



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201734210 A

(43)公開日：中華民國 106 (2017) 年 10 月 01 日

(21)申請案號：105141650

(22)申請日：中華民國 105 (2016) 年 12 月 15 日

(51)Int. Cl.：

*C12N15/12 (2006.01)**C12N15/86 (2006.01)**C12N7/00 (2006.01)**C07K14/005 (2006.01)**A61K39/21 (2006.01)*

(30)優先權：2015/12/15

歐洲專利局

15200138.4

2016/10/17

歐洲專利局

16194124.0

(71)申請人：傑森疫苗防護公司 (荷蘭) JANSSEN VACCINES & PREVENTION B.V. (NL)
荷蘭(72)發明人：蘭格迪克 約翰納斯 佩特魯斯 瑪利亞 LANGEDIJK, JOHANNES PETRUS
MARIA (NL)

(74)代理人：陳長文

申請實體審查：無 申請專利範圍項數：15 項 圖式數：7 共 137 頁

(54)名稱

人類免疫缺陷病毒抗原、載體、組成物、及其使用方法

HUMAN IMMUNODEFICIENCY VIRUS ANTIGENS, VECTORS, COMPOSITIONS, AND
METHODS OF USE THEREOF

(57)摘要

本文描述了合成的 HIV 包膜蛋白、載體及其組成物，以及用於誘導針對人類免疫缺陷病毒(HIV)感染的保護性免疫之方法。編碼該合成的 HIV 包膜蛋白的病毒表現載體可用於疫苗中以提供針對 HIV 之改善保護性免疫。

Synthetic HIV envelope proteins, vectors and compositions thereof, and methods for inducing protective immunity against human immunodeficiency virus (HIV) infection are described. Viral expression vectors encoding the synthetic HIV envelope proteins can be used in vaccines to provide improved protective immunity against HIV.

指定代表圖：

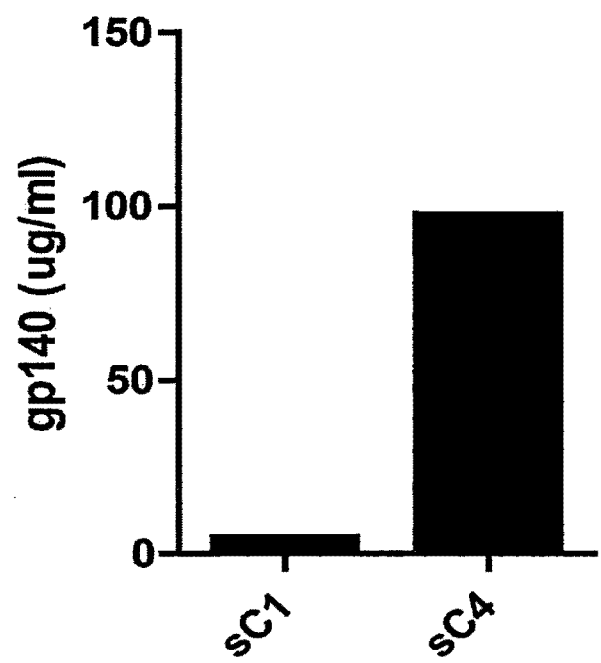


圖 2

201734210

發明摘要

※ 申請案號： 105141650

※ 申請日： 105/12/15

※IPC 分類：

C12N 15/12 (2006.01)**C12N 15/86** (2006.01)**C12N 7/00** (2006.01)**C07K 14/005** (2006.01)**A61K 39/21** (2006.01)

【發明名稱】(中文/英文)

人類免疫缺陷病毒抗原、載體、組成物、及其使用方法 /Human Immunodeficiency Virus Antigens, Vectors, Compositions, and Methods of Use Thereof

【中文】

本文描述了合成的HIV包膜蛋白、載體及其組成物，以及用於誘導針對人類免疫缺陷病毒（HIV）感染的保護性免疫之方法。編碼該合成的HIV包膜蛋白的病毒表現載體可用於疫苗中以提供針對HIV之改善保護性免疫。

【英文】

Synthetic HIV envelope proteins, vectors and compositions thereof, and methods for inducing protective immunity against human immunodeficiency virus (HIV) infection are described. Viral expression vectors encoding the synthetic HIV envelope proteins can be used in vaccines to provide improved protective immunity against HIV.

【代表圖】

【本案指定代表圖】：第（ 2 ）圖。

【本代表圖之符號簡單說明】：

（無）

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

（無）

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

【0001】 人類免疫缺陷病毒抗原、載體、組成物、及其使用方法/[0001]

Human Immunodeficiency Virus Antigens, Vectors, Compositions, and
Methods of Use Thereof

【發明背景】

【0002】 人類免疫缺陷病毒 (HIV) 影響全世界數百萬人，並且即使在廣泛的抗逆轉錄病毒治療的時代，藉由有效疫苗預防 HIV 仍然具有非常高的優先順序。HIV-1 係該病毒中最常見的並且致病的菌株，其中超過 90% 的 HIV/AIDS 病例來源於 HIV-1 組 M 的感染。該 M 組進一步細分為分枝 (clade) 或亞型。有效疫苗理想地能夠引發強大的細胞應答和廣泛中和性抗體兩者，該等廣泛中和性抗體能夠中和來自不同分枝的 HIV-1 菌株。

【0003】 HIV-1 的高遺傳變異性使得 HIV-1 疫苗的開發成為前所未有的挑戰。為了改進潛在 T-細胞表位的覆蓋並改善細胞應答，其他人描述並且開發了來源於 HIV 組抗原 (Gag)、聚合酶 (Pol)、和包膜蛋白 (Env) 蛋白的“花葉 (mosaic)” HIV-1 Gag、Pol 和 Env 抗原，試圖提供潛在 T-細胞表位的最大覆蓋 (例如，巴魯什 (Barouch) 等人，自然醫學 (*Nat Med*) 2010, 16: 319-323)。該等花葉抗原在長度和結構域結構上與野生型，即天然存在的 HIV-1 抗原相似。

【0004】 例如，在疫苗中描述並且使用的花葉 HIV 抗原包括

巴魯什 (Barouch) 等人 (同上), 以及 WO 2010/059732 中描述的那些, 如:

(a) Gag 花葉抗原, 該 Gag 花葉抗原包括:

(a)(i) 具有如在本文 SEQ ID NO: 1 中所示的胺基酸序列的第一花葉 Gag 序列 (“mos1Gag”), 以及

(a)(ii) 具有如在本文 SEQ ID NO: 2 中所示的胺基酸序列的第二花葉 Gag 序列 (“mos2Gag”);

(b) Pol 花葉抗原, 該 Pol 花葉抗原包括:

(b)(i) 具有如在本文 SEQ ID NO: 3 中所示的胺基酸序列的第一花葉 Pol 序列 (“mos1Pol”), 以及

(b)(ii) 具有如在本文 SEQ ID NO: 4 中所示的胺基酸序列的第二花葉 Pol 序列 (“mos2Pol”); 以及

(c) Env 花葉抗原, 該 Env 花葉抗原包括:

(c)(i) 具有如在本文 SEQ ID NO: 5 中所示的胺基酸序列的第一花葉 Env 序列 (“mos1Env”), 以及

(c)(ii) 具有如在本文 SEQ ID NO: 6 中所示的胺基酸序列的第二花葉 Env 序列 (“mos2Env”)。

【0005】 將編碼該等抗原的序列選殖在載體, 例如, 如重組腺病毒載體 (例如重組腺病毒血清型 26 (rAd26)) 中, 並且該等重組載體先前用作疫苗以產生對該等抗原的免疫應答 (參見例如巴魯什等人, 同上; 以及 WO 2010/059732)。例如, 典型地將

mos1Gag和mos1Pol花葉抗原序列組合到Gag和Pol的融合蛋白（“mos1GagPol”）中，並且將其編碼序列選殖進入第一Ad26載體（“rAd26.mos1GagPol”）中；並且將mos2Gag和mos2Pol抗原序列組合到Gag和Pol的另一個融合蛋白（“mos2GagPol”）中，並且將其編碼序列選殖進入第二Ad26載體（“rAd26.mos2GagPol”）中。典型地將編碼mos1Env和mos2Env的構建體選殖進入分開的Ad26載體（分別為“rAd26.mos1Env”和“rAd26.mos2Env”）中。

【0006】 一組如上所述的這樣的花葉抗原給出了M組HIV-1分離株的良好的全域覆蓋，其中編碼花葉1抗原序列的rAd26載體（例如，rAd26.mos1GagPol和rAd26.mos1Env）有利於分枝B和CRF01 HIV-1亞型，並且編碼花葉2抗原序列的rAd26載體（例如，rAd26.mos2GagPol和rAd26.mos2Env）有利於分枝C菌株。在rAd26載體中表現的花葉HIV-1 Gag、Pol和Env抗原可以用於提高獼猴中抗原特異性T淋巴細胞應答的寬度和深度，而與共有序列或天然序列HIV-1抗原相比不會損害細胞應答和體液應答的大小（巴魯什等人，同上；以及WO 2010/059732）。

【0007】 然而，在對上述該等疫苗組分的進一步開發努力後，發現rAd26.mos2Env在非人類靈長動物中顯示出非最佳的細胞表面表現和免疫應答，但是此外與其他rAd26載體（如rAd26.mos1Env）相比，其在製造過程期間顯示出迄今未報告的、意想不到的和不可預測的非最佳遺傳穩定性。因此，包含rAd26.mos2Env的疫苗可以導致針對分枝C HIV-1亞型的非最佳免

疫應答，因為該mos2Env花葉抗原有利於分枝C HIV-1菌株。因此，對於抗HIV的疫苗中的mos2Env抗原的替代物存在需要，該替代物可用於誘導針對HIV-1分枝C的改善的免疫應答。

【發明簡述】

【0008】 本發明涉及新穎的合成的人類免疫缺陷病毒（HIV）包膜蛋白，當與先前描述的mos2Env抗原相比時該等包膜蛋白具有改進的細胞表面表現和遺傳穩定性。本發明還涉及以下組成物及方法，該等組成物及方法使用此類新穎的合成的HIV包膜蛋白和/或其編碼序列以誘導抗HIV-1，特別是HIV-1分枝C的增加的免疫應答，較佳的是當該等組成物與其他HIV抗原組合使用時。

【0009】 在一個概述方面中，本發明涉及一種核酸，該核酸編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一個或多個選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P（即，在SEQ ID NO: 8的位置529處的Ile取代為Pro），(ii) K480E（即，在SEQ ID NO: 8的位置480處的Lys取代為Glu），以及(iii) EK479-480RRRR（即，在SEQ ID NO: 8的位置479-480處的Glu-Lys被四個連續Arg殘基替換）、I529P、A471C（即，在SEQ ID NO: 8的位置471處的Ala取代為Cys）以及T575C（即，在SEQ ID NO: 8的位置575處的Thr取代為Cys）的組合。在一個實施方式中，該合成的HIV包膜蛋白進一步包含信號序列，例如具有選自下組的胺基酸序列的信號序列，該組由以下各項組成：SEQ ID NO: 9-12。在一個實施方式中，該信號序列具有SEQ ID NO: 9的胺基酸序列。

【0010】 在某些實施方式中，該合成的HIV包膜蛋白進一步包括跨膜結構域，較佳的是具有SEQ ID NO: 13的胺基酸序列的跨膜結構域。

【0011】 在某些實施方式中，該合成的HIV包膜蛋白進一步包括細胞質結構域的片段，較佳的是包括SEQ ID NO: 14的胺基酸序列、或其胺基酸1-4（即，NRVR）的細胞質結構域的片段。在其中該合成的HIV包膜蛋白進一步包括跨膜結構域和細胞質結構域的片段的實施方式中，較佳的是該蛋白還包括被融合到SEQ ID NO: 8的羧基-末端（C-末端）和該跨膜區域的胺基-末端（N-末端）的SEQ ID NO: 37的胺基酸序列。

【0012】 在另一個實施方式中，該合成的HIV包膜蛋白包括三聚結構域，例如，包括SEQ ID NO: 15（GCN4）或SEQ ID NO: 16（foldon結構域）的胺基酸序列的三聚結構域。在一個較佳的實施方式中，該三聚結構域包括SEQ ID NO: 15的胺基酸序列。具有三聚結構域的該等實施方式對於基於本文提供的胞外結構域序列（如包含SEQ ID NO: 8的胺基酸序列的胞外結構域序列）的可溶性（即，非膜結合的）的合成的HIV包膜蛋白係有用的，其中該三聚結構域位於該合成的HIV包膜蛋白的C-末端。

【0013】 在又其他的實施方式中，該合成的HIV包膜蛋白包括具有以下突變：EK479-480RRRR、I529P、A471C和T575C的SEQ ID NO: 8。6個連續精胺酸殘基（SEQ ID NO: 8的天然序列中的位置478和481已經是Arg殘基）的引入產生另外的優化的弗林蛋白酶（furin）切割位點，從而獲得改進的加工的（即切割的）胞外結

構域。I529P、A471C和T575C這三種突變被稱為SOSIP突變，其中最後兩個突變導致在新產生的半胱胺酸殘基之間引入可能的二硫鍵。總的來說，該等突變導致可溶性的、三聚的、合成的HIV包膜蛋白，而不需要三聚結構域。

【0014】 在一個較佳的實施方式中，本發明涉及一種核酸，該核酸編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包括SEQ ID NO: 17、SEQ ID NO: 18的胺基酸序列，或SEQ ID NO: 19的aa 1-686。最較佳的是，由該核酸編碼的該合成的HIV包膜蛋白包括SEQ ID NO: 18的胺基酸序列或由SEQ ID NO: 18的胺基酸序列組成。

【0015】 在另一概述方面中，本發明涉及一種載體，該載體包括編碼根據本發明的實施方式的合成的HIV包膜蛋白的核酸。在一個實施方式中，該載體係病毒載體。在一個較佳的實施方式中，該病毒載體係腺病毒載體。在一個較佳的實施方式中，該腺病毒載體係腺病毒26載體。

