体如腺病毒的和痘病毒载体具有容纳巨大克隆容量的优点,具有在类型广泛的宿主细胞中表达多个核苷酸序列的潜能。表达多个核苷酸序列可以有利于改善由所编码的多肽提供的治疗效力(例如联合体液免疫及细胞免疫)。有利的是产生单个重组载体,而非产生分别改造以表达每种想要的核苷酸序列的多个重组载体,至少对于生产步骤和行政批准而言是有利的。

[0003] 例如,就乳头状瘤病毒感染而言,有意义的将是表达来自几种乳头状瘤病毒基因型的免疫原性多肽,旨在拓宽或增强宿主的免疫应答,尤其在面临多重感染(例如 HPV-16 和 HPV-18)风险的对象中。然而,编码此类免疫原性多肽的核苷酸序列在相关 HPV 基因型之间是高度同源的。例如,在核苷酸水平显示 63% 的整体同源性的 HPV-16E6 和 HPV-18E6 序列反而包含具有超过 75% 极高同源性的可以危及从单个载体表达 HPV-16 和 HPV-18 基因的特殊区域。

[0004] 另外,当表达病毒源多肽时,同源核苷酸序列也可以因病毒基因组的整体组织作用而产生。病毒通过生物学机制(如内部翻译起始或移码)使用相同核苷酸序列来编码两种不同蛋白质是相当频繁的,例如相同 DNA 的序列在多于一个可读框中翻译。例如,在 HPV-16 基因组中,毗邻的 E1 和 E2 基因具有 59 个核苷酸的重叠区,其在不同的可读框中被翻译。换句话说,E1 基因的最后 59 个核苷酸与 E2 基因的头 59 个核苷酸重叠。

[0005] 然而,预期载体中同源序列的存在不利地影响载体的稳定性,尤其在载体生产步骤期间,这导致基因序列丢失,原因在于同源序列之间出现的重组事件。因此,在单个载体中表达 HPV-16E1 和 E2 基因涉及 59 个核苷酸共同部分的存在,其中所述共同部分可能潜在地造成同源重组事件并且最终造成 E1 与 E2 同源序列之间所包含的序列丢失。这类不希望的同源重组事件也可以在相同载体中表达 HPV-16 和 HPV-18 基因序列时发生。这种不稳定性问题可以导致载体原液不可用,尤其对于人类临床试验而言。

[0006] 在这个方面,W092/16636 提出将同源核苷酸序列以彼此相反的方向插入重组载体中,从而降低重组事件的可能性。然而,该策略是就痘苗病毒载体进行描述的,没有对其他重组载体(如腺病毒)进行描述。另外,以相反方向排列并非总是可能的,原因在于可能的启动子干扰和构建限制。

[0007] 本领域需要产生能够在宿主细胞或对象中表达从同一种生物体或从密切相关的生物体所获得的核苷酸序列的重组载体,其中所述生物在天然环境下含有高度同源的部分。本发明通过提供一种新策略满足这种需求,所述策略被设计为旨在以如下方式使重组

事件的可能性最小化,即利用遗传密码的简并性改变同源核苷酸序列两者之一或两者以使得它们的同源性比修饰前更少,同时不改变或不显著改变所编码的氨基酸序列。本发明允许避开同源重组的有害作用,其中所述有害作用可以在同源序列之间出现,尤其在载体生产步骤期间,并且导致同源序列之间所含有的核苷酸序列丢失。已经发现:本发明的载体在表达天然环境下共同拥有 59 个核苷酸的 100%同源部分的 E1 和 E2 乳头状瘤病毒基因方面出人意料地有效并且在载体生产步骤期间出人意料地稳定。还发现:本发明的载体在表达从密切相关的 HPV-16 和 HPV-18 基因型中获得的 E6 和 E7 基因方面出人意料地有效。

[0008] 这个技术问题通过提供如权利要求书中定义的实施方案加以解决。

[0009] 本发明的其他和另外方面、特征和优点将因以下对本发明当前优选实施方案的描述而显而易见。这些实施方案出于公开目的给出。

[0010] 因此,在第一方面,本发明提供载体,其至少包含编码第一多肽的第一核酸分子和编码第二多肽的第二核酸分子,其中

[0011] - 所述第一和第二核酸分子分别获得自第一和第二天然核酸序列,其中所述第一和第二天然核酸序列在一 40 个或更多个连续核苷酸的部分显示大约 80%或大于 80%的同源性百分比,和

[0012] - 在该载体中包含的所述第一核酸分子和 / 或所述第二核酸分子被修饰,以使得所述的同源性百分比减低至小于 75%。

[0013] 如整个申请书范围内通篇所用,术语“一(“a”和“an”)”以如此含义使用,从而它们意指所指向的化合物或步骤的“至少一个”、“至少第一个”、“一个或多个”或“多个”,除非上下文另外申明。例如,术语“一细胞”包括多个细胞,包括其混合物。

[0014] 如论在本文中何处使用,术语“和 / 或”包括“和”、“或”和“与该术语相联的要素的全部或任何其他组合”的意思。例如,“第一核酸分子和 / 或第二核酸分子”意指第一核酸分子或第二核酸分子或第一以及第二核酸分子。

[0015] 如本文中所用的术语“约”或“大约”意指在给定值或区间的 5%范围内,优选在 4%范围内并且更优选在 2%范围内。如本文中所用,当用来定义产物、组合物和方法时,术语“包括”意图指该产品、组合物和方法包括所指的组分或步骤,但不排除其他组分或步骤。“基本上由……组成”应当指排除任何具有重要意义的其他组分或步骤。因此,基本上由所提及组分组成的组合物将不排除痕量杂质和可药用载体。“由……组成”应当指将其他组分或步骤的更多痕量要素排除。例如,当多肽不含有除所提及氨基酸序列之外的任何氨基酸时,则该多肽“由一个氨基酸序列组成”。当多肽“基本上由一个氨基酸序列组成”时,这种氨基酸序列仅连同几个额外氨基酸残基,一般连同约 1 个至约 50 个或左右的额外残基一起存在。多肽“包含”一个氨基酸序列,当该氨基酸序列至少是此多肽的最终氨基酸序列的部分时。此种多肽可以具有几个直至几千个额外的氨基酸残基。此类额外的氨基酸残基可以在多肽转运方面发挥作用、促进肽产生或纯化、延长半寿期及其他作用。同理可以适用于核苷酸序列。

[0016] 如本文中所用,“载体”可以是能够送递并在宿主细胞或对象中表达至少第一和第二核酸分子的任意介质。载体可以是染色体外的(例如附加体)或整合的(用于并入宿主染色体)、自主复制或非自我复制的,多拷贝或低拷贝的,双链或单链的、裸露的或与其他分子复合的(例如与脂质或聚合物复合以形成颗粒结构(如脂质体、脂质复合体或纳米粒子))

的载体、包裹在病毒衣壳内的载体和固定在固相粒子上的载体等)。术语“载体”的定义也包括已经被修饰以允许偏好地靶向特定宿主细胞的载体。靶向载体的一个特征性特点是在其表面存在配体,其中所述配体能够识别并结合至细胞的且表面暴露的组分如细胞特异性标记(例如 HPV 感染的细胞)、组织特异性标记或肿瘤特异性标记。该配体可以被遗传地插入在载体表面存在的多肽中(例如,如 W094/10323 和 W002/96939 中描述的腺病毒纤突、五邻体、pIX 或如 EP 1 146 125 中描述的痘苗病毒 p14 基因产物)。

[0017] 在本发明的上下文中,术语“核酸”、“核酸分子”、“多核苷酸”和“核苷酸序列”以互换方式使用并且定义了多脱氧核糖核苷酸(DNA)或多核糖核苷酸(RNA)分子或其任意组合的任意长度的聚合物。该定义包括单链或双链的、线性或环状的、天然存在的或合成性核苷酸。另外,此类多核苷酸可以包含非天然存在性核苷酸(例如甲基化核苷酸和核苷酸类似物,如 US5,525,711、US 4,711,955 或 EPA 302 175 中所描述的那些)以及化学修饰(例如见 WO 92/03568 ;US 5,118,672),以提高核酸的体内稳定性、增强其递送或减低从宿主对象中的清除速率。若存在,可以在聚合之前或之后赋予修饰。

[0018] 术语“多肽”、“肽”和“蛋白质”在本文中以互换方式用来指包含由肽键连接的 9 个或更多个氨基酸的氨基酸残基聚合物。该聚合物可以是线性或环状的。在本发明的上下文中,“多肽”可以包含作为 L 立体异构物(天然存在形式)或 D 立体异构物的氨基酸并且可以包含除 20 种天然存在性氨基酸之外的氨基酸,如 β -丙氨酸、鸟氨酸或甲硫氨酸亚砷或在一个或多个 α -氨基、 α -羧基或侧链上被修饰(例如通过附接甲基、甲酰、乙酰基、糖基、磷酸基等)的氨基酸。作为一般提示,若氨基酸聚合物长(例如多于 50 个氨基酸残基),则它优选称作多肽或蛋白质。因此,“肽”指约 9 个至约 50 个氨基酸长度的片段。在本发明的上下文中,肽优选地包含天然存在性(或天然)蛋白质的所选区域,例如含有表位的其免疫原性片段。

[0019] 如本文中定义的术语“多肽”包括天然多肽以及修饰的多肽。如本文中所用,术语“天然”指从自然界来源中回收的区别于由人在实验室中以人工方式修饰或改造的材料。例如,天然多肽由存在于野生型生物或细胞的基因组中的基因编码。相反,修饰的多肽由这样的核酸分子编码,其中所述核酸分子已经在实验室中被修饰,从而与天然多肽的区别在于例如一个或多个氨基酸的插入、缺失或置换或这些可能性的任意组合。当构思了几个修饰时,它们可以涉及连续残基和/或非连续残基。由本发明构思的修饰的实例可以导致由该天然多肽显示的生物活性的改变。可以通过常规方法,如通过结构性和功能性分析法,鉴定对给定生物学活性关键的氨基酸,并且本领域技术人员可以轻易地确定能够减低或取消这种生物学活性的突变类型。此类修饰可以通过常规技术如位点定向诱变法进行。备选地,可以通过化学合成在自动化过程中产生编码修饰的多肽的合成性核酸分子(例如,如所附实施例部分中所述,从重叠的合成性寡核苷酸中装配出来)。

[0020] 如本文中所用的术语“获得”指从自然界中找到、分离、纯化或衍生的材料。“分离”意指从其自然环境中移出。“纯化”指该材料基本上不含与其天然关联的至少一种其他组分。“衍生”指与天然材料相比的一个或多个修饰(尤其突变,如置换、缺失和/或插入)。分离、纯化和修饰技术在本领域中是常规的并且取决于待获得的材料(例如从自然界的一个来源克隆核酸分子可以通过使用限制性酶、通过 PCR 或通过化学合成进行)。

[0021] 如本文中所用,术语“同源性”通常表述百分比并且指在至少 40 个连续核苷酸(例

如大约 40、45、50、55、57、58、59、60、70 或甚至更多个连续核苷酸) 部分的范围内彼此仍保留给定程度的相同性的核苷酸序列。“至少 80%”指大约 80%或大于 80% (例如超过 80% 的任意值, 有利地至少 85%, 合意地至少 87%, 优选至少 90%, 更优选至少 95%, 仍更优选至少 97% 至多到 100% 的序列同源性)。“小于 75%”指低于 75 的任意值, 例如大约 74、72、70、68、65、62、60% 或甚至更小。两个核苷酸序列之间的同源性百分比是由所述序列共有的相同位置数目的函数, 同时考虑了需要导入用于最佳比对的空位数和每个空位的长度。在本领域中多种计算机程序和数学算法可用于确定核苷酸序列之间的相同性百分比, 如 GCG Wisconsin 软件包和在国家生物技术信息中心 (National Center for Biotechnology Information, NCBI) 可公开获得并在印刷出版物中描述 (例如 Altschul 等人, 1990, J. Mol. Biol. 215, 403-410) 的基本局部比对搜索工具 (BLAST) 程序。

[0022] 作为起点, 可以使用第一和第二核酸分子之间在修饰之前的序列比对结果, 旨在揭示共同拥有 80% 或大于 80% 同源性百分比的具有 40 个或更多个连续核苷酸的一个或多个部分, 例如“同源”部分。在具体实施方案, 至少在 40 个或更多个 (例如大约 40、45、50、55、57、58、59、60、70 个或甚至更多个) 连续核苷酸的所述同源部分中修饰第一核酸分子或第二核酸分子或第一以及第二核酸分子的密码子使用模式 (例如借助密码子使用模式的简并性), 以使得所述的同源性百分比减低至小于 75% (例如大约 74、72、70、68、65、62、60% 或甚至更小)。

[0023] 尽管甲硫氨酸和色氨酸残基各自由独特的核酸三联体 (例如密码子) 编码, 然而可以使用不同的密码子来编码其余 18 种氨基酸 (遗传密码的简并性)。例如, 氨基酸由以下密码子编码: 丙氨酸 (Ala 或 A) 由密码子 GCA、GCC、GCG 和 GCU 编码; 半胱氨酸 (C 或 Cys) 由密码子 UGC 和 UGU; 天冬氨酸 (D 或 Asp) 由密码子 GAC 和 GAU 编码; 谷氨酸 (E 或 Glu) 由密码子 GAA 和 GAG 编码; 苯丙氨酸 (F 或 Phe) 由密码子 UUC 和 UUU 编码; 甘氨酸 (G 或 Gly) 由密码子 GGA、GGC、GGG 和 GGU 编码; 组氨酸 (H 或 His) 由密码子 CAC 和 CAU 编码; 异亮氨酸 (I 或 Ile) 由密码子 AUA、AUC 和 AUU 编码; 赖氨酸 (K 或 Lys) 由密码子 AAA 和 AAG 编码; 亮氨酸 (L 或 Leu) 由密码子 UUA、UUG、CUA、CUC、CUG 和 CUU 编码; 甲硫氨酸 (M 或 Met) 密码子 AUG 编码; 天冬酰胺 (N 或 Asn) 由密码子 AAC 和 AAU 编码; 脯氨酸 (P 或 Pro) 由密码子 CCA、CCC、CCG 和 CCU 编码; 谷氨酰胺 (Q 或 Gln) 由密码子 CAA 和 CAG 编码; 精氨酸 (R 或 Arg) 由密码子 AGA、AGG、CGA、CGC、CGG 和 CGU 编码; 丝氨酸 (S 或 Ser) 由密码子 AGC、AGU、UCA、UCC、UCG 和 UCU 编码; 苏氨酸 (T 或 Thr) 由密码子 ACA、ACC、ACG 和 ACU 编码; 缬氨酸 (V 或 Val) 由密码子 GUA、GUC、GUG 和 GUU 编码; 色氨酸 (W 或 Trp) 由密码子 UGG 编码和酪氨酸 (Y 或 Tyr) 由密码子 UAC 和 UAU 编码。

[0024] 减低在所述第一和第二核酸分子内存在的一个或多个同源部分中的同源性百分比可以通过利用遗传密码的简并性并修饰第一核酸分子和 / 或第二核酸分子中的密码子使用模式实现。修饰密码子使用模式一般通过将一个或多个“天然性”密码子替换为另一个密码子来进行。例如, 将编码 Arg 的 AGA 密码子替换为编码 Arg 的 CGC 密码子会降低密码子的 3 个位置中 2 个位置的同源性。不需要简并全部天然密码子, 因为可以用局部替换充分地减低同源性。另外, 修饰密码子使用模式可以在整个核酸分子范围内进行或可以限定于在修饰前存在的同源部分。希望地是, 在本发明的上下文中, 简并性在第一核酸分子中进行并且限定于同源部分。优选地, 在核苷酸水平修饰密码子使用模式并且所述修饰是在

氨基酸水平上沉默的,例如,当可能时,将每个“天然”密码子替换为编码相同氨基酸的密码子,以致此类修饰在编码的多肽中不翻译。更优选地,当可能时,以这样的方式修饰密码子使用模式,以致将第一和第二核酸分子之间的同源部分限制在小于9个或8个连续核苷酸,有利地至小于7个连续核苷酸,优选地至小于6个连续核苷酸,并且更优选至小于5个连续核苷酸。对密码子使用模式的修饰可以通过本领域技术人员已知的众多方式进行,如位点定向诱变(例如使用法国 Amersham, Les Ullis 的 Sculptor™ 体外诱变系统)、PCR 诱变、DNA 改组和通过化学合成技术(例如产生合成性核酸分子)。

[0025] 当本发明的载体包含多于两种核酸分子时,则可以修饰在该载体中所包含的从天然核酸序列获得的任意核酸分子(其中所述天然核酸序列在一40个或更多个连续核苷酸的部分与从中获得另一个核酸分子的至少一个其他天然核酸序列显示大约80%或大于80%的同源性百分比),以使得所述的同源性百分比减低至小于75%,即使得该载体中所包含的核酸分子对无一会包含显示大于75%的相同性百分比的40个或更多个连续核苷酸部分。

[0026] 可以使用从中获得所述核酸分子的每个(每对)天然性序列之间的序列比对结果,旨在揭示出显示80%或大于80%同源性百分比的一个或多个部分。随后,修饰一个或多个天然序列的序列,尤其通过简化密码子使用,从而减低至少同源部分中的同源性百分比至小于75%。最后,该载体中所包含的核酸分子均应当不包含与所述载体中所包含的任何其余核酸分子显示大于75%相同性百分比的40个或更多个(例如45、50、55、57、58、59、60、70或甚至更多个)连续核苷酸部分。

[0027] 如上所述,由该载体中包含的核酸分子编码的多肽可以具有或不具有与天然多肽相同的氨基酸序列。尤其,除了使用简并密码子以致于至少在核酸分子的同源部分中减低同源性的突变之外,该载体中所包含的所述核酸分子也可以包含导致或不导致修饰所编码多肽的氨基酸序列的额外突变。

[0028] 本发明的载体包括病毒载体以及非病毒载体(例如质粒DNA)。合适的非病毒载体包括质粒如 pREP4、pCEP4(Invitrogene)、pCI(Promega)、pCDM8(Seed,1987, Nature 329,840)、pVAX 和 pgWiz(Gene Therapy System Inc;Himoudi 等人,2002, J. Virol. 76, 12735-12746)。“病毒载体”在本文中根据本领域认可的意思使用。它指包含病毒来源的至少一个元件的任意载体,所述的元件包括完整病毒基因组、其部分或如下文所述的修饰的病毒基因组以及其产生的病毒粒子(例如包裹到病毒衣壳中以产生感染性病毒粒子的病毒载体)。本发明的病毒载体可以是复制型的或可以是遗传上失能的,从而是复制缺陷型的或复制受损的。如本文中所用的术语“有复制能力”包括选择复制型和条件复制型病毒载体,其中所述病毒被改造成在特定宿主细胞(例如肿瘤细胞)中更好地或选择性地复制。病毒载体可以从多种不同的病毒获得并且尤其从选自自由逆转录病毒、腺病毒、腺相关病毒(AAV)、痘病毒、疱疹病毒、麻疹病毒和泡沫病毒组成的组中的病毒获得。

[0029] 在一个实施方案中,本发明的载体是腺病毒载体(综述,见“Adenoviral vectors for gene therapy”,2002,编辑 D. Curiel 和 J. Douglas, Academic Press)。它可以从任意的人或动物腺病毒衍生。可以在本发明的环境中使用任意的血清型和亚群。可以更具体地提及亚群 A(例如血清型 12、18 和 31)、亚群 B(例如血清型 3、7、11、14、16、21、34 和 35)、亚群 C(例如血清型 1、2、5 和 6)、亚群 D(例如血清型 8、9、10、13、15、17、19、20、22-30、32、33、

36-39 和 42-47)、亚群 E(血清型 4) 和亚群 F(血清型 40 和 41)。特别优选的是人腺病毒 2(Ad2)、5(Ad5)、6(Ad6)、11(Ad11)、24(Ad24) 和 35(Ad35)。此类腺病毒从美国典型培养物保藏中心(American Type Culture Collection(ATCC, Rockville, Md.)) 可获得并且已经是描述其序列、组织和产生方法的无数出版物的主题,因而允许技术人员使用它们(见例如 US 6,133,028;US 6,110,735;WO 02/40665;WO 00/50573;EP 1016711;Vogels 等人,2003, J. Virol. 77,8263-8271)。

[0030] 本发明的腺病毒载体可以是复制型的。复制型的腺病毒载体的众多实例是本领域技术人员轻易可得的(见例如 Hernandez-Alcoceba 等人,2000, Human Gene Ther. 11, 2009-2024;Nemunaitis 等人,2001, Gene Ther. 8,746-759;Alemany 等人,2000, Nature Biotechnology 18,723-727;W000/24408;US5,998,205, W099/25860, US5,698,443, W000/46355, W000/15820 和 W001/36650)。

[0031] 备选地,本发明的腺病毒载体可以是复制缺陷型的(见例如 W094/28152)。优选的复制缺陷型腺病毒载体是 E1 缺陷的(例如 US6,136,594 和 US 6,013,638),其中 E1 缺失从大约第 459 位置延伸至第 3328 位置或从大约第 459 位置延伸至第 3510 位置(通过参考在 GeneBank 中以登录号 M73260 并在 Chroboczek 等人,1992, Virology 186,280-285 中披露的人腺病毒 5 型的序列)。可以通过缺失腺病毒基因组的额外部分(例如在非必需 E3 区域或在其他必需 E2、E4 区域中,如 Lusky 等人,1998, J. Virol. 72,2022-2032 中所述)进一步改善克隆容量和安全性。

[0032] 第一和第二核酸分子可以独立地插入本发明腺病毒载体的任意位置中,如 Chartier 等人(1996, J. Virol. 70,4805-4810) 中描述,并且相对于插入区域的天然转录方向在有义和 / 或反义方向上独立地安置。例如,它们可以被同时插入而替代 E1 区域,或备选地,插入一个核酸分子替代 E1 区域并且插入另一个核酸分子替代 E3 区域。

[0033] 在另一个实施方案中,本发明的载体是痘病毒载体(见例如 Cox 等人“Viruses in Human Gene Therapy”中, J. M. Hos 编辑, Carolina Academic Press)。该载体可以从痘病毒科(poxviridae) 的任意成员、尤其金丝雀痘病毒(例如,如 W095/27780 中所述的 ALVAC)、禽痘病毒(例如,如 Paoletti 等人,1995, Dev. Biol. Stand. 84,159-163 中所述的 TROVAC) 或痘苗病毒获得,优选后者。合适的痘苗病毒可以选自自由哥本哈根株(Goebel 等人,1990, Virology 179,247-266 和 517-563;Johnson 等人,1993, Virology 196,381-401)、Wyeth 株、NYVAC(见 W092/15672 和 Tartaglia 等人,1992, Virology 188,217-232) 和高度减毒的改良 Ankara(MVA) 株(Mayr 等人,1975, Infection 3,6-16) 构成的组。此类载体和产生方法在本领域技术人员可获得的众多文献中描述(例如 Paul 等人,2002, Cancer gene Ther. 9, 470-477;Piccini 等人,1987, Methods of Enzymology 153,545-563;US 4,769,330;US 4,772,848;US 4,603,112;US 5,100,587 和 US 5,179,993)。在本发明中使用的第一和第二核酸分子优选地插入痘病毒基因组的非必需基因座中,旨在重组痘病毒仍保持活力和感染性。非必需区域是非编码的基因间区或其失活或缺失不明显破坏病毒生长、复制或感染的任意基因。也可以构思在病毒必需基因座中的插入,只要在病毒粒子产生期间提供缺损功能,例如通过使用携带互补序列的辅助细胞系,其中所述互补序列与痘病毒基因组中所缺失的那些序列相对应。

[0034] 当使用哥本哈根痘苗病毒时,至少第一和第二核酸分子优选地插入胸苷激酶基

因 (tk) 中 (Hruby 等人, 1983, *Proc. Natl. Acad. Sci. USA* 80, 3411-3415; Weir 等人, 1983, *J. Virol.* 46, 530-537)。然而, 其他插入位点也是适合的, 例如在血凝素基因 (Guo 等人, 1989, *J. Virol.* 63, 4189-4198) 中、在 KIL 基因座中、在 u 基因中 (Zhou 等人, 1990, *J. Gen. Virol.* 71, 2185-2190) 或在痘苗病毒基因组的左侧端, 其中已经在文献中报道了在左侧端的多种自发或工程化缺失 (Altenburger 等人, 1989, *Archives Virol.* 105, 15-27; Moss 等人, 1981, *J. Virol.* 40, 387-395; Panicali 等人, 1981, *J. Virol.* 37, 1000-1010; Perkus 等人, 1989, *J. Virol.* 63, 3829-3836; Perkus 等人, 1990, *Virol.* 179, 276-286; Perkus 等人, 1991, *Virol.* 180, 406-410)。

[0035] 当使用 MVA 时, 可以在 MVA 基因组中存在的已鉴定的缺失 I 至 VII 的任一者中以及在 D4R 基因座中独立地插入至少第一和第二核酸分子 (Antoine 等人, 1998, *Virology* 244, 365-396), 不过优选在缺失 II 和 / 或 III 中的插入 (Meyer 等人, 1991, *J. Gen. Virol.* 72, 1031-1038; Sutter 等人, 1994, *Vaccine* 12, 1032-1040)。

[0036] 当使用禽痘病毒时, 尽管可以考虑在胸苷激酶基因中插入, 然而至少第一和第二核酸分子优选地导入位于 ORF7 和 9 之间的基因间区内 (见例如 EP 314569 和 US 5, 180, 675)。

[0037] 在本发明的另一个实施方案中, 至少第一和第二核酸分子独立地编码能够在表现或易于表现病理状况的对象中提供治疗活性或保护活性的多肽。如本文中所用的术语“对象”指脊椎动物, 特别是哺乳动物物种的成员, 尤其是家养动物、牲畜、体育动物和包括人在内的灵长类。此种多肽优选地选自免疫原性多肽和抗肿瘤多肽组成的组。