【0016】 本發明的另一個概述方面涉及一種組成物，較佳的是疫苗組成物，所述組成物包含免疫原性（immunogenically）有效量的根據本發明的實施方式之載體（vector）、和運載體（carrier），其中編碼該合成的HIV包膜蛋白的核酸可操作地連接至啟動子序列。在一個實施方式中，該組成物包含腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼包括SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白。

【0017】 在另一概述方面中，本發明涉及一種用於在對其有

需要的受試者體內誘導針對人類免疫缺陷病毒（HIV）的免疫應答之疫苗組合。該疫苗組合包括第一組成物，該第一組成物包括免疫原性有效量的載體，較佳的是腺病毒載體，更較佳的是腺病毒26載體，該載體編碼具有SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白；第二組成物，該第二組成物包括免疫原性有效量的第二載體，較佳的是第二腺病毒載體，更較佳的是第二腺病毒26載體，該第二腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽；以及視情況至少一種另外的組成物，該另外的組成物包括免疫有效量的至少一種選自由以下各項組成的組：一種載體，該載體編碼具有選自由SEQ ID NO: 1-4、28和29組成之群組的胺基酸序列的抗原性多肽，和包含免疫原性有效量的分離的HIV抗原性多肽的多肽，該分離的HIV抗原性多肽包括但不限於具有SEQ ID NO: 7的胺基酸序列的殘基30-708的多肽或具有SEQ ID NO: 36的殘基30-724的多肽，其中該第一組成物、第二組成物和視情況另外的組成物係存在於同一組成物或一種或多種不同組成物中。

【0018】 本發明的又另一個概述方面涉及在對其有需要的受試者體內誘導針對人類免疫缺陷病毒（HIV）的免疫應答之方法，該等方法包括向該受試者給予根據本發明的實施方式之組成物或疫苗組合。本發明還涉及誘導針對HIV的免疫應答的多種方法，該等方法包括使用根據本發明的實施方式之組成物或疫苗組合來引發並增強該免疫應答。

【0019】 本發明的又另外的方面涉及一種合成的HIV包膜蛋

白，該合成的HIV包膜蛋白包括SEQ ID NO: 8的、或具有一個或多個選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，(iii) EK479-480RRRR、I529P、A471C和T575C的組合。在一個實施方式中，該等合成的HIV包膜蛋白包括具有EK479-480RRRR、I529P、A471C和T575C的突變的SEQ ID NO: 8。在另一個實施方式中，該等合成的HIV包膜蛋白包括SEQ ID NO: 18的胺基酸序列的殘基30-704或30-711。在又另一個實施方式中，該等合成的HIV包膜蛋白包括SEQ ID NO: 19的胺基酸序列的殘基30-686。

【0020】 本發明的另一方面涉及一種細胞，較佳的是分離的細胞，該細胞包括根據本發明的實施方式之載體。

【附圖的幾個視圖的簡述】

【0021】 當連同附圖一起閱讀時，會更好地理解先前概述以及以下本發明的詳述。應理解的是本發明不限於附圖中示出的準確實施方式。

【0022】 在該等附圖中：

【0023】 圖1A-1C係HIV包膜蛋白結構的示意性表示；圖1A示出了全長HIV包膜蛋白；圖1B示出根據本發明的實施方式的可溶性單鏈HIV包膜蛋白結構，其中該跨膜結構域（TM）被GCN4三聚結構域替換，並且該弗林蛋白酶切割位點係突變的（sC4）；圖1C示出根據本發明的實施方式的膜結合HIV包膜蛋白之結構，該結構包括跨膜結構域和細胞質結構域的片段（C4D7）；

【0024】 圖2示出了可溶性sC1包膜蛋白的表現水平，該sC1包膜蛋白基於具有另外的C-末端三聚結構域的mos2Env花葉抗原序列、和根據本發明的實施方式的可溶性合成的HIV包膜蛋白（sC4）；表現藉由使用針對gp120的多株抗體的定量西方墨點法測量；將編碼sC1或sC4的質粒暫態表現兩次，並且每次轉染藉由光密度測定法定量兩次；相比於顯示出相對高的表現水平的sC4合成的HIV包膜蛋白，該sC1蛋白顯示出非常低的表現水平；

【0025】 圖3A和3B示出在如藉由ELISA測定而確定的可溶性CD4存在（淺灰色）和不存在（深灰色）的情況下，合成的HIV包膜蛋白與單株抗體17b（mAb17b）之結合；圖3A示出sC1的結合；圖3B示出sC4的結合；

【0026】 圖4係來自sC1蛋白和sC4合成的HIV包膜蛋白的天然聚丙烯醯胺凝膠電泳之西方墨點之圖像；

【0027】 圖5示出了使用抗gp120多株抗體（GP120）藉由表現該等蛋白質的細胞的FACS分析，並且藉由結合至廣泛中和性抗體PG9（PG9）和PG16（PG16）的膜結合的C1、C1D7、C4和C4D7合成的HIV包膜蛋白的相對細胞表面表現水平，該廣泛中和性抗體PG9和PG16係四級結構依賴性的並且優先結合正確折疊的Env三聚體；

【0028】 圖6係在多次病毒傳代後腺病毒載體的穩定性之圖示，該等腺病毒載體含有編碼本發明的合成的HIV包膜蛋白之序列，該等序列包括全長C4（FLC4）、C4D7和sC4；在PER.C6細胞中產生重組腺病毒26載體；在用於轉染和斑塊純化的最初3次傳代

後，選擇5個斑塊並以T25形式擴增10代，導致總病毒傳代數（vpn）為13；示出了如藉由E1轉基因盒聚合酶鏈式反應（PCR）測定的vpn 3、5、10和13之後的穩定性；例如，3/5係指在5個測試的斑塊中的3個斑塊係穩定的，並且5/5係指在5個測試的斑塊中的5個斑塊是穩定的；並且

【0029】 圖7A和7B示出了在兔子中的基於TZM-bl細胞的中和測定中針對HIV-1包膜假型病毒顆粒（EVP）之病毒中和效價；在第1週、第8週、第14週和第20週，針對EVP VSV-G（陰性對照）和MW965.26（Tier 1A分枝C）測量高腺病毒載體給藥組的log10-轉化的IC₅₀值；每個點表示單獨的兔的log10-轉化的IC₅₀值，組平均值由水平線表示；HD：測試的最高稀釋度（上部實線）；LD：測試的最低稀釋度（下部實線）；LOB：背景的限制值，編譯的陰性樣品的95百分位值（虛線）；在相應的線處設置超過LD或HD閾值的Log10 IC₅₀值；對每個時間點使用聯合分級的Dunn方法與對照進行單向非參數比較；該等圖中示出了以下統計顯著差異： $* = P < 0.05$ ， $** = P < 0.01$ ，以及 $*** = P < 0.001$ ；圖7A示出了VSV-G（陰性對照）的結果；並且圖7B示出了MW965.26（Tier 1A分枝C）的結果。

【發明詳述】

【0030】 在背景和整個說明書中引用或描述了各種出版物、文章和專利；該等參考文獻中的每一篇均藉由引用以其整體結合在此。包括在本說明書中的文件、法案、材料、裝置、物品或類似物的討論用於提供本發明背景之目的。這種討論不承認任何或

所有該等事項形成關於所揭露或要求保護的任何發明的先前技術的一部分。

【0031】 除非另外指明，本文所用的所有科學技術術語具有如熟習該項技術者通常理解的相同含義。否則，本文使用的某些術語具有如說明書中示出的含義。本文引用的所有專利、公開專利申請和出版物如同完全列出一樣藉由引用結合在此。必須注意，除非上下文另外明確指示，否則如在此和所附申請專利範圍中所使用，單數形式“一/一個/一種（a/an）”和“該”包括複數指示物。

【0032】 如在此所使用，“受試者”係指將向其給予或已經給予根據本發明的實施方式的載體、組成物或組合疫苗的任何動物，較佳的是哺乳動物，最較佳的是人類。如在此所用的術語“哺乳動物”涵蓋任何哺乳動物。哺乳動物的實例包括但不限於牛、馬、羊、豬、貓、狗、小鼠、大鼠、兔、豚鼠、猴、人等，更較佳的是人。

【0033】 本發明總體上涉及合成的HIV包膜蛋白、編碼該等合成的HIV的包膜蛋白的核酸和載體、以及用編碼該等合成的HIV的包膜蛋白的載體或該等合成的HIV的包膜蛋白單獨地或與一種或多種另外的編碼一種或多種另外的HIV抗原性多肽的載體組合地，和/或與一種或多種另外的分離的HIV抗原性多肽組合地誘導針對HIV的免疫應答之方法。

【0034】 人類免疫缺陷病毒(HIV)係慢病毒屬(Lentivirinae)之成員，其係逆轉錄病毒科(Retroviridae)的一部分。兩種HIV

感染人類：HIV-1和HIV-2。HIV-1係最常見的HIV病毒的菌株，並且已知比HIV-2更致病。如在此所使用，術語“人類免疫缺陷病毒”和“HIV”係指但不限於HIV-1和HIV-2。

【0035】 HIV被分類為具有高度遺傳趨異的多個分枝。如在此所使用，術語“HIV分枝”或“HIV亞型”係指根據它們的遺傳相似性程度進行分類的相關人類免疫缺陷病毒。目前存在三組HIV-1分離株：M、N和O。M組（主要菌株）由至少從A至J的十個分枝組成。O組（外部菌株）可以由相似數目的分枝組成。N組係未分類於M或O組中的新HIV-1分離株。

【0036】 如在此所使用，術語“HIV抗原性多肽”、“HIV抗原蛋白”、和“HIV免疫原”係指能夠在受試者中誘導針對HIV的免疫應答，例如體液和/或細胞介導的應答的多肽。該抗原性多肽可以是HIV的蛋白質、其片段或表位、或其多個HIV蛋白或部分的組合，其可以在受試者中誘導針對HIV的免疫應答或產生免疫，例如保護性免疫。

【0037】 較佳的是，抗原性多肽能夠在宿主中產生保護性免疫應答，例如誘導針對病毒性疾病或感染的免疫應答，和/或在受試者中產生針對病毒性疾病或感染的免疫（即接種），其保護受試者免受病毒性疾病或感染。例如，該抗原性多肽可以包含來自猿猴免疫缺陷病毒（SIV）或HIV的蛋白或其片段，如HIV或SIV包膜gp160蛋白，HIV或SIV基質/衣殼蛋白，以及HIV或SIV *gag*、*pol*和*env*基因產物。

【0038】 HIV抗原性多肽可以是任何HIV-1或HIV-2抗原或其

片段。HIV抗原的實例包括但不限於 *gag*、*pol*、和 *env* 基因產物，該等基因產物編碼結構蛋白和必需酶。*Gag*、*pol*、和 *env* 基因產物被合成為多蛋白，該等多蛋白被進一步加工成多個其他蛋白產物。該 *gag* 基因的初級蛋白產物係病毒結構蛋白 *gag* 多蛋白，其被進一步加工成 MA、CA、SP1、NC、SP2 以及 P6 蛋白產物。該 *pol* 基因編碼病毒酶（Pol，聚合酶），並且該初級蛋白產物被進一步加工成 RT、RNA 酶 H、IN、以及 PR 蛋白產物。該 *env* 基因編碼結構蛋白，尤其病毒體包膜的糖蛋白。該 *env* 基因的初級蛋白產物係 gp160，gp160 被進一步加工成 gp120 和 gp41。HIV 抗原的其他實例包括基因調節蛋白 Tat 和 Rev；輔助蛋白 Nef、Vpr、Vif 和 Vpu；衣殼蛋白、核衣殼蛋白、以及 p24 病毒蛋白。

【0039】 在某些實施方式中，該 HIV 抗原性多肽包括 HIV Gag、Env、或 Pol 抗原、或任何抗原部分或其表位或組合，較佳的是 HIV-1 Gag、Env、或 Pol 抗原或任何抗原部分或其表位或組合。

【0040】 HIV 抗原性多肽還可以是花葉 HIV 抗原。如在此所使用，“花葉抗原”係指自天然序列的片段組裝的重組蛋白。花葉抗原類似於天然抗原，但被優化以最大化在天然序列中發現的潛在 T 細胞表位的覆蓋，這提高了免疫應答的寬度和覆蓋。用於本發明之花葉 HIV 抗原較佳的是花葉 Gag、Pol、和/或 Env 抗原，並且更較佳的是花葉 HIV-1 Gag、Pol 和/或 Env 抗原。如在此所使用，“花葉 HIV Gag、Pol、和/或 Env 抗原”尤其是指包含來源於 HIV 的一個或多個 Gag、Pol 和/或 Env 多蛋白序列的多個表位的花葉抗原。

【0041】 在一個實施方式中，用於本發明之花葉 HIV 抗原係

具有來源於*gag*基因產物的序列的表位之花葉HIV Gag抗原 (SEQ ID NO: 1、2中提供了實例)；係具有來源於*pol*基因產物的序列的表位之花葉HIV Pol抗原 (SEQ ID NO: 3、4中提供了實例)；或係具有來源於*env*基因產物的序列的表位之花葉HIV Env抗原 (SEQ ID NO: 5、6中提供了實例；還有例如在SEQ ID N: 8、17、18、19中的本發明的新穎抗原可以被認為是花葉HIV Env抗原)。在某些實施方式中，用於本發明的花葉HIV抗原可以包括來源於*gag*、*pol*、和/或*env*基因產物的表位的組合。說明性的和非限制的實例包括具有來源於*env*和*pol*基因產物的序列的表位的花葉Env-Pol抗原；具有來源於*gag*和*pol*基因產物的序列的表位的花葉Gag-Pol抗原；以及具有來源於*gag*和*env*基因產物的序列的表位的花葉Gag-Env抗原。*gag*、*pol*、和*env*基因產物的序列可以來源於一種或多種分枝。

【0042】 可以用於本發明中的花葉HIV Gag、Pol和/或Env抗原的實例包括描述於例如US 20120076812；巴魯什等人，自然醫學 (*Nat Med*) 2010，16: 319-323；以及巴魯什等人，細胞 (*Cell*) 155: 1-9，2013中的那些，所有該等文獻都以其整體藉由引用結合在此。較佳的是，用於本發明的花葉HIV Gag、Pol、和/或Env抗原包括但不限於mos1Gag (SEQ ID NO: 1)、mos2Gag (SEQ ID NO: 2)、mos1Pol (SEQ ID NO: 3)、mos2Pol (SEQ ID NO: 4)、mos1Env (SEQ ID NO: 5)、mos2Env (SEQ ID NO: 6)、mos1GagPol (SEQ ID NO: 28)、mos2GagPol (SEQ ID NO: 29)、以及其組合。

【0043】 如本文所使用，術語“HIV包膜蛋白”、“env蛋

白”、和“Env”各自係指在HIV病毒體的包膜上表現，並且使得HIV靶向並附著於HIV感染的細胞的質膜上的蛋白質或其片段或衍生物，該等蛋白質或片段或衍生物可以在對其有需要的受試者中誘導針對HIV的免疫應答或產生針對HIV的免疫。該HIV *env*基因編碼先質蛋白gp160，該先質蛋白gp160被蛋白質裂解地切割成兩種成熟的包膜糖蛋白，即gp120和gp41。該切割反應由宿主細胞蛋白酶（弗林蛋白酶）在逆轉錄病毒包膜糖蛋白先質中的高度保守的序列處介導。更具體地，gp160三聚化為(gp160)₃，並且然後經歷切割成為兩個非共價締合的gp120和gp41。病毒進入隨後由gp120/gp41異源二聚體的三聚體介導。Gp120係受體結合片段，並且結合到靶細胞上的CD4受體上，該靶細胞具有這樣的受體如，例如T-輔助細胞。非共價結合至gp120的Gp41係融合片段並且提供HIV進入細胞的第二步驟。Gp41最初埋在該病毒包膜內，但當gp120與CD4受體結合時，gp120改變其構象，從而導致gp41暴露，其中它可以幫助與宿主細胞融合。Gp140係三聚gp160，即(gp160)₃的未切除的胞外結構域，其被用作天然狀態的切割的病毒刺突的替代物。

【0044】 根據本發明的實施方式，“HIV包膜蛋白”可以是gp160、gp140、gp120、gp41蛋白，其組合，融合物，截短物或衍生物。例如，“HIV包膜蛋白”可以包括與gp41蛋白非共價締合的gp120蛋白。它還可以包括穩定的三聚gp140蛋白，該三聚gp140蛋白可以具有或可以被修飾以包括三聚結構域，該三聚結構域穩定gp140的三聚體。三聚結構域的實例包括但不限於T4-fibritin “foldon”三聚結構域；來源於GCN4的捲曲螺旋三聚結構域；以

及作為三聚體標籤的大腸桿菌天冬胺酸轉胺甲醯酶的催化亞基。

“HIV包膜蛋白”還可以是截短的HIV包膜蛋白，該截短的HIV包膜蛋白包括但不局限於包含在胞外結構域（即，延伸至細胞外空間的結構域）中的C-末端截短、在gp41中的截短物（如在gp41的跨膜結構域中的截短物，或在gp41的細胞質結構域中的截短物）的包膜蛋白。“HIV包膜蛋白”可以進一步係天然存在的HIV包膜蛋白的衍生物，該HIV包膜蛋白的衍生物具有例如在弗林蛋白酶切割位點中的序列突變、和/或所謂的SOSIP突變。

【0045】 較佳的是，“HIV包膜蛋白”係“合成的HIV包膜蛋白。”如在此所使用，術語“合成的HIV包膜蛋白”係指非天然存在的HIV包膜蛋白，其在對其有需要的受試者中針對一種或多種天然存在的HIV菌株誘導免疫應答或產生免疫。花葉HIV Env蛋白係合成的HIV Env蛋白的實例，並且本發明提供了新穎的合成的HIV Env抗原，例如包含SEQ ID NO: 8、17、18、或19的那些。

合成的HIV包膜蛋白及其編碼序列

【0046】 本發明的實施方式涉及新穎的合成的HIV包膜蛋白和編碼該等的核酸分子。

【0047】 在一個實施方式中，本發明涉及 合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及(iii) EK479-480RRRR、I529P、A471C和T575C的組合。SEQ ID NO: 8包括合成的成熟gp120和合成的截短gp41，該合成的截短gp41不具有跨膜區域，也不具有細

胞質結構域。SEQ ID NO: 8係非天然存在的序列，其包含來自該mos2Env花葉抗原（SEQ ID NO: 6）的序列和其他HIV包膜蛋白序列的嵌合體。包含 SEQ ID NO: 8的新穎的合成的Env抗原的序列被優化以提供針對HIV分枝C的廣泛覆蓋和增強的T-細胞應答（當與該mos2Env抗原（SEQ ID NO: 6）相比時）。在某些實施方式中，另外的胺基酸可以添加至SEQ ID NO: 8中或本文所定義的它的變體之一中。

【0048】 在某些實施方式中，該合成的HIV包膜蛋白進一步包括信號序列。該合成的HIV包膜蛋白使用信號序列合成，所述信號序列在其轉運到內質網（ER）的腔內期間從新生多肽鏈切割。原則上，可以使用任何已知的信號序列。較佳的是使用HIV Env信號序列或其變體。本領域中已經使用針對HIV Env蛋白的不同信號序列（參見例如，WO 2014/107744）。在某些實施方式中，該信號序列包括SEQ ID NO: 9、SEQ ID NO: 10、SEQ ID NO: 11或SEQ ID NO: 12。在一個較佳的實施方式中，該信號序列包括SEQ ID NO: 9。

【0049】 在某些實施方式中，該合成的HIV包膜蛋白進一步包括跨膜結構域。該跨膜結構域將合成的HIV包膜蛋白錨定到該ER膜上，並且有助於膜組裝和HIV包膜的功能。較佳的是，該跨膜結構域包括SEQ ID NO: 13。

【0050】 在另一個實施方式中，該合成的HIV包膜蛋白包括具有截短的細胞質結構域的gp41。該gp41在其羧基末端具有異常長的細胞質結構域，通常約150個胺基酸（愛德華茲（Edwards）

等人，病毒學雜誌 (*J. Virology*)，2002，76: 2683-2691)。據報導該細胞質結構域的截短誘導HIV-1 Env蛋白的胞外結構域中的保守區域的暴露 (*Id.*)。本發明的合成的HIV包膜中的截短的細胞質結構域 可以在從一個至約140個胺基酸的範圍內，如全長細胞質結構域的1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、110、120、130、或140個胺基酸。在某些實施方式中，藉由給定的胺基酸直到C-末端之後的截短，該截短的細胞質結構域來源於SEQ ID NO: 17的胺基酸704-862 (即，來源於本發明的C4分子的細胞質結構域)。在一個較佳的實施方式中，該合成的HIV包膜蛋白包括截短的細胞質結構域，該截短的細胞質結構域 具有HIV gp41細胞質結構域的1至10個胺基酸殘基，更較佳的是4至8個胺基酸殘基，並且最較佳的是7個胺基酸殘基。合成的HIV包膜蛋白的細胞質結構域或其片段位於該細胞外結構域(細胞外結構域 (ectodomain)) 的C-末端，並且當該合成的HIV包膜蛋白還包括跨膜結構域時，該細胞質結構域或其片段位於該跨膜結構域的C-末端。參見，例如，圖1A和1C。在一個特定實施方式中，該合成的HIV包膜蛋白包括 具有截短的細胞質結構域的gp41，該截短的細胞質結構域具有SEQ ID NO: 14的胺基酸序列或其片段，如其殘基1-4 (即NRVR)。已經描述了並且可使用其他截短的細胞質結構域 (例如Schiernle等人，美國科學院院刊 (*PNAS*)，1997；亞伯拉罕 (Abrahamyan) 等人，病毒學雜誌 (*J Virol*) 2005)。

【0051】 在其中該合成的HIV包膜蛋白進一步包括跨膜結構域和細胞質結構域的片段的實施方式中，較佳的是該蛋白還包括SEQ ID NO: 37的胺基酸序列，該SEQ ID NO: 37包含SEQ ID NO:

18的殘基655-682，其中SEQ ID NO: 37的胺基酸序列被融合到SEQ ID NO: 8的C-末端和該跨膜結構域的N-末端。

【0052】 在本發明的一個特別較佳的實施方式中，該合成的HIV包膜蛋白進一步包括跨膜結構域，例如具有SEQ ID NO: 13的胺基酸序列的那些和截短的細胞質結構域或細胞質結構域的片段例如具有SEQ ID NO: 14的胺基酸序列或SEQ ID NO: 14的殘基1-4（即，NRVR）的那些。最較佳的是，該合成的HIV包膜蛋白包括SEQ ID NO: 18的胺基酸序列，具有或沒有信號序列（即，SEQ ID NO: 18的胺基酸殘基1-29）或由其組成。

【0053】 在另一個實施方式中，該合成的HIV包膜蛋白包括替代Env跨膜區域的三聚結構域。該三聚結構域增加了Env三聚結構的穩定性。較佳的是，該合成的HIV包膜蛋白包括gp140多肽，該gp140多肽被修飾以包括穩定gp140的三聚體的三聚結構域。三聚結構域的實例包括但不限於T4-fibritin “foldon” 三聚結構域，如包含SEQ ID: 16的胺基酸序列的那些；來源於GCN4的捲曲螺旋三聚結構域，如包含SEQ ID: 15的胺基酸序列的那些；作為三聚體標籤的大腸桿菌天冬胺酸轉胺甲醯酶的催化亞基；或基於軟骨基質蛋白的三聚基序。如果存在時，該三聚結構域通常位於該細胞外結構域的C-末端（參見圖1B）。在某些較佳的實施方式中，其中該合成的HIV包膜蛋白包括三聚結構域，該合成的HIV包膜蛋白包括SEQ ID NO: 19的胺基酸序列，具有或沒有信號序列（即，SEQ ID NO: 19的胺基酸殘基1-29）。具有三聚結構域的該等實施方式主要用於該合成的HIV包膜蛋白的可溶性胞外結構域變體。在本發明的

此類可溶性變體的某些實施方式中，有可能對該弗林蛋白酶切割位點進行突變（例如在SEQ ID NO: 8的位置480處的Lys突變為Glu）以使這個切割位點失活，從而使得該蛋白質將是一條單鏈；它與三聚結構域，尤其是與SEQ ID NO: 19的GCN4三聚結構域良好地結合。

【0054】 不使用三聚結構域的該合成的HIV包膜蛋白的此類可溶性細胞外結構域的可替代版本也是本發明的實施方式，並且可以藉由將優化該弗林蛋白酶切割位點的突變（在位置479-480處的Gly-Lys二肽被四個Arg殘基替換）以及所謂的SOSIP突變（在位置529處的I > P突變，並且在位置471和575之間藉由置換在SEQ ID NO: 8中各自具有Cys殘基的那些位置上的相應Ala和Thr而引入二硫鍵）進行合併來從SEQ ID NO: 8製備。這產生一種蛋白質，該蛋白質具有以下突變的組合的SEQ ID NO: 8的胺基酸序列：EK479-480RRRR、I529P、A471C和T575C。

【0055】 進一步增加本發明的合成的HIV包膜蛋白（包含SEQ ID NO: 8）的三聚體含量的一可能的修飾係位置529處的Ile到Pro的修飾。這對於可溶性變體和膜結合變體兩者都可以是有效的。

【0056】 載體

【0057】 本發明的另一個概述方面涉及包含編碼合成的HIV包膜蛋白的核酸的載體。根據本發明的實施方式，該等載體可以包含本文所述的任何合成的HIV包膜蛋白。在本發明的一較佳的實施方式中，該載體包含編碼以下合成的HIV包膜蛋白的核酸，該HIV包膜蛋白包括SEQ ID NO: 8、SEQ ID NO: 17、SEQ ID NO: 18、

或SEQ ID NO: 19的胺基酸序列，並且更較佳的是SEQ ID NO:18的胺基酸序列。

【0058】 根據本發明的實施方式，編碼該合成的HIV包膜蛋白的核酸可操作地連接至啟動子，這意味著該核酸在啟動子的控制下。該啟動子可以是同源啟動子（即，來源於相同的作為載體的遺傳來源）或異源啟動子（即，來源於不同載體或遺傳來源）。適合的啟動子的實例包括巨細胞病毒（CMV）啟動子和勞氏肉瘤病毒（RSV）啟動子。較佳的是，該啟動子位於表現盒內的核酸的上游。可以可操作地連接至編碼該合成的HIV包膜蛋白的核酸的示例性CMV啟動子序列示於SEQ ID NO: 24中。

【0059】 根據本發明的實施方式，載體可以是表現載體。表現載體包括但不限於用於重組蛋白表現的載體和用於將核酸遞送至受試者中用於在該受試者的組織中表現的載體，例如病毒載體。適合用於本發明的病毒載體的實例包括但不限於腺病毒載體、腺相關病毒載體、痘病毒載體、MVA載體、腸系病毒載體、委內瑞拉馬腦炎病毒載體、塞姆利基森林病毒（Semliki Forest Virus vector）載體、菸草花葉病毒載體、慢病毒載體等。該載體還可以是非病毒載體。非病毒載體的實例包括但不限於質粒、細菌人工染色體、酵母人工染色體、噬菌體等。

【0060】 在本發明的某些實施方式中，該載體係腺病毒載體。根據本發明的腺病毒屬於腺病毒科，並且較佳的是屬於哺乳動物腺病毒（Mastadenovirus）屬的一種。它可以是人類腺病毒，但還可以是感染其他物種的腺病毒，包括但不限於牛腺病毒（例