[0038] “免疫原性”多肽指这样的多肽, 所述多肽能够在表达该多肽的对象中诱导、刺激、发育或促进免疫系统。这种免疫应答可以是体液应答或细胞应答或体液应答和细胞应答。体液应答激发针对所讨论多肽的抗体产生, 而细胞应答激发 T- 辅助细胞和 / 或 CTL 应答和 / 或刺激细胞因子产生。一般, 可以通过本领域中作为标准的多种测定法体外或体内地评价多肽的免疫原属性 (对可用于评价免疫应答的启动和激活的技术的一般描述, 见例如 JWiley & Sons Inc, National Institute of Health 编辑的 Coligan 等人, *Current Protocols in Immunology* 最新版)。例如, 检测可以是比色法、荧光测定或放射法, 并且合适的技术包括 ELISA、蛋白质印迹法、放射免疫测定法和免疫沉淀测定法。细胞免疫的测量可以通过测量由激活的效应细胞 (包括从 CD4+ 和 CD8+ T- 细胞衍生的那些效应器细胞) 分泌的细胞因子谱 (例如通过 ELISpot 量化产生 IFN γ 的细胞)、通过确定免疫效应细胞的激活状态免疫 (例如借助传统 [^{3} H] 胸苷摄取的 T 细胞增殖测定法)、通过分析致敏对象中的抗原特异性 T 淋巴细胞 (例如在细胞毒性测定法中的肽特异性裂解) 进行。多肽的免疫原属性也可能在合适的动物模型中通过 ELISpot、基于四聚物的分析技术或用于分析 T 细胞介导的免疫的其他标准技术评价。合适的免疫原性多肽可以从乙型肝炎病毒 (HBV) (如 EP 414 374; EP 304 578 或 EP 198 474 中描述的 S、preS2 或 preS1- 多肽); 丙型肝炎病毒 (HCV) (例如 Core (C)、包膜糖蛋白 E1、E2、非结构多肽 NS2、NS3、NS4, 或 NS5 或其任意组合); 人免疫缺陷病毒 (HIV) (例如 gp120 或 gp160) 和乳头状瘤病毒 (如此后所述) 获得。

[0039] “抗肿瘤”多肽指能够导致抑制或净缩减肿瘤细胞扩张的多肽。多肽的抗肿瘤属性可以在适宜的动物模型中或在治疗的对象中由实际肿瘤大小在一段时间后缩小而确定。

多种方法可以用来估计肿瘤大小,所述方法包括放射学方法(例如,单光子和正电子发射计算机断层摄影术;通常见“Nuclear Medicine in Clinical Oncology”, Winkler, C. (编辑) Springer-Verlag, New York, 1986), 使用常规成像试剂(例如柠檬酸镓-67)的方法、免疫学方法(例如放射标记的针对特定肿瘤标记的单克隆抗体)以及超声波方法(见“Ultrasonic Differential Diagnosis of Tumors”, Kossoff 和 Fukuda, (编辑), Igaku-Shoin, New York, 1984)。备选地,可以基于肿瘤标记存在的减少而确定多肽的抗肿瘤属性。实例包括用于检测前列腺癌的 PSA 和用于检测结直肠癌和某些乳腺癌的 CEA。另外,可以在合适的动物模型中,例如使用以代表性人癌细胞系注射的小鼠,确定多肽的抗肿瘤属性。在明显的肿瘤已经发育后,用本发明的载体注射小鼠,并随后监测肿瘤生长速率的降低和存活增加。另外,多种体外方法可以用来预测体内肿瘤抑制作用。合适的抗肿瘤多肽包括肿瘤相关抗原(TAA)如 MUC-I (W092/07000; Acres 等人, 2005, Exp. Rev. Vaccines 4(4))、BRCA-I、BRCA-2 (Palma 等人, 2006, Critical Reviews in Oncology/haematology 27, 1-23)、癌胚抗原 CEA (Conroy 等人, 1995, Gene Ther ; 2, 59-65)、MAGE (W099/40188; De Plaen 等人, 1994, Immunogenetics 40, 360-369)、MART-I、gp100 (Bakker 等人, 1994, J. Exp. Med. 179, 1005-9)、PRAME、BAGE、Lage (又称作 NY Eos 1)、SAGE、HAGE (W099/53061)、GAGE (Robbins 和 Kawakami, 1996, Current Opinions in Immunol. 8, 628-36) 和前列腺特异性抗原 (PSA) (Ferguson 等人, 1999, Proc. Natl. Acad. Sci. USA. 96, 3114-9; W098/12302, W098/20117 和 W000/04149) 以及来自具有肿瘤诱导潜能的病毒(例如乳头状瘤病毒)中的病毒多肽。

[0040] 在本发明的另一个实施方案中,所述至少第一和第二核酸分子从同一种生物体或从密切相关的生物体获得。

[0041] 如本文中所用,术语“生物体”包括微生物,其优选地具有致病潜能,以及高等真核生物。术语“微生物”指真菌、细菌、原虫和病毒。病毒的代表性实例包括,而限于 HIV (HIV-I 或 HIV-2)、人疱疹病毒(例如 HSV1 或 HSV2)、巨细胞病毒 (CMV)、Epstein Barr 病毒 (EBV)、肝炎病毒(例如甲型肝炎病毒 (HAV)、HBV、HCV 和戊型肝炎病毒)、黄病毒(例如黄热病毒)、带状疱疹病毒 (VZV)、副粘病毒、呼吸道合胞病毒、副流感病毒、麻疹病毒、流感病毒和乳头状瘤病毒(如上文定义)。合适细菌的代表性实例包括,不限于奈瑟氏菌属 (Neisseria) (例如淋病奈瑟氏菌 (N. gonorrhea) 和脑膜炎奈瑟氏菌 (N. meningitidis)); 博德特菌属 (Bordetella) (例如百日咳博德特菌 (B. pertussis)、副百日咳博德特菌 (B. parapertussis) 和支气管炎博德特菌 (B. bronchiseptica))、分枝杆菌属 (Mycobacteria) (例如结核分枝杆菌 (M. tuberculosis)、牛分枝杆菌 (M. bovis)、麻风分枝杆菌 (M. leprae)、鸟分枝杆菌 (M. avium)、副结核分枝杆菌 (M. paratuberculosis)、耻垢分枝杆菌 (M. smegmatis)); 军团菌属 (Legionella) (例如嗜肺军团菌 (L. pneumophila)); 埃希氏菌属 (Escherichia) (例如肠毒性大肠杆菌、肠出血性大肠杆菌、肠致病性大肠杆菌); 志贺氏菌属 (Shigella) (例如宋内志贺氏菌 (S. sonnei)、痢疾志贺氏菌 (S. dysenteriae)、福氏志贺氏菌 (S. flexneri)); 沙门氏菌属 (Salmonella) (例如伤寒沙门氏菌 (S. typhi)、副伤寒沙门氏菌 (S. paratyphi)、猪霍乱沙门氏菌 (S. choleraesuis)、肠炎沙门氏菌 (S. enteritidis)); 李斯特菌属 (Listeria) (例如单核细胞增多李斯特菌 (L. monocytogenes)); 螺杆菌属 (Helicobacter) (例如幽门螺杆菌 (H. pylori)); 假单胞菌

属 (*Pseudomonas*) (例如铜绿假单胞菌 (*P. aeruginosa*)); 葡萄球菌属 (*Staphylococcus*) (例如金黄色葡萄球菌 (*S. aureus*)、表皮葡萄球菌 (*S. epidermidis*)); 肠球菌属 (*Enterococcus*) (例如粪肠球菌 (*E. faecalis*)、屎肠球菌 (*E. faecium*)); 芽孢杆菌属 (*Bacillus*) (例如炭疽芽孢杆菌 (*B. anthracis*)); 棒状杆菌 (*Corynebacterium*) (例如白喉杆菌 (*C. diphtheriae*)) 和衣原体 (*Chlamydia*) (例如沙眼衣原体 (*C. trachomatis*)、肺炎衣原体 (*C. pneumoniae*)、鹦鹉热衣原体 (*C. psittaci*))。寄生虫的代表性实例包括, 不限于疟原虫属 (*Plasmodium*) (例如恶性疟原虫 (*P. falciparum*)); 弓形虫属 (*Toxoplasma*) (例如刚地弓形虫 (*T. gondii*)); 利什曼原虫 (*Leshmania*) (例如硕大利什曼原虫 (*L. major*)); 肺囊虫属 (*Pneumocystis*) (例如卡氏肺囊虫 (*P. carinii*)) 和血吸虫属 (*Schistosoma*) (例如曼氏血吸虫 (*S. mansoni*))。真菌的代表性实例包括, 不限于假丝酵母属 (*Candida*) (例如白色假丝酵母 (*C. albicans*)) 和曲霉属 (*Aspergillus*)。高等真核生物优选地是包括人的哺乳动物。

[0042] “同一种生物体”定义为源自共同祖先并已经遵循相同进化途径的生物。代表性实例包括具有相同血清型或基因型的多种病毒分离株。例如, 将两个 HPV-16 分离株划分在这个类别中。“密切相关的生物体”定义为源自共同祖先但在进化期间已经趋异的生物。代表性实例包括具有不同血清型或基因型的病毒。例如, 将 HPV-16 和 HPV-18 划分在这个类别中。

[0043] 在优选的实施方案中, 从中获得至少第一和第二核酸分子的生物体是乳头状瘤病毒并且每种核酸分子编码一种乳头状瘤病毒多肽。“乳头状瘤病毒”可以定义为属于乳头状瘤病毒亚科的病毒并且该术语包括非人物种来源的动物乳头状瘤病毒以及人乳头状瘤病毒 (HPV), 其中所述非人物种包括, 但不限于牛、马、兔、绵羊、犬、非人灵长类和啮齿类。目前已经鉴定了多于 100 种 HPV 基因型 (Van Ranst 等人, 1992, *J. Gen. Virol.* 73, 2653; DeVilliers 等人, 2004, *Virology* 324, 17-27), 其中根据致癌潜能, 已经所述基因型划分为“低危” (LR) 和“高危” (HR) 血清型”。LR HPV 在感染对象中引起良性肿瘤, 而 HR 具有恶性进展的高风险。

[0044] 一般而言, 乳头状瘤病毒具有一个被蛋白质衣壳包围的约 7900 碱基对双链环状 DNA (见例如 Pfister, 1987, 在 *The papovaviridae: The Papillomaviruses*, Salzman 和 Howley 编辑, Plenum Press, New York, 第 1-38 页)。它们的基因组由三个功能区: 早期区 (E)、晚期区 (L) 和长调控区 (LCR) 组成。LCR 含有转录调节序列如增强子和启动子。晚期区编码 L1 和 L2 结构蛋白, 它们分别是主要衣壳蛋白和次要衣壳蛋白, 而早期区编码在细胞核中优势存在的控制病毒复制、转录和细胞转化的调节蛋白 (E1-E7)。更具体地, E1 蛋白是具有 ATP 依赖性解螺旋酶活性的 DNA 结合磷蛋白 (Desaintes 和 Demeret, 1996, *Semin. Cancer Biol.* 7, 339-347; Wilson 等人, 2002, *Virus Gene* 24, 275-290)。E2 蛋白是调节病毒基因转录并控制 DNA 复制的多功能性 DNA 结合磷蛋白 (Bechtold 等人, 2003, *J. Virol.* 77, 2021-8)。E4 编码的蛋白质结合并破坏细胞浆角蛋白网络并在病毒成熟中发挥作用。E5 蛋白的功能仍存在争论并且该蛋白质的表达往往在宿主染色体中病毒整合期间丧失。HR HPV 基因型的编码 E6 和 E7 的基因产物往往参与感染细胞的致癌性转化 (Kanda 等人, 1988, *J. Virol.* 62, 610-3; Vousden 等人, 1988, *Oncogene Res.* 3, 1-9; Bedell 等人, 1987, *J. Virol.* 61, 3635-40), 假定通过这些病毒蛋白分别与细胞肿瘤抑制蛋白基因产物 p53 和

视网膜母细胞瘤 (Rb) 蛋白的结合作用 (这在 B. N. Fields, D. M. Knipe 和 P. M. Howley (编辑), Virology, 第 3 版. Lippincott-Raven Press, New York, N. Y 中的 Howley, 1996, Papillomaviruses and their replication, 第 2045-2076 页综述)。参与天然 HPV-16E6 多肽与 p53 结合的氨基酸残基已经清晰地定义为从第 118 至 122 位残基的 (+1 是首个 Met 残基, 或从优选使用的第二 Met 残基开始的从第 111 至 115 位残基) (Crook 等人, 1991, Cell 67, 547-556) 并且参与天然 HPV-16E7 多肽与 Rb 结合的那些氨基酸残基位于第 21 至 26 位残基 (Munger 等人, 1989, EMBO J. 8, 4099-4105; Heck 等人, 1992, Proc. Natl. Acad. Sci. USA 89, 4442-4446)。

[0045] 优选地, 至少第一和第二核酸分子独立地从选自 HPV-16、HPV-18、HPV-30、HPV-31、HPV-33、HPV-35、HPV-39、HPV-45、HPV-51、HPV-52、HPV-56、HPV-58、HPV-59、HPV-66、HPV-68、HPV-70 和 HPV-85 组成的组中的高危型乳头状瘤病毒获得。

[0046] 如本文中所述的“乳头状瘤病毒多肽”指由源自乳头状瘤病毒基因组 / 来源的核酸分子编码的本领域确认的多肽。如与术语“多肽”相关联的上文定义, “乳头状瘤病毒多肽”包括天然的、修饰的乳头状瘤病毒多肽和其肽。乳头状瘤病毒来源包括, 但不限于生物样品 (例如从已经暴露于乳头状瘤病毒的对象中收集的生物样品、组织切片、尸检标本和组织培养物)、培养的细胞 (例如在 ATCC 获得的 CaSki 细胞) 以及在保藏机构、商业目录中获得或在文献中描述的重组材料。众多乳头状瘤病毒基因组的核苷酸序列和所编码多肽的氨基酸序列已经在文献中描述并且可在专业数据库例如 Genbank 中获得。就一般信息而言, HPV-16 基因组在 Genbank 中以登录号 NC 01526 和 K02718 描述; HPV-18 以登录号 NC 001357 和 X05015 描述; HPV-31 以登录号 J04353 描述; HPV-33 以登录号 M12732 描述; HPV-35 以登录号 NC 001529 描述; HPV-39 以登录号 NC 001535 描述; HPV-45 以登录号 X74479 描述; HPV-51 以登录号 NC 001533 描述; HPV-52 以登录号 NC 001592 描述; HPV-56 以登录号 X74483 描述; HPV-58 以登录号 D90400 描述; HPV-59 以登录号 NC 001635 描述; HPV-68 以登录号 X67160 和 M73258 描述; HPV-70 以登录号 U21941 描述并且 HPV-85 以登录号 AF131950 描述。

[0047] 由第一和 / 或第二核酸分子编码的乳头状瘤病毒多肽可以是早期、晚期乳头状瘤病毒多肽或其任意组合。早期乳头状瘤病毒多肽包括 E1、E2、E4、E5、E6 和 E7, 而晚期多肽可以是 L1 或 L2。数目众多的乳头状瘤病毒血清型的早期和晚期多肽的核苷酸和氨基酸序列在技术人员可获得的文献中描述。

[0048] 希望是, 至少第一和第二核酸分子独立地编码选自 E1、E2、E6 和 E7 组成的组中的早期多肽。出于说明目的, 天然 HPV-16 E1、E2、E6 和 E7 多肽的氨基酸序列分别在 SEQ ID NO: 1-4 中给出。然而, 本发明不限于这些示例性序列。实际上, 所述核苷酸和氨基酸序列可以在不同的乳头状瘤病毒分离株之间变动, 并且本发明的范围包括这种天然的遗传变异以及非天然的修饰, 如下文所述的那些修饰。在此后给出合适的修饰的乳头状瘤病毒多肽的示例描述 (例如在 SEQ ID NO: 5-8 和 64-65 中), 然而 (例如, 针对源自其他乳头状瘤病毒基因型的多肽, 通过序列比较) 调整本文中所述的修饰, 属于本领域技术人员的能力范围。

[0049] 在本发明中使用的合适乳头状瘤病毒 E1 多肽包括刺激病毒复制无效的突变体, 即它们刺激病毒复制的能力在统计上显著地低于相应的天然 E1 多肽 (例如小于 75%, 有利地小于 50%, 优选小于 10% 并且更优选小于 5%)。对于一般指南而言, 负责刺激病毒复

制的结构域位于 E1 的中央部分（例如 Hugues 和 Romanos, 1993, *Nucleic Acids Res.* 21, 5817-23）。在本领域技术人员可获得的文献中描述复制缺陷型 E1 多肽的代表性实例，例如在 Yasugi 等人（1997, *J. Virol.* 71, 5942-51）中。在本发明上下文中优选的修饰包括将 HPV-16 E1 多肽第 482 位置处的 Gly 残基置换为另一种残基（优选置换为 Asp 残基）（例如见 SEQ ID NO :5）或将 HPV-18 E1 多肽第 489 位置处的 Gly 残基置换为另一种残基（优选置换为 Asp 残基）（例如见 SEQ ID NO :6）。

[0050] 在本发明中使用的合适 E2 多肽包括与天然 E2 多肽相比，在转录激活作用和 / 或复制活性方面缺陷的突变体（例如小于 75%，有利地小于 50%，优选地小于 10% 并且更优选地小于 5%）。对于一般指南，负责转录激活作用和刺激复制的结构域位于 E2 的氨基末端部分（见 Dorf 等人, 1985, *Virology*, 145, 181-185; Kennedy 等人, 1991, *J. Virol.* 65, 2093-2097; Cole 等人, 1987, *J. Mol. Biol.* 193, 599-608; McBride 等人, 1989, *Proc. Natl. Acad. Sci. USA*, 86, 510-514）并且可以使用标准方法轻易地确定 E2 复制和转录活性的降低或缺乏（见例如 Sakai 等人, 1996, *J. Virol.* 70, 1602-1611）。在本领域技术人员可获得的文献中描述用在本发明中的合适缺陷型 E2 突变体，例如在 Demeret 等人（1995, *Nucleic Acids Res.* 23, 4777-4784）、Sakai 等人（1996, *J. Virol.* 70, 1602-1611）、Brokaw 等人（1996, *J. Virology* 70, 23-29）和 Ferguson 等人（1996, *J. Virology* 70, 4193-4199）中描述。在本发明上下文中优选的修饰包括将 HPV-16E2 第 39 位置处的 Glu 残基优选地置换为 Ala 残基（E39A）和 / 或将第 73 位置处的 Ile 残基优选地置换为 Ala 残基（I73A）（例如见 SEQ ID NO :7）。出于说明目的，在 HPV-16 E2 的第 39 和 73 位置处的 Glu 和 Ile 残基分别对应于 HPV-18 E2 的第 43 和 77 位置处的 Glu 和 Ile 残基（例如见 SEQ ID NO :8）。

[0051] 在本发明中使用的合适 E6 多肽包括在与细胞的肿瘤抑制基因产物 p53 结合方面无效的非致癌性突变体。非致癌性 E6 多肽的代表性实例在例如 Pirn 等人（1994, *Oncogene* 9, 1869-1876）和 Crook 等人（1991, *Cell* 67, 547-556）中描述。在本上下文中优选的修饰包括 HPV-16 E6 中残基 118 至 122（CPEEK）的缺失（例如见 SEQ ID NO :64）或 HPV-18E6 中 113 至 117（NPAEK）的缺失。

[0052] 在本发明中使用的合适 E7 多肽包括在与细胞的肿瘤抑制基因产物 Rb 结合方面无效的非致癌性突变体。非致癌性 E7 多肽的代表性实例在例如 Munger 等人（1989, *EMBO J.* 8, 4099-4105）、Heck 等人（1992, *Proc. Natl. Acad. Sci. USA* 89, 4442-4446）和 Phelps 等人（1992, *J. Virol.* 66, 2148-2427）中描述。在本上下文中优选的修饰包括 HPV-16 E7 中残基 21 至 26（DLICYE）的缺失（例如见 SEQ ID NO :65）或 HPV-18 E7 中 24 至 28（DLLCH）的缺失。

[0053] 此外，由至少第一和 / 或第二核酸分子编码的多肽（例如乳头状瘤病毒多肽）还可以包含额外的修饰，其中所述额外修饰有益于所编码多肽的加工、稳定性和 / 或溶解度的额外修饰，例如阻抑潜在切割位点、阻抑潜在位点和 / 或呈递在表达细胞的表面。例如，编码的多肽可以包含锚定于表达细胞的质膜内部的合适信号。实际上，先前已知膜呈递作用改善了 I 类 MHC 和 / 或 II 类 MHC 呈递，从而导致宿主免疫系统识别作用增强（见例如 W099/0388）。由于天然的早期乳头状瘤病毒多肽（E1、E2、E6 和 E7）是细胞核蛋白质（尽管不能确切地鉴定到常见核定位信号），可能有益的是使之定位在质膜处，旨在改善其免疫原潜力并且因此改善它们在宿主对象中的治疗效力。

[0054] 多肽在宿主表达细胞表面的高效膜呈递可以通过将该多肽融合于信号肽和膜锚定肽来实现。此类肽是本领域已知的。简而言之，信号肽通常存在于膜呈递或分泌的多肽的氨基端并且启动它们进入内质网 (ER)。信号肽包含 15 至 35 个基本上疏水的氨基酸，所述氨基酸随后由位于 ER 的特异性内肽酶去除以产生成熟的多肽。膜锚定肽通常是本质上高度疏水的并且起到将多肽锚定于细胞膜中的作用（见例如 Branden 和 Tooze, 1991, 在 *Introduction to Protein Structure* 中, 第 202-214 页, NY Garland）。可以在本发明上下文中使用的信号肽和膜锚定性肽的选择是繁多的。它们可以独立地从任意的分泌或膜锚定的多肽（例如细胞多肽或病毒多肽）如狂犬病病毒糖蛋白、HIV 病毒包膜糖蛋白或麻疹病毒 F 蛋白获得或可以是合成肽。插入信号肽的优选位点是用于翻译起始的密码子的氨基端下游并且插入膜锚定性肽的优选位点是羧基端，例如紧邻终止密码子的上游。根据需要，可以使用接头肽来将信号肽和 / 或膜锚定肽与编码的多肽连接。

[0055] 合适在本发明中使用的膜锚定性和缺陷型 E1 多肽的代表性实例在 SEQ ID NO: 5 (定义如实施例部分中所述 HPV-16 SS-E1*-TMR 多肽) 和在 SEQ ID NO: 6 (定义如实施例部分中所述 HPV-18 SS-E1*-TMF 多肽) 给出。合适在本发明中使用的膜锚定性和缺陷型 E2 多肽的代表性实例在 SEQ ID NO: 7 (定义如实施例部分中所述 HPV-16 SS-E2*-TMR 多肽) 和在 SEQ ID NO: 8 (定义如实施例部分中所述 HPV-18 SS-E2*-TMF 多肽) 给出。合适在本发明中使用的膜锚定性和非致癌性 E6 和 E7 多肽的代表性实例分别在 SEQ ID NO: 64 (定义如实施例部分中所述 HPV-16 SS-E6*-TMR 多肽) 和在 SEQ ID NO: 65 (定义如实施例部分中所述 HPV-16 SS-E7*-TMF 多肽) 给出。

[0056] 在特别优选的实施方案中，至少第一核酸分子和第二核酸分子编码从同一种 HPV 血清型中获得的两种不同乳头状瘤病毒多肽。

[0057] 在这个实施方案的优选方面，第一核酸分子编码 E1 多肽并且第二核酸分子编码 E2 多肽或反之亦然。希望是，编码 E1 和 E2 的核酸分子从 HPV-16 或从 HPV-18 获得。优选地，第一核酸分子编码包含 SEQ ID NO: 5 所示的氨基酸序列或基本上由该氨基酸序列组成、或由该氨基酸序列组成的多肽并且第二核酸分子编码包含 SEQ ID NO: 7 所示的氨基酸序列或基本上由该氨基酸序列组成、或由该氨基酸序列组成的多肽。优选地，第一核酸分子编码包含 SEQ ID NO: 6 所示的氨基酸序列或基本上由该氨基酸序列组成、或由该氨基酸序列组成的多肽并且第二核酸分子编码包含 SEQ ID NO: 8 所示的氨基酸序列或基本上由该氨基酸序列组成、或由该氨基酸序列组成的多肽。

[0058] 在天然环境（例如 HPV-16 或 HPV-18 基因组）中，编码 E1 的序列的 3' 部分与编码 E2 的序列的 5' 部分重叠 59 个核苷酸。预计这 59 个 100% 同源核苷酸的存在不利地影响同时表达编码 E1 和 E2 的核酸分子的载体的稳定性。同源重组可以在这些共同部分之间发生并且导致共同部分之间所含有的核苷酸序列丢失。

[0059] 根据本发明，在编码 E1 和 E2 的核酸分子中存在于修饰之前的 59 核苷酸重叠部分之间的 100% 同源性可以通过在所述核酸分子之一中简化密码子使用模式而减低至小于 75%。简并的序列的代表性实例在 SEQ ID NO: 9 中给出，其中在重叠 59 个核苷酸的 E1/E2 内的同源性减低至 69%（如图 1 中所示）并且编码 HPV-16 E1 和 E2 多肽的本发明优选载体包含 SEQ ID NO: 9 所示的核苷酸序列。相同策略可以适用于 HPV-18 编码 E1 和 E2 的序列中存在的重叠部分。可以在编码 E1 的第一核酸分子中导入此类简并的序列以替换天然重

叠的 59 个核苷酸（例如分别是 SEQ ID NO:10 和 11）。

[0060] 因此,本发明的优选载体包含第一核酸分子,其中所述第一核酸分子包含 SEQ ID NO:10 所示的核苷酸序列（其编码 SEQ ID NO:5 的 HPV-16 E1 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成;和第二核酸分子,其中所述第二核酸分子包含 SEQ ID NO:12 所示的核苷酸序列（其编码 SEQ ID NO:7 的 HPV-16 E2 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成。因此,本发明的另一个优选载体包含第一核酸分子,其中所述第一核酸分子包含 SEQ ID NO:11 所示的核苷酸序列（其编码 SEQ ID NO:6 的 HPV-18 E1 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成;和第二核酸分子,其中所述第二核酸分子包含 SEQ ID NO:13 所示的核苷酸序列（其编码 SEQ ID NO:8 的 HPV-18 E2 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成。更优选地,本发明的载体是 MVA 载体,第一（编码 E1 的）核酸分子置于痘苗病毒 7.5K 启动子的控制下并且第二（编码 E2 的）核酸分子在痘苗病毒 H5R 启动子的控制下,并且第一和第二核酸分子均插入所述 MVA 载体的缺失 III 中。