如牛腺病毒3、BAdV3)、犬腺病毒(例如CAdV2)、豬腺病毒(例如PAdV3或PAdV5)、或猿猴腺病毒(其包括猴腺病毒和猿腺病毒,如黑猩猩腺病毒或大猩猩腺病毒)。較佳的是,該腺病毒係人類腺病毒(HAdV、或AdHu)、或猿猴腺病毒如黑猩猩或大猩猩腺病毒(ChAd、AdCh、或SAdV)。在本發明中,人類腺病毒係指如果稱為Ad而不指明物種,例如該簡短符號“Ad26”表示與HadV26相同,該HadV26係人腺病毒血清型26。還如在此所使用的,符號“rAd”係指重組腺病毒,例如,“rAd26”係指重組人類腺病毒26。

【0061】 已使用人腺病毒進行大多數高級研究,並且根據本發明的某些方面,人腺病毒係較佳的。在某些較佳的實施方式中,根據本發明的重組腺病毒係基於人腺病毒。在較佳的實施方式中,該重組腺病毒基於人腺病毒血清型5、11、26、34、35、48、49、50、52等。根據本發明的一特別較佳的實施方式,腺病毒係人腺病毒血清型26。這種血清型的優點係在人群體中的低血清陽性率和/或低預先存在的中和抗體效價,以及在臨床試驗中用於人類受試者的經驗。

【0062】 猿猴腺病毒總體上也具有在人群體中的低血清陽性率和/或低預先存在的中和抗體效價,並且已經報導了大量使用黑猩猩腺病毒載體的工作(例如,US 6083716;WO 2005/071093;WO 2010/086189;WO 2010085984;法裡納(Farina)等人,2001,病毒學雜誌(*J Virol*) 75: 11603-13 [13];科恩(Cohen)等人,2002,普通病毒學雜誌(*J Gen Virol*) 83: 151-55 [69];科賓格(Kobinger)

等人，2006，病毒學（*Virology*）346: 394-401 [70]；Tatsis等人，2007，分子治療（*Molecular Therapy*）15: 608-17 [71]；還參見由Bangari和Mittal，2006，疫苗（*Vaccine*）24: 849-62 [72]的綜述；以及由拉莎諾（Lasaro）和埃特爾（Ertl），2009，分子治療（*Mol Ther*）17: 1333-39 [73]的綜述）。因此，在其他的實施方式中，根據本發明的重組腺病毒係基於猿猴腺病毒，例如黑猩猩腺病毒。在某些實施方式中，該重組腺病毒係基於猿猴腺病毒類型1、7、8、21、22、23、24、25、26、27.1、28.1、29、30、31.1、32、33、34、35.1、36、37.2、39、40.1、41.1、42.1、43、44、45、46、48、49、50或SA7P。

【0063】 較佳的是，該腺病毒載體係複製缺陷型重組病毒載體，如rAd26、rAd35、rAd48、rAd5HVR48等。

【0064】 在本發明的一個較佳的實施方式中，該等腺病毒載體包括來自罕見血清型的衣殼蛋白，包括Ad26。在典型實施方式中，該載體係rAd26病毒。“腺病毒衣殼蛋白”係指在腺病毒（例如，Ad26、Ad35、rAd48、rAd5HVR48載體）的衣殼上的蛋白質，該腺病毒衣殼蛋白參與確定特定的腺病毒的血清型和/或趨向性。腺病毒衣殼蛋白典型地包括纖維、五鄰體蛋白和/或六鄰體蛋白。如在此所使用的用於特定腺病毒的“衣殼蛋白”（如“Ad26衣殼蛋白”）可以是例如包括至少一部分Ad26衣殼蛋白的嵌合的衣殼蛋白。在某些實施方式中，該衣殼蛋白係Ad26的整個衣殼蛋白。在某些實施方式中，該六鄰體、五鄰體和纖維都是Ad26的。

【0065】 熟習該項技術者將認識到來源於多個血清型的元件

可以被組合在單一的重組腺病毒載體中。因此，可以產生組合了來自不同血清型的所希望性質的嵌合腺病毒。因此，在一些實施方式中，本發明的嵌合腺病毒可以將第一血清型的預先存在的免疫性的缺失與以下特徵相結合：如溫度穩定性、組裝、錨定、產量、重定向或改善的感染、靶細胞中DNA的穩定性等。

【0066】 在某些實施方式中，本發明中有用的重組腺病毒載體主要或完全來源於Ad26（*即*，該載體係rAd26）。在某些實施方式中，該腺病毒係複製缺陷的，例如，因為它包含基因組的E1區域中的缺失。對於來源於非類群C的腺病毒（如Ad26或Ad35）的腺病毒，典型地是將腺病毒的E4-orf6編碼序列與人類亞群C（如Ad5）的腺病毒的E4-orf6交換。這允許在表現Ad5的E1基因的熟知的補充細胞系中此類腺病毒的繁殖，如例如293細胞或PER.C6細胞等（參見，例如哈文格（Havenga）等人，2006，普通病毒學雜誌（*J Gen Virol*）87: 2135-43；WO 03/104467）。然而，這樣的腺病毒將不能在不表現Ad5的E1基因的非補充細胞中複製。

【0067】 重組腺病毒載體的製備在本領域中是熟知的。例如在WO 2007/104792中和在阿賓克（Abbink）等人，(2007) 病毒學（*Virol*）81(9): 4654-63中描述了rAd26載體的製備。Ad26的示例性基因組序列發現於GenBank登錄號EF 153474中和WO 2007/104792的SEQ ID NO: 1中。用於本發明的載體的實例例如包括描述於WO 2012/082918中的那些，將其揭露內容以其整體藉由引用結合在此。

【0068】 典型地，使用一核酸來產生本發明中有用的載體，

該核酸包括整個重組腺病毒基因組（例如，質粒、粘粒、或桿狀病毒載體）。因此，本發明還提供了分離的核酸分子，該等分離的核酸分子編碼本發明的該等腺病毒載體。本發明的該等核酸分子可以處於RNA形式或處於藉由選殖獲得或合成產生的DNA形式。該DNA可以是雙股或單股的。

【0069】 本發明中有用的該等腺病毒載體典型地是複製缺陷的。在該等實施方式中，該病毒藉由使對病毒複製關鍵的區域（例如E1區域）缺失或失活而變得複製缺陷。可以藉由在所述區域內例如插入感興趣的基因，如編碼合成的HIV包膜蛋白的基因（通常連接至啟動子），或編碼HIV抗原性多肽的基因（通常連接至啟動子）來使該等區域基本上缺失或失活。在一些實施方式中，本發明的該等載體可以包含其他區域如E2、E3或E4區域中的缺失，或該等區域中的一個或多個內的連接至啟動子的異源基因的插入。對於E2-和/或E4-突變的腺病毒，一般使用E2-和/或E4補充細胞系來產生重組腺病毒。該腺病毒的E3區域中的突變不需要被細胞系補充，因為對於複製E3係需要的。

【0070】 通常使用包裝細胞系來生產足夠量的用於在本發明中使用的腺病毒載體。包裝細胞係包含在複製缺陷型載體中缺失或失活的那些基因的細胞，因此允許了該病毒在該細胞中複製。適合用於在該E1區域具有缺失的腺病毒的包裝細胞系包括，例如，PER.C6、911、293、以及E1 A549。

【0071】 根據本發明的實施方式，並且如以上所指出的，本文所述的任何合成的HIV包膜蛋白可以在本發明的該等載體中表

現。鑒於遺傳密碼的簡並性，技術人員將充分意識到可以根據本領域中完全常規的方法來設計編碼相同蛋白質的若干種核酸序列。編碼該合成的HIV包膜蛋白的核酸可以視情況是密碼子優化的，以確保在所治療的宿主（例如，人類）中的適當的表現。密碼子優化係廣泛應用於本領域中的技術。SEQ ID NO: 25、26和27中提供了一些編碼本發明的合成的HIV包膜蛋白的序列的非限制性實例。典型地，編碼該合成的HIV包膜蛋白的核酸被選殖入該腺病毒基因組的E1和/或E3區域。

【0100】 在本發明的一個較佳的實施方式中，該載體係腺病毒載體，並且更較佳的是rAd26載體，最較佳的是在該腺病毒基因組的E1區域內具有至少一個缺失的rAd26載體，例如如在阿賓克（Abbink），病毒學雜誌（*J Virol*），2007. 81(9)：第4654-63頁中描述的，將其藉由引用結合在此。

【0072】 本發明還提供了細胞，較佳的是包含本文所述的任何載體的分離的細胞。該等細胞可以用於生產重組蛋白質，或用於生產病毒顆粒。

【0073】 因此，本發明的實施方式還涉及製造合成的HIV抗原性多肽的方法。該方法包括用表現載體轉染宿主細胞，所述表現載體包含可操作地連接至啟動子的編碼該合成的HIV抗原性多肽的核酸；在適合於表現該合成的HIV抗原性多肽的條件下培養該轉染的細胞，並從該細胞分離該合成的HIV抗原性多肽。可以藉由本領域中已知的任何方法（包括親和層析等）從該細胞中分離或收集該合成的HIV抗原性多肽。鑒於本揭露，用於重組蛋白表現的

技術將是熟習該項技術者所熟知的。

【0074】 本發明還包括用於生產編碼本發明的合成的HIV抗原性多肽的載體的方法，該方法包括培養包含該載體的細胞以在所述培養期間繁殖和增殖該載體，並從該細胞培養物（例如，從細胞、從培養基或兩者）中分離編碼本發明的合成的HIV抗原性多肽的載體。可以根據本領域已知的方法進一步純化該載體。

【0075】 在某些實施方式中，本發明提供了根據本發明的實施方式的載體，該載體包含編碼合成的HIV抗原性多肽的核酸，並且在某些示例性實施方式中，該核酸具有選自下組的核苷酸序列，該組以下各項組成：SEQ ID NO: 25、26和27。

【0076】 組成物

【0077】 在另一個概述方面中，本發明涉及包含一種載體和一種運載體的組成物，該載體包含編碼合成的HIV包膜蛋白的核酸。根據本發明的實施方式，本文所述的任何載體可以包括在該組成物中。較佳的是，該載體係病毒載體，更較佳的是腺病毒載體，並且甚至更較佳的是腺病毒26載體。在一較佳的實施方式中，組成物包含腺病毒載體，較佳的是編碼合成的HIV包膜蛋白的腺病毒26載體，該合成的HIV包膜蛋白包含SEQ ID NO: 8、SEQ ID NO: 18、或SEQ ID NO: 19的胺基酸序列，並且更較佳的是SEQ ID NO: 18的胺基酸序列。

【0078】 在一個方面，本發明提供了包含一種或多種載體的組合疫苗，該等載體一起包含編碼以下各項的核酸序列：(i) 合成

的HIV包膜蛋白，其包含SEQ ID NO: 8（例如SEQ ID NO: 18或19）的胺基酸序列，和（ii）第二HIV包膜蛋白，其包含SEQ ID NO: 5的胺基酸序列。該等載體可以各自在單獨的組成物中，或可被結合在單一的組成物中。旨在向一個受試者給予在一個或多個載體中的這兩種核酸，這將導致對HIV的免疫應答，其比給予任一單獨的載體時將獲得的免疫應答更廣泛。這兩種核酸序列還可以存在於一個單一的載體上。

【0079】 根據本發明的實施方式，組成物包含免疫原性有效量的載體，如病毒載體。如在此所使用的，“免疫原性有效量”或“免疫有效量”係指足以在對其有需要的受試者中誘導所需免疫效應或免疫應答的組成物的量。在一個實施方式中，免疫原性有效量係指足以在對其有需要的受試者中誘導免疫應答的量。在另一個實施方式中，免疫原性有效量係指足以在對其有需要的受試者中產生免疫，例如，提供對抗疾病例如病毒感染的保護效應的量。免疫原性有效量可以根據多種因素而變化，諸如受試者的身體狀況、年齡、體重、健康等；具體應用，是否誘導免疫應答或提供保護性免疫；給予的特定重組載體；由所給予的重組載體編碼的免疫原或抗原性多肽；給予的特異性抗原性多肽；以及需要免疫的具體疾病，例如病毒感染。考慮到本揭露，藉由熟習該項技術者可以容易地確定免疫原性有效量。

【0080】 作為一般指導，當參考重組病毒載體如腺病毒載體使用時，免疫原性有效量可以在從約 10^8 個病毒顆粒至約 10^{12} 個病毒顆粒，例如 10^8 、 10^9 、 10^{10} 、 10^{11} 、或 10^{12} 個病毒顆粒的範圍內。

免疫原性有效量可以以單一組成物或多個組成物，例如1、2、3、4、5、6、7、8、9或10個組成物（例如，片劑，膠囊或注射劑）給予，其中所述多個膠囊或注射劑的給予共同以免疫有效量提供給受試者。一般而言，當參考多肽（如分離的抗原性多肽）使用時，免疫原性有效量可以在從例如約0.3至約3000微克（ μg ）的範圍內，例如1-1000 μg ，例如10-500 μg ，例如約10、50、100、150、200、250、300、350、400、450或500 μg 。作為非限制性實例，有可能將編碼本發明的合成的HIV Env抗原（具有SEQ ID NO: 8）的載體的給予與Env多肽，例如250 μg 的具有SEQ ID NO: 7的胺基酸30-708的HIV分枝C Env三聚體蛋白的給予進行合併。在所謂的初免-加強方案中，還有可能向受試者給予免疫原性有效量，並且隨後向同一受試者給予另一劑量的免疫原性有效量。初免-加強方案的一般概念係疫苗領域的技術人員熟知的。如果需要，可以視情況向該方案中加入另外的加強給予。

【0081】 本發明的組成物進一步包括運載體。運載體可以包括一種或多種藥學上可接受的賦形劑，如粘合劑、崩散劑、蓬鬆劑、助懸劑、乳化劑、潤濕劑、潤滑劑、調味劑、甜味劑、防腐劑、染料、增溶劑以及塗層。該運載體或其他物質的確切性質可以取決於給藥途徑，例如肌內、皮下、口服、靜脈內、皮膚、粘膜內（例如，腸）、鼻內或腹膜內途徑。對於液體可注射製劑（例如，懸浮液和溶液），適合的運載體和添加劑包括水、乙二醇、油、醇、防腐劑、著色劑等。對於固體口服製劑（例如散劑、膠囊、囊片、軟膠囊和片劑），適合的運載體和添加劑包括澱粉、糖、稀釋劑、粒化劑、潤滑劑、粘合劑、崩散劑等。對於鼻噴霧劑/吸入

劑混合物，該水性溶液/懸浮液可以包括作為適合的運載體和添加劑的水、乙二醇、油、軟化劑、穩定劑、潤濕劑、防腐劑、芳香劑、調味劑等。

【0082】 本發明的組成物可以配製成適合於給予受試者以便於給予並改善功效的任何物質，包括但不限於口服（腸內）給予和腸胃外注射。腸胃外注射包括靜脈內注射或輸注、動脈內注射、皮下注射、肌內注射以及關節內注射。還可以配製本發明的組成物用於其他給藥途徑，包括經粘膜、眼、直腸、長效植入、舌下給藥、舌下、從口腔粘膜繞過門脈循環、吸入或鼻內。

【0083】 根據本發明的某些實施方式中，組成物包含免疫原性有效量的純化的或部分純化的腺病毒載體（如腺病毒26載體），該腺病毒載體包含編碼本發明的合成的HIV包膜蛋白的核酸。所述組成物可以根據本領域熟知的方法被配製成疫苗（也稱為“免疫原性組成物”）。

【0084】 本發明的組成物可以進一步視情況包括佐劑以增強免疫應答。術語“佐劑”和“免疫刺激劑”在此可互換地使用，並且被定義為引起免疫系統刺激的一種或多種物質。在本文中，佐劑用於增強對該等載體的免疫應答，該等載體編碼本發明的合成的HIV包膜蛋白和/或與編碼本發明的合成的HIV包膜蛋白的載體組合使用的HIV抗原多肽。

【0085】 適合用於本發明的佐劑應當係在人中潛在安全的、耐受性良好的並且有效的佐劑，如例如QS-21、Detox-PC、MPL-SE、MoGM-CSF、TiterMax-G、CRL-1005、GERBU、TER醯胺、PSC97B、

Adjumer、PG-026、GSK-I、GcMAF、B-aletine、MPC-026、Adjuvax、CpG ODN、Betafectin、鋁鹽（例如AdjuPhos）、Adjuplex、以及MF59。鑒於本揭露，可以藉由熟習該項技術者熟知的技術確定製劑中每種組分的最佳比例。

【0086】 在一較佳的實施方式中，該佐劑係鋁鹽，如AdjuPhos。

【0087】 免疫原性組成物的製備和使用係熟習該項技術者所熟知的。液體藥物組成物一般包括液體載體例如水、石油、動物或植物油、礦物油或合成油。還可以包括生理鹽水溶液、右旋糖或其他糖溶液或二醇，例如乙二醇、丙二醇或聚乙二醇。

【0088】 例如，重組腺病毒載體可以儲存於緩衝液中，該緩衝液還可以用於腺病毒世界標準（Adenovirus World Standard）（霍根森（Hoganson）等人，2002，生物工藝雜誌（Bioprocessing J）1: 43-8）：20 mM Tris pH 8，25 mM NaCl，2.5%甘油。適用於向人類給予的另一種有用的腺病毒配製物緩衝液係20 mM Tris、2 mM MgCl₂、25 mM NaCl、蔗糖10% w/v、聚山梨酯-80 0.02% w/v。適用於重組腺病毒的另一種配製物緩衝液包括10-25 mM檸檬酸鹽緩衝劑pH 5.9-6.2，4%-6%（w/w）羥丙基-β-環糊精（HBCD）、70-100 mM NaCl、0.018%-0.035%（w/w）聚山梨酯-80、以及視情況包括0.3%-0.45%（w/w）乙醇。顯而易見地，可以使用許多其他緩衝液，並且已知針對存儲和針對純化載體的藥用給藥的適合的配製物的若干實例。

【0089】 根據本發明的實施方式，本發明的組成物可以與一

種或多種編碼一種或多種另外的HIV抗原性多肽的另外的載體，和/或一種或多種分離的HIV抗原性多肽一起使用。該等另外的載體和/或HIV抗原性多肽可以存在於包含本發明的合成的HIV Env蛋白的同一組成物中。它們還可以存在於一種或多種不同的組成物中，所述不同的組成物可以與包含本發明的合成的HIV包膜蛋白的組成物在疫苗組合中一起使用。較佳的是，該一種或多種另外的載體係病毒載體，如腺病毒載體，並且最較佳的是腺病毒26載體。鑒於本揭露，該一種或多種另外的載體可以編碼任何熟習該項技術者已知的HIV抗原性多肽。

【0090】 在一個實施方式中，組成物或疫苗組合進一步包含第二腺病毒載體，較佳的是腺病毒26載體，該第二腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽。此類實施方式的優點係增加免疫應答的寬度（覆蓋來自分枝B和C的菌株）。

【0091】 在另一個實施方式中，本發明的組成物或疫苗組合進一步包含編碼HIV抗原性多肽的腺病毒載體，較佳的是腺病毒26載體，該HIV抗原性多肽包含SEQ ID NO: 28（mos1 GagPol）的胺基酸序列。

【0092】 在另一個實施方式中，本發明的組成物或疫苗組合進一步包含編碼HIV抗原性多肽的腺病毒載體，較佳的是腺病毒26載體，該HIV抗原性多肽包含SEQ ID NO: 29（mos2 GagPol）的胺基酸序列。

【0093】 在一個具體實施方式中，本發明的組成物或疫苗組合進一步包含第二腺病毒載體，較佳的是腺病毒26載體，該第二

腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽；以及一種或多種另外的腺病毒載體，較佳的是腺病毒26載體，該一種或多種另外的載體編碼一種或多種包含選自由SEQ ID NO: 28或SEQ ID NO: 29組成之群組的胺基酸序列的HIV抗原性多肽。例如，根據本發明的實施方式的組成物或疫苗組合可以包含四種腺病毒載體，較佳的是腺病毒26載體，其中第一載體編碼包含SEQ ID NO: 8（例如SEQ ID NO: 18）的胺基酸序列的合成的HIV包膜蛋白；第二載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽；第三載體編碼包含SEQ ID NO: 28的胺基酸序列的合成的HIV包膜蛋白；並且第四載體編碼包含SEQ ID NO: 29的胺基酸序列的合成的HIV包膜蛋白。

【0094】 在一些實施方式中，該組成物或疫苗組合進一步包含一種或多種分離的HIV抗原性多肽。鑒於本揭露，熟習該項技術者已知的任何HIV抗原性多肽可以被進一步包括在本發明的組成物或疫苗組合中，包括但不限於HIV包膜蛋白（例如，gp160、gp140、gp120、或gp41），較佳的是穩定化的三聚gp140蛋白，如穩定化的分枝C或分枝A gp140蛋白。在一個較佳的實施方式中，該分離的HIV抗原性多肽係一穩定化的HIV分枝C三聚gp140蛋白，如包含SEQ ID NO: 7的胺基酸序列的殘基30-708的蛋白（SEQ ID NO: 7的殘基1-29在該信號序列中）。可以使用的替代的或另外的HIV Env多肽（除了分枝C gp140蛋白或單獨的）係花葉Env三聚體蛋白，該葉Env三聚體蛋白例如具有如SEQ ID NO: 36的胺基酸30-724中所揭露的胺基酸序列（對應於WO 2014/107744的SEQ ID NO: 2，SEQ ID NO: 36的殘基1-29在該信號序列中）。

【0095】 根據本發明的一個具體實施方式，HIV抗原性蛋白質可以是本發明的合成的HIV包膜蛋白。因此，本發明的合成的包膜蛋白可以以分離的和/或純化的形式在對其有需要的受試者中使用從而例如針對HIV誘導免疫應答或提供保護性免疫。任何本文所述的包含SEQ ID N: 8的胺基酸序列的合成的包膜蛋白可以用作處於分離和/或純化形式的HIV抗原性蛋白。在一個較佳的實施方式中，當以分離的形式作為HIV抗原性蛋白使用時，該合成的包膜蛋白包含SEQ ID NO: 18的胺基酸序列的殘基30-711或SEQ ID NO: 19的胺基酸序列的殘基30-686，並且更較佳的是包含SEQ ID NO: 18的胺基酸序列的殘基30-704。該分離的HIV抗原性多肽還可以包含具有以下突變：EK479-480RRRR、I529P、A471C和T575C的SEQ ID NO: 8。

【0096】 本發明的實施方式還涉及組成物或疫苗組合，該等組成物或疫苗組合包含分離的合成的HIV包膜蛋白，該分離的合成的HIV包膜蛋白包含SEQ ID NO: 8的胺基酸序列。可以使用任何本文所述的合成的HIV包膜蛋白。在本發明的具體實施方式中，該分離的合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列的殘基30-704或30-711，SEQ ID NO: 19的胺基酸序列的殘基30-686，或具有以下突變：EK479-480RRRR、I529P、A471C和T575C的SEQ ID NO: 8的胺基酸序列。此類組成物或疫苗組合可以進一步包含一種或多種表現載體，例如，腺病毒載體，如腺病毒26載體，該等載體編碼一種或多種另外的HIV抗原性多肽，如本發明的合成的HIV包膜蛋白，或其他HIV抗原性蛋白如列於SEQ ID NO: 4、5、7、28或29中的那些或其片段。

【0097】 本發明還涉及生產本發明的組成物或疫苗組合之方法。根據本發明的實施方式，生產組成物或組合的方法包括將包含編碼本發明的合成的HIV包膜蛋白的核酸的載體與運載體、以及視情況編碼一種或多種另外的HIV抗原性多肽的一種或多種另外的載體和/或一種或多種分離的HIV抗原性多肽進行組合。熟習該項技術者將熟悉用於製備這樣的組成物的常規技術。

【0098】 疫苗和疫苗組合

【0099】 根據本發明之實施方式，組成物可以是疫苗。如在此所使用的，術語“疫苗”係指組成物，該組成物包含編碼本發明的能夠對受試者或接種疫苗的受試者提供保護性免疫或保護性免疫應答的合成的HIV包膜蛋白的表現載體，較佳的是病毒載體。根據本發明的實施方式，當向受試者給予該組成物時，該表現載體表現該編碼的合成的HIV包膜蛋白，並且將該表現的合成的HIV包膜蛋白提供給受試者的免疫系統，從而誘導所需的應答以產生免疫，或誘導免疫應答。

【0100】 因此，在另一概述方面中，本發明提供了用於在受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之疫苗。根據本發明的實施方式，該疫苗包含組成物，該組成物包含免疫原性有效量的表現載體，該表現載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的胺基酸序列，並且較佳的是SEQ ID NO: 18的胺基酸序列。較佳的是，該表現載體係病毒載體，更較佳的是腺病毒載體，並且最較佳的是腺病毒26載體。

【0101】 根據本發明的實施方式，當參考本文所述的該等方

法和組成物使用時，“誘導免疫應答”涵蓋用於預防目的針對感染（如HIV感染）提供保護性免疫和/或接種受試者，以及用於治療目的在對其有需要的受試者中引起針對感染（如HIV感染）的所希望的免疫應答或效應。較佳的是，本發明的該等方法係用於預防目的，如用於提供保護性免疫。該免疫應答可以是細胞免疫應答和/或體液免疫應答。

【0102】 如在此所使用的，術語“保護性免疫”或“保護性免疫應答”係指該接種的受試者能夠控制對其進行疫苗接種的致病因數的感染。通常，已經發展了“保護性免疫應答”的受試者僅發展輕度至中度的臨床症狀或根本沒有症狀。通常，對某種試劑具有“保護性免疫應答”或“保護性免疫”的受試者不會由於所述試劑的感染而死亡。

【0103】 根據本發明的實施方式，疫苗組成物可以進一步包含編碼一種或多種另外的HIV抗原性多肽的一種或多種另外的載體，例如病毒載體，如腺病毒載體，較佳的是腺病毒26載體，和/或一種或多種分離的HIV抗原性多肽。該合成的HIV包膜蛋白、另外的載體和/或一種或多種分離的HIV抗原性多肽可以被配製在該疫苗中的同一組成物或一種或多種不同組成物中。

【0104】 本發明還涉及用於在對其有需要的受試者中使用一種或多種載體與分離的抗原性多肽組合來引發和加強對一種或多種HIV分枝的免疫應答之疫苗組合。因此，在另一概述方面中，本發明提供了用於在受試者中誘導針對HIV的免疫應答的疫苗組合。根據本發明的實施方式，該疫苗組合包含：

(i) 第一組成物，該第一組成物包含免疫原性有效量的表現載體、和運載體，該表現載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(a) I529P, (b) K480E, 以及(c) EK479-480RRRR、I529P、A471C和T575C的組合；和

(ii) 第二組成物，該第二組成物包含免疫原性有效量的分離的HIV抗原性多肽和運載體，

其中該第一和第二組成物之一用於引發免疫並且另一種組成物用於加強免疫。

【0105】 根據本發明的實施方式，該疫苗組合視情況進一步包含免疫原性有效量的一種或多種另外的表現載體，該一種或多種另外的表現載體編碼一種或多種另外的HIV抗原性多肽。該一種或多種另外的表現載體可以包括在該第一組成物或該第二組成物中，或該一種或多種另外的表現載體可以包括在有待與該第一和/或第二組成物一起給予的一種或多種另外的組成物中。