[0061] 本发明也涉及载体,其中所述载体包含编码 HPV-16 E1 多肽的第一核酸分子、编码 HPV-16 E2 多肽的第二核酸分子、编码 HPV-18 E1 多肽的第三核酸分子和编码 HPV-18 E2 多肽的第四核酸分子,其中所述第一、第二、第三和第四核酸分子不包含显示 75%或大于 75%同源性百分比的 40 个或更多个连续核苷酸的部分。优选地,所述 HPV-16 E1 多肽包含 SEQ ID NO:5 所示的氨基酸序列,所述 HPV-16 E2 多肽包含 SEQ ID NO:7 所示的氨基酸序列,所述 HPV-18 E1 多肽包含 SEQ ID NO:6 所示的氨基酸序列和 / 或所述 HPV-18 E2 多肽包含 SEQ ID NO:8 所示的氨基酸序列。

[0062] 在天然环境中,编码 HPV-16 和 HPV-18 E1 的序列包含显示 80%或大于 80%同源性百分比的 40 个或更多个连续核苷酸的几个部分。对于编码 HPV-16 和 HPV-18 E2 的序列而言也是如此。此外,毗邻的编码 E1 和 E2 的序列在 HPV-16 和 HPV-18 基因组中的一个大约 59 核苷酸的部分内重叠。在这种情况下,建议修饰编码的 HPV-18 E1 和 E2 的核酸分子的序列,从而减低与其 HPV-16 对应物的同源性至小于 75%,尤其在两种血清型共有的同源部分中。为此目的,可以比对 HPV-16 和 HPV-18 E1 和 E2 基因的核苷酸序列并且可以在核苷酸水平设计修饰旨在减低同源性至小于 8、7、6 或优选 5 个连续核苷酸。另外,可以进一步修饰 HPV-18 E1 序列以减低与重叠 HPV-18 E2 序列 5' 末端的 59 核苷酸部分的同源性至小于 75%。优选地,修饰了密码子使用,但是所述修饰在氨基酸水平不翻译,除了产生如本文中定义的修饰之外,例如产生缺陷型酶功能的修饰。分别在 SEQ ID NO:11 和 SEQ ID NO:13 中给出可以适宜地用作第三和第四核酸分子的编码 HPV-18 E1- 和 HPV-18 E1 的“简并”核苷酸序列的代表性实例。13)。本发明的优选载体包含第一核酸分子,其中所述第一核酸分子包含 SEQ ID NO:10 所示的核苷酸序列（其编码 SEQ ID NO:5 所示的 HPV-16 E1 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成;第二核酸分子,其中所述第二核酸分子包含 SEQ ID NO:12 所示的核苷酸序列（其编码 SEQ ID NO:7 所示的 HPV-16 E2 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成;第三核酸分子,其中所述第三核酸分子包含 SEQ ID NO:11 所示的核苷酸序列（其编码 SEQ ID NO:6 所示的 HPV-18 E1 多肽）或基本上由该核苷酸序列组成、或由该核苷酸序列组成;和第四核酸分子,其中所述第四核酸分子包含 SEQ ID NO:13 所示的核苷酸序列（其编码 SEQ ID NO:8 所示的 HPV-18

E2 多肽)或基本上由该核苷酸序列组成、或由该核苷酸序列组成。优选地,该载体是 MVA 载体,第一、第二、第三和第四核酸分子导入该 MVA 载体的缺失 III 中,第一和第三(编码 E1 的)核酸分子按照相反的方向布置,各自处在痘苗病毒 p7.5K 启动子的控制下,并且第二和第四(编码 E2 的)核酸分子按照相反的方向布置,各自处在痘苗病毒 pH5R 启动子的控制下。

[0063] 在另一个特别优选的实施方案中,至少第一核酸分子和第二核酸分子编码从同一种 HPV 血清型(例如密切相关的 HPV 血清型如 HPV-16、HPV-18、HPV-33 和 / 或 HPV-52 获得的两种不同的乳头状瘤病毒多肽。

[0064] 在该实施方案的第一方面,从密切相关的生物体获得的同一种多肽优选地是 E2 多肽。编码的 E2 多肽优选地被修饰,从而是膜锚定的并且对于病毒复制无效,如本文中定义。在天然环境中,多种基因型的编码 E2 的序列在核苷酸水平显示高度同源性,尤其在最保守的区域中。预计这些同源序列的存在不利地影响共表达两种或多种(例如 3、4 种或甚至更多种)E2 基因(例如来自 HR HPV(如 HPV-16、HPV-18、HPV-33 和 HPV-52)的 E2 基因)的载体的稳定性。同源重组可以在这些共同部分之间发生并且导致共同部分之间所含有的核苷酸序列丢失,并且因此导致基因沉默。

[0065] 根据本发明,在本发明载体中包含的编码 E2 多肽的核酸分子可以通过简化密码子使用模式进行修饰,从而减低同源性至小于 75%,尤其在高度同源的部分中。在 SSEQ ID NO:13、66、67、68 和 69 中给出编码 E2 多肽的简并核酸分子的代表性实例。更具体地,SEQ ID NO:13 编码膜呈递和复制缺陷型的 HPV-18 E2 多肽,其中所述多肽的核苷酸序列已经如此设计,从而减低与其编码 E2 的对应物的同源性至小于 8 或 7 个连续核苷酸。SEQ ID NO:66 和 67 均编码复制缺陷型 HPV-33 E2 多肽(它在 SEQ ID NO:67 中还是膜呈递的),其中所述多肽的核苷酸序列已经如此设计,从而减低与编码 E2 的其余对应物的同源性至小于 8 或 7 个连续核苷酸。SEQ ID NO:68 和 69 均编码复制缺陷型 HPV-52E2 多肽(它在 SEQ ID NO:69 中还是膜呈递的),其中所述多肽的核苷酸序列已经如此设计,从而减低与编码 E2 的其余对应物的同源性至小于 8 或 7 个连续核苷酸。然而,本发明不限于这些示例性序列并且可以基于这个原则设计编码如上文定义的 E2 乳头状瘤病毒多肽的简并核酸分子的备选形式。

[0066] 本发明的优选载体包含编码如本文中定义的 HPV-16E2 多肽(例如包含 SEQ ID NO:7 所示氨基酸序列的膜呈递和复制缺陷型的 E2 多肽)的第一核酸分子;编码如本文中定义的 HPV-18E2 多肽(例如包含 SEQ ID NO:8 所示氨基酸序列的膜呈递和复制缺陷型的 E2 多肽)的第二核酸分子;编码如本文中定义的 HPV-33E2 多肽(例如包含 SEQ ID NO:70 所示氨基酸序列的膜呈递和复制缺陷型的 E2 多肽)的第三核酸分子;编码如本文中定义的 HPV-52 E2 多肽(例如包含 SEQ ID NO:71 所示氨基酸序列的膜呈递和复制缺陷型的 E2 多肽)的第四核酸分子。更优选地,第一核酸分子包含 SEQ ID NO:12 所示的核苷酸序列或基本上由该核苷酸序列组成;第二核酸分子包含 SEQ ID NO:13 所示的核苷酸序列或基本上由该核苷酸序列组成;第三核酸分子包含 SEQ ID NO:67 所示的核苷酸序列或基本上由该核苷酸序列组成;和/或第四核酸分子包含 SEQ ID NO:69 所示的核苷酸序列或基本上由该核苷酸序列组成。更优选地,本发明的载体是 MVA 载体并且编码 E2 的第四核酸分子插入缺失 III 中。甚至更优选地,第一和第二核酸分子处在痘苗病毒 H5R 启动子的控制下并且置

于彼此倒置的方向上,而第三和第四核酸分子处在痘苗病毒 p7.5K 启动子的控制下并且置于彼此倒置的方向上。

[0067] 在该实施方案的另一个方面,从密切相关的生物体获得的同一种多肽优选地是 E6 多肽、E7 多肽或 E6 和 E7 多肽。E6 和 E7 可以独立地表达或表达为融合多肽。编码的 E2 多肽优选地被修饰,从而是膜锚定和非致癌的,如本文中定义。

[0068] 在天然环境中,HPV-16 和 HPV-18 的天然 E6 序列在核苷酸水平具有 63%的同源性而 HPV-16 和 HPV-18 的天然 E7 序列彼此 57%同源。然而,在两种情况下,HPV-16 和 HPV-18 天然序列共享了显示 80%或大于 80%同源性的 40 个或更多个连续核苷酸的几个部分(见图 2)。预计这些同源序列的存在不利地影响共表达 HPV-16 和 HPV-18 E6 和 / 或 E7 基因的载体的稳定性。同源重组可以在这些同源部分之间发生并且导致同源部分之间所含有的核苷酸序列丢失,并且因此导致基因沉默。

[0069] 根据本发明,HPV-16 和 / 或 HPV-18 E6 和 E7 多肽的核酸分子可以通过简化密码子使用模式进行修饰,从而减低同源性至小于 75%,尤其在同源部分中。在 SEQ ID NO:14 中给出编码 HPV-18 E6 多肽的简并核酸分子的代表性实例并且在 SEQ ID NO:15 中给出编码 HPV-18 E7 多肽的简并的修饰核酸分子的代表性实例。更具体地,已经如此设计 SEQ ID NO:14 和 SEQ ID NO:15,从而减低与 HPV-16 对应物的同源性至小于 8、7、6 个或优选 5 个连续核苷酸,同时编码 HPV-18 的膜锚定和非致癌性 E6 和 E7 多肽。然而,可以基于这个原则设计编码如上文定义的 E6 和 / 或 E7 乳头状瘤病毒多肽的简并核酸分子的备选形式。

[0070] 本发明的优选载体包含编码如本文中定义的(例如膜锚定的和非致癌性)HPV-16 E6 多肽的第一核酸分子和编码如本文中定义的(例如膜锚定的和非致癌性)HPV-18 E6 多肽的第二核酸分子,其中所述第二核酸分子包含 SEQ ID NO:14 所示的核苷酸序列或基本上由该核苷酸序列组成。本发明的优选载体包含编码如本文中定义的(例如膜锚定的和非致癌性)HPV-16 E6 多肽的第一核酸分子和编码如本文中定义的(例如膜锚定的和非致癌性)HPV-18 E6 多肽的第二核酸分子,其中所述第二核酸分子包含 SEQ ID NO:15 所示的核苷酸序列或基本上由该核苷酸序列组成。更优选地,本发明的载体是 MVA 载体,第一核酸分子置于痘苗病毒 7.5K 启动子的控制下并且第二核酸分子在痘苗病毒 H5R 启动子的控制下,并且第一和第二核酸分子均插入所述 MVA 载体的缺失 III 中。

[0071] 本发明也涉及载体,所述载体包含编码 HPV-16 E6 多肽与 HPV-16 E7 多肽的融合物的第一核酸分子和编码 HPV-18 E6 多肽与 HPV-18 E7 多肽的融合物的第二核酸分子,其中所述第一和第二核酸分子不包含显示大约 75%或大于 75%的同源性百分比的 40 个或更多个连续核苷酸的部分。

[0072] 本发明也涉及载体,其中所述载体包含编码 HPV-16 E6 多肽的第一核酸分子、编码 HPV-18 E6 多肽的第二核酸分子、编码 HPV-16 E7 多肽的第三核酸分子和编码 HPV-18 E7 多肽的第四核酸分子,其中所述第一、第二、第三和第四核酸分子不包含显示 75%或大于 75%同源性百分比的 40 个或更多个连续核苷酸的部分。优选地,第二核酸分子包含 SEQ ID NO:14 或基本上由 SEQ ID NO:14 组成、或由 SEQ ID NO:14 组成和 / 或第四核酸分子包含 SEQ ID NO:15 或基本上由 SEQ ID NO:15 组成、或由 SEQ ID NO:15 组成。更优选地,本发明的载体是 MVA 载体,第一和第二核酸分子各自以倒置方向置于痘苗病毒 7.5K 启动子的控制下并且第三和第四核酸分子各自以倒置方向置于痘苗病毒 H5R 启动子的控制下,并且所

述第一、第二、第三和第四核酸分子插入所述 MVA 载体的缺失 III 中。

[0073] 在另一个方面,本发明还提供一种基本上分离的核酸分子,其包含在任意的 SEQ ID NO :9、10、11、12、13、14、15、66、67、68 或 69 所示的核苷酸序列、基本上由所述核苷酸序列组成或由所述核苷酸序列组成。

[0074] 在本发明的另一个实施方案中,本发明载体中所包含的第一和第二和(若存在)第三和第四核酸分子处于适合在宿主细胞或对象表达所编码多肽的形式,这意指将它们置于对其表达必需的调节序列的控制下。

[0075] 如本文中所用,术语“调节序列”指允许、有助于或调节给定宿主细胞中核酸分子表达(包括该核酸或其衍生物之一(例如 mRNA)的复制、重复、转录、剪接、翻译、稳定性和/或运输至此宿主细胞)的任意序列。在本发明的上下文中,调节序列与待表达的核酸分子“有效连接”,例如它们处于允许在宿主细胞或对象内表达的一种功能性联系中。此类调节序列是本领域熟知(见例如 Goeddel,1990, Gene Expression Technology :Methods in Enzymology 185, Academic Press, San Diego)。本领域技术人员应当理解,调节序列的设计可能取决于这类因素,如载体类型、宿主细胞、想要的表达水平等。

[0076] 启动子是特别重要的并且本发明包括了指指导核酸分子在多种类型宿主细胞中表达的组成型启动子和指导仅在某些宿主细胞中(例如组织特异性调节序列)或应答于特定事件或外源因素时(例如因温度、营养添加物、激素或其他配体)表达的那些启动子的用途。用于真核系统中组成型表达的合适启动子包括病毒启动子,如 SV40 启动子、巨细胞病毒(CMV)立即早期启动子或增强子(Boshart 等人,1985, Cell 41,521-530)、腺病毒早期和晚期启动子、单纯疱疹病毒(HSV)-1 的胸苷激酶(TK)启动子和逆转录病毒长末端重复序列(例如 MoMuLV 和罗氏肉瘤病毒(RSV)LTR)以及细胞的启动子如磷酸甘油激酶(PGK)启动子(Hitzeman 等人,1983, Science 219,620-625;Adra 等人,1987, Gene 60,65-74)。用于驱动痘病毒载体中核酸分子表达的合适启动子包括痘苗病毒的 7.5K、H5R、TK、p28、p11 或 K1L 启动子。备选地,可以使用人造启动子如 Chakrabarti 等人(1997, Bio 技术 23,1094-1097),Hammond 等人(1997, J. Virological Methods 66,135-138)和 Kumar 和 Boyle(1990, Virology 179,151-158)中描述的那些人造启动子以及早期与晚期痘病毒启动子之间的嵌合启动子。

[0077] 诱导型启动子受外源性供应的化合物调节,并且包括,不限于锌诱导型金属硫蛋白(MT)启动子(Mc Ivor 等人,1987, Mol. Cell Biol. 7,838-848)、地塞米松(Dex)-诱导型小鼠乳腺肿瘤病毒(MMTV)启动子、T7 聚合酶启动子系统(WO 98/10088)、蜕皮激素昆虫启动子(No 等人,1996, Proc. Natl. Acad. Sci. USA 93,3346-3351),四环素抑制型启动子(Gossen 等人,1992, Proc. Natl. Acad. Sci. USA 89,5547-5551)\四环素诱导型启动子(Kim 等人,1995, J. Virol. 69,2565-2573)、RU486 诱导型启动子(Wang 等人,1997, Nat. Biotech. 15,239-243 和 Wang 等人,1997, Gene Ther. 4,432-441)、雷帕霉素诱导型启动子(Magari 等人,1997, J. Clin. Invest. 100,2865-2872)和来自大肠杆菌的 lac、TRP 和 TAC 启动子。在本发明上下文中使用的调节序列也可以是组织特异性的,以驱动核酸分子在需要治疗益处的特定组织中表达。合适的启动子可以取自肿瘤细胞中偏好性表达的基因。此类基因可以通过展示和比较基因组杂交法鉴定(见例如 US 5,759,776 和 5,776,683)。

[0078] 本领域技术人员将认识到,控制本发明载体中所包含的核酸分子表达的调节元

件还可以包含用于宿主细胞或对象中正确起始、调节和 / 或终止转录（例如聚腺苷酸转录中止序列）、mRNA 运输（例如核定位信号序列），加工（例如剪接信号）、稳定性（例如内含子和非编码性 5' 和 3' 序列）和翻译（例如三重体前导序列、核糖体结合位点、Shine-Dalgarno 序列等）的额外元件。

[0079] 在另一个方面，本发明提供包含上述载体的感染性病毒粒子。这里将不试图详细地描述已知用于产生感染性病毒粒子的多种方法。一般，此类病毒粒子通过如此方法产生，所述方法包括步骤 (a) 在适宜的细胞系中导入病毒载体，(b) 在合适的条件下培养该细胞系，从而允许产生所述感染性病毒粒子，从所述细胞系的培养中回收产生的感染性病毒粒子并且任选地纯化所述回收的感染性病毒粒子。

[0080] 当病毒载体具有缺陷性时，一般通过使用辅助病毒在互补细胞系中产生感染性颗粒，其中所述辅助病毒以反式方式提供无功能的病毒基因。例如，用于弥补缺失 E1 的腺病毒载体的合适细胞系包括 293 细胞 (Graham 等人, 1997, J. Gen. Virol. 36, 59-72)、PER-C6 细胞 (Fallaux 等人, 1998, Human Gene Ther. 9, 1909-1917) 和 HER96 细胞。适于增殖痘病毒载体的细胞是禽细胞，并且最优选地是从鸡胚制备的原代鸡胚成纤维细胞 (CEF)，其中所述鸡胚从受精卵获得。生产者细胞可以在常规发酵反应器、烧瓶和培养皿中，在适宜的温度、pH 和氧含量条件下培养。

[0081] 感染性病毒粒子可以从培养上清液或从裂解后的细胞回收。它们可以根据标准技术（如在例如 W096/27677、W098/00524、W098/22588、W098/26048、W000/40702、EP1016700 和 W000/50573 中描述的层析、超离）进一步纯化。

[0082] 在另一个方面，本发明提供包含上述核酸分子、载体或感染性病毒粒子的宿主细胞。如本文中所述的术语“宿主细胞”定义了可以或已经是本发明载体或感染性病毒粒子的受者的任意细胞和此类细胞的子代。该术语应当广义地理解，从而包括分离的细胞，一组细胞以及细胞的特定组织形式，例如在组织或器官中。此类细胞可以是原代、转化的或培养的细胞。

[0083] 本发明上下文中的宿主细胞包括原核细胞（例如大肠杆菌、芽孢杆菌属、李斯特菌属）、低等真核细胞如酵母（例如酿酒酵母 (*Saccharomyces cerevisiae*)、粟酒裂殖酵母 (*Saccharomyces pombe*) 或巴斯德毕赤酵母 (*Pichia pastoris*) 和其他真核细胞如昆虫细胞、植物和高等真核细胞，特别优选哺乳动物细胞（例如人或非人细胞）。合适宿主细胞的代表性实例包括，但不限于 BHK（乳仓鼠肾）细胞、MDCK 细胞 (Madin-Darby 犬肾细胞系)、CRFK 细胞 (Crandell 猫肾细胞系)、CV-1 细胞（非洲猴肾细胞系）、COS（例如 COS-7）细胞、中国仓鼠卵巢 (CHO) 细胞、小鼠 NIH/3T3 细胞、HeLa 细胞和 Vero 细胞。术语“宿主细胞”也包括能够弥补本发明复制缺陷型载体（例如腺病毒载体）的至少一个缺损功能的互补细胞，如上文所提及的那些细胞。

[0084] 宿主细胞可以用于通过重组手段产生由本发明的载体或感染性颗粒中所包含核酸分子编码的多肽。此类技术是本领域熟知（见例如 Ausubel, Current Protocols in Molecular Biology, John Wiley, 1987-2002；和 Sambrook 等人, Molecular Cloning, A Laboratory Manual 最新版, Cold Spring Harbor Laboratory Press）。

[0085] 在另一个方面，本发明提供包含上述核酸分子、载体、感染性病毒粒子或宿主细胞（如本文也称作“活性物质”）或其任意组合的组合物。有利地，该组合物是包含治疗有效

量的活性物质和可药用介质的药物组合物。

[0086] 如本文中所述的术语“可药用介质”意图包括与药物施用兼容的任意和全部的载体、溶剂、稀释剂、赋形剂、分散介质、包衣剂、抗菌剂和抗真菌剂和吸收延迟剂等。如本文中所述，“治疗有效量”是足以减缓通常与对象中想要治疗或预防的病理状况关联的一种或多种症状的剂量。当涉及预防性用途时，该术语意指足以预防或延迟对象中建立病理状况的剂量。例如，治疗有效量可以是足以在治疗的对象中诱导或增强受免疫应答的那个量或是足以减轻、改进、稳定、逆转、延缓或延迟病理状况进展（例如对象中损害或肿瘤的尺寸缩小或退化损害，感染的对象中病毒感染的逆转）的那个量。

[0087] 希望地，本发明的组合物包含在所用的剂量和浓度上无毒的一种或多个载体和/或稀释剂。此类载体和/或稀释剂优选地选自通常用来配制处于全身性或粘膜施用的单位剂量或多剂量形式的组合物的那些载体和/或稀释剂。合适载体可以溶剂、含有例如水、乙醇、多元醇（例如丙三醇、丙二醇、液体聚乙二醇等）的分散介质、植物油或其合适的混合物。稀释剂优选地是等渗的、低渗的或略微高渗的，并且具有相对低的离子强度。合适稀释剂的代表性实例包括无菌水、生理盐水（例如氯化钠）、Ringer's 溶液、葡萄糖、海藻糖或蔗糖溶液、Hank's 溶液和其他生理平衡盐水溶液（见例如最新版 Remington: The Science and Practice of Pharmacy, A. Gennaro, Lippincott, Williams & Wilkins）。适当地调节并缓冲本发明组合物的 pH，旨在适合用在人或动物中，优选地在生理性或略微碱性的 pH（在大约 pH 7.5 至大约 pH 9 之间，特别优选大约 8 或 8.5 的 pH）。合适的缓冲液包括磷酸盐缓冲液（例如 PBS）、碳酸氢钠缓冲液和/或 Tris 缓冲液。

[0088] 本发明的组合物可以处于多种形式，例如冷冻的、固体（例如干燥粉碎的或冻干形式）或液体（例如含水的）。活性物质外加任意额外所需成分的固体组合物可以从经受真空干燥和冷冻干燥的其先前经无菌过滤的溶液获得。根据需要，它可以贮存在通过使用前添加合适媒介物而立即重构的无菌安瓿中。

[0089] 特别优选的组合物（尤其当活性物质是腺病毒载体）在 1M 蔗糖、150mM NaCl、1mM MgCl₂、54mg/l Tween 80、10mM Tris, pH 8.5 中配制。另一种优选的组合物在 10mg/ml 甘露醇、1mg/ml HSA、20mM Tris, pH 7.2 和 150mM NaCl 中配制。此类制剂特别适合在至少 2 个月时间范围内在冷冻（例如 -70°C、-40°C、-20°C）或冷藏（例如 4°C）温度保护本发明组合物的稳定性。

[0090] 该组合物也可以含有一种或多种可药用赋形剂用于提供希望的药用属性或药代动力学属性，例如包括调节或维持 pH、渗透性、黏性、透明度、颜色、无菌性、稳定性、释放至或吸收进入人或动物对象。稳定组分的代表性实例包括聚山梨酯 80、L-精氨酸、聚乙烯吡咯烷酮、海藻糖和可以用来获得溶解度、稳定性和半寿期的希望属性的聚合物如聚乙二醇 (Davis 等人, 1978, Enzyme Eng. 4, 169-173; Burnham 等人, 1994, Am. J. Hosp. Pharm. 51, 210-218)。增稠剂包括羧甲基纤维素钠、山梨醇和葡聚糖。该组合物也可以含有本领域已知促进穿透或跨粘膜屏障或在特定器官中运输的物质。例如，适用于阴道施用的组合物可以逐步地包含用来增加粘膜孔径的一种或多种吸收增强剂。

[0091] 另外，本发明的组合物包含适合给人类全身性或粘膜性施用的一种或多种佐剂。术语“佐剂”指这样的化合物，其具有增强针对特定抗原的免疫应答的能力。佐剂可以被递送至抗原部位处或其附近。体液免疫的增强一般由针对该抗原所产生的抗体的滴度显著提

高(通常大于10倍)表现。细胞免疫的增可以通过阳性皮肤试验、细胞毒T细胞测定、IFN γ 或IL-2的ELIspot测定法测量。优选地,在本发明中施用的佐剂能够刺激针对活性物质的免疫,尤其通过toll样受体(TLR)如TLR-7、TLR-8和TLR-9。有用佐剂的代表性实例包括,不限于铝胶、矿物油乳液如弗氏完全佐剂和不完全佐剂(IFA)、脂多糖或其衍生物(Ribi等人,1986, *Immunology and Immunopharmacology of Bacterial Endotoxins*, Plenum Publ. Corp., NY, p407-19)、皂甙如QS21(Sumino等人,1998, *J. Virol.* 72, 4931-9; W098/56415)、咪唑喹啉化合物如咪唑莫特(Suader, 2000, *J. Am Acad Dermatol.* 43, S6-S11)、1H-咪唑并(4,5-c)奎诺酮-4-胺衍生物(AldaraTM)和相关化合物(Smorlesi, 2005, *Gene Ther.* 12, 1324-32)、胞嘧啶磷酸鸟苷寡脱氧核苷酸如CpG(Chu等人,1997, *J. Exp. Med.* 186:1623; Tritel等人,2003, *J. Immunol.* 171:2358-2547)和阳离子肽如IC-31(Kritsch等人,2005, *J. Chromatogr Anal. Technol Biomed Life Sci* 822, 263-70)。