【0106】 如在此所使用的，術語“共遞送”、“共給予”或“與.....一起給予”係指同時給予兩種或更多種組分，如病毒表現載體和分離的抗原性多肽，或多種病毒表現載體。“同時給予”可以是至少在同一天給予兩種或更多種組分。當兩種組分“一起給予”時，它們可以在短時間段內（例如24、20、16、12、8或4小時，或1小時或更短時間內）順序地在分開的組成物中給予，或者它們可以同時在單一組成物中給予。

【0107】 在本發明的疫苗組合的具體實施方式中，該第一組成物包含腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列；並且該分離的HIV抗原性多肽包含SEQ ID NO: 7的胺基酸序列的殘基30-708或SEQ ID NO: 36的殘基30-724。在一個具體實施方式中，該第一組成物進一步包含腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽。在另一個具體實施方式中，該第一組成物進一步包含一種或多種另外的腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼一種或多種另外的HIV抗原性多肽，該HIV抗原性多肽包含選自下組的胺基酸序列，該組由以下各項組成：SEQ ID NO: 28和29。

【0108】 本發明的另一個概述方面涉及試劑盒（kit），該試劑盒包含根據本發明的實施方式之疫苗組合。

【0109】 在上文和下面的說明性實施方式中詳細討論了可以在本發明的疫苗組合中使用的合成的HIV包膜蛋白、表現載體、另外的表現載體、由表現載體編碼的HIV抗原性多肽、以及分離的HIV抗原性多肽等的其他實施方式。

【0110】 用於誘導針對HIV感染的保護性免疫之方法

【0111】 本發明還涉及在對其有需要的受試者中誘導針對一種或多種HIV分枝的免疫應答之方法。本文所述的方法包括將一種或多種表現載體與一種或多種分離的抗原性多肽組合使用來引發和加強免疫應答之方法。

【0112】 在一個概述方面中，在受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答的方法包括向受試者給予包含免疫有效量的表現載體的組成物，該表現載體包含編碼合成的HIV包膜蛋白的核酸，該合成的HIV包膜蛋白包含SEQ ID NO: 8的胺基酸序列。本文所述的任何組成物可以用於在受試者中誘導針對HIV的免疫應答的方法。較佳的是，該組成物包括腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體包含編碼合成的HIV包膜蛋白的核酸，該合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列。該組成物可以進一步包含編碼一種或多種另外的HIV抗原性多肽的一種或多種另外的載體和/或一種或多種另外的分離的HIV抗原性多肽。

【0113】 在另一概述方面中，在受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之方法包括：

(i) 向該受試者給予第一組成物，該第一組成物包含免疫原性有效量的表現載體、和運載體，該表現載體編碼花葉HIV包膜蛋白，該花葉HIV包膜蛋白具有SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(a) I529P，(b) K480E，以及(c) EK479-480RRRR、I529P、A471C和T575C的組合；

(ii) 向該受試者給予第二組成物，該第二組成物包含免疫原性有效量的分離的HIV抗原性多肽、和運載體；並且

(iii) 視情況，向該受試者給予免疫原性有效量的一種或多種另外的表現載體，該另外的表現載體編碼一種或多種另外的HIV

抗原性多肽，

其中步驟 (i) 和 (ii) 以任一順序進行，其中步驟之一用於引發免疫並且另一步驟用於加強免疫。根據本發明的實施方式，該視情況、有效量的一種或多種另外的表現載體與該第一組成物或該第二組成物一起給予。在一個較佳的實施方式中，該視情況有效量的一種或多種另外的表現載體與該第一組成物一起給予。

【0114】 在誘導免疫應答的方法的一具體實施方式中，該第一組成物包含一種腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼包含SEQ ID NO: 8的胺基酸序列的合成的HIV包膜蛋白；和第二腺病毒載體，較佳的是腺病毒26載體，該第二腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽；該第二組成物包含分離的HIV抗原性多肽，該分離的HIV抗原性多肽具有SEQ ID NO: 7的胺基酸序列的殘基30-708或SEQ ID NO: 36的殘基30-724；並且該一種或多種另外的表現載體係腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼一種或多種另外的HIV抗原性多肽，該HIV抗原性多肽包含選自下組的胺基酸序列，該組由以下各項組成：SEQ ID NO: 28和29；其中向該受試者一起給予該第一組成物與該一種或多種另外的表現載體一次或多次以用於引發免疫，並且向該受試者給予該第二組成物一次或多次以用於加強免疫。

【0115】 通常肌內、皮內或皮下給予包含該等表現載體和/或抗原性多肽的免疫原性組成物。然而，也可以設想其他給藥模式，如靜脈內、直腸、皮膚、口服、鼻內等。免疫原性組成物的肌內

給予可以藉由使用針頭注射該等表現載體（例如，腺病毒載體）和/或抗原性多肽的懸浮液來實現。一替代方案係使用無針注射裝置來給予該組成物（使用例如Biojector™）或含有該疫苗的冷凍乾燥粉末。

【0116】 對於肌肉、靜脈內、皮膚或皮下注射，或病痛部位的注射，該載體將處於腸胃外可接受的水性溶液形式，該水性溶液為無熱原的並且具有適當的pH、等滲性以及穩定性。同樣地，該分離的抗原性多肽將處於腸胃外可接受的溶液形式，該溶液具有適當的pH、等滲性以及穩定性。熟習該項技術者完全能夠使用例如等滲媒介物，如氯化鈉注射液、林格氏注射液（Ringer's Injection）、乳酸鹽林格氏注射液來製備適合的溶液。根據需要，可以包括防腐劑、穩定劑、緩衝劑、抗氧化劑和/或其他添加劑。還可以使用緩釋配製物。

【0117】 通常，根據本發明的實施方式的疫苗組成物的給予將具有預防目的，以在感染或症狀發展之前產生針對HIV抗原的免疫應答。在其他實施方式中，該等表現載體（例如，腺病毒載體）、和/或HIV抗原性多肽可以被給予用於曝露後的預防用品。

【0118】 向受試者給予包含該等表現載體（例如，腺病毒載體）和/或抗原性多肽的該等免疫原性組成物，在受試者中產生抗HIV免疫應答。足以誘導可檢測的免疫應答的組成物的量被定義為“免疫原性有效劑量”或“免疫原性有效量。”在本發明的一典型實施方式中，該免疫應答係保護性免疫應答。

【0119】 給予的實際量以及給予的速率和時程將取決於正治

療的疾病的性質和嚴重度。治療處方（例如劑量等的決定）係在全科醫生和其他醫師、或在獸醫環境中是獸醫的職責範圍內的，並且典型地考慮了有待治療的失調、個體患者的狀況、遞送部位、給藥方法以及醫師已知的其他因素。以上提及的該等技術和方案的實例可以在雷明頓藥學大全（Remington's Pharmaceutical Sciences），第16版，Osol, A.編著，1980中找到。

【0120】 在產生腺病毒載體和任選將這種顆粒配製成組成物後，可以將該等載體給予個體，特別是人或其他靈長類動物。對非人哺乳動物的遞送不需要用於治療目的，而是可以用於在實驗背景中使用，例如在對由本發明的該等腺病毒載體表現的合成的HIV包膜蛋白的免疫應答的機制的研究中使用。

【0121】 在揭露的方法的一個實施方式中，使用編碼一種或多種HIV抗原性多肽的一種或多種腺病毒載體來引發免疫應答。一種或多種分離的HIV抗原性多肽可以與一種或多種腺病毒載體一起用於引發免疫。引發免疫可以僅施用一次，但也可以視情況施用多次，例如，在時間0初次引發施用，隨後是在約第4-14週，例如，第4、5、6、7、8、9、10、11、12、13或14週，或在初始引發施用之間、之後的任何時間另一次引發施用。可以將一種或多種分離的HIV抗原性多肽視情況與一種或多種另外的腺病毒或編碼一種或多種另外的HIV抗原性多肽的其他載體一起使用以增強免疫應答。加強免疫也可以施用一次或多次，例如，首先在約第18-36週，例如，第18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、或36週，或在初始引發施用

之間、之後的任何時間，隨後在約第36-52週，例如，第36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59或60週，或在初始引發施用之間、之後的任何時間另一次加強施用。監測由免疫誘導的免疫應答。

【0122】 該等揭露的方法的實施方式還考慮更短的初免-加強方案，意味著在初始引發施用之後的約22-26週施用該最終加強免疫。可以在第0週給予該引發免疫。該加強免疫可以施用多次，例如，首先在約第7-9週或第11-13週，或在約第4、5、6、7、8、9、10、11、12、13、或14週，或在初始引發施用之間、之後的任何時間，隨後在約第22-26週、或在約第16、17、18、19、20、21、22、23、24、25、26、27、或28週，或在初始引發施用之間、之後的任何時間另一次加強施用。在某些實施方式中，將一種或多種分離的HIV抗原性多肽與一種或多種腺病毒載體一起給予用於引發和/或加強免疫。

【0123】 熟習該項技術者容易理解，可以基於施用後測量的免疫應答來調整用於引發和加強施用的方案。例如，通常在給予引發組成物後數週或數月，例如，約2-3週或4週、或8週、或16週、或20週、或24週、或28週、或30週、或32週，或給予引發組成物後一至兩年給予該等加強組成物。

【0124】 根據本發明的實施方式，佐劑可以與該分離的HIV抗原性多肽一起作為引發和/或加強免疫的一部分來給予。鑒於本揭露，可以使用任何佐劑，並且在某些實施方式中，該佐劑係鋁

鹽，如AdjuPhos。

【0125】 在本發明的一較佳的實施方式中，在本文揭露的該等方法中使用的該等腺病毒載體包括rAd26載體。較佳的是，編碼合成的HIV包膜蛋白（包含SEQ ID NO: 18或SEQ ID NO: 19的，最較佳的是SEQ ID NO: 18的胺基酸序列）的rAd26載體，單獨地或與一種或多種另外的rAd26載體（編碼一種或多種另外的HIV抗原性多肽，如具有SEQ ID NO: 5的胺基酸序列的mos1Env）組合用於引發免疫應答，並且分離的HIV抗原性多肽（如包含SEQ ID NO: 7的胺基酸序列的殘基30-708或SEQ ID NO: 36的殘基30-724的那些）被用於加強免疫應答，或者反之亦然。

【0126】 在一個示例性實施方式中，編碼包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白的rAd26載體與編碼具有SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽的rAd26載體組合用於引發免疫應答。一種或多種另外的rAd26載體還可以與其他rAd26載體一起給予以引發免疫應答，該一種或多種另外的rAd26載體編碼一種或多種另外的具有選自由SEQ ID NOs: 1-4、28和29組成之群組的胺基酸序列的HIV抗原性多肽。在某些實施方式中，在給予任何加強免疫之前給予該引發施用兩次。然後給予分離的HIV抗原性多肽，如包含SEQ ID NO: 7(較佳的是)的胺基酸序列的殘基30-708的那些，或包含SEQ ID NO: 36的胺基酸序列的殘基30-724的那些，或這樣的分離的HIV抗原性多肽的至少兩種的組合，以加強免疫應答，並且較佳的是給予超過一次。較佳的是，在加強免疫中，進一步給予佐劑與該分離的HIV抗原性多肽。

【0127】 在一具體實施方式中，藉由給予在腺病毒載體（較佳的是rAd26載體）上編碼的四種HIV抗原來引發免疫應答，編碼的這四種抗原為：(i) 包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白，(ii) 具有SEQ ID NO: 5的胺基酸序列的多肽，(iii) 具有SEQ ID NO: 28的胺基酸序列的多肽，以及 (iv) 具有SEQ ID NO: 29的胺基酸序列的多肽。這四種抗原各自可以在單獨的腺病毒載體（較佳的是rAd26載體）上編碼，以約1、2、3、4、5、6、7、8、9、或 10×10^{10} 個病毒顆粒（vp），例如約 5×10^{10} vp（對於所有載體一起）的總劑量給予。該等載體可以例如以1 : 1 : 1 : 1比率預先混合，。可以在初始引發施用後，例如，在初始引發施用後第8、9、10、11、12、13、14、15或16週時重複該引發施用。在此實施方式中，藉由將用於該引發施用的同一腺病毒載體疫苗與分離的HIV Env gp140蛋白（例如分枝C gp140蛋白（包含SEQ ID NO: 7的胺基酸序列的殘基30-708）或花葉gp140蛋白（包含SEQ ID NO: 36的胺基酸序列的殘基30-724），或者分枝C gp140蛋白和花葉gp140蛋白）一起，以約50-300 μ g蛋白的總劑量，例如分枝C gp140蛋白的50、100、150、200、250、或300微克或之間的任何量，或例如花葉gp140蛋白的50、100、150、200、250、或300微克或之間的任何量，或例如分枝C gp140蛋白和花葉gp140蛋白的組合（例如以1 : 1比率，混合在一起活單獨地給予）的50、100、150、200、250、或300微克或之間的任何量給予來加強免疫應答。較佳的是，該gp140蛋白與佐劑，例如磷酸鋁一起給予。可以在約第18、19、20、21、22、23、24、25、26、27、28、29或30週，或在該初始引發施用之間、之後的任何時間進行該腺病毒加上gp140蛋白給予

以加強免疫應答。可以例如在約第42、43、44、45、46、47、48、49、50、51、52、53或54週，或在該初始引發施用之間、之後的任何時間重複該加強施用。藉由肌內途徑較佳的是進行根據這個實施方式的所有給予。

【實施方式】

【0128】 實施方式1係編碼合成的HIV包膜蛋白的核酸，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及(iii) EK479-480RRRR、I529P、A471C和T575C的組合。

【0129】 實施方式2係根據實施方式1所述之核酸，其中該合成的HIV包膜蛋白進一步包含信號序列，例如包含選自下組的胺基酸序列的信號序列，該組由以下各項組成：SEQ ID NO: 9至12，較佳的是SEQ ID NO: 9。

【0130】 實施方式3係根據實施方式1或2所述之核酸，其中該合成的HIV包膜蛋白進一步包含跨膜結構域，例如具有SEQ ID NO: 13的胺基酸序列的跨膜結構域，較佳的是該合成的HIV包膜蛋白進一步包含融合到SEQ ID NO: 8的C-末端和該跨膜結構域的N-末端的SEQ ID NO: 37。