[0092] 本发明的核酸分子,载体,感染性颗粒或组合物可以通过多种施用模式施用,包括全身、局部和粘膜施用。全身施用可以通过任意方式进行,例如通过皮下、皮内、肌内、静脉内、腹膜内、血管内、动脉内注射。注射可以用常规注射器或注射针或本领域可获得的任何其他装置进行。粘膜施用可以通过口腔、鼻、呼吸道内、肺内、阴道内或肛门内途径进行。局部施用可以使用经皮肤手段(例如贴剂等)进行。特别优选病毒载体和感染性颗粒作为活性物质的肌内或皮下施用。

[0093] 适宜剂量可以根据已知因素变动,如具体活性物质的药代动力学特征、年龄、健康和待治疗的对象的重量、待治疗的病理状况、症状的特征和范围、同时治疗的种类、治疗的频率、预防或治疗的需求和/或所需的作用。剂量还将根据所选的具体施用途径计算。对确定适宜治疗剂量必需的计算的进一步修正由开业医生在相关环境影响下常规地进行。就一般指导而言,腺病毒粒子的合适剂量是从约 10^5 至约 10^{13} iu(感染单位),希望是从约 10^7 至约 10^{12} iu并且优选地从约 10^8 至约 10^{11} iu。痘苗病毒粒子的合适剂量是从约 10^4 至约 10^{10} pfu(蚀斑形成单位),希望是从约 10^5 至约 10^9 pfu并且优选地从约 10^6 至约 5×10^8 pfu。载体质粒可以在 $10 \mu\text{g}$ 与 20mg 之间并且优选在 $100 \mu\text{g}$ 与 2mg 之间的剂量上施用。另外,施用可以根据标准方法、剂量或方案,在几小时、几日和/或几周范围内以单个剂量或备选地以多个剂量进行。此外,施用可以通过大丸剂注射或连续输注方式进行。例如,对象可以用上述核酸分子、载体、感染性颗粒或组合物的至少2次(例如从2次至10次)施用进行治疗。优选地,第一系列施用在从几日至4周的时间段内依次地实施,随后在1至6个月范围内实施第一系列施用(例如1或2次施用),随后是所述第一系列的最后施用。第二系列的每次施用之间的时间段可以是几日至4周。在优选的实施方案中,第一系列施用包括以周间隔的三次依次施用并且第二系列施用包括在4至6个月内的一次施用,随后是所述第一系列。作为一般指导,对于MVA载体,优选的施用途径是以包含于 10^6 与 5×10^6 pfu之间的MVA颗粒/粒子剂量的皮下方式。

[0094] 本发明的核酸分子、载体、感染性颗粒、宿主细胞或组合物可以导入对象中,以治疗或预防多种病理状况,包括遗传疾病、先天疾病和获得性疾病。本发明也涉及本发明核酸分子、载体、感染性颗粒、宿主细胞或组合物用于制备用于治疗或预防此类病理状况的药物的用途。它特别适宜于治疗或预防感染性疾病(例如病毒性和/或细菌性感染)、癌症和免疫缺陷病。如本文中所用的术语“癌症”包括因不想要的细胞增殖所致的任意癌变状况,包

括弥散或局限性肿瘤、转移、癌变息肉和肿瘤前损伤（例如瘤形成）。

[0095] 在本发明上下文中构思的感染性疾病包括与如上所述致病微生物感染相关的任意病况。在本发明上下文中构思的癌症包括，而不限于恶性胶质瘤、肉瘤、黑素瘤、肥大细胞瘤、癌以及乳腺癌、前列腺癌、睾丸癌、卵巢癌、子宫内膜癌、子宫颈癌（尤其与乳头状瘤病毒感染相关的那些子宫颈癌）、肺癌（例如包括巨细胞、小细胞、鳞状细胞和腺癌）、肾癌、膀胱癌、肝癌、结肠癌、直肠癌、胰腺癌、胃癌、胃肠癌、口腔癌、喉癌、脑和 CNS 癌、皮肤癌（例如黑素瘤和非黑素瘤）、血癌（淋巴瘤、白血病，尤其当它们在实体组织中发育时）、骨癌、视网膜母细胞瘤和甲状腺癌。

[0096] 在优选的实施方案中，本发明用于预防性或治疗性治疗与乳头状瘤病毒（尤其 HR HPV）感染相关的病况，如持续性感染、恶变前和恶性病变。“持续性感染”指在没有实现病毒根除的感染对象中乳头状瘤病毒感染的无症状期。一般观察不到临床体征。恶变前损害的实例包括，不限于可以在多种组织中检测到的低度、中度或高度的上皮内瘤形成如 CIN（颈上皮内瘤形成）、外阴上皮内瘤形成（VIN）、肛门上皮内瘤形成（AIN）、阴茎上皮内瘤形成（PIN）和阴道上皮内瘤形成（VaIN）。恶性病变的实例包括，不限于子宫颈癌、肛门癌、阴道癌、阴茎癌和口腔癌。编码乳头状瘤病毒多肽的本发明核酸分子、载体、感染性颗粒、宿主细胞或组合物特别定义用于治疗恶变前损害、尤其 CIN2/3 损害或恶性病变，尤其子宫颈癌。在另一个实施方案中，本发明也可以用于预防性或治疗性治疗与肝炎病毒（例如 HBV 或 HCV）感染相关的病况，如持续性感染、慢性或 fulgurant hepatitis 和肝癌。

[0097] 活性物质可以单独施用或根据与常规治疗用药程式（例如放疗、化疗和 / 或手术）结合使用。使用标准药物、剂量和治疗法，根据标准方法送递常规治疗用药程式至动物或人对象中，并且此类用药程式可以在施用本发明活性物质之前、期间和 / 或之后进行。例如，为处理与相关 HCV 感染的病况，本发明的方法或用途优选地与例如蛋白酶抑制剂（例如丝氨酸蛋白酶抑制剂如 Vertex 公司的 VX950）、聚合酶抑制剂、解螺旋酶抑制剂、抗纤维变性药、核苷类似物、TLR 激动剂、siRNA、反义寡核苷酸、抗 HCV 抗体、免疫调节剂、治疗性疫苗和在治疗 HCV 相关性肝癌（例如通常通过肝动脉中化学栓塞法施用的阿霉素或阿霉素碘油与动物胶的混合物）中常规使用的抗肿瘤药联合。一个特别合适的组合包括用 PEG 化的 IFN- α （IFN- α 2a 或 IFN- α 2b）和 / 或利巴韦林治疗，优选地持续 24 至 48 周，此后施用本发明的活性物质。为治疗与乳头状瘤病毒感染相关的病况，本发明的方法或用途可以与烧蚀术如电圈切除术联合。本发明的方法或用途也可以与免疫刺激物如细胞因子（例如 IL-2、IL-7、IL-15、IL-18、IL-21、IFN γ ）或自杀基因产物（例如在 Caruso 等人，1993, Proc. Natl. Acad. Sci. USA 90, 7024-28 中描述的 HSV-1 胸苷激酶；在 WO 99/54481 中描述的 FCU-I）或表达此类多肽的载体联合实施。

[0098] 在另一个实施方案中，本发明的方法或用途根据包括依次施用激发组合物和增强组合物的激发 / 增强治疗用药程式实施。一般，激发性和增强组合物使用包含或编码至少一个共同抗原结构域的不同媒介物。本发明的方法或用途可以包括施用激发性组合物 1 至 10 次，随后施用增强性组合物 1 至 10 次。希望是，注射间隔期是 1 日至 12 个月。优选的用药程式包括独立地依次施用激发物 3 或 4 次，由 3 至 10 日（例如 1 周）的间隔时间分隔，随后在终末施用激发物后 1 至几周施用增强物 1 或 2 次。另外，激发性和增强性组合物可以在相同部位或在备选部位通过相同施用途或不同施用途施用。例如，可以通过粘

膜途径施用基于多肽的组合物,而优选地注射基于载体的组合物,例如皮下注射 MVA 载体、肌肉注射 DNA 质粒和腺病毒载体。本发明的载体、感染性颗粒或组合物可以用来激发或增强或激发并增强治疗的对象免疫应答。在一个实施方案中,用本发明的质粒载体进行激发并且用本发明的痘苗病毒感染性颗粒增强。在另一个实施方案中,用本发明的腺病毒感染性颗粒进行激发并且用本发明的痘苗病毒感染性颗粒增强。仍在另一个实施方案中,用本发明的痘苗病毒感染性颗粒进行激发并且用本发明的腺病毒感染性颗粒增强。

[0099] 本发明已经以说明性方式进行描述,并且应当理解已经使用的术语意图本质上是描述性而非限制性的。显然,本发明的众多修改和变体在上述教授内容的影响下是可能的。因此应当理解在所附权利要求书的范围内,本发明可以按照不同于本文具体所描述的方式实施。

[0100] 以上所提及的专利、出版物和数据库条目的全部公开内容出版物特别地通过引用方式以相同程度完整并入本文,如同专门且个别地说明通过引用方式并入每份这种单独的专利、出版物和条目。

[0101] 附图简述

[0102] 图 1 说明 (A) 存在于天然 HPV-16 E1 序列末尾的 59 个核苷酸 (a) 与存在于天然 HPV-16 E2 序列开始处的 59 个核苷酸 (b) 之间的序列比对结果以及 (B) 存在于天然 HPV-16 E1 序列末尾或天然 HPV-16 E2 序列开始处的 59 个核苷酸 (a) 与 SEQ ID NO:9(b) 之间的序列比对结果。

[0103] 图 2 说明 (A) 在 HPV-16 与 HPV-18 E6- 编码序列之间和 (B) 在 HPV-16E6- 编码序列与 SEQ ID NO:14 之间的序列比对 / 比对结果。

[0104] 以下实施例起到说明本发明的作用。

实施例

[0105] 下文所述的构建根据详述于 Maniatis 等人 (1989, Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY) 中的一般基因工程技分子克隆技术或当使用试剂盒时根据制造商推荐实施。PCR 扩增技术是本领域技术人员已知的 (见例如 1990 年 Academic Press 出版社 Innis、Gelfand、Sninsky 和 White 出版的《PCR 方案 - 方法和应用指南》(PCR protocols-A guide to methods and applications))。携带氨苄青霉素抗性基因的重组质粒在补充有 100 μ g/ml 抗生素的琼脂或液体培养基上的大肠杆菌 C600 (Stratagene)、BJ5183 (Hanahan, 1983, J. Mol. Biol. 166, 557-580) 和 NM522 中复制。重组痘苗病毒的构建根据以上所提及文献和 Mackett 等人 (1982, Proc. Natl. Acad. Sci. USA 79, 7415-7419) 和 Mackett 等人 (1984, J. Virol. 49, 857-864) 中所属领域内的常规技术进行。使用大肠杆菌的选择基因 gpt (黄嘌呤鸟嘌呤磷酸核糖转移酶) (Falkner 和 Moss, 1988, J. Virol. 62, 1849-1854) 来促进重组痘苗病毒的选择。

[0106] 实施例 1: 构建表达 HPV-16 E1 和 E2 基因 (MVATG17410) 的 MVA 载体

[0107] a) 构建编码 HPV-16 E2 基因的重组 MVA 载体 (MVATG 17408)

[0108] 克隆 HP V16 E2 基因

[0109] 从分离自 CaSki 细胞 (ATCC CRL-1550) 的基因组 DNA 中克隆编码 HPV-16 E2 的核苷酸序列。使用引物 OTG16809 (SEQ ID NO:16) 和 OTG16810 (SEQ ID NO:17) 扩增 E2 基

因。所得片段由 Bam HI 和 EcoRI 消化并插入由相同酶限制性处理过的 pGEX2T (Amersham Biosciences), 产生 pTG17239。对克隆的 E2 基因测序显示与 (Genbank NC-01526 中描述的) HPV16 E2 原型序列相比, 存在 5 个突变。两个突变是沉默的, 并且使用 QuickChange 位点定向诱变 (Stratagene) 校正三个非沉默突变 (T210I, S219P, K310T), 产生 pTG 17268。

[0110] 修饰 HPV-16 E2 多肽

[0111] 并入 pTG 17268 的 E2 核苷酸序列通过位点定向诱变法修饰, 旨在产生名为 edE2* 的 HPV-16 E2 变体 (E39A 和 I73A)。更具体的, 通过将位置 39 中的 Glu 残基替换为 Ala 来取消 E2 的复制功能, 并且通过将位置 73 中的 He 残基替换为 Ala 来取消其反式激活功能。所得质粒 pTG17318 包含编码 HPV-16 E2* 的修饰的序列。

[0112] HPV-16 E2* 通过在氨基端融合于信号肽并且在羧基融合于衍生自狂犬病病毒 (PG 株; Genbank ay009097) 糖蛋白的膜锚定序列而进一步修饰, 从而指导 HPV-16 E2* 在表达宿主细胞中呈递在质膜表面上。使用以下引物: OTG17500 (SEQ ID NO: 18)、OTG17501 (SEQ ID NO: 19)、OTG17502 (SEQ ID NO: 20)、OTG17503 (SEQ ID NO: 21)、OTG17504 (SEQ ID NO: 22) 和 OTG17505 (SEQ ID NO: 23), 通过三重 PCR 再装配出编码名为 SS-E2*-TMR 的膜呈递的 E2 缺陷型变体的核苷酸序列 (SEQ ID NO: 12)。将再装配的序列插入 pBS- 衍生性载体 (Stratagene), 以产生 pTG17360, 并随后克隆在痘苗病毒转移质粒中 e pH5R 启动子下游 (Rosel 等, 1986, J Virol. 60, 436-449), 产生 pTG17408。

[0113] 设计转移质粒旨在允许通过同源重组于 MVA 基因组的缺失 III 中插入待转移的核苷酸序列。它源自质粒 pTG1E (在 Braun 等人, 2000, Gene Ther. 7, 1447-57 中描述), 其中向所述 pTG1E 克隆包围 MVA 缺失 III 的侧翼序列 (BRG3 和 BRD3), 其中所述侧翼序列通过 PCR 从 MVATGN33. 1 DNA 获得 (Sutter 和 Moss, 1992, Proc. Natl. Acad. Sci. USA 89, 10847-51)。该转移质粒也含有 (Lausanne 大学 R. Wittek 友好提供的) 早晚期痘苗病毒人造启动子 p11K7.5 的控制下的维多利亚水母 (*Aequorea victoria*) 增强型绿色荧光蛋白 (eGFP 基因, 从 Clontech 公司的 pEGP-C1 分离) 与大肠杆菌黄嘌呤鸟嘌呤磷酸核糖转移酶基因 (gpt 基因) 之间的融合。黄嘌呤鸟嘌呤磷酸核糖转移酶的合成能够使 GPT⁺ 重组 MVA 在含有霉酚酸、黄嘌呤和次黄嘌呤的选择培养基 (Falkner 等, 1988, J. Virol. 62, 1849-54) 中形成蚀斑并且 eGFP 能够使得重组 MVA 蚀斑可见到。选择标记 eGFP-GPT 置于处于相同方向的同源序列两个之间。当实现了克隆性选择时, 在无选择作用下允许 eGFP-GPZ 重组 MVA 生长时, 通过几次传代轻易消除该选择标记。

[0114] 构建表达 HPV-16 SS-E2*-TMR 基因的重组 MVA

[0115] 通过在感染 MVATGN33. 1 (MOI 0.1 pfu/ 细胞) 并以 pTG17408 (根据标准磷酸钙 DNA 沉淀法) 转染的原代鸡胚成纤维细胞 (CEF) 中同源重组开展 MVATG 17408 病毒的产生。通过在含有霉酚酸、黄嘌呤和次黄嘌呤的选择培养基存在下的三轮蚀斑纯化进行病毒的选择。如上所述, 随后通过非选择培养基中传代消除选择标记。通过 PCR 验证不存在亲代 MVA 混杂。

[0116] 通过蛋白质印迹法进行 E2 表达的分析。CEF 用 MOI 0.2 的 MVATG17408 感染并且 24 小时后, 收获细胞。使用商业化抗 E2 单克隆抗体 TVG271 (Abeam) 进行蛋白质印迹分析。检测到具有 55kD 表观分子量的蛋白质的表达, 同时 E2*-TMR 的理论分子量是 48.9kDa。用内切糖苷酶 F 处理细胞提取物后, 观察到重组蛋白的大小降低, 表明从 MVATG17408 表达的

E2*TMR 多肽是 N-糖基化的。

[0117] b) 构建编码 HPV-16E1 基因的重组 MVA, 其中所述 HPV-16E1 基因在与 HPV-16 E2 基因 (MVATG 17409) 交叠的部分中是简并的

[0118] 从 CaSki 细胞 DNA (ATCC CRL-1550) 克隆了编码 HPV-16 E1 多肽的核苷酸序列。更具体地, E1 基因分 E1a (nt 1-1102) 和 E1b (nt 1001 to 1950) 两个部分扩增。使用引物 OTG16811 (SEQ ID NO:24) 和 OTG 16814 (SEQ ID NO:25) 来扩增 E1a 片段, 其中所述 E1a 片段由 BamHI 和 EcoRI 消化并插入由相同酶限制性处理的 pGEX2T, 产生 pTG17240。使用 OTG16813 (SEQ ID NO:26) 和 OTG 16812 (SEQ ID NO:27) 产生 E1b 片段, 并由 BamHI 和 EcoRI 消化, 随后插入 pGEX2T, 产生 pTG 17241。与 (Genbank NC-01526 中描述的) HPV 16 E1 原型序列相比, 测序显示存在 4 个突变。一个突变是沉默的, 并且三个存在于 E1a 中的非沉默突变 (K130Q, N185T 和 T220S) 由位点定向诱变诱变法校正。随后通过克隆在 BsrGI 和 EcoRI 消化的 pTG 17241 中的已校正 E1a 片段而重新装配出完整的 E1 基因。所得的质粒命名为 pTG17289。

[0119] 在 HPV-16 基因组中, E1 基因的最后 59 个核苷酸与 E2 基因的头 59 个核苷酸相同。为了在编码 E1 和 E2 的 MVA 载体的生产步骤期间规避不稳定性问题, 通过密码子使用修饰法修饰 E1 编码序列的这个部分, 从而减低与 E2 编码序列的序列同源性。使用简并引物 OTG17408 (SEQ ID NO:28) 和 OTG17409 (SEQ ID NO:29), 通过扩增 E1 基因的 3' 末端获得简并序列 (SEQ ID NO:9)。所得扩增片段由 NsiI 和 BglIII 消化并插入由相同酶限制性处理过的 pTG 17289, 产生 pTG17340。

[0120] HPV-16 简并 E1 序列也借助位点定向诱变法突变, 通过将 HPV-16 E1 的位置 482 中的 Gly 残基置换为 Asp 残基 (G482D, 本文中又命名为 E1*), 旨在消除所编码 E1 多肽的复制功能, 从而产生 pTG17373。

[0121] 通过与衍生自狂犬病病毒 (ERA 分离株; 在 Genbank NM38452 中描述) 糖蛋白的信号肽和膜锚定肽融合, 还修饰了 HPV-16 deg* 序列, 从而指导所编码的多肽在细胞质表面表达。使用以下引物 OTG 17560 (SEQ ID NO:30)、OTG17561 (SEQ ID NO:31)、OTG17562 (SEQ ID NO:32)、OTG17563 (SEQ ID NO:33)、OTG17564 (SEQ ID NO:34) 和 OTG17565 (SEQ ID NO:35), 通过三重 PCR 重构 SS-E1deg*-TMR 序列 (SEQ ID NO:10)。所得的片段插入 pBS- 衍生的载体 (Stratagene), 产生 pTG17404。SS-E 1deg*-TMR 序列随后克隆在痘苗病毒转移质粒中 p7.5K 启动子下游 (Cochran 等, 1985, J. Virol. 54, 30-7), 产生 pTG17409。

[0122] MVATG17409 病毒的产生通过如上所述的同源重组在 CEF 中进行。

[0123] c) 构建编码 HPV-16 E1 和 E2 基因的重组 MVA 载体 (MVATG 17410)

[0124] 从 pTG1740 分离已测序并受 p7.5K 启动子控制的 SS-E 1deg*-TMR 序列并插入 pTG17408, 产生 pTG17410。

[0125] MVATG17410 病毒的产生通过如上所述的同源重组在 CEF 中进行。

[0126] 实施例 2: 构建编码 HPV-18 E1 和 E2 基因 (MVATG17582) 的 MVA 载体

[0127] 将 HPV-18 E1 和 E2 基因重构为合成基因并且设计寡核苷酸从而 (i) 旨在降低由天然 HPV-16 和 HPV-18 序列所共有的同源部分之间的同源性百分比至小于 75% (比对 HPV-16 和 HPV-18 的 E1 和 E2 基因的序列并设计寡核苷酸从而旨在减低同源性至小于 5 个连续核苷酸) (ii) 降低在天然 HPV-18 E1 序列 3' 端和在 HPV-18 E2 序列 5' 端均存在的 59 个核

核苷酸部分之间的同源性至小于 75% 并且 (ii) 导入取消 HPV-18 E1 和 E2 基因产物之酶功能的突变 (E1 :G489D, E2 :E43A 和 I77A)。

[0128] HPV-18 degE1* 序列通过装配 50 个寡核苷酸重构并克隆在 pBS 载体中, 产生 pTG 17473。使用引物 OTG15315 (SEQ ID NO :36)、OTG17881 (SEQ ID NO :37)、OTG17882 (SEQ ID NO :38)、OTG17883 (SEQ ID NO :39)、OTG17884 (SEQ ID NO :40) 和 OTG17885 (SEQ ID NO :41), 通过三重 PCR 随后将 E1 序列融合于来自麻疹病毒 F 蛋白的信号序列克隆 (SS-18E1deg*-TMF)。所得的片段 (SEQ ID NO :11) 克隆在 MVA 转移载体中处于 p7.5K 启动子的控制下, 以产生 pTG17521。

[0129] HPV-18degE2* 序列通过装配 26 个寡核苷酸重构并克隆在 pBS 载体中, 产生 pTG 17498。使用引物 OTG17875 (SEQ ID NO :42)、OTG17876 (SEQ ID NO :43)、OTG17877 (SEQ ID NO :44)、OTG17878 (SEQ ID NO :45)、OTG17879 (SEQ ID NO :46) 和 OTG17880 (SEQ ID NO :47), 通过三重 PCR 进行与狂犬病病毒 (ERA 株 ;Genbank n° M38452) 糖蛋白的信号肽和膜锚定肽的融合。将 SS-18E2*-TMR 盒插入痘苗病毒转移质粒中 p7H5R 启动子下游, 产生 pTG17552。最后, 从 pTG17521 分离 p7.5K-SS-E1deg*-TMF 盒并插入 pTG17552, 产生 pTG17582。

[0130] 重组 MVATG17521、MVATG17552 和 MVATG17582 的产生如上所述进行。

[0131] 实施例 3 : 构建表达 HPV-16 和 HPV-18 E1 和 E2 基因的多价 MVA 载体 (MVATG17583)

[0132] 将 p7.5K-SS-18E1deg*-TMF 盒和 pH5R-SS-18E2*-TMR 盒导入 pTG17410 (含有 p7.5K-SS-16E1deg*-TMR 盒和 pH5R-SS-16E2*-TMR) 并且将所得转移质粒命名为 pTG17583。MVATG 17583 的产生如上所述进行。

[0133] 实施例 4 : 构建表达 HPV16 和 HPV18 E6 和 E7 基因的多价重组病毒

[0134] MVATG 16327 是表达 HPV-16 和 HPV-18 E6 和 E7 多肽的膜锚定型和非致癌性变体的重组 MVA 病毒。将 E6 和 E7 核苷酸序列突变, 旨在消除其致癌属性 (E6* 和 E7*), 并且融合于编码适宜信号肽和膜锚定肽的序列 (E6*tm, E7*tm)。更具体地, HPV-18 E7* 分别在其氨基端和羧基端与麻疹病毒 F 糖蛋白的信号肽和膜锚定肽融合, 其中 HPV-16 E6*、HPV-16 E7* 和 HPV-18 E6* 与从狂犬病病毒糖蛋白的信号肽和膜锚定肽衍生的信号肽和膜锚定肽融合。另外, 通过密码子使用修饰法进一步修饰 HPV-18 E6 和 E7 核苷酸序列, 从而减低与其 HPV16 对应物的同源性。为此目的, 比对天然 HPV 16 和 HPV 18 基因的序列并且进行密码子简并以减低同源性至小于 5 个连续核苷酸。在该载体中, HPV-16 和 HPV-18 E6 序列在彼此相反的方向上均置于 p7.5K 启动子的控制下, 其中 HPV-16 和 HPV-18 E7 序列由 H5R 启动子驱动并且将全部表达盒插入 MVA 基因组的切除 III 区域。

[0135] a) 构建 HPV-16E7*tm 表达盒

[0136] 如 W099/03885 中所述, 从 Caski 细胞分离并修饰 HPV-16 E7 基因, 从而编码一种非致癌性和膜呈递的 E7 多肽 (16E7*tmR)。通过缺失氨基酸残基 21-26 (Δ DLYCYE) 进行非致癌性突变, 并且通过将 E7* 突变的序列分别在其 5' 端和 3' 端与编码从狂犬病病毒 (ERA 株 ;Genbank 登录号 M38452) 糖蛋白克隆的信号肽和膜锚定性肽的序列融合进行膜呈递。将所得的序列克隆处在早期晚期 pH5R 启动子 (Rose1 等, 1986. J. Viro1. 60, 436-9) 的控制下并将该盒导入 pBS 衍生的载体, 产生 pTG16161。

[0137] b) 克隆 HPV-16 E6*tm 和 HPV-18 E6*tm 表达盒

[0138] 如 W099/03885 中所述, 分离并修饰 HPV-16 E6 基因, 从而编码一种非致癌性和膜

呈递的 E6 多肽。通过缺失氨基酸残基 118-122 (Δ DLYCYE) 进行非致瘤性突变, 并且通过将 E6* 突变的序列分别在其 5' 端和 3' 端与编码从狂犬病病毒 (PG strain; Genbank 登录号 ay009097) 糖蛋白衍生的信号肽和膜锚定性肽的序列融合进行膜呈递。这通过在含有被一个 BamHI 位点隔开的信号肽和膜锚定性肽序列的载体中插入所述 E6* 突变序列进行, 从而产生 pTG 16097。

[0139] 通过装配寡核苷酸 OTG1 5174 (SEQ ID NO:48)、OTG1 5175 (SEQ ID NO:49)、OTG1 5176 (SEQ ID NO:50)、OTG15177 (SEQ ID NO:51)、OTG15178 (SEQ ID NO:52)、OTG15179 (SEQ ID NO:53)、OTG15180 (SEQ ID NO:54) 和 OTG15181 (SEQ ID NO:55), 产生了合成性 HPV-18E6 序列。如此设计寡核苷酸, 从而引起编码氨基酸残基 113-117 (非致瘤性突变 Δ NPAEK) 密码子缺失并且简化密码子使用, 旨在减低与 HPV-16E6 基因的同源性 (简并的序列)。所得的合成性序列随后分别在其 5' 端和 3' 端与编码从狂犬病病毒糖蛋白基因衍生的信号肽和膜锚定性肽的序列融合以提供如 SEQ ID NO:14 所示的序列, 从而产生 pTG16160。将 HPV-16 E6*tmR 和 HPV-18 degE6*tmR 序列以相反方向插入, 每个序列均处在 p7.5K 启动子的控制下。将表达盒随后导入 pTG16161 以产生 pTG 16215。

[0140] c) 克隆 HPV-18 E7*tmF 表达盒

[0141] 通过过装配寡核苷酸 OTG14773 (SEQ ID NO:56)、OTG14774 (SEQ ID NO:57)、OTG14775 (SEQ ID NO:58)、OTG14776 (SEQ ID NO:59)、OTG14777 (SEQ ID NO:60) and OTG14778 (SEQ ID NO:61), 产生了合成性 HPV-18 E7 序列。如此设计寡核苷酸, 从而引起编码氨基酸残基 24-28 (非致瘤性突变 Δ DLLCH) 密码子缺失并且简化密码子使用, 旨在减低与 HPV-16 E7 基因的同源性 (简并的序列)。所得的合成性序列随后分别在其 5' 端和 3' 端与从麻疹病毒 F 蛋白基因衍生的信号肽和膜锚定性肽的编码序列融合 (在 EP 0305229 中描述)。将所得的序列 (SEQ ID NO:15) 克隆处在 pH5R 启动子的控制下并将表达盒导入 pBS 衍生的载体以产生 pTG 16015。

[0142] d) 构建转移质粒 pTG 16327

[0143] 转移质粒 pTG6019 (其在 W099/03885 的实施例 2 中描述) 含有位于 MVA 缺失 III 侧翼的同源序列。此质粒如下修饰。通过引物 OTG 15040 (SEQ ID NO:62) 和 OTG15041 (SEQ ID NO:63) 杂交所获得的合成性多接头被导入由 BamU1 和 Sad 消化的 pTG6019, 产生 pTG 16007。从 (在 EP 1 146 125 的实施例 2 中描述的) pTG14033 分离含有置于早期晚期 pH5R 启动子控制下的编码大肠杆菌 gpt 的表达盒的 SacI-SacI 片段并且导入经 SacI 消化的 pTG16007, 产生 pTG16093。黄嘌呤鸟嘌呤磷酸核糖转移酶的合成能够使 GPT⁺ 重组 MVA 在含有霉酚酸、黄嘌呤和次黄嘌呤的选择培养基 (Falkner 等人, 1988. J. Virol. 62, 1849-54) 中形成蚀斑。选择标记 GPT 置于处于相同方向的同源序列两个之间。当实现了克隆性选择时, 在无选择作用下允许 eGPP-GPZ 重组 MVA 生长时, 通过几次传代轻易消除该选择标记。

[0144] 从 pTG16015 分离含有 HPV-18degE7*TMF 表达盒的 HindIII-SmaI 片段并导入经相同酶消化的 pTG 16093, 产生 pTG16105。另一方面, pTG16215 由 SalI 和 EcoR 消, 用 T4DNA 聚合酶处理, 并且将含有 HPV-16 E7*tmR、HPV-16 E6*tmR 和 HPV-18 degE6*TMF 表达盒的所得片段导经 SmaI 消化的 pTG16105, 产生 pTG16327 (图 2)。

[0145] e) MVATG16327 的产生

[0146] 通过在原代鸡胚成纤维细胞 (CEF) 中同源重组开展 MVATG16327 的产生。为此目

的,在事先以 MOI 0.1pfu/ 细胞感染 MVATGN33.1 的 CEF 上根据标准磷酸钙 DNA 沉淀法转染 pTG 16327。通过在含有霉酚酸、黄嘌呤和次黄嘌呤的选择培养基存在下,在 CEF 上的 3 轮蚀斑纯化进行病毒的选择。随后通过非选择培养基中传代消除选择标记。通过 PCR 验证不存在亲代 MVA 混杂。

[0147] 通过蛋白质印迹法进行基因表达的分析。CEF 用 MOI 0.2 的 MVATG16327 感染并且 24 小时后,收获细胞。使用分别针对 HPV-16 与 HPV-18 E6 和 E7 蛋白的兔多克隆抗体,开展蛋白质印迹分析。结果显示全部 HPV 多肽正确地由 MVATG16327 中表达。

[0148] f) 研究 MVATG16327 的遗传稳定性

[0149] 在以 MOI 10^{-2} pfu/ 细胞和 10^{-4} pfu/ 细胞感染的 CEF 上将 MVATG16327 传 5 代。对从第 5 代研究用原液分离的 100 个病毒克隆评价遗传稳定性。使用两种方法:确定表达盒的结构的 PCR 扩增法,和借助蛋白质印迹法的抗原检测。PCR 分析的结果显示 99% 克隆含有目的表达盒。免疫检测法显示 97% 的克隆表达 4 种抗原:HPV-16 和 HPV-18 的 E6*tm 和 E7*tm 多肽。

[0150] 这些分析表明在 5 次传代后,从 MVATG 16327 衍生的 97% 的克隆是一致的,表明该构建体的良好遗传稳定性。

[0151] 实施例 5:构建表达 HPV-16、HPV-18、HPV-33 和 HPV-52 E2 基因的多价 MVA 载体

[0152] 编码 HPV-33 E2 多肽的合成基因由 Geneart (Regensburg, Germany) 合成。如此设计该合成性序列,从而 (i) 减低与源自 HPV-16, HPV-18 和 HPV-52 的 E2 基因的同源性百分比至小于 75% (若可能,缩小同源部分至小于 6 个连续核苷酸) 和 (ii) 导入取消 HPV-33 基因产物之酶功能的突变 (E39A 和 I73A)。

[0153] HPV-33 degE2* 序列随后与编码狂犬病病毒 (ERA 株, Genbank n° M38452) 糖蛋白的信号肽和膜锚定性肽的核苷酸序列融合。使用引物 TG18962 (SEQ ID NO: 72)、OTG18963 (SEQ ID NO: 73)、OTG18964 (SEQ ID NO: 74)、OTG18965 (SEQ ID NO: 75)、OTG18966 (SEQ ID NO: 76) and OTG18967 (SEQ ID NO: 77), 通过三重 PCR 进行该步骤。将编码 SS-33degE2*-TMR 多肽的所得片段 (SEQ ID NO: 67) 克隆在 MVA 转移载体中处于 p7.5K 启动子的控制下,并且如上所述产生病毒粒子。

[0154] 编码 HPV-52 E2 多肽的合成基因由 Geneart (Regensburg, Germany) 合成。如此设计该合成性序列,从而 (i) 减低与源自 HPV-16、HPV-18 和 HPV-33 的 E2 基因的同源性百分比至小于 75% (优选地缩小同源部分至小于 6 个连续核苷酸) 和 (ii) 导入取消 HPV-52 基因产物之酶功能的突变 (E39A 和 I73A)。

[0155] 使用引物 OTG18968 (SEQ ID NO: 78)、OTG18969 (SEQ ID NO: 79)、OTG18970 (SEQ ID NO: 80)、OTG18971 (SEQ ID NO: 81)、OTG18972 (SEQ ID NO: 82) 和 OTG18973 (SEQ ID NO: 83), 通过三重 PCR 随后将合成性 HPV-52 E2*deg 序列融合于编码麻疹病毒 F 蛋白的信号肽和膜锚定性肽的核苷酸序列 (产生 SS-52E2*deg-TM)。

[0156] 将编码 SS-52E2*deg-TMF 多肽的所得片段 (SEQ ID NO: 69) 克隆在 MVA 转移载体中 p7.5K 启动子下游,并且如上所述产生病毒粒子。

[0157] 将编码膜呈递的且酶缺陷型 HPV-18 E2 多肽的 pH5R-SS-18E2*-TMR 盒 (从 pTG17552 分离)、编码膜呈递的且酶缺陷型 HPV-33 E2 多肽的 p7.5K-SS-33degE2*-TMR 盒、编码膜呈递的且酶缺陷型 HPV-52 E2 多肽的 7.5K-SS-52degE2*-TMF 盒导入 (含有

pH5R-SS-16E2*-TMR 盒的)TG17408 并且如上所述产生病毒粒子。

[0001]

序列表

<110> 特兰斯吉恩股份有限公司

<120> 用于多基因表达的载体

<130> TG183 EXT

<150> EP07360019.9

<151> 2007-05-15

<160> 83

<170> PatentIn version 3.3

<210> 1

<211> 649

<212> PRT

<213> Human papillomavirus type 16

<400> 1

```

Met Ala Asp Pro Ala Gly Thr Asn Gly Glu Glu Gly Thr Gly Cys Asn
1      5      10      15
Gly Trp Phe Tyr Val Glu Ala Val Val Glu Lys Lys Thr Gly Asp Ala
20      25      30
Ile Ser Asp Asp Glu Asn Glu Asn Asp Ser Asp Thr Gly Glu Asp Leu
35      40      45
Val Asp Phe Ile Val Asn Asp Asn Asp Tyr Leu Thr Gln Ala Glu Thr
50      55      60
Glu Thr Ala His Ala Leu Phe Thr Ala Gln Glu Ala Lys Gln His Arg
65      70      75      80
Asp Ala Val Gln Val Leu Lys Arg Lys Tyr Leu Gly Ser Pro Leu Ser
85      90      95
Asp Ile Ser Gly Cys Val Asp Asn Asn Ile Ser Pro Arg Leu Lys Ala
100     105     110
Ile Cys Ile Glu Lys Gln Ser Arg Ala Ala Lys Arg Arg Leu Phe Glu
115     120     125
Ser Glu Asp Ser Gly Tyr Gly Asn Thr Glu Val Glu Thr Gln Gln Met
130     135     140
Leu Gln Val Glu Gly Arg His Glu Thr Glu Thr Pro Cys Ser Gln Tyr
145     150     155     160
Ser Gly Gly Ser Gly Gly Cys Ser Gln Tyr Ser Ser Gly Ser Gly
165     170     175
Gly Glu Gly Val Ser Glu Arg His Thr Ile Cys Gln Thr Pro Leu Thr
180     185     190
Asn Ile Leu Asn Val Leu Lys Thr Ser Asn Ala Lys Ala Ala Met Leu
195     200     205
Ala Lys Phe Lys Glu Leu Tyr Gly Val Ser Phe Ser Glu Leu Val Arg
210     215     220
Pro Phe Lys Ser Asn Lys Ser Thr Cys Cys Asp Trp Cys Ile Ala Ala
225     230     235     240
Phe Gly Leu Thr Pro Ser Ile Ala Asp Ser Ile Lys Thr Leu Leu Gln
245     250     255
Gln Tyr Cys Leu Tyr Leu His Ile Gln Ser Leu Ala Cys Ser Trp Gly
260     265     270
Met Val Val Leu Leu Leu Val Arg Tyr Lys Cys Gly Lys Asn Arg Glu
275     280     285
Thr Ile Glu Lys Leu Leu Ser Lys Leu Leu Cys Val Ser Pro Met Cys
290     295     300
Met Met Ile Glu Pro Pro Lys Leu Arg Ser Thr Ala Ala Ala Leu Tyr

```

[0002]

```

305          310          315          320
Trp Tyr Lys Thr Gly Ile Ser Asn Ile Ser Glu Val Tyr Gly Asp Thr
          325          330          335
Pro Glu Trp Ile Gln Arg Gln Thr Val Leu Gln His Ser Phe Asn Asp
          340          345          350
Cys Thr Phe Glu Leu Ser Gln Met Val Gln Trp Ala Tyr Asp Asn Asp
          355          360          365
Ile Val Asp Asp Ser Glu Ile Ala Tyr Lys Tyr Ala Gln Leu Ala Asp
          370          375          380
Thr Asn Ser Asn Ala Ser Ala Phe Leu Lys Ser Asn Ser Gln Ala Lys
          385          390          395          400
Ile Val Lys Asp Cys Ala Thr Met Cys Arg His Tyr Lys Arg Ala Glu
          405          410          415
Lys Lys Gln Met Ser Met Ser Gln Trp Ile Lys Tyr Arg Cys Asp Arg
          420          425          430
Val Asp Asp Gly Gly Asp Trp Lys Gln Ile Val Met Phe Leu Arg Tyr
          435          440          445
Gln Gly Val Glu Phe Met Ser Phe Leu Thr Ala Leu Lys Arg Phe Leu
          450          455          460
Gln Gly Ile Pro Lys Lys Asn Cys Ile Leu Leu Tyr Gly Ala Ala Asn
          465          470          475          480
Thr Gly Lys Ser Leu Phe Gly Met Ser Leu Met Lys Phe Leu Gln Gly
          485          490          495
Ser Val Ile Cys Phe Val Asn Ser Lys Ser His Phe Trp Leu Gln Pro
          500          505          510
Leu Ala Asp Ala Lys Ile Gly Met Leu Asp Asp Ala Thr Val Pro Cys
          515          520          525
Trp Asn Tyr Ile Asp Asp Asn Leu Arg Asn Ala Leu Asp Gly Asn Leu
          530          535          540
Val Ser Met Asp Val Lys His Arg Pro Leu Val Gln Leu Lys Cys Pro
          545          550          555          560
Pro Leu Leu Ile Thr Ser Asn Ile Asn Ala Gly Thr Asp Ser Arg Trp
          565          570          575
Pro Tyr Leu His Asn Arg Leu Val Val Phe Thr Phe Pro Asn Glu Phe
          580          585          590
Pro Phe Asp Glu Asn Gly Asn Pro Val Tyr Glu Leu Asn Asp Lys Asn
          595          600          605
Trp Lys Ser Phe Phe Ser Arg Thr Trp Ser Arg Leu Ser Leu His Glu
          610          615          620
Asp Glu Asp Lys Glu Asn Asp Gly Asp Ser Leu Pro Thr Phe Lys Cys
          625          630          635          640
Val Ser Gly Gln Asn Thr Asn Thr Leu
          645

```

<210> 2

<211> 365

<212> PRT

<213> Human papillomavirus type 16

<400> 2

```

Met Glu Thr Leu Cys Gln Arg Leu Asn Val Cys Gln Asp Lys Ile Leu
1          5          10          15
Thr His Tyr Glu Asn Asp Ser Thr Asp Leu Arg Asp His Ile Asp Tyr
          20          25          30
Trp Lys His Met Arg Leu Glu Cys Ala Ile Tyr Tyr Lys Ala Arg Glu
          35          40          45
Met Gly Phe Lys His Ile Asn His Gln Val Val Pro Thr Leu Ala Val
          50          55          60
Ser Lys Asn Lys Ala Leu Gln Ala Ile Glu Leu Gln Leu Thr Leu Glu
65          70          75          80

```

[0003]

```

Thr Ile Tyr Asn Ser Gln Tyr Ser Asn Glu Lys Trp Thr Leu Gln Asp
      85                      90                      95
Val Ser Leu Glu Val Tyr Leu Thr Ala Pro Thr Gly Cys Ile Lys Lys
      100                    105                    110
His Gly Tyr Thr Val Glu Val Gln Phe Asp Gly Asp Ile Cys Asn Thr
      115                    120                    125
Met His Tyr Thr Asn Trp Thr His Ile Tyr Ile Cys Glu Glu Ala Ser
      130                    135                    140
Val Thr Val Val Glu Gly Gln Val Asp Tyr Tyr Gly Leu Tyr Tyr Val
      145                    150                    155                    160
His Glu Gly Ile Arg Thr Tyr Phe Val Gln Phe Lys Asp Asp Ala Glu
      165                    170                    175
Lys Tyr Ser Lys Asn Lys Val Trp Glu Val His Ala Gly Gly Gln Val
      180                    185                    190
Ile Leu Cys Pro Thr Ser Val Phe Ser Ser Asn Glu Val Ser Ser Pro
      195                    200                    205
Glu Ile Ile Arg Gln His Leu Ala Asn His Pro Ala Ala Thr His Thr
      210                    215                    220
Lys Ala Val Ala Leu Gly Thr Glu Glu Thr Gln Thr Thr Ile Gln Arg
      225                    230                    235                    240
Pro Arg Ser Glu Pro Asp Thr Gly Asn Pro Cys His Thr Thr Lys Leu
      245                    250                    255
Leu His Arg Asp Ser Val Asp Ser Ala Pro Ile Leu Thr Ala Phe Asn
      260                    265                    270
Ser Ser His Lys Gly Arg Ile Asn Cys Asn Ser Asn Thr Thr Pro Ile
      275                    280                    285
Val His Leu Lys Gly Asp Ala Asn Thr Leu Lys Cys Leu Arg Tyr Arg
      290                    295                    300
Phe Lys Lys His Cys Thr Leu Tyr Thr Ala Val Ser Ser Thr Trp His
      305                    310                    315                    320
Trp Thr Gly His Asn Val Lys His Lys Ser Ala Ile Val Thr Leu Thr
      325                    330                    335
Tyr Asp Ser Glu Trp Gln Arg Asp Gln Phe Leu Ser Gln Val Lys Ile
      340                    345                    350
Pro Lys Thr Ile Thr Val Ser Thr Gly Phe Met Ser Ile
      355                    360                    365

```

<210> 3

<211> 158

<212> PRT

<213> Human papillomavirus type 16

<400> 3

```

Met His Gln Lys Arg Thr Ala Met Phe Gln Asp Pro Gln Glu Arg Pro
1      5                      10                      15
Arg Lys Leu Pro Gln Leu Cys Thr Glu Leu Gln Thr Thr Ile His Asp
      20                    25                    30
Ile Ile Leu Glu Cys Val Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu
      35                    40                    45
Val Tyr Asp Phe Ala Phe Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly
      50                    55                    60
Asn Pro Tyr Ala Val Cys Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile
      65                    70                    75                    80
Ser Glu Tyr Arg His Tyr Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu
      85                    90                    95
Gln Gln Tyr Asn Lys Pro Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn
      100                    105                    110
Cys Gln Lys Pro Leu Cys Pro Glu Glu Lys Gln Arg His Leu Asp Lys
      115                    120                    125
Lys Gln Arg Phe His Asn Ile Arg Gly Arg Trp Thr Gly Arg Cys Met

```

[0004]

130	135	140
Ser Cys Cys Arg Ser Ser Arg Thr Arg Arg Glu Thr Gln Leu		
145	150	155

<210> 4
 <211> 98
 <212> PRT
 <213> Human papillomavirus type 16

<400> 4
 Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
 1 5 10 15
 Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
 20 25 30
 Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
 35 40 45
 Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
 50 55 60
 Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
 65 70 75 80
 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
 85 90 95
 Lys Pro

<210> 5
 <211> 737
 <212> PRT
 <213> artificial sequence

<220>
 <223> HPV-16 membrane-presented and replication-defective E1
 polypeptide

<400> 5
 Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu
 1 5 10 15
 Cys Phe Gly Lys Phe Pro Ile Ala Asp Pro Ala Gly Thr Asn Gly Glu
 20 25 30
 Glu Gly Thr Gly Cys Asn Gly Trp Phe Tyr Val Glu Ala Val Val Glu
 35 40 45
 Lys Lys Thr Gly Asp Ala Ile Ser Asp Asp Glu Asn Glu Asn Asp Ser
 50 55 60
 Asp Thr Gly Glu Asp Leu Val Asp Phe Ile Val Asn Asp Asn Asp Tyr
 65 70 75 80
 Leu Thr Gln Ala Glu Thr Glu Thr Ala His Ala Leu Phe Thr Ala Gln
 85 90 95
 Glu Ala Lys Gln His Arg Asp Ala Val Gln Val Leu Lys Arg Lys Tyr
 100 105 110
 Leu Gly Ser Pro Leu Ser Asp Ile Ser Gly Cys Val Asp Asn Asn Ile
 115 120 125
 Ser Pro Arg Leu Lys Ala Ile Cys Ile Glu Lys Gln Ser Arg Ala Ala
 130 135 140
 Lys Arg Arg Leu Phe Glu Ser Glu Asp Ser Gly Tyr Gly Asn Thr Glu
 145 150 155 160
 Val Glu Thr Gln Gln Met Leu Gln Val Glu Gly Arg His Glu Thr Glu
 165 170 175
 Thr Pro Cys Ser Gln Tyr Ser Gly Gly Ser Gly Gly Gly Cys Ser Gln
 180 185 190

[0005]

Tyr	Ser	Ser	Gly	Ser	Gly	Gly	Glu	Gly	Val	Ser	Glu	Arg	His	Thr	Ile	195	200	205
Cys	Gln	Thr	Pro	Leu	Thr	Asn	Ile	Leu	Asn	Val	Leu	Lys	Thr	Ser	Asn	210	215	220
Ala	Lys	Ala	Ala	Met	Leu	Ala	Lys	Phe	Lys	Glu	Leu	Tyr	Gly	Val	Ser	225	230	235
Phe	Ser	Glu	Leu	Val	Arg	Pro	Phe	Lys	Ser	Asn	Lys	Ser	Thr	Cys	Cys	245	250	255
Asp	Trp	Cys	Ile	Ala	Ala	Phe	Gly	Leu	Thr	Pro	Ser	Ile	Ala	Asp	Ser	260	265	270
Ile	Lys	Thr	Leu	Leu	Gln	Gln	Tyr	Cys	Leu	Tyr	Leu	His	Ile	Gln	Ser	275	280	285
Leu	Ala	Cys	Ser	Trp	Gly	Met	Val	Val	Leu	Leu	Leu	Val	Arg	Tyr	Lys	290	295	300
Cys	Gly	Lys	Asn	Arg	Glu	Thr	Ile	Glu	Lys	Leu	Leu	Ser	Lys	Leu	Leu	305	310	315
Cys	Val	Ser	Pro	Met	Cys	Met	Met	Ile	Glu	Pro	Pro	Lys	Leu	Arg	Ser	325	330	335
Thr	Ala	Ala	Ala	Leu	Tyr	Trp	Tyr	Lys	Thr	Gly	Ile	Ser	Asn	Ile	Ser	340	345	350
Glu	Val	Tyr	Gly	Asp	Thr	Pro	Glu	Trp	Ile	Gln	Arg	Gln	Thr	Val	Leu	355	360	365
Gln	His	Ser	Phe	Asn	Asp	Cys	Thr	Phe	Glu	Leu	Ser	Gln	Met	Val	Gln	370	375	380
Trp	Ala	Tyr	Asp	Asn	Asp	Ile	Val	Asp	Asp	Ser	Glu	Ile	Ala	Tyr	Lys	385	390	395
Tyr	Ala	Gln	Leu	Ala	Asp	Thr	Asn	Ser	Asn	Ala	Ser	Ala	Phe	Leu	Lys	405	410	415
Ser	Asn	Ser	Gln	Ala	Lys	Ile	Val	Lys	Asp	Cys	Ala	Thr	Met	Cys	Arg	420	425	430
His	Tyr	Lys	Arg	Ala	Glu	Lys	Lys	Gln	Met	Ser	Met	Ser	Gln	Trp	Ile	435	440	445
Lys	Tyr	Arg	Cys	Asp	Arg	Val	Asp	Asp	Gly	Gly	Asp	Trp	Lys	Gln	Ile	450	455	460
Val	Met	Phe	Leu	Arg	Tyr	Gln	Gly	Val	Glu	Phe	Met	Ser	Phe	Leu	Thr	465	470	475
Ala	Leu	Lys	Arg	Phe	Leu	Gln	Gly	Ile	Pro	Lys	Lys	Asn	Cys	Ile	Leu	485	490	495
Leu	Tyr	Gly	Ala	Ala	Asn	Thr	Asp	Lys	Ser	Leu	Phe	Gly	Met	Ser	Leu	500	505	510
Met	Lys	Phe	Leu	Gln	Gly	Ser	Val	Ile	Cys	Phe	Val	Asn	Ser	Lys	Ser	515	520	525
His	Phe	Trp	Leu	Gln	Pro	Leu	Ala	Asp	Ala	Lys	Ile	Gly	Met	Leu	Asp	530	535	540
Asp	Ala	Thr	Val	Pro	Cys	Trp	Asn	Tyr	Ile	Asp	Asp	Asn	Leu	Arg	Asn	545	550	555
Ala	Leu	Asp	Gly	Asn	Leu	Val	Ser	Met	Asp	Val	Lys	His	Arg	Pro	Leu	565	570	575
Val	Gln	Leu	Lys	Cys	Pro	Pro	Leu	Leu	Ile	Thr	Ser	Asn	Ile	Asn	Ala	580	585	590
Gly	Thr	Asp	Ser	Arg	Trp	Pro	Tyr	Leu	His	Asn	Arg	Leu	Val	Val	Phe	595	600	605
Thr	Phe	Pro	Asn	Glu	Phe	Pro	Phe	Asp	Glu	Asn	Gly	Asn	Pro	Val	Tyr	610	615	620
Glu	Leu	Asn	Asp	Lys	Asn	Trp	Lys	Ser	Phe	Phe	Ser	Arg	Thr	Trp	Ser	625	630	635
Arg	Leu	Ser	Leu	His	Glu	Asp	Glu	Asp	Lys	Glu	Asn	Asp	Gly	Asp	Ser	645	650	655
Leu	Pro	Thr	Phe	Lys	Cys	Val	Ser	Gly	Gln	Asn	Thr	Asn	Thr	Leu	Tyr	660	665	670
Val	Leu	Leu	Ser	Ala	Gly	Thr	Leu	Ile	Ala	Leu	Met	Leu	Ile	Ile	Phe			