【0131】 實施方式4係根據實施方式3所述之核酸，其中該合成的HIV包膜蛋白進一步包含一個細胞質結構域的片段，較佳的是包含SEQ ID NO: 14的胺基酸序列或其胺基酸殘基1-4（即，NRVR）

的細胞質結構域的片段。

【0132】 實施方式5係如以上實施方式1-4中任一項所述之核酸，其中該合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列。

【0133】 實施方式6係根據實施方式1或2所述之核酸，其中該合成的HIV包膜蛋白 (a) 進一步包含選自下組的三聚結構域，該組由以下各項組成：GCN4、fibrin(foldon結構域)，例如具有SEQ ID NO: 15或SEQ ID NO: 16的，較佳的是SEQ ID NO: 15的胺基酸序列的三聚結構域，或 (b) 包含具有以下突變：EK479-480RRRR、I529P、A471C和T575C的組合的SEQ ID NO: 8。

【0134】 實施方式7係根據實施方式6所述之核酸，其中該合成的HIV包膜蛋白包含SEQ ID NO: 19的胺基酸序列。

【0135】 實施方式8係根據實施方式5所述之核酸，其中該合成的HIV包膜蛋白由SEQ ID NO: 18的胺基酸序列組成。

【0136】 實施方式9係根據實施方式7所述之核酸，其中該合成的HIV包膜蛋白由SEQ ID NO: 19的胺基酸序列組成。

【0137】 實施方式10係一種載體，該載體包含實施方式1-9中任一項所述之核酸，其中所述核酸可操作地連接至啟動子序列。

【0138】 實施方式11係根據實施方式10所述之載體，係病毒載體，較佳的是腺病毒載體，並且更較佳的是腺病毒26載體。

【0139】 實施方式12係分離的細胞，該分離的細胞包含實施方式10或實施方式11所述的載體。

【0140】 實施方式13係一種組成物，該組成物包含免疫原性有效量的實施方式10或申請專利範圍11所述的載體、和運載體。

【0141】 實施方式14係一種疫苗組合，該疫苗組合包括第一組成物，該第一組成物包括免疫原性有效量的腺病毒載體，較佳的是腺病毒26載體，該載體編碼具有SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白；第二組成物，該第二組成物包括免疫原性有效量的第二腺病毒載體，較佳的是第二腺病毒26載體，該第二腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列的HIV抗原性多肽；以及視情況至少一種另外的組成物，該另外的組成物包括免疫有效量的選自由以下各項組成之群組的至少一種：載體，該載體編碼具有選自由SEQ ID NO: 1-4、28和29組成之群組的胺基酸序列的抗原性多肽，和分離的HIV抗原性多肽，該分離的HIV抗原性多肽具有SEQ ID NO: 7的胺基酸序列的殘基30-708或具有SEQ ID NO: 36的殘基30-724，其中該第一組成物、第二組成物和另外的組成物係存在於同一組成物或一種或多種不同組成物中。

【0142】 實施方式15係一種在對其有需要的受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之方法，該方法包括向該受試者給予實施方式13所述的組成物或實施方式14所述的疫苗組合。

【0143】 實施方式16係實施方式13所述的組成物或實施方式14所述之疫苗組合，該組成物或疫苗組合包含一種腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列；第

二腺病毒載體，較佳的是腺病毒26載體，該第二腺病毒載體編碼HIV抗原性多肽，該HIV抗原性多肽包含SEQ ID NO: 5的胺基酸序列；一種或多種另外的腺病毒載體，該一種或多種另外的腺病毒載體編碼一種或多種另外的抗原性多肽，該抗原性多肽包含選自下組的胺基酸序列，該組由以下各項組成：SEQ ID NO: 1-4、28和29；以及分離的HIV抗原性多肽，該分離的HIV抗原性多肽包含SEQ ID NO: 7的胺基酸序列的殘基30-708或SEQ ID NO: 36的殘基30-724，用於誘導針對人類免疫缺陷病毒（HIV）的免疫應答。

【0144】 實施方式17係合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及(iii) EK479-480RRRR、I529P、A471C和T575C的組合。

【0145】 實施方式18係實施方式17所述的合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含具有突變EK479-480RRRR、I529P、A471C和T575C的SEQ ID NO: 8的胺基酸序列，或SEQ ID NO: 18的胺基酸序列的殘基30-704或SEQ ID NO: 19的殘基30-686。

【0146】 實施方式19係實施方式13所述之組成物，該組成物進一步包含編碼一種或多種另外的HIV抗原性多肽的一種或多種另外的表現載體（該表現載體）和/或一種或多種分離的HIV抗原性多肽。

【0147】 實施方式20係實施方式13所述之組成物，該組成物包含腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼由

SEQ ID NO: 18的胺基酸序列組成的合成的HIV包膜蛋白。

【0148】 實施方式21係根據實施方式20所述之組成物，該組成物進一步包含第二腺病毒載體，較佳的是腺病毒26載體，該第二腺病毒載體編碼HIV抗原性多肽，該HIV抗原性多肽包含SEQ ID NO: 5的胺基酸序列；以及視情況一種或多種另外的腺病毒載體，較佳的是腺病毒26載體，該一種或多種另外的腺病毒載體編碼一種或多種另外的HIV抗原性多肽，該HIV抗原性多肽包含SEQ ID NO: 1-4、28和29的胺基酸序列。

【0149】 實施方式22係在對其有需要的受試者中產生針對人類免疫缺陷病毒（HIV）的免疫應答之方法，該方法包括向該受試者給予根據實施方式19、20、或21中任一項所述的組成物。

【0150】 實施方式23係生產組成物或疫苗組合之方法，該方法包括將實施方式10或實施方式11所述的載體與運載體、以及視情況編碼一個或多個另外的HIV抗原性多肽的一種或多種另外的載體和/或一種或多種分離的HIV抗原性多肽連同運載體一起進行組合。

【0151】 實施方式24係用於在受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之疫苗組合，該疫苗組合包含：

(i) 第一組成物，該第一組成物包含免疫原性有效量的表現載體、和運載體，該表現載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i)

I529P，(ii) K480E，以及 (iii) EK479-480RRRR、I529P、A471C 和T575C的組合；和

(ii) 第二組成物，該第二組成物包含免疫原性有效量的分離的HIV抗原性多肽、和運載體，

其中該第一和第二組成物之一用於引發免疫並且另一種組成物用於加強免疫，並且

其中該疫苗組合視情況進一步包含免疫原性有效量的一種或多種另外的表現載體，該一種或多種另外的表現載體編碼一種或多種另外的HIV抗原性多肽，並且該一種或多種另外的表現載體包括在該第一或該第二組成物中，或一種或多種要與該第一或該第二組成物一起使用的另外的組成物中。

【0152】 實施方式25係根據實施方式24所述的疫苗組合，其中該第一組成物包含腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列；該分離的HIV抗原性多肽包含SEQ ID NO: 7的胺基酸序列的殘基30-708，或SEQ ID NO: 36的殘基30-724；並且該一種或多種另外的表現載體係腺病毒載體，較佳的是腺病毒26載體，該一種或多種另外的載體編碼一種或多種另外的HIV抗原性多肽，該另外的HIV抗原性多肽包含選自下組的胺基酸序列，該組由以下各項組成：SEQ ID NO: 1-5、28和29。

【0153】 實施方式26係用於在對其有需要的受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之方法，該方法包括：

(i) 向該受試者給予第一組成物，該第一組成物包含免疫原性有效量的表現載體、和運載體，該表現載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及 (iii) EK479-480RRRR、I529P、A471C和T575C的組合；

(ii) 向該受試者給予第二組成物，該第二組成物包含免疫原性有效量的分離的HIV抗原性多肽、和運載體；並且

(iii) 視情況，向該受試者給予免疫原性有效量的一種或多種另外的表現載體，該一種或多種另外的表現載體編碼一種或多種另外的HIV抗原性多肽，

其中步驟 (i) 和 (ii) 係以任一順序進行，該等步驟之一用於引發免疫並且另一步驟用於加強免疫，並且較佳的是，該視情況、有效量的一種或多種另外的表現載體與該第一組成物一起給予。

【0154】 實施方式27係根據實施方式26所述之方法，其中該第一組成物包含一種腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼合成的HIV包膜蛋白，該合成的HIV包膜蛋白具有SEQ ID NO: 18的胺基酸序列；和第二腺病毒載體，較佳的是腺病毒26載體，該第二腺病毒載體編碼HIV抗原性多肽，該HIV抗原性多肽包含SEQ ID NO: 5的胺基酸序列；該第二組成物包含分離的HIV抗原性多肽，該分離的HIV抗原性多肽具有SEQ ID NO: 7的胺基酸序列的殘基30-708、或SEQ ID NO: 36的殘基30-724；並且該視情況一種或多種另外的表現載體係腺病毒載體，較佳的是腺病

毒26載體，該一種或多種另外的表現載體編碼一種或多種另外的HIV抗原性多肽，該另外的HIV抗原性多肽包含選自下組的胺基酸序列，該組由以下各項組成：SEQ ID NO: 28和29；其中向該受試者給予該第一組成物，視情況連同該一種或多種另外的表現載體一次或多次以用於引發免疫，並且向該受試者給予該第二組成物一次或多次以用於加強免疫。

【0155】 實施方式28係合成的HIV包膜蛋白，該合成的HIV包膜蛋白由SEQ ID NO: 18或SEQ ID NO: 19的胺基酸序列（具有或沒有該信號序列）組成。

【0156】 實施方式29係一種疫苗組合，該疫苗組合包含一種或多種載體，該等載體一起包含編碼以下各項的核酸序列：(i) 包含SEQ ID NO: 8的胺基酸序列的第一合成的HIV包膜蛋白和 (ii) 包含SEQ ID NO: 5的胺基酸序列的第二HIV包膜蛋白。

【0157】 實施方式30係根據實施方式29所述之疫苗組合，其中該第一合成的HIV包膜蛋白包含SEQ ID NO: 18的胺基酸序列。

【0158】 實施方式31係一種疫苗組合，該疫苗組合包含以下組分：

(i) 編碼由SEQ ID NO: 18的胺基酸序列組成的合成的HIV包膜蛋白的Ad26載體；和

(ii) 編碼由SEQ ID NO: 5的胺基酸序列組成的HIV包膜蛋白的Ad26載體。

【0159】 實施方式32係根據實施方式31所述之疫苗組合，該

疫苗組合進一步包含以下組分：

(iii) 編碼由SEQ ID NO: 28的胺基酸序列組成的HIV抗原的Ad26載體。

【0160】 實施方式33係根據實施方式31或32所述之疫苗組合，該疫苗組合進一步包含以下組分：

(iv) 編碼由SEQ ID NO: 29的胺基酸序列組成的HIV抗原的Ad26載體。

【0161】 實施方式34係根據實施方式31-33中任一項所述之疫苗組合，該疫苗組合進一步包含以下組分：

(v) 分離的HIV抗原性多肽，該分離的HIV抗原性多肽具有SEQ ID NO: 7的胺基酸序列的殘基30-708、或SEQ ID NO: 36的胺基酸序列的殘基30-724，視情況進一步包含佐劑。

【0162】 實施方式35係一種在對其有需要的受試者中誘導針對人類免疫缺陷病毒（HIV）的免疫應答之方法，該方法包括：(a) 向該受試者給予：(i) 編碼包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白的rAd26載體；(ii) 編碼包含SEQ ID NO: 5的胺基酸序列的抗原的rAd26載體；(iii) 編碼包含SEQ ID NO: 28的胺基酸序列的抗原的rAd26載體；以及 (iv) 編碼包含SEQ ID NO: 29的胺基酸序列的抗原的rAd26載體；較佳的是，其中以約1:1:1:1的比率，以約 $1-10 \times 10^{10}$ 病毒顆粒（vp）（例如 5×10^{10} vp）的總劑量給予該等rAd26載體；

(b) 在步驟 (a) 之後的約第10-14週，例如在第12週重複步驟

(a)；

(c) 向該受試者給予：(i) 編碼包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白的rAd26載體；(ii) 編碼包含SEQ ID NO: 5的胺基酸序列的抗原的rAd26載體；(iii) 編碼包含SEQ ID NO: 28的胺基酸序列的抗原的rAd26載體；(iv) 編碼包含SEQ ID NO: 29的胺基酸序列的抗原的rAd26載體；(v) 具有SEQ ID NO: 7的胺基酸30-708的序列的分離的HIV gp140蛋白；以及(vi) 磷酸鋁佐劑；較佳的是，其中以約1 : 1 : 1 : 1的比率、以約 $1-10 \times 10^{10}$ 病毒顆粒 (vp) (例如 5×10^{10} vp) 的總劑量給予該等rAd26載體，並且其中以約50-300微克 (例如250微克) 的劑量給予該分離的HIV gp140蛋白；在步驟 (a) 之後的約第20-28週，例如在第24週；以及 (d) 在步驟 (a) 之後的約第42-54週，例如在第48週重複步驟 (c)。

【0163】 實施方式36係一種在對其有需要的受試者中誘導針對人類免疫缺陷病毒 (HIV) 的免疫應答之方法，該方法包括：(a) 向該受試者給予：(i) 編碼包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白的rAd26載體；(ii) 編碼包含SEQ ID NO: 5的胺基酸序列的抗原的rAd26載體；(iii) 編碼包含SEQ ID NO: 28的胺基酸序列的抗原的rAd26載體；以及(iv) 編碼包含SEQ ID NO: 29的胺基酸序列的抗原的rAd26載體；較佳的是，其中以約1 : 1 : 1 : 1的比率，以約 $1-10 \times 10^{10}$ 病毒顆粒 (vp) (例如 5×10^{10} vp) 的總劑量給予該等rAd26載體；