[0006]

<400>	6														
Met	Gly	Leu	Lys	Val	Asn	Val	Ser	Ala	Ile	Phe	Met	Ala	Val	Leu	Leu
1				5					10					15	
Thr	Leu	Gln	Thr	Pro	Thr	Gly	Gln	Ile	His	Trp	Gly	Ala	Asp	Pro	Glu
			20				25					30			
Gly	Thr	Asp	Gly	Glu	Gly	Thr	Gly	Cys	Asn	Gly	Trp	Phe	Tyr	Val	Gln
		35				40					45				
Ala	Ile	Val	Asp	Lys	Lys	Thr	Gly	Asp	Val	Ile	Ser	Asp	Asp	Glu	Asp
	50					55					60				
Glu	Asn	Ala	Thr	Asp	Thr	Gly	Ser	Asp	Met	Val	Asp	Phe	Ile	Asp	Thr
65					70					75					80
Gln	Gly	Thr	Phe	Cys	Glu	Gln	Ala	Glu	Leu	Glu	Thr	Ala	Gln	Ala	Leu
			85					90						95	
Phe	His	Ala	Gln	Glu	Val	His	Asn	Asp	Ala	Gln	Val	Leu	His	Val	Leu
			100				105						110		
Lys	Arg	Lys	Phe	Ala	Gly	Gly	Ser	Thr	Glu	Asn	Ser	Pro	Leu	Gly	Glu
		115				120						125			
Arg	Leu	Glu	Val	Asp	Thr	Glu	Leu	Ser	Pro	Arg	Leu	Gln	Glu	Ile	Ser
	130					135					140				
Leu	Asn	Ser	Gly	Gln	Lys	Lys	Ala	Lys	Arg	Arg	Leu	Phe	Thr	Ile	Ser
145					150					155					160
Asp	Ser	Gly	Tyr	Gly	Cys	Ser	Glu	Val	Glu	Ala	Thr	Gln	Ile	Gln	Val
			165						170					175	
Thr	Thr	Asn	Gly	Glu	His	Gly	Gly	Asn	Val	Cys	Ser	Gly	Gly	Ser	Thr
			180				185						190		
Glu	Ala	Ile	Asp	Asn	Gly	Gly	Thr	Glu	Gly	Asn	Asn	Ser	Ser	Val	Asp
		195				200						205			
Gly	Thr	Ser	Asp	Asn	Ser	Asn	Ile	Glu	Asn	Val	Asn	Pro	Gln	Cys	Thr
	210					215					220				
Ile	Ala	Gln	Leu	Lys	Asp	Leu	Leu	Lys	Val	Asn	Asn	Lys	Gln	Gly	Ala
225					230					235					240
Met	Leu	Ala	Val	Phe	Lys	Asp	Thr	Tyr	Gly	Leu	Ser	Phe	Thr	Asp	Leu
			245						250					255	
Val	Arg	Asn	Phe	Lys	Ser	Asp	Lys	Thr	Thr	Cys	Thr	Asp	Trp	Val	Thr
			260					265					270		
Ala	Ile	Phe	Gly	Val	Asn	Pro	Thr	Ile	Ala	Glu	Gly	Phe	Lys	Thr	Leu
		275				280						285			
Ile	Gln	Pro	Phe	Ile	Leu	Tyr	Ala	His	Ile	Gln	Cys	Leu	Asp	Cys	Lys
	290					295					300				
Trp	Gly	Val	Leu	Ile	Leu	Ala	Leu	Leu	Arg	Tyr	Lys	Cys	Gly	Lys	Ser
305					310					315					320

[0007]

Arg	Leu	Thr	Val	Ala	Lys	Gly	Leu	Ser	Thr	Leu	Leu	His	Val	Pro	Glu	
				325					330					335		
Thr	Cys	Met	Leu	Ile	Gln	Pro	Pro	Lys	Leu	Arg	Ser	Ser	Val	Ala	Ala	
			340					345					350			
Leu	Tyr	Trp	Tyr	Arg	Thr	Gly	Ile	Ser	Asn	Ile	Ser	Glu	Val	Met	Gly	
		355				360						365				
Asp	Thr	Pro	Glu	Trp	Ile	Gln	Arg	Leu	Thr	Ile	Ile	Gln	His	Gly	Ile	
	370					375					380					
Asp	Asp	Ser	Asn	Phe	Asp	Leu	Ser	Glu	Met	Val	Gln	Trp	Ala	Phe	Asp	
385				390						395					400	
Asn	Glu	Leu	Thr	Asp	Glu	Ser	Asp	Met	Ala	Phe	Glu	Tyr	Ala	Leu	Leu	
			405						410					415		
Ala	Asp	Ser	Asn	Ser	Asn	Ala	Ala	Ala	Phe	Leu	Lys	Ser	Asn	Cys	Gln	
			420					425						430		
Ala	Lys	Tyr	Leu	Lys	Asp	Cys	Ala	Thr	Met	Cys	Lys	His	Tyr	Arg	Arg	
		435					440						445			
Ala	Gln	Lys	Arg	Gln	Met	Asn	Met	Ser	Gln	Trp	Ile	Arg	Phe	Arg	Cys	
	450					455					460					
Ser	Lys	Ile	Asp	Glu	Gly	Gly	Asp	Trp	Arg	Pro	Ile	Val	Gln	Phe	Leu	
465					470					475					480	
Arg	Tyr	Gln	Gln	Ile	Glu	Phe	Ile	Thr	Phe	Leu	Gly	Ala	Leu	Lys	Ser	
			485						490					495		
Phe	Leu	Lys	Gly	Thr	Pro	Lys	Lys	Asn	Cys	Leu	Val	Phe	Cys	Gly	Pro	
			500					505					510			
Ala	Asn	Thr	Asp	Lys	Ser	Tyr	Phe	Gly	Met	Ser	Phe	Ile	His	Phe	Ile	
		515					520					525				
Gln	Gly	Ala	Val	Ile	Ser	Phe	Val	Asn	Ser	Thr	Ser	His	Phe	Trp	Leu	
	530					535					540					
Glu	Pro	Leu	Thr	Asp	Thr	Lys	Val	Ala	Met	Leu	Asp	Asp	Ala	Thr	Thr	
545					550					555					560	
Thr	Cys	Trp	Thr	Tyr	Phe	Asp	Thr	Tyr	Met	Arg	Asn	Ala	Leu	Asp	Gly	
			565						570					575		
Asn	Pro	Ile	Ser	Ile	Asp	Arg	Lys	His	Lys	Pro	Leu	Ile	Gln	Leu	Lys	
			580					585						590		
Cys	Pro	Pro	Ile	Leu	Leu	Thr	Thr	Asn	Ile	His	Pro	Ala	Lys	Asp	Asn	
		595					600					605				
Arg	Trp	Pro	Tyr	Leu	Glu	Ser	Arg	Ile	Thr	Val	Phe	Glu	Phe	Pro	Asn	
	610					615					620					
Ala	Phe	Pro	Phe	Asp	Lys	Asn	Gly	Asn	Pro	Val	Tyr	Glu	Ile	Asn	Asp	
625					630					635					640	
Lys	Asn	Trp	Lys	Cys	Phe	Phe	Glu	Arg	Thr	Trp	Ser	Arg	Leu	Asp	Leu	
			645						650					655		
His	Glu	Glu	Glu	Glu	Asp	Ala	Asp	Thr	Glu	Gly	Asn	Pro	Phe	Gly	Thr	
			660					665					670			
Phe	Lys	Leu	Arg	Ala	Gly	Gln	Asn	His	Arg	Pro	Leu	Gly	Leu	Ser	Ser	
		675					680					685				
Thr	Ser	Ile	Val	Tyr	Ile	Leu	Ile	Ala	Val	Cys	Leu	Gly	Gly	Leu	Ile	
	690					695					700					
Gly	Ile	Pro	Ala	Leu	Ile	Cys	Cys	Cys	Arg	Gly	Arg	Cys	Asn	Lys	Lys	
705					710					715					720	
Gly	Glu	Gln	Val	Gly	Met	Ser	Arg	Pro	Gly	Leu	Lys	Pro	Asp	Leu	Thr	
			725						730					735		
Gly	Thr	Ser	Lys	Ser	Tyr	Val	Arg	Ser	Leu							
			740					745								

<210> 7

<211> 453

<212> PRT

<213> artificial sequence

[0008]

<220>

<223> HPV-16 membrane-anchored and replication-defective E2 polypeptide

<400> 7

```

Met Val Pro Leu Ala Leu Leu Leu Val Pro Leu Leu Gly Phe Ser Leu
1      5      10      15
Cys Phe Gly Lys Phe Pro Ile Glu Thr Leu Cys Gln Arg Leu Asn Val
      20      25      30
Cys Gln Asp Lys Ile Leu Thr His Tyr Glu Asn Asp Ser Thr Asp Leu
      35      40      45
Arg Asp His Ile Asp Tyr Trp Lys His Met Arg Leu Ala Cys Ala Ile
      50      55      60
Tyr Tyr Lys Ala Arg Glu Met Gly Phe Lys His Ile Asn His Gln Val
      65      70      75      80
Val Pro Thr Leu Ala Val Ser Lys Asn Lys Ala Leu Gln Ala Ala Glu
      85      90      95
Leu Gln Leu Thr Leu Glu Thr Ile Tyr Asn Ser Gln Tyr Ser Asn Glu
      100     105     110
Lys Trp Thr Leu Gln Asp Val Ser Leu Glu Val Tyr Leu Thr Ala Pro
      115     120     125
Thr Gly Cys Ile Lys Lys His Gly Tyr Thr Val Glu Val Gln Phe Asp
      130     135     140
Gly Asp Ile Cys Asn Thr Met His Tyr Thr Asn Trp Thr His Ile Tyr
      145     150     155     160
Ile Cys Glu Glu Ala Ser Val Thr Val Val Glu Gly Gln Val Asp Tyr
      165     170     175
Tyr Gly Leu Tyr Tyr Val His Glu Gly Ile Arg Thr Tyr Phe Val Gln
      180     185     190
Phe Lys Asp Asp Ala Glu Lys Tyr Ser Lys Asn Lys Val Trp Glu Val
      195     200     205
His Ala Gly Gly Gln Val Ile Leu Cys Pro Thr Ser Val Phe Ser Ser
      210     215     220
Asn Glu Val Ser Ser Pro Glu Ile Ile Arg Gln His Leu Ala Asn His
      225     230     235     240
Pro Ala Ala Thr His Thr Lys Ala Val Ala Leu Gly Thr Glu Glu Thr
      245     250     255
Gln Thr Thr Ile Gln Arg Pro Arg Ser Glu Pro Asp Thr Gly Asn Pro
      260     265     270
Cys His Thr Thr Lys Leu Leu His Arg Asp Ser Val Asp Ser Ala Pro
      275     280     285
Ile Leu Thr Ala Phe Asn Ser Ser His Lys Gly Arg Ile Asn Cys Asn
      290     295     300
Ser Asn Thr Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr Leu
      305     310     315     320
Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr Ala
      325     330     335
Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His Lys Ser
      340     345     350
Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln Phe
      355     360     365
Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly Phe
      370     375     380
Met Ser Ile Tyr Val Leu Leu Ser Ala Gly Thr Leu Ile Ala Leu Met
      385     390     395     400
Leu Ile Ile Phe Leu Ile Thr Cys Cys Lys Arg Val Asp Arg Pro Glu
      405     410     415
Ser Thr Gln Arg Ser Leu Arg Gly Thr Gly Arg Asn Val Ser Val Thr
      420     425     430
Ser Gln Ser Gly Lys Phe Ile His Ser Trp Glu Ser Tyr Lys Ser Gly
      435     440     445

```

[0009]

Gly Glu Thr Gly Leu
450

<210> 8

<211> 453

<212> PRT

<213> artificial sequence

<220>

<223> HPV-18 membrane-anchored and replication-defective E2
polypeptide

<400> 8

Met	Val	Pro	Gln	Ala	Leu	Leu	Phe	Val	Pro	Leu	Leu	Val	Phe	Pro	Leu
1			5					10					15		
Cys	Phe	Gly	Lys	Phe	Pro	Ile	Gln	Thr	Pro	Lys	Glu	Thr	Leu	Ser	Glu
		20						25					30		
Arg	Leu	Ser	Cys	Val	Gln	Asp	Lys	Ile	Ile	Asp	His	Tyr	Glu	Asn	Asp
		35					40					45			
Ser	Lys	Asp	Ile	Asp	Ser	Gln	Ile	Gln	Tyr	Trp	Gln	Leu	Ile	Arg	Trp
	50					55				60					
Ala	Asn	Ala	Ile	Phe	Phe	Ala	Ala	Arg	Glu	His	Gly	Ile	Gln	Thr	Leu
65					70					75					80
Asn	His	Gln	Val	Val	Pro	Ala	Tyr	Asn	Ile	Ser	Lys	Ser	Lys	Ala	His
			85					90						95	
Lys	Ala	Ala	Glu	Leu	Gln	Met	Ala	Leu	Gln	Gly	Leu	Ala	Gln	Ser	Arg
			100					105					110		
Tyr	Lys	Thr	Glu	Asp	Trp	Thr	Leu	Gln	Asp	Thr	Cys	Glu	Glu	Leu	Trp
		115					120					125			
Asn	Thr	Glu	Pro	Thr	His	Cys	Phe	Lys	Lys	Gly	Gly	Gln	Thr	Val	Gln
	130					135					140				
Val	Tyr	Phe	Asp	Gly	Asn	Lys	Asp	Asn	Cys	Met	Thr	Tyr	Val	Ala	Trp
145					150					155					160
Asp	Ser	Val	Tyr	Tyr	Met	Thr	Asp	Ala	Gly	Thr	Trp	Asp	Lys	Thr	Ala
			165					170						175	
Thr	Cys	Val	Ser	His	Arg	Gly	Leu	Tyr	Val	Lys	Glu	Gly	Tyr	Asn	
		180						185				190			
Thr	Phe	Tyr	Ile	Glu	Phe	Lys	Ser	Glu	Cys	Glu	Lys	Tyr	Gly	Asn	Thr
		195					200					205			
Gly	Thr	Trp	Glu	Val	His	Phe	Gly	Asn	Asn	Val	Ile	Asp	Cys	Asn	Asp
	210					215					220				
Ser	Met	Cys	Ser	Thr	Ser	Asp	Asp	Thr	Val	Ser	Ala	Thr	Gln	Leu	Val
225					230					235					240
Lys	Gln	Leu	Gln	His	Thr	Pro	Ser	Pro	Tyr	Ser	Ser	Thr	Val	Ser	Val
				245					250					255	
Gly	Thr	Ala	Lys	Thr	Tyr	Gly	Gln	Thr	Ser	Ala	Ala	Thr	Arg	Pro	Gly
		260						265					270		
His	Cys	Gly	Leu	Ala	Glu	Lys	Gln	His	Cys	Gly	Pro	Val	Asn	Pro	Leu
		275					280					285			
Leu	Gly	Ala	Ala	Thr	Pro	Thr	Gly	Asn	Asn	Lys	Arg	Arg	Lys	Leu	Cys
	290					295					300				
Ser	Gly	Asn	Thr	Thr	Pro	Ile	Ile	His	Leu	Lys	Gly	Asp	Arg	Asn	Ser
305					310					315					320
Leu	Lys	Cys	Leu	Arg	Tyr	Arg	Leu	Arg	Lys	His	Ser	Asp	His	Tyr	Arg
				325					330					335	
Asp	Ile	Ser	Ser	Thr	Trp	His	Trp	Thr	Gly	Ala	Gly	Asn	Glu	Lys	Thr
			340					345					350		
Gly	Ile	Leu	Thr	Val	Thr	Tyr	His	Ser	Glu	Thr	Gln	Arg	Thr	Lys	Phe
		355					360					365			
Leu	Asn	Thr	Val	Ala	Ile	Pro	Asp	Ser	Val	Gln	Ile	Leu	Val	Gly	Tyr

[0010]

370	375	380
Met Thr Met Tyr Val Leu Leu Ser Ala Gly Ala Leu Thr Ala Leu Met		
385	390	395
Leu Ile Ile Phe Leu Met Thr Cys Cys Arg Arg Val Asn Arg Ser Glu		400
	405	410
Pro Thr Gln His Asn Leu Arg Gly Thr Gly Arg Glu Val Ser Val Thr		415
	420	425
Pro Gln Ser Gly Lys Ile Ile Ser Ser Trp Glu Ser His Lys Ser Gly		430
	435	440
Gly Glu Thr Arg Leu		445
450		

<210> 9

<211> 59

<212> DNA

<213> artificial sequence

<220>

<223> portion of 59 nucleotides of HPV-16 E1-encoding sequences
degenerated to decrease homology with the overlapping
portion in HPV-16 E2-encoding sequences

<400> 9

atggtgattc attacctaca ttcaagtgcg tatctggtca gaacacaaat actttgtga 59

<210> 10

<211> 2214

<212> DNA

<213> artificial sequence

<220>

<223> Nucleotide sequence encoding a membrane-anchored and
replication-defective HPV-16 E1 polypeptide degenerated
in the 59 nucleotides portion overlapping with the native
HPV-16 E2-encoding sequence.

<400> 10

atggtaccgc	aagccctgct	attcgtaoct	ttattggtct	ttcccctctg	tttcggtaag	60
tttcctatag	ctgatcctgc	aggtaccaat	ggggaagagg	gtacgggatg	taatggatgg	120
ttttatgtag	aggctgtagt	ggaaaaaaaa	acaggggatg	ctatatcaga	tgacgagaac	180
gaaaatgaca	gtgatacagg	tgaagatttg	gtagatttta	tagtaaatga	taatgattat	240
ttaacacagg	cagaaacaga	gacagcacat	gcgttgttta	ctgcacagga	agcaaaaacaa	300
catagagatg	cagtacagggt	tctaaaacga	aagtattttg	gtagtccact	tagtgatatt	360
agtggatgtg	tagacaataa	tattagtcct	agattaaaag	ctatatgtat	agaaaaacaa	420
agtagagctg	caaaaaggag	attatttgaa	agcgaagaca	gcgggtatgg	caatactgaa	480
gtggaaactc	agcagatggt	acaggtagaa	gggcgccatg	agactgaaac	accatgtagt	540
cagtatagtg	gtggaagtgg	gggtggttgc	agtcagtaca	gtagtggaag	tggggggagag	600
gggtgttagtg	aaagacacac	tatatgccaa	acaccactta	caaatatatt	aaatgtacta	660
aaaactagta	atgcaaaggc	agcaatgtta	gcaaaattta	aagagttata	cgggggtgagt	720
ttttcagaat	tagtaagacc	atttaaaagt	aataaatcaa	cgtgttgcca	ttggtgtatt	780
gctgcatttg	gacttacacc	cagtatagct	gacagtataa	aaacactatt	acaacaatat	840
tgtttatatt	tacacattca	aagtttagca	tgttcatggg	gaatggttgt	gttactatta	900
gtaagatata	aatgtggaaa	aaatagagaa	acaattgaaa	aattgctgtc	taaactatta	960
tgtgtgtctc	caatgtgtat	gatgatagag	cctccaaaat	tgcgtagtac	agcagcagca	1020
ttatattggt	ataaaacagg	tatatcaaat	attagtgaag	tgtatggaga	cacgccagaa	1080
tgatatacaa	gacaaacagt	attacaacat	agttttaatg	attgtacatt	tgaattatca	1140
cagatggtag	aatgggccta	cgataatgac	atagtagacg	atagtgaat	tgcatataaaa	1200
tatgcacaat	tggcgacac	taatagtaat	gcaagtgcct	ttctaaaaag	taattcacag	1260
gcaaaaattg	taaaggattg	tgcaacaatg	tgtagacatt	ataaacgagc	agaaaaaaaa	1320

[0011]

caaagtga	tgagtcaatg	gataaaatat	agatgtgata	gggtagatga	tggagggtgat	1380
tggagcaaa	ttgttatgtt	tttaagggtat	caagggtgtag	agtttatgtc	atttttaact	1440
gcattaaaa	gatttttgca	aggcatacct	aaaaaaaatt	gcatattact	atatggtgca	1500
gctaacacag	ataaatcatt	atgttggtatg	agtttaataga	aatttctgca	agggtctgta	1560
atatgttttg	taaattctaa	aagccatttt	tggttacaac	cattagcaga	tgccaaaata	1620
ggtatgtag	atgatgctac	agtgcctgt	tggaaactata	tagatgacaa	tttaagaaat	1680
gcattggatg	gaaatttagt	ttctatggat	gtaaagcata	gaccattggt	acaactaaaa	1740
tgccctccat	tattaattac	atctaacatt	aatgctggta	cagattctag	gtggccttat	1800
ttacataata	gattggtggt	gtttacattt	cctaatagag	ttccatttga	cgaaaacgga	1860
aatccagtgt	atgagcttaa	tgataagaac	tggaaatcct	ttttctcaag	gacgtggtcc	1920
agattaagtt	tgacgagga	cgaggacaag	gaaaacgatg	gtgattcatt	acctacattc	1980
aagtgcgtat	ctggtcagaa	cacaaatact	ttgtacgtac	tgctatcggc	aggcacgttg	2040
atcgactaa	tgcttatcat	cttcctaata	acctgctgca	agcgggttga	tagggccgaa	2100
agtacccaaa	ggtccttgag	aggtagcgga	cgcaacgtat	cggtaacgtc	gcaaagcggc	2160
aagttcatta	gcagttggga	gtcgacacaaa	tcaggtggag	agaccgcct	gtga	2214

<210> 11

<211> 2241

<212> DNA

<213> artificial sequence

<220>

<223> nucleotide sequence encoding a membrane-anchored and replication-defective HPV-18 E1 polypeptide degenerated to decrease homology with E1-encoding HPV-16 sequence

<400> 11

atgggtctca	agggtgaacgt	ctctgccata	ttcatggcag	tactgttaac	tctccaaaca	60
cccaccggtc	aaatccattg	gggcgcagac	ccagaaggca	cagacggaga	aggcacgggt	120
tgcaacggct	ggttctacgt	acaagctatt	gtagacaaga	agaccggaga	tgtaatttct	180
gacgatgagg	acgagaatgc	aacagacaca	gggtcggata	tgggtgactt	cattgataca	240
caaggaacat	tttgtgaaca	agccgagcta	gaaactgctc	aggcattggt	ccatgcgcag	300
gaggtccaca	atgatgcaca	agtgttgcat	gttttaaagc	ggaagtttgc	aggagggcagc	360
acagaaaaca	gtccattagg	ggagcggctg	gaggtggata	cagagttaag	cccacgggtta	420
caagaaatat	ctttaaataag	tgggcagaaa	aaggctaaga	ggcggctggt	tacaatatca	480
gatagtggct	acggctgttc	tgaggtggaa	gcaacacaga	ttcaggtaac	tacaaatggc	540
gaacatggcg	gcaatgtatg	cagtggcgcg	agtagcgagg	ctatagacaa	cggagggcaca	600
gagggcaaca	acagcagtg	agcgggtaca	agcgacaata	gcaatataga	aaatgtaaat	660
ccacaatgta	ccatagcaca	attaaaagac	ttgttaaaag	taaacaataa	acaaggagct	720
atgcttgcag	tattcaagga	cacatatggg	ctatcattta	cagatttagt	tagaaatttc	780
aagagtgaca	aaaccacatg	tacagactgg	gttacagcta	tattcggagt	aaacccaaca	840
atcgacaga	gatttaagac	tctaatacag	ccatttatat	tgtatgccca	tatacaatgt	900
ctagactgta	agtggggtgt	attaatatta	gccctgttgc	gttacaagtg	cggtaaagag	960
agactaacag	ttgctaaagg	tttaagtacg	ttgttacacg	tacctgaaac	ttgcatgtta	1020
attcaaccac	ctaagttacg	aagtagtggt	gctgcactat	actggtacag	aactggaatt	1080
tctaacataa	gcgaggtaat	gggtgacaca	cctgagtggg	ttcagagact	tactattata	1140
cagcatggaa	tagacgatag	caatttcgat	ttgtcagaaa	tggttcagtg	ggcatttgac	1200
aacgagctga	cagatgaaag	cgatatggca	tttgaatacg	ccttattagc	tgacagcaac	1260
agcaacgcag	ctgcattttt	aaagagcaat	tgccaagcta	aatattttaa	agactgtgcc	1320
actatgtgca	aacactatag	gcgtgcccag	aaacgacaga	tgaatatgtc	acagtggatt	1380
cgatttaggt	gttcaaaaat	agacgaagg	ggagactgga	gaccaatagt	gcaattcctg	1440
cgataccaac	aaatagaatt	cataacattc	ttaggagcct	tgaaatcatt	cttaaaagga	1500
acccccaaga	agaactgttt	agtattttgt	ggaccagcaa	atactgacaa	gtcatatttc	1560
ggaatgagct	ttatacactt	tatacaagga	gcagttatat	cattcgtgaa	ctccactagt	1620
cacttctggc	tggaaaccgtt	aacagacact	aagggtggcca	tgctagacga	cgcaacgacc	1680
acgtgctgga	catactttga	tacctatatg	aggaacgcgt	tagacggcaa	tccaataagt	1740
attgatagaa	aacacaaacc	tttaatacag	cttaagtgtc	cgccaatact	actaaccaca	1800
aatatacatc	cagcaaaagga	taatagtatg	ccatacttag	aaagtagaat	aaagtagatt	1860
gaattcccaa	atgcattccc	gttcgatata	aatggcaacc	ctgtatacga	aataaacgac	1920
aaaaattgga	agtgtttctt	tgaaagaaca	tgggtcaagg	tagattttaca	tgaagaagaa	1980

[0012]

gaagatgctg	atacagaggg	taatccattt	ggtactttca	aattacgagc	tggacagaat	2040
cacaggcctc	ttggtttatc	gagcactagc	atagtctaca	tcctgattgc	agtgtgtctt	2100
ggagggttga	tagggatccc	cgctttaata	tgttgctgca	gggggcgttg	taacaaaaag	2160
ggagaacaag	ttggtatgtc	aagaccaggc	ctaaagcctg	atcttacggg	aacatcaaaa	2220
tcctatgtaa	ggtcgtctg	a				2241

<210> 12

<211> 1362

<212> DNA

<213> artificial sequence

<220>

<223> nucleotide sequence encoding a membrane-anchored and
replication-defective HPV-16 E2 polypeptide

<400> 12

atggtaccac	aagcgctggt	acttgtccca	ctgcttggtt	tctctttatg	ttttggaaaa	60
ttcccaatag	agactctttg	ccaacgttta	aatgtgtgtc	aggacaaaat	actaacacat	120
tatgaaaatg	atagtacaga	cctacgtgac	catatagact	attggaaaca	catgcgccta	180
gcatgtgcta	tttattacaa	ggccagagaa	atgggattta	aacatattaa	ccaccagggtg	240
gtgccaacgc	tggctgtatc	aaagaataaa	gcattacaag	cagctgaact	gcaactaacg	300
ttagaaaaca	tatataactc	acaatatagt	aatgaaaagt	ggacattaca	agacgttagc	360
cttgaagtgt	atttaactgc	accaacagga	tgtataaaaa	aacatggata	tacagtggaa	420
gtgcagtttg	atggagacat	atgcaataca	atgcattata	caaactggac	acatatatat	480
atgtgtgaag	aagcatcagt	aactgtggta	gagggtcaag	ttgactatta	tggtttatat	540
tatgttcatg	aaggaatacg	aacatatatt	gtgcagttta	aagatgatgc	agaaaaatat	600
agtaaaaaata	aagtatggga	agttcatgcy	ggtggtcagg	taatatattg	tcctacatct	660
gtgttttagca	gcaacgaagt	atcctctcct	gaaattatta	ggcagcactt	ggccaaccac	720
cccgcgcgca	cccataccaa	agccgtcgcc	ttgggcaccg	aagaaacaca	gacgactatc	780
cagcgaccaa	gatcagagcc	agacaccgga	aacccttgcc	acaccactaa	gttggtgcac	840
agagactcag	tggacagtgc	tccaatcctc	actgcattta	acagctcaca	caaaggacgg	900
attaactgta	atagtaacac	tacacccata	gtacatttta	aagggtgatgc	taatacttta	960
aaatgtttta	gatatagatt	taaaaagcat	tgtacattgt	atactgcagt	gtcgtctaca	1020
tggcattgga	caggacataa	tgtaaaacat	aaaagtgcaa	ttgttacact	tacatatgat	1080
agtgaatggc	aacgtgacca	atttttgtct	caagttaaaa	taccaaaaac	tattacagtg	1140
tctactggat	ttatgtctat	atatgttctt	ctctctgctg	gaactttaat	agctttaatg	1200
tttaataatg	tcttaataac	gtgctgtaaa	agggtagacc	gtccagagtc	aactcagcgc	1260
agccttaggg	gtactgggag	aaatgtttcc	gtgacatcac	agagtggaaa	atztatctcg	1320
tcttgggaat	ctcataagag	tggaggcgaa	acacgtcttt	ga		1362

<210> 13

<211> 1362

<212> DNA

<213> artificial sequence

<220>

<223> nucleotide sequence encoding a membrane-anchored and
replication-defective HPV-18 E2 polypeptide degeneratde to
reduce homology with native HPV-16 E2-encoding sequence

<400> 13

atggttcctc	aggctctcct	gtttgtaccc	cttctggttt	ttccattgtg	ttttgggaaa	60
ttccctattc	agacaccgaa	ggaaaccctt	tcggaacgat	taagttgcgt	gcaagataag	120
atcatagacc	actacgagaa	cgacagtaaa	gacatagaca	gccaaatata	gtactggcaa	180
ctaatacgtt	gggcaaatgc	aatattcttt	gcagcaaggg	aacatggcat	acagacatta	240
aatcatcagg	tagtcccagc	ctataacatt	tcgaaaagta	aggcacataa	agctgccgag	300
ctccaaatgg	ccctacaagg	ccttgacaaa	agtcgatata	aaaccgagga	ttggactctg	360
caggacacat	gcgaggaact	atggaataca	gaacctactc	actgctttta	gaaaggtggc	420
caaaccgtac	aagtatat	cgacggcaac	aaagacaatt	gtatgacct	tgtagcatgg	480

[0013]

gacagtgtgt	attatatgac	tgatgcagga	acatgggaca	aaaccgctac	ctgtgtaagt	540
cacaggggat	tgtactacgt	aaaggagggg	tacaacacgt	tttatataga	attcaaaagt	600
gaatgtgaga	agtatgggaa	cacaggtacg	tgaggaggtac	atthttgggaa	taatgtcatt	660
gattgtaatg	actctatgtg	cagtaccagt	gacgacacgg	tctccgctac	tcagcttggt	720
aaacagctac	agcacacccc	ctcaccggtat	tccagcaccg	tgtccgtggg	aaccgcaaag	780
acctacggcc	agacgtcggc	tgctacacga	cctggccact	gtggactcgc	ggagaagcag	840
cattgtggac	ctgtcaaccc	acttctcggg	gcagctacac	ctacaggcaa	caacaagaga	900
cgaaaactct	gcagtggtaa	tacgacgcct	ataatacact	tgaagggaga	cagaaacagt	960
ttgaagtgtc	tacggtacag	gttgcgaaaa	catagcgacc	actatagaga	tatatcatcc	1020
acctggcact	ggaccggtgc	aggcaatgaa	aaaacaggaa	tactgactgt	aacctaccat	1080
agcgaaacac	aaagaacaaa	attcttaaat	actgttgcaa	ttccagatag	tgtacaaata	1140
ttggtgggat	acatgacaat	gtatgtatta	ctgagtgcag	gggccctgac	tgcttgatg	1200
ttgataatth	tcctgatgac	atgttgtaga	agagtcaatc	gatcagaacc	tacgcaacac	1260
aatctcagag	ggacagggag	ggaggtgtca	gtcactcccc	aaagcgggaa	gatcataatc	1320
tcatgggaat	cacacaagag	tgggggtgag	accagactgt	ga		1362

<210> 14

<211> 741

<212> DNA

<213> artificial sequence

<220>

<223> Nucleotide sequence encoding a membrane-anchored and non-oncogenic HPV-18 E6 polypeptide degenerated to reduce homology with the native HPV-16 E6-encoding sequence

<400> 14

atgggtaccgc	aagccctgct	attcgtagct	ttattgggtc	ttcccctctg	tttcggtaag	60
tttcctatag	gatctatggc	gcgctttgag	gatccaacac	ggcgacccta	caagctacct	120
gatctgtgca	cggaactgaa	cacttcaactg	caagacatag	aaataacctg	tgtatatattg	180
aagacagtat	tggaaacttac	agaggtatth	gaattttgcat	ttaaagacct	atthtgggtg	240
tatcgtgaca	gtatacccca	tgcgcgatgc	cataagtgtg	tagattttta	ctctagaatc	300
agagaattaa	ggcactattc	agactctgtg	tacggagaca	cattggaaaa	actaactaac	360
actgggttat	acaattttatt	aataagatgc	ctgcgggtgc	agaaaccggt	gcttagacac	420
cttaatgaaa	aacgacgatt	tcacaacata	gctgggcact	atagaggcca	gtgccattcg	480
tgctgcaacc	gagcacgaca	ggaacgactc	caacgacgca	gggagacaca	agtaagatcc	540
tacgtactgc	tatcggcagg	cacgttgatc	gcactaatgc	ttatcatctt	cctaataacc	600
tgctgcaagc	gggttgatag	gcccgaaggt	accgaaaggt	ccttgagagg	taccggacgc	660
aacgtatcgg	taacgtcgcg	aagcggcaag	ttcattagca	gctggggagtc	gcacaaatca	720
gggtggagaga	cccgcctgtg	a				741

<210> 15

<211> 585

<212> DNA

<213> artificial sequence

<220>

<223> Nucleotide sequence encoding a membrane-anchored and non-oncogenic HPV-18 E7 polypeptide degenerated to reduce homology with the native HPV-16 E7-encoding sequence

<400> 15

atgggtctca	aggtgaacgt	ctctgccata	ttcatggcag	tactgttaac	tctccaaaca	60
cccaccggtc	aatccattg	gggcagatct	atgcacggac	ctaaggcaac	actgcaagac	120
attgtattgc	atthtagagcc	ccaaaatgaa	attccgggtg	cacagthtaag	cgactcagag	180
gaagaaaacg	acgagattga	cggagthtaac	catcaacatt	taccagccc	acgagctgaa	240
ccacaacgct	acacaatgth	gtgtatgtgc	tgtaaatgcg	aagccagaat	tgagctggta	300
gtagagagct	cagcagacga	ccttcgagca	ttccagcagc	tatttctgaa	caccctgtcc	360
tttgtctgtc	cgtgggtgtgc	atcccagcag	ggatctgggt	tatcgagcac	tagcatagtc	420

[0014]

tacatcctga ttgcagtgtg tcttggaggg ttgatagga tccccgcttt aatatgttgc	480
tgcagggggc gttgtaacaa aaagggagaa caagttggta tgtcaagacc aggcctaaag	540
cctgatctta cgggaacatc aaaatcctat gtaaggtcgc tctga	585

<210> 16
 <211> 34
 <212> DNA
 <213> artificial sequence

<220>
 <223> sense primer to amplify HPV-16 E2 sequence from CaSki cells

<400> 16 aaaccggat ccatggagac tctttgcaa cggt	34
---	----

<210> 17
 <211> 49
 <212> DNA
 <213> artificial sequence

<220>
 <223> antisense primer to isolate HPV-16 E2 sequence from caSki cells

<400> 17 aaaccgaat tcaagcttag atcttcatat agacataaat ccagtagac	49
--	----

<210> 18
 <211> 33
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer for reassembling HPV16 SS-E2muté-TMR sequence

<400> 18 aaaccggat ccatggtacc acaagcgctg tta	33
---	----

<210> 19
 <211> 53
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer for reassembling HPV-16 SS-E2muté-TMR sequence

<400> 19 tctctttatg ttttggaaaa ttccaatag agactctttg ccaacgttta aat	53
---	----

<210> 20
 <211> 53
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer for reassembling HPV-16 SS-E2muté-TMR sequence

[0015]

<400> 20
attttaaactg tggcaaagag tctctattgg gaattttcca aaacataaag aga 53

<210> 21
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer for reassembling HPV-16 SS-E2muté-TMR sequence

<400> 21
cagtgtctac tggatttatg tctatatatg ttcttctctc tgctggaac 49

<210> 22
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer for reassembling HPV-16 SS-E2muté-TMR sequence

<400> 22
gttcacgac agagaagaac atatataagc ataatccag tagacactg 49

<210> 23
<211> 44
<212> DNA
<213> artificial sequence

<220>
<223> primer for reassembling HPV-16 SS-E2muté-TMR sequence

<400> 23
aaaccagat cttcaaagac gtgtttcgcc tccactctta tgag 44

<210> 24
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer for isolating HPV-16 E1 sequence from CaSki cells

<400> 24
aaaccggat ccatggctga tctgcaggt acca 34

<210> 25
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer for isolating HPV-16 E1 sequence from CaSki cells

[0016]

<400> 25
aaacccgaat tccattatcg taggccatt gtac 34

<210> 26
<211> 32
<212> DNA
<213> artificial sequence

<220>
<223> primer for isolating HPV-16 E1 sequence from CaSki cells

<400> 26
aaacccggat ccgagacacg ccagaatgga ta 32

<210> 27
<211> 50
<212> DNA
<213> artificial sequence

<220>
<223> primer for isolating HPV-16 E1 sequence from CaSki cells

<400> 27
aaacccgaat tcaagcttag atcttcataa tgtgttagta tttgtcctg 50

<210> 28
<211> 88
<212> DNA
<213> artificial sequence

<220>
<223> primer for generating E1 degenerated sequence

<400> 28
aaaccagat cttcaciaaag tatttgtgtt ctgaccagat acgcacttga atgtaggtaa 60
tgaatcacca tcgttttctt tgtcctcg 88

<210> 29
<211> 21
<212> DNA
<213> artificial sequence

<220>
<223> primer for generating degenerated HPV-16 E1 sequence

<400> 29
gatgctacag tgccctgttg g 21

<210> 30
<211> 35
<212> DNA
<213> artificial sequence

<220>
<223> primer for reconstituting sequence encoding
HPV-16 SS-E1*deg-TMR

[0017]

<400> 30
 aaaccaagg atccatggta ccgcaagccc tgcta 35

<210> 31
 <211> 52
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer for reconstituting sequence encoding
 HPV-16 SS-E1*deg-TMR

<400> 31
 ttcccctctg tttcggtaag tttcctatag ctgacacctgc aggtaccaat gg 52

<210> 32
 <211> 52
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer for reconstituting sequence encoding
 HPV-16 SS-E1*deg-TMR

<400> 32
 ccattggtac ctgcaggatc agctatagga aacttaccga aacagagggg aa 52

<210> 33
 <211> 50
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer oTG17563 for reconstituting sequence encoding
 HPV-16 SS-E1

<400> 33
 tatctggtca gaacacaaat actttgtacg tactgctatc ggcaggcacg 50

<210> 34
 <211> 50
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer oTG17564 for reconstituting sequence encoding
 HPV-16 SS-E1*deg-TMR

<400> 34
 cgtgcctgcc gatagcagta cgtacaaagt atttgtgttc tgaccagata 50

<210> 35
 <211> 42
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer oTG17565 for reconstituting sequence encoding

[0018]

SS-E1*deg-TMR

<400> 35 aaacccaaag atcttcacag gcgggtctct ccacctgatt tg	42
<210> 36 <211> 33 <212> DNA <213> artificial sequence	
<220> <223> primer oTG15315 for reconstituting sequence encoding HPV-18 SS6E1*deg-TMF	
<400> 36 ggggagatct atgggtctca aggtgaacgt ctc	33
<210> 37 <211> 39 <212> DNA <213> artificial sequence	
<220> <223> primer oTG17881 for reconstituting sequence encoding HPV-18 SS-E1*deg-TMF	
<400> 37 gtgccttctg ggtctgcgcc ccaatggatt tgaccggtg	39
<210> 38 <211> 37 <212> DNA <213> artificial sequence	
<220> <223> primer oTG17882 for reconstituting sequence encoding HPV-18 SS-E1*deg-TMF	
<400> 38 ggtcaaatcc attggggcgc agaccagaa ggcacag	37
<210> 39 <211> 42 <212> DNA <213> artificial sequence	
<220> <223> primer oTG17883 for reconstituting sequence encoding HPV-18 SS-E1*deg-TMF	
<400> 39 cagaatcaca ggcctcttg tttatcgagc actagcatag tc	42
<210> 40 <211> 39 <212> DNA	

[0019]

<213> artificial sequence
 <220>
 <223> primer oTG17884 for reconstituting sequence encoding
 HPV-18 SS-E1*deg-TMF
 <400> 40
 gctagtgtctc gataaaccaa gaggcctgtg attctgtcc 39
 <210> 41
 <211> 32
 <212> DNA
 <213> artificial sequence
 <220>
 <223> primer oTG17885 for reconstituting sequence encoding HPV-18
 SS-E1*deg-TMF
 <400> 41
 gggggcggcc gctcagagcg accttacata gg 32
 <210> 42
 <211> 31
 <212> DNA
 <213> artificial sequence
 <220>
 <223> sense primer oTG17875 for reconstituting sequence
 encoding HPV-18 SS-E2*deg-TMR
 <400> 42
 ggggagatct atggttcctc aggctctcct g 31
 <210> 43
 <211> 39
 <212> DNA
 <213> artificial sequence
 <220>
 <223> sense primer oTG17876 for reconstituting sequence
 encoding HPV-18 SS-E2*deg-TMR
 <400> 43
 gttttgggaa attccctatt cagacaccga aggaaaccc 39
 <210> 44
 <211> 43
 <212> DNA
 <213> artificial sequence
 <220>
 <223> antisense primer oTG17877 for reconstituting sequence
 encoding HPV-18 SS-E2*deg-TMR
 <400> 44
 gtttccttcg gtgtctgaat aggaatttc caaaacaca atg 43

[0020]

<210> 45
<211> 39
<212> DNA
<213> artificial sequence

<220>
<223> sense primer oTG17878 for reconstituting sequence
encoding HPV-18 SS-E2*deg-TMR

<400> 45
gtgggataca tgacaatgta tgtattactg agtgcaggg 39

<210> 46
<211> 38
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG17879 for reconstituting sequence
encoding HPV-18 SS-E2*deg-TMR

<400> 46
ctgcactcag taatacatatc attgtcatgt atcccacc 38

<210> 47
<211> 29
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG17880 for reconstituting sequence
encoding HPV-18 SS-E2*deg-TMR

<400> 47
gggggcggcc gctcacagtc tgggtctcac 29

<210> 48
<211> 83
<212> DNA
<213> artificial sequence

<220>
<223> sense primer oTG15174 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 48
ggggagatct atggcgcgct ttgaggatcc aacacggcga ccctacaagc tacctgatct 60
gtgcacggaa ctgaacactt cac 83

<210> 49
<211> 84
<212> DNA
<213> artificial sequence

<220>
<223> sense primer oTG15175 for reconstituting sequence

[0021]

encoding HPV-18 E6*deg

<400> 49

gtattggaac ttacagaggt atttgaattt gcatttaaag acctatttgt ggtgtatcgt 60
gacagtatac cccatgccgc atgc 84

<210> 50

<211> 85

<212> DNA

<213> artificial sequence

<220>

<223> sense primer oTG15176 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 50

aggcactatt cagactctgt gtacggagac acattggaaa aactaactaa cactgggtta 60
tacaatttat taataagatg cctgc 85

<210> 51

<211> 72

<212> DNA

<213> artificial sequence

<220>

<223> sense primer oTG15177 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 51

ccttaatgaa aaacgacgat ttcacaacat agctgggcac tatagaggcc agtgccattc 60
gtgctgcaac cg 72

<210> 52

<211> 76

<212> DNA

<213> artificial sequence

<220>

<223> antisense primer oTG15178 for reconstituting sequence
encoding HPV-18 E2*deg

<400> 52

ggggagatct tacttgtgtc tccctgcgtc gttggagtcg ttcctgtcgt gctcgggttc 60
agcacgaatg gactg 76

<210> 53

<211> 78

<212> DNA

<213> artificial sequence

<220>

<223> antisense primer oTG15179 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 53

gttgtgaaat cgtcgttttt cattaagggtg tetaagcaac ggtttctggc accgcaggca 60
tcttattaat aaattgta 78

[0022]

<210> 54
<211> 82
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG15180 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 54
cacagagtct gaatagtgcc ttaattctct gattctagag taaaaatcta tacacttatg 60
gcatgcggca tggggtatac tg 82

<210> 55
<211> 86
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG15181 for reconstituting sequence
encoding HPV-18 E6*deg

<400> 55
caaatacctc tgtaagttcc aatactgtct tacaatatac acaggttatt tctatgtctt 60
gcagtgaagt gttcagttcc gtgcac 86

<210> 56
<211> 70
<212> DNA
<213> artificial sequence

<220>
<223> sense primer oTG14773 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 56
aaaccagat ctatgcacgg acctaaggca aactgcaag acattgtatt gcatttagag 60
ccccaaaatg 70

<210> 57
<211> 76
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG14774 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 57
aatctcgtcg ttttcttcct ctgagtcgct taactgtgca accggaattt cattttgggg 60
ctctaaatgc aataca 76

<210> 58
<211> 76
<212> DNA
<213> artificial sequence

[0023]

<220>
<223> sense primer oTG14775 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 58
actcagagga agaaaacgac gagattgacg gagttaatca tcaacattta ccagcccgac 60
gagctgaacc acaacg 76

<210> 59
<211> 76
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG14776 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 59
ctagctcaat tctggcttcg catttacagc acatacacia cattgtgtga cgttggtggt 60
cagctcgtcg ggctgg 76

<210> 60
<211> 76
<212> DNA
<213> artificial sequence

<220>
<223> sense primer oTG14777 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 60
taaattgcgaa gccagaattg agctagtagt agagagctca gcagacgacc ttcgagcatt 60
ccagcagcta tttctg 76

<210> 61
<211> 79
<212> DNA
<213> artificial sequence

<220>
<223> antisense primer oTG14778 for reconstituting sequence
encoding HPV-18 E7*deg

<400> 61
aaacccggat ccctgctggg atgcacacca cggacagaca aaggacaggg tgttcagaaa 60
tagctgctgg aatgctcga 79

<210> 62
<211> 20
<212> DNA
<213> artificial sequence

<220>
<223> primer sens polylinker

<400> 62

[0024]

cctgcagaag cttccccggg

20

<210> 63

<211> 28

<212> DNA

<213> artificial sequence

<220>

<223> anstisense primer polylinker

<400> 63

gatcccccg gaagcttctg caggagct

28

<210> 64

<211> 243

<212> PRT

<213> artificial sequence

<220>

<223> HPV-16 membrane-presented and non-oncogenic E6 variant
(SS-16E6*-TMF)

<400> 64

Met	Gly	Leu	Lys	Val	Asn	Val	Ser	Ala	Ile	Phe	Met	Ala	Val	Leu	Leu
1				5					10					15	
Thr	Leu	Gln	Thr	Pro	Thr	Gly	Gln	Ile	His	Trp	Gly	Met	His	Gln	Lys
			20					25					30		
Arg	Thr	Ala	Met	Phe	Gln	Asp	Pro	Gln	Glu	Arg	Pro	Arg	Lys	Leu	Pro
		35					40					45			
Gln	Leu	Cys	Thr	Glu	Leu	Gln	Thr	Thr	Ile	His	Asp	Ile	Ile	Leu	Glu
	50					55					60				
Cys	Val	Tyr	Cys	Lys	Gln	Gln	Leu	Leu	Arg	Arg	Glu	Val	Tyr	Asp	Phe
65					70					75				80	
Ala	Phe	Arg	Asp	Leu	Cys	Ile	Val	Tyr	Arg	Asp	Gly	Asn	Pro	Tyr	Ala
			85						90					95	
Val	Cys	Asp	Lys	Cys	Leu	Lys	Phe	Tyr	Ser	Lys	Ile	Ser	Glu	Tyr	Arg
			100					105					110		
His	Tyr	Cys	Tyr	Ser	Leu	Tyr	Gly	Thr	Thr	Leu	Glu	Gln	Gln	Tyr	Asn
		115					120					125			
Lys	Pro	Leu	Cys	Asp	Leu	Leu	Ile	Arg	Cys	Ile	Asn	Cys	Gln	Lys	Pro
		130				135					140				
Leu	Gln	Arg	His	Leu	Asp	Lys	Lys	Gln	Arg	Phe	His	Asn	Ile	Arg	Gly
145					150					155				160	
Arg	Trp	Thr	Gly	Arg	Cys	Met	Ser	Cys	Cys	Arg	Ser	Ser	Arg	Thr	Arg
			165					170					175		
Arg	Glu	Thr	Gln	Leu	Gly	Leu	Ser	Ser	Thr	Ser	Ile	Val	Tyr	Ile	Leu
			180				185						190		
Ile	Ala	Val	Cys	Leu	Gly	Gly	Leu	Ile	Gly	Ile	Pro	Ala	Leu	Ile	Cys
		195					200					205			
Cys	Cys	Arg	Gly	Arg	Cys	Asn	Lys	Lys	Gly	Glu	Gln	Val	Gly	Met	Ser
		210				215					220				
Arg	Pro	Gly	Leu	Lys	Pro	Asp	Leu	Thr	Gly	Thr	Ser	Lys	Ser	Tyr	Val
225					230					235				240	
Arg	Ser	Leu													

<210> 65

<211> 185

<212> PRT

[0025]

<213> artificial sequence

<220>

<223> HPV-16 membrane-presented and non-oncogenic E7 variant
(SS-16E7*-TMR)

<400> 65

```

Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu
1           5           10           15
Cys Phe Gly Lys Phe Pro Ile Gly Ser Met His Gly Asp Thr Pro Thr
          20           25           30
Leu His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Gln Leu Asn
          35           40           45
Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala
          50           55           60
Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys
65           70           75           80
Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg
          85           90           95
Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile
          100          105          110
Cys Ser Gln Lys Pro Arg Ser Tyr Val Leu Leu Ser Ala Gly Ala Leu
          115          120          125
Thr Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys Arg Arg Val
          130          135          140
Asn Arg Ser Glu Pro Thr Gln His Asn Leu Arg Gly Thr Gly Arg Glu
145          150          155          160
Val Ser Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser Trp Glu Ser
          165          170          175
His Lys Ser Gly Gly Glu Thr Arg Leu
          180          185

```

<210> 66

<211> 1062

<212> DNA

<213> artificial sequence

<220>

<223> nucleotide sequence encoding a replication-defective
HPV-33 E2 polypeptide (degenerated sequence)

<400> 66

```

atggaggaaa tatcagcacg cttgaatgca gtccaagaga aaattctaga tcttttacgaa      60
gcagataaaa ctgatttacc atctcaaat gaacactgga aattgatacg catggcctgc      120
gctttattgt atacagccaa acagatgggc ttttcacatt tatgtcacca agtggtacct      180
tctttgttag catccaaaac caaagcgttt caagtgcgg aactacagat ggcattagag      240
acattaagta aatcacagta tagcacaagc caatggacgt tgcaacagac aagcttagag      300
gtttggcttt gtgaaccacc aaaatgtttt aaaaagcaag gagaaacagt aactgtgcaa      360
tatgacaatg acaaaaaaaaa taccatggac tatactaact ggggtgaaat atacattata      420
gaggaagata catgtactat ggttacaggg aaagtagatt atataggtat gtattacata      480
cataactgtg aaaaggtata ctttaaatat tttaaggagg atgctgccaa atactctaaa      540
acacaaatgt gggaagtcca tgtaggtggc caggttattg tttgccctac gtctatatct      600
agcaatcaaa tatccactac tgagactgct gacatacaga cagacaacga taaccgacca      660
ccacaagcag cggccaaacg acgacgacct gcagacacta ctgacaccgc ccagcccctt      720
acaaagctgt tctgtgcaga cccgcgcttg gataatagaa cagcacgtac agcaactaac      780
tgcacaaata agcagcggac tgtgtgtagt tctaacgttg caccaatagt gcatttgaaa      840
ggcgaatcaa atagcttaaa gtgtttgaga tacagattaa aaccttataa agagtgtgac      900
agttctatgt cttcaacttg gcaactggact agtgacaaca aaaatagtaa aaatggcata      960
gtaaccgtga catttgtaac tgaacagcaa caacaaatgt tcttgggtac cgtaaagata     1020
cctcctactg tgcagataag taccggattc atgaccttat aa                       1062

```

[0026]

<210> 67
 <211> 1326
 <212> DNA
 <213> artificial sequence

<220>

<223> nucleotide sequence encoding a membrane-presented and replication-defective HPV-33 E2 polypeptide (SS-33E2*-TMR) (degenerated sequence)

<400> 67

atggtaccgc	aagccctgct	attcgtacct	ttattggctct	ttccccctctg	tttcggtaag	60
tttcctatag	aggaaatatac	agcacgcttg	aatgcagtc	aagagaaaat	tctagatctt	120
tacgaagcag	ataaaaactga	tttaccatct	caaattgaac	actggaaatt	gatacgcagt	180
gcctgcgctt	tattgtatac	agccaaacag	atgggctttt	cacatttatg	tcaccaagt	240
gtaccttctt	tgtagcatc	caaaaccaa	gcgtttcaag	tagcggaact	acagatggca	300
ttagagacat	taagtaaatac	acagtatagc	acaagccaat	ggacgttgca	acagacaagc	360
ttagagggtt	ggctttgtga	accaccaaaa	tgtttttaaaa	agcaaggaga	aacagtaact	420
gtgcaatatg	acaatgacaa	aaaaaatacc	atggactata	ctaactgggg	tgaaatatac	480
attatagagg	aagatacatg	tactatgggt	acagggaaag	tagattatat	aggtatgtat	540
tacatacata	actgtgaaaa	ggtatacttt	aaatatttta	aggaggatgc	tgccaaatac	600
tctaaaacac	aaatgtggga	agtccatgta	ggtggccagg	ttattgtttg	ccctacgtct	660
atatctagca	atcaaatac	cactactgag	actgctgaca	tacagacaga	caacgataac	720
cgaccaccac	aagcagcggc	caaacgacga	cgacctgcag	acactactga	caccgcccag	780
ccccttacia	agctgttctg	tgacagcccc	gccttgata	atagaacagc	acgtacagca	840
actaaatgca	caaataagca	gcggactgtg	tgtagttcta	acgttgccac	aatagtgcac	900
ttgaaaggcg	aatcaaatac	cttaaagtgt	ttgagataca	gattaaaacc	ttataaagag	960
ttgtacagtt	ctatgtcttc	aacttggcac	tggtactagt	acaacaaaaa	tagtaaaaat	1020
ggcatagtaa	ccgtgacatt	tgtaactgaa	cagcaacaac	aaatgttctt	gggtaccgta	1080
aagatacctc	ctactgtgca	gataagtacc	ggattcatga	ccttatacgt	actgctatcg	1140
gcaggcacgt	tgatcgcaat	aatgcttatc	atcttcctaa	taacctgctg	caagcggggt	1200
gataggcccg	aaagtaccca	aaggctcctg	agaggtaccg	gacgcaacgt	atcggtaacg	1260
tcgcaaagcg	gcaagttcat	tagcagttgg	gagtcgcaca	aatcaggtgg	agagaccgcg	1320
ctgtga						1326

<210> 68
 <211> 1107
 <212> DNA
 <213> artificial sequence

<220>

<223> nucleotide sequence encoding a replication-defective HPV-52 E2 polypeptide (52degE2*) (degenerated sequence)

<400> 68

atggaatcga	taccggcacg	gttaaaccgct	gtgcaggaaa	agataactcga	tctatatgag	60
gctgacagca	atgatctaaa	cgcacaaatc	gagcattgga	agttgactcg	aatggcttgt	120
gttttggttt	ataaagcaaa	ggaactggga	ataactcata	taggccatca	agtagtgctt	180
ccaatggcag	tgtctaaggc	aaaggcctgc	caagccgcag	agcttcaatt	ggctttggag	240
gcattgaaca	aaactcaata	cagtacagat	ggctggacct	tacagcaaac	aagtctagaa	300
atgtggcgtg	cagagccaca	aaaatacttc	aagaagcacg	ggtacacaat	aacagtccaa	360
tacgataatg	ataaaaacaa	cactatggat	tacacaaatt	ggaaggaaat	ttatttactt	420
ggtgagtgtg	aatgcacaat	tgtagaagga	caagtggatt	actatgggtt	atactattgg	480
tgtgatggag	aaaaaatcta	tttcgtaaaa	tttagtaacg	acgcaaagca	atattgtgta	540
acaggagtct	gggaggtgca	cgtgggcggt	caagtaatcg	tgtgtccagc	atcggtatca	600
agtaacgagg	tttctactac	agaaacagct	gtccacctat	gcaccgaaac	ctccaagacc	660
tccgcagtgt	ccgtgggtgc	caaagacaca	cacctacaac	caccacagaa	gcgacgtcga	720
ccagatgtca	cagattccag	aaacaccaag	tacccaaca	accttttgcg	gggacaacaa	780

[0027]

tccgttgaca	gcactacacg	gggactcgta	actgccactg	agtgcactaa	taaaggctcg	840
gttgcacata	caacttgtac	tgctcctatt	attcacctaa	agggtgaccc	caacagcttg	900
aaatgcctaa	ggtatagggt	aaaaacacat	aaaagtttat	atgttcaa	ttcatctacg	960
tggcattgga	cgagtaaatga	atgtacaaat	aataaactag	gtattgtaac	aataacgtac	1020
agtgatgaga	cacagcgta	acagttttta	aaaactgtca	aaatcccaa	taccgtccaa	1080
gttatacaag	gtgtcatgtc	attgtaa				1107

<210> 69

<211> 1374

<212> DNA

<213> artificial sequence

<220>

<223> nucleotide sequence encoding a membrane-presented and
replication-defective HPV-52 E2 polypeptide (SS-52E2*-TMF)
(degenerated sequence)

<400> 69

atgggtctca	aggagaacgt	ctctgccata	ttcatggcag	tactgttaac	tctccaaaca	60
cccaccggtc	aaatccattg	gggcgaatcg	ataccggcac	ggtaaacgc	tgtgcaggaa	120
aagatactcg	atctatatga	ggctgacagc	aatgatctaa	acgcacaaat	cgagcattgg	180
aagttgactc	gaatggcttg	tgttttgttt	tataaagcaa	aggaactggg	aataactcat	240
ataggccatc	aagtagtgcc	tccaatggca	gtgtctaagg	caaaggcctg	ccaagccgca	300
gagcttcaat	tggctttgga	ggcattgaac	aaaactcaat	acagtacaga	tggctggacc	360
ttacagcaaa	caagtctaga	aatgtggcgt	gcagagccac	aaaaataactt	caagaagcac	420
gggtacacaa	taacagtcca	atacgataat	gataaaaaca	acactatgga	ttacacaaat	480
tgggaaggaaa	tttattttact	tggtagtggt	gaatgcacaa	ttgtagaagg	acaagtggat	540
tactatgggt	tatactattg	gtgtgatgga	gaaaaaatct	atttcgtaaa	atttagtaac	600
gacgcaaagc	aatatttgtgt	aacaggagtc	tgggaggtgc	acgtgggcgg	tcaagtaatc	660
gtgtgtccag	catcgggtatc	aagtaacgag	gtttctacta	cagaaacagc	tgtccaccta	720
tgcaccgaaa	cctccaagac	ctccgcagtg	tccgtgggtg	ccaaagacac	acacctacaa	780
ccaccacaga	agcgacgtcg	accagatgtc	acagattcca	gaaacaccaa	gtaccccaac	840
aaccttttgc	ggggacaaca	atccgttgac	agcactacac	ggggactcgt	aactgccact	900
gagtgcacta	ataaaggctcg	ggttgcacat	acaacttgta	ctgtccttat	tattcaccta	960
aagggtgacc	ccaacagctt	gaaatgccta	aggtataggg	taaaaacaca	taaaagttaa	1020
tatgttcaaa	tttcatctac	gtggcattgg	acgagtaatg	aatgtacaaa	taataaacta	1080
ggtattgtaa	caataacgta	cagtgatgag	acacagcgtc	aacagttttt	aaaaactgtc	1140
aaaatcccaa	ataccgtcca	agttatacaa	gggtgcatgt	cattgggttt	atcgagcact	1200
agcatagtct	acatcctgat	tgcagtgtgt	cttggagggt	tgatagggat	ccccgtttta	1260
atatgttgct	gcagggggcg	ttgtaacaaa	aaggggagaac	aagttgggtat	gtcaagacca	1320
ggcctaaagc	ctgatcttac	gggaacatca	aaatcctatg	taaggtcgct	ctga	1374

<210> 70

<211> 441

<212> PRT

<213> artificial sequence

<220>

<223> membrane-presented and replication-defective HPV-33 E2
polypeptide (SS-33E2*-TMR)

<400> 70

Met	Val	Pro	Gln	Ala	Leu	Leu	Phe	Val	Pro	Leu	Leu	Val	Phe	Pro	Leu
1			5					10						15	
Cys	Phe	Gly	Lys	Phe	Pro	Ile	Glu	Glu	Ile	Ser	Ala	Arg	Leu	Asn	Ala
			20					25					30		
Val	Gln	Glu	Lys	Ile	Leu	Asp	Leu	Tyr	Glu	Ala	Asp	Lys	Thr	Asp	Leu
			35				40					45			
Pro	Ser	Gln	Ile	Glu	His	Trp	Lys	Leu	Ile	Arg	Met	Ala	Cys	Ala	Leu

[0028]

50	55	60
Leu Tyr Thr Ala Lys Gln Met Gly Phe Ser His Leu Cys His Gln Val		
65	70	75
Val Pro Ser Leu Leu Ala Ser Lys Thr Lys Ala Phe Gln Val Ala Glu		80
	85	90
Leu Gln Met Ala Leu Glu Thr Leu Ser Lys Ser Gln Tyr Ser Thr Ser		95
	100	105
Gln Trp Thr Leu Gln Gln Thr Ser Leu Glu Val Trp Leu Cys Glu Pro		110
	115	120
Pro Lys Cys Phe Lys Lys Gln Gly Glu Thr Val Thr Val Gln Tyr Asp		125
	130	135
Asn Asp Lys Lys Asn Thr Met Asp Tyr Thr Asn Trp Gly Glu Ile Tyr		140
145	150	155
Ile Ile Glu Glu Asp Thr Cys Thr Met Val Thr Gly Lys Val Asp Tyr		160
	165	170
Ile Gly Met Tyr Tyr Ile His Asn Cys Glu Lys Val Tyr Phe Lys Tyr		175
	180	185
Phe Lys Glu Asp Ala Ala Lys Tyr Ser Lys Thr Gln Met Trp Glu Val		190
	195	200
His Val Gly Gly Gln Val Ile Val Cys Pro Thr Ser Ile Ser Ser Asn		205
	210	215
Gln Ile Ser Thr Thr Glu Thr Ala Asp Ile Gln Thr Asp Asn Asp Asn		220
225	230	235
Arg Pro Pro Gln Ala Ala Ala Lys Arg Arg Arg Pro Ala Asp Thr Thr		240
	245	250
Asp Thr Ala Gln Pro Leu Thr Lys Leu Phe Cys Ala Asp Pro Ala Leu		255
	260	265
Asp Asn Arg Thr Ala Arg Thr Ala Thr Asn Cys Thr Asn Lys Gln Arg		270
	275	280
Thr Val Cys Ser Ser Asn Val Ala Pro Ile Val His Leu Lys Gly Glu		285
	290	295
Ser Asn Ser Leu Lys Cys Leu Arg Tyr Arg Leu Lys Pro Tyr Lys Glu		300
305	310	315
Leu Tyr Ser Ser Met Ser Ser Thr Trp His Trp Thr Ser Asp Asn Lys		320
	325	330
Asn Ser Lys Asn Gly Ile Val Thr Val Thr Phe Val Thr Glu Gln Gln		335
	340	345
Gln Gln Met Phe Leu Gly Thr Val Lys Ile Pro Pro Thr Val Gln Ile		350
	355	360
Ser Thr Gly Phe Met Thr Leu Tyr Val Leu Leu Ser Ala Gly Thr Leu		365
	370	375
Ile Ala Leu Met Leu Ile Ile Phe Leu Ile Thr Cys Cys Lys Arg Val		380
385	390	395
Asp Arg Pro Glu Ser Thr Gln Arg Ser Leu Arg Gly Thr Gly Arg Asn		400
	405	410
Val Ser Val Thr Ser Gln Ser Gly Lys Phe Ile Ser Ser Trp Glu Ser		415
	420	425
His Lys Ser Gly Gly Glu Thr Arg Leu		430
	435	440

<210> 71

<211> 457

<212> PRT

<213> artificial sequence

<220>

<223> membrane-presented and replication-defective HPV-52 E2 polypeptide (SS-52E2*-TMF)

<400> 71

[0029]

Met	Gly	Leu	Lys	Val	Asn	Val	Ser	Ala	Ile	Phe	Met	Ala	Val	Leu	Leu
1				5					10					15	
Thr	Leu	Gln	Thr	Pro	Thr	Gly	Gln	Ile	His	Trp	Gly	Glu	Ser	Ile	Pro
			20					25					30		
Ala	Arg	Leu	Asn	Ala	Val	Gln	Glu	Lys	Ile	Leu	Asp	Leu	Tyr	Glu	Ala
		35						40				45			
Asp	Ser	Asn	Asp	Leu	Asn	Ala	Gln	Ile	Glu	His	Trp	Lys	Leu	Thr	Arg
	50					55					60				
Met	Ala	Cys	Val	Leu	Phe	Tyr	Lys	Ala	Lys	Glu	Leu	Gly	Ile	Thr	His
65					70					75				80	
Ile	Gly	His	Gln	Val	Pro	Pro	Met	Ala	Val	Ser	Lys	Ala	Lys	Ala	
			85					90					95		
Cys	Gln	Ala	Ala	Glu	Leu	Gln	Leu	Ala	Leu	Glu	Ala	Leu	Asn	Lys	Thr
			100					105					110		
Gln	Tyr	Ser	Thr	Asp	Gly	Trp	Thr	Leu	Gln	Gln	Thr	Ser	Leu	Glu	Met
		115					120					125			
Trp	Arg	Ala	Glu	Pro	Gln	Lys	Tyr	Phe	Lys	Lys	His	Gly	Tyr	Thr	Ile
	130					135					140				
Thr	Val	Gln	Tyr	Asp	Asn	Asp	Lys	Asn	Asn	Thr	Met	Asp	Tyr	Thr	Asn
145					150					155				160	
Trp	Lys	Glu	Ile	Tyr	Leu	Leu	Gly	Glu	Cys	Glu	Cys	Thr	Ile	Val	Glu
			165					170					175		
Gly	Gln	Val	Asp	Tyr	Tyr	Gly	Leu	Tyr	Tyr	Trp	Cys	Asp	Gly	Glu	Lys
		180					185					190			
Ile	Tyr	Phe	Val	Lys	Phe	Ser	Asn	Asp	Ala	Lys	Gln	Tyr	Cys	Val	Thr
	195					200					205				
Gly	Val	Trp	Glu	Val	His	Val	Gly	Gly	Gln	Val	Ile	Val	Cys	Pro	Ala
	210				215					220					
Ser	Val	Ser	Ser	Asn	Glu	Val	Ser	Thr	Thr	Glu	Thr	Ala	Val	His	Leu
225				230						235				240	
Cys	Thr	Glu	Thr	Ser	Lys	Thr	Ser	Ala	Val	Ser	Val	Gly	Ala	Lys	Asp
			245					250					255		
Thr	His	Leu	Gln	Pro	Pro	Gln	Lys	Arg	Arg	Arg	Pro	Asp	Val	Thr	Asp
		260				265						270			
Ser	Arg	Asn	Thr	Lys	Tyr	Pro	Asn	Asn	Leu	Leu	Arg	Gly	Gln	Gln	Ser
	275					280					285				
Val	Asp	Ser	Thr	Thr	Arg	Gly	Leu	Val	Thr	Ala	Thr	Glu	Cys	Thr	Asn
	290				295					300					
Lys	Gly	Arg	Val	Ala	His	Thr	Thr	Cys	Thr	Ala	Pro	Ile	Ile	His	Leu
305				310						315				320	
Lys	Gly	Asp	Pro	Asn	Ser	Leu	Lys	Cys	Leu	Arg	Tyr	Arg	Val	Lys	Thr
			325					330					335		
His	Lys	Ser	Leu	Tyr	Val	Gln	Ile	Ser	Ser	Thr	Trp	His	Trp	Thr	Ser
		340					345					350			
Asn	Glu	Cys	Thr	Asn	Asn	Lys	Leu	Gly	Ile	Val	Thr	Ile	Thr	Tyr	Ser
	355					360					365				
Asp	Glu	Thr	Gln	Arg	Gln	Gln	Phe	Leu	Lys	Thr	Val	Lys	Ile	Pro	Asn
	370				375						380				
Thr	Val	Gln	Val	Ile	Gln	Gly	Val	Met	Ser	Leu	Gly	Leu	Ser	Ser	Thr
385				390						395				400	
Ser	Ile	Val	Tyr	Ile	Leu	Ile	Ala	Val	Cys	Leu	Gly	Gly	Leu	Ile	Gly
			405				410						415		
Ile	Pro	Ala	Leu	Ile	Cys	Cys	Cys	Arg	Gly	Arg	Cys	Asn	Lys	Lys	Gly
		420				425						430			
Glu	Gln	Val	Gly	Met	Ser	Arg	Pro	Gly	Leu	Lys	Pro	Asp	Leu	Thr	Gly
	435					440					445				
Thr	Ser	Lys	Ser	Tyr	Val	Arg	Ser	Leu							
	450					455									

[0030]

<210> 72
<211> 36
<212> DNA
<213> artificial sequence

<220>
<223> priler oTG18962 for reconstituting sequence encoding
SS-33E2*-TMR

<400> 72
cccaaaggat ccaccatggt accgcaagcc ctgcta 36

<210> 73
<211> 52
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18963 for reconstituting sequence encoding
SS-33E2*-TMR

<400> 73
ttccccctctg tttcggtaag tttcctatag aggaaatattc agcacgcttg aa 52

<210> 74
<211> 52
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18964 for reconstituting sequence encoding
SS-33E2*-TMR

<400> 74
ttcaagcgtg ctgatatttc ctctatagga aacttaccga aacagagggg aa 52

<210> 75
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18965 for reconstituting sequence encoding
SS-33E2*-TMR

<400> 75
gataagtacc ggattcatga ccttatacgt actgctatcg gcaggcacg 49

<210> 76
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18966 for reconstituting sequence encoding
SS-33-E2*-TMR

[0031]

<400> 76
cgtgcctgcc gatagcagta cgtataaggt catgaatccg gtacttatac 49

<210> 77
<211> 56
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18967 for reconstituting sequence encoding
SS-33E2*-TMR

<400> 77
aaaacccccgc atgcgcggcc gcaagctatc acaggcgggt ctctccacct gatttg 56

<210> 78
<211> 37
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18968 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 78
aaacccgaga tctaccatgg gtctcaaggt gaacgtc 37

<210> 79
<211> 46
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18969 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 79
cccaccggtc aaatccattg gggcgaatcg ataccggcac ggttaa 46

<210> 80
<211> 46
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18970 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 80
ttaaccgtgc cggatatgat tcgccccaat ggatttgacc ggtggg 46

<210> 81
<211> 43
<212> DNA
<213> artificial sequence

[0032]

<220>
<223> primer oTG18971 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 81
gttatacaag gtgtcatgtc attgggttta tcgagcacta gca 43

<210> 82
<211> 43
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18972 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 82
tgctagtgtc cgataaaccc aatgacatga caccttgtat aac 43

<210> 83
<211> 53
<212> DNA
<213> artificial sequence

<220>
<223> primer oTG18973 for reconstituting sequence encoding
SS-52E2*-TMF

<400> 83
aagcttgcta gccaccggtg gggccgcggc cgctcagagc gaccttacat agg 53

E6 HPV16 (a) 与 E6 HPV18 (b)

```

a  ATGCACCAAAAGAGAACTGCAATGTTTCAGGACCCACAGGAGCGACCCAG  50
      |  ||      ||| |||| |||      |||||
b  .....ATGGCGCGCTTTGAGGATCCAACACGGCGACCCTA  35
      .      .      .      .
a  AAAGTTACCACAGTTATGCACAGAGCTGCAAACAACACTATACATGATATAA  100
      ||| |||| | | |||| || ||| | | | | | |
b  CAAGCTACCTGATCTGTGCACGGAAGTGAACACTTCACTGCAAGACATAG  85
      .      .      .      .
a  TATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATAT  150
      | ||  |||| | | |||| | |  ||  | ||||| |
b  AAATAACCTGTGTATATTGCAAGACAGTATTGGAAGTTACAGAGGTATTT  135
      .      .      .      .
a  GACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATA  200
      || |||| | | |||| | || ||||| | | | |
b  GAATTTGCATTTAAAGATTTATTTGTGGTGTATAGAGACAGTATACCCCA  185
      .      .      .      .
a  TGCTGTATGTGATAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATA  250
      |||| | | |||| | | ||||| ||||| | | |
b  TGCTGCATGCCATAAATGTATAGATTTTTATTCTAGAATTAGAGAATTAA  235
      .      .      .      .
a  GACATTATTGTTATAGTTTGTATGGAACAACATTAGAACAGCAATACAAC  300
      ||||| | | | | ||||| |||| | | | | |
b  GACATTATTCAGACTCTGTGTATGGAGACACATTGGAAAACTAACTAAC  285
      .      .      .      .
a  AAACCGTTGTGTGATTTGTTAATTAGGTGTATTAAGTGTCAAAGCCACT  350
      |  ||| |  ||| |||| |||| |  || | | | |
b  ACTGGGTTATACAATTTATTAATAAGGTGCCTGCGGTGCCAGAAACCGTT  335
      .      .      .      .
a  GTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAAAGATTCCATA  400
      |  ||| | |||| |  |||| ||  |  ||| | | ||| |
b  GAATCCAGCAGAAAACTTAGACACCTTAATGAAAAACGACGATTCACA  385
      .      .      .      .
a  ATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCA  450
      | |||  || | |  || | ||  || | | |||  | ||
b  ACATAGCTGGGCACTATAGAGGCCAGTGCCATTCGTGCTGCAACCGAGCA  435
      .      .      .      .
a  AG.....AACACGTAGAGAAACCCAGCTGTAA  477
      |  ||| ||||| ||  | |||
b  CGACAGGAACGACTCCAACGACGCAGAGAAACACAAGTATAA  477

```

图 2A

HPV16 E6 (a) 与 SEQ ID NO:14 (b)

```

a .....ATGCACCAAAAGAGAACTGCAATGTTTCAGGACCCACAGG 40
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
b .....ATGGCGCGCTTTGAGGATCCAACAC 100

a AGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAAACAACATA 90
  |||||  |||  ||||  |  |  |||||  ||  |||  |  |  |  |
b GGCGACCCTACAAGCTACCTGATCTGTGCACGGAAGTGAACACTTCACTG 150

a CATGATATAATATTAGAATGTGTGTACTGCAAGCAACAGTTACTGCGACG 140
  ||  ||  |||  |  ||  |||||  ||  ||  |||  |  ||  |
b CAAGACATAGAAATAACCTGTGTATATTGTAAGACAGTATTGGAACCTAC 200

a TGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATG 190
  |||||  |||  |||||  |||  ||  |||  |  ||  |||  |  ||
b AGAGGTATTTGAATTTGCATTTAAAGACCTATTTGTGGTGTATCGTGACA 250

a GGAATCCATATGCTGTATGTGATAAATGTTTAAAGTTTTATTCTAAAATT 240
  |  |  ||  ||||  |  |||  ||||  |||  ||  |  |||||  ||||  |||
b GTATACCCCATGCCGCATGCCATAAGTGTATAGATTTTTACTCTAGAATC 300

a AGTGAGTATAGACATTATTGTTATAGTTTGTATGGAACAACATTAGAACA 290
  ||  ||  |  ||  ||  ||||  |  |  ||||  |||  |||||  |||  |
b AGAGAATTAAGGCACTATTCAGACTCTGTGTACGGAGACACATTGGAAAA 350

a GCAATACAACAAACCGTTGTGTGATTTGTTAATTAGGTGTATTAAGTGTG 340
  |  |  ||||  |||  |  ||||  |||||  ||  ||  |  ||  |
b ACTAATAACACTGGGTTATACAATTTATTAATAAGATGCCTGCGGTGCC 400

a AAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAA 390
  |  ||  ||  |  |  |  ||  |||||  ||  |  ||||  |  |
b AGAAACCGTT.....GCTTAGACACCTTAATGAAAAACGA 435

a AGATTCCATAATATAAGGGTTCGGTGGACCGGTCGATGTATGTCTTGTTG 440
  ||||  ||  ||  |||  ||  |  |  |  ||  |  ||  ||  ||  ||
b CGATTTCACAACATAGCTGGGCACTATAGAGGCCAGTGCCATTCGTGCTG 485

a C.....AGATCATCAAGAACACGTAGAGAAACCCAGCTGTAA..... 477
  |  |  ||  |  ||  ||||  |  |  ||  |||  |
b CAACCGAGCACGACAGGAACGACTCCAACGACGCAGGGAGACACAAGTAA 535

```

图 2B