(b) 在步驟 (a) 之後的約第10-14週，例如在第12週重複步驟 (a)；

(c) 向該受試者給予：(i) 編碼包含SEQ ID NO: 18的胺基酸序列的合成的HIV包膜蛋白的rAd26載體；(ii) 編碼包含SEQ ID NO: 5的胺基酸序列的抗原的rAd26載體；(iii) 編碼包含SEQ ID NO: 28的胺基酸序列的抗原的rAd26載體；(iv) 編碼包含SEQ ID NO: 29的胺基酸序列的抗原的rAd26載體；(v) 具有SEQ ID NO: 7的胺基酸30-708的序列的分離的HIV gp140蛋白；(vi) 具有SEQ ID NO: 36的胺基酸30-724的序列的分離的HIV gp140蛋白；以及 (vii) 磷酸鋁佐劑；較佳的是，其中以約1：1：1：1的比率、以約 $1-10 \times 10^{10}$ 病毒顆粒 (vp) (例如 5×10^{10} vp) 的總劑量給予該等rAd26載體，並且其中以約1：1的比率、以約50-300微克 (例如250微克) 的總劑量給予該等分離的HIV gp140蛋白；在步驟 (a) 之後的約第20-28週，例如在第24週；以及 (d) 在步驟 (a) 之後的約第42-54週，例如在第48週重複步驟 (c)。

【實例】

實例1：HIV包膜抗原序列之設計

【0164】 設計了幾個HIV包膜抗原序列，其與花葉HIV抗原mos2Env (SEQ ID NO: 6；先前也描述於WO 2010/059732中) 具有序列相似性。該等新設計的膜結合的序列基於來自HIV包膜蛋白的完全天然野生型序列、或mos2Env序列和野生型HIV包膜蛋白序列的嵌合體 (的組合)。除了全長包膜蛋白序列 (參見圖1A)，還設計了具有細胞質結構域的C-末端截短的序列 (參見例如，圖1C)。還參見例如，Schiernle等人，美國科學院院刊 (PNAS) 1997；亞伯拉罕 (Abrahamyan) 等人，病毒學雜誌 (J Virol) 2005；愛德

華茲 (Edwards) 等人，病毒學雜誌 (*J. Virology*)，2002，76: 2683-2691。在跨膜 (TM) 區域之前，還藉由C-末端截短來製備可溶性變體，該跨膜區域被三聚結構域 (如GCN4三聚結構域) 替換 (參見例如，圖1B)。該等可溶性變體藉由弗林蛋白酶切割位點的突變進一步轉化為單鏈變體，從而抑制該包膜蛋白的細胞外結構域加工成gp120和gp41亞基。

【0165】 在產生的和測試的所有構建體中，基於C4的構建體具有最佳性質，例如良好的可製造性、折疊、免疫原性等，並且選擇該等基於C4的構建體用於進一步研究。在進一步研究中還產生並測試了具有代替跨膜結構域的GCN4三聚結構域的C4構建體的可溶性變體 (sC4，圖1B) 和包含細胞質結構域的7個胺基酸片段的變體 (C4D7，圖1C)。C4、sC4、以及C4D7的胺基酸序列分別示於SEQ ID NO: 17、19、和18中。編碼該等的序列分別示於SEQ ID NO: 25、27、和26中。構建體C1具有基於mos2Env序列 (SEQ ID NO: 6) 的細胞外結構域序列。還分別產生了具有代替跨膜結構域的GCN4三聚結構域的構建體C1的可溶性變體 (sC1) 和包含細胞質結構域的7個胺基酸片段的變體 (C1D7)，它們類似於如圖1B和1C中示出的sC4和C4D7。出於比較目的，將構建體C1及其變體用於進一步研究中，因為它們基本上基於先前技術的mos2Env序列。C1、sC1以及C1D7的胺基酸序列分別示於SEQ ID NO: 31、30、和32中。編碼該等的核酸序列分別示於SEQ ID NO: 34、33、和35中。測試的其他構建體不如基於構建體C4的那些最優，並且沒有進入進一步發展。

實例2：合成的HIV包膜蛋白之表現和折疊

【0166】 測量合成的HIV包膜蛋白的表現水平、折疊、和細胞表面表現。

【0167】 表現水平

【0168】 用編碼如實例1所述的可溶性合成的HIV包膜蛋白sC1和sC4的質粒暫態轉染HEK293F細胞。使用定量西方墨點法（QWB）在上清液中測量該可溶性蛋白的表現水平。結果示於圖2中。sC1的低表現水平（其基本上對應於具有添加的跨膜結構域的mos2Env）與我們最近對mos2Env的見解一致。如藉由該等結果所證明的，本發明的sC4變體顯示比該sC1變體（對照）顯著更高的表現水平。

【0169】 蛋白質折疊

【0170】 藉由測量可溶性合成的HIV包膜蛋白與已知結合該HIV包膜蛋白的共受體結合位點的抗體（MAb 17b）的結合來測試蛋白質折疊，該HIV包膜蛋白僅在結合CD4後藉由酶聯免疫吸附測定（ELISA）而暴露。具體地，在存在預先的sC4與CD4的結合以及不存在預先的sC4與CD4的結合的情況下，測試了純化的sC4與MAb17b的結合。使用純化的sC1作為對照。在沒有預先的CD4結合到該包膜蛋白的情況下，MAb 17b與sC4的結合係部分解折疊的或預先觸發的包膜蛋白（即，在不存在CD4結合時採取“開放”構象的不穩定Env）的指示。該ELISA測定的結果示於圖3A和3B中。

【0171】 如圖3B中所示，在存在預先與CD4結合的情況下，

sC4顯示了與MAb 17b的強結合，但在不存在預先與CD4結合的情況下，sC4顯示沒有與MAb 17b的可檢測的結合。相比之下，如在圖3A中所示，在存在與不存在預先與CD4結合的情況下，sC1顯示與MAb 17b的低得多的結合。結果表明，在不暴露預先與CD4結合的共受體結合位點的情況下，sC4具有正確的折疊模式。

【0172】 還藉由 sC1 和 sC4 的天然聚丙烯醯胺凝膠電泳 (PAGE) 來分析蛋白質折疊，以評價可溶性蛋白質變體的四級結構，以及原聚體之間可能的不正確的二硫鍵形成。在非變性凝膠上電泳後，藉由西方墨點分析來檢測凝膠中的蛋白質。如圖4中的結果所示，大部分sC4以三聚體狀態存在，這係正確的四級結構。

【0173】 總之，蛋白質折疊實驗的結果表明該sC4可溶性合成的HIV包膜蛋白具有希望的折疊特性 (folding profile)，這與現有的mos2Env抗原 (由sC1表示) 的折疊特性相比係被改進的。

【0174】 細胞表面表現

【0175】 還研究了HIV包膜蛋白C1 (全長)、C4 (全長，參見圖1A)、C1D7和C4D7的膜結合變體的細胞表面表現。僅用編碼eGFP的質粒 (陰性對照，NC)，或用編碼eGFP的質粒連同編碼HIV包膜蛋白變體的表現構建體暫態轉染HEK293T細胞。轉染後兩天，在暴露於針對gp120的多種多株和單株抗體、以及二級抗體後，對細胞進行螢光啟動細胞分選 (FACS) 分析，並且然後檢查包膜蛋白細胞表面表現水平。藉由使用抗gp120多株抗體測定總表現水平，並藉由評估廣泛中和性抗體PG9和PG16 (它們係四級結構依賴性的，並優先結合正確折疊的包膜三聚體) 的相對結合，

來評估該等包膜變體的質量。

【0176】 細胞表面表現實驗的結果示於圖5中。如使用抗gp120抗體測量的截短變體C1D7和C4D7的表面表現水平分別遠高於其全長對應物C1和C4的表面表現水平。這證實了來自Env的羧基-末端的144個殘基的缺失增加了包膜表面表現水平。與全長C1相比，本發明的全長C4構建體還顯示出提高的PG9和PG16結合，這表明該C4包膜序列在細胞表面上被適當地折疊（即，三聚體）。

【0177】 結果還表明，C1D7變體，其基本上是具有添加的跨膜結構域和細胞質結構域的7個胺基酸的Mos2Env，可以在HEK293T細胞上進行表面表現。這與當轉染至A549細胞時不能在表面上以可檢測的水平表現的Ad26.mos2Env中的可溶性構建體形成對比。然而，與PG9和PG16的相對結合在背景之上幾乎檢測不到，這表明該C1D7包膜序列折疊不良，並且可能不作為細胞表面上的完整三聚體存在。

【0178】 總的來說，與C1和C1D7相比，該C4D7包膜變體具有最佳的抗體結合特性，具有比其全長對應物C4更高的gp120表現，和大於15倍的增加的PG9和PG16結合（圖5）。

實例3：編碼HIV包膜序列的載體之穩定性

【0179】 我們實驗室中的先前工作（未公開）表明編碼該mos2Env抗原序列的腺病毒26（Ad26）載體顯示具有相對高的VP/IU比率（表明腺病毒產物批次的較低質量），並且此外這樣的載體顯示穩定性問題。因此，重要的是在腺病毒背景下測試本發

明的合成的HIV包膜蛋白構建體的穩定性。

【0180】 在PER.C6細胞中產生編碼如以上實例1中所述的本發明的HIV抗原序列C4、C4D7和sC4的重組Ad26(rAd26)載體(分別稱為“rAd26.C4”、“rAd26.C4D7”、和“rAd26.sC4”)。挑選載體植株(斑塊)並按比例放大以產生研究批次。將最多5個病毒植株(斑塊)按比例擴大至T25形式，並以T25形式連續傳代10代(傳代1-3係轉染和斑塊純化步驟，隨後以T25形式傳代10次，得到總共13次傳代)。藉由E1轉基因盒PCR分析在病毒傳代數(vpn)3、5、10和13時評估遺傳穩定性，隨後在vpn 13時測序。結果示於圖6中。

【0181】 編碼全長C4(rAd26.C4)的rAd26載體顯示差的生長特徵，如藉由在2-3天中沒有完全細胞病變效應(CPE)；遺傳不穩定性，如藉由E1轉基因盒區域的缺失所確定的；或其組合所確定的(圖6)。由於差的生長特徵和觀察到的遺傳不穩定性，不進一步追蹤該編碼全長C4的載體。

【0182】 相比之下，對於編碼C4D7(rAd26.C4D7)和sC4(rAd26.sC4)的該等rAd26載體，所有繁殖的斑塊在實驗過程中保持遺傳穩定(圖6)。因此，就腺病毒載體背景中的穩定性而言，該等新穎的sC4和C4D7構建體勝過原始mos2Env構建體。直到vpn 13的遺傳穩定性試驗代表了在載體的工業規模製備中使用的幾個傳代之外的繁殖。

實例4：腺病毒載體中的HIV包膜序列的表現和體內抗原性

【0183】 在體內在載體轉導的A549細胞（人細胞系）中單獨地或與編碼mos1Env（SEQ ID NO: 5）的重組Ad26載體（下文稱為“rAd26.mos1Env”）組合來評估rAd26.C4D7和rAd26.sC4的表現和抗原性（數據未顯示）。流式細胞術分析證明所有抗原在用作為對照的該等單一包膜抗原的 2×10^4 個病毒顆粒（vp），或用藉由腺病毒轉導的2種組合的Env抗原的 1×10^4 vp轉導的細胞培養物中表現。所有轉導物另外含有編碼mos1GagPol（“rAd26.mos1GagPol”）和mos2GagPol（“rAd26.mos2GagPol”）的單劑量（ 1×10^4 vp）的腺病毒載體（巴魯什等人，自然醫學（Nat Med）2010, 16: 319-323），使得所評估的載體組合顯示與用於臨床前和臨床用途的不同腺病毒載體相同的相對比例。較佳的是，將編碼本發明的合成的HIV包膜蛋白的該等載體與編碼用於臨床用途的mos1GagPol和mos2GagPol抗原的載體進行組合。

【0184】 rAd26.mos1Env和rAd26.C4D7的組合產生如藉由單株抗體結合測定的評估表位的最大覆蓋。特別地，藉由用Ad26.C4D7轉化而貢獻的PG16表位的暴露有希望用於疫苗使用，因為PG16代表識別HIV-1 Env的V1/V2環區域的廣泛中和性單株抗體（沃克（Walker）等人，科學（Science），2009）。因此，與僅由mos1Env產生的免疫應答相比，來源於C4序列的本發明的合成的HIV包膜蛋白增加了針對該HIV包膜蛋白的免疫應答的寬度。在RV144研究中，針對包膜蛋白區域的疫苗誘導的抗體應答已經顯示與來自HIV-1感染的保護相關（海恩斯（Haynes）等人，新英格蘭醫學雜誌（N Engl J Med.）2012），因此本發明的合成的HIV包膜

蛋白係包括在HIV疫苗方案中的物有希望的候選。

實例5：編碼合成的HIV包膜蛋白的載體之免疫原性

【0185】 在兔中測試Ad26載體背景下的本發明的合成的HIV包膜蛋白序列，以確定該等構建體是否是rAd26.mos2Env構建體的免疫原性替代物。

【0186】 單獨地，並與編碼本發明的合成的HIV包膜蛋白的腺病毒載體（rAd26.C4D7和rAd26.sC4；包含SEQ ID NO: 8，具體地分別是SEQ ID NO: 18和19）組合地測試編碼mos1Env的腺病毒載體（rAd26.mos1Env；SEQ ID NO: 5）的免疫原性。在所有情況下，還給予編碼mos1GagPol和mos2GagPol抗原的腺病毒26載體（分別是rAd26.mos1GagPol [SEQ ID NO: 28]和rAd26.mos2GagPol [SEQ ID NO: 29]）。更確切地說，將rAd26.mos1Env單獨的免疫原性（三價疫苗：rAd26.mos1GagPol、rAd26.mos2GagPol和rAd26.mos1Env）與rAd26.mos1Env和rAd26.C4D7或rAd26.sC4之一的組合的免疫原性進行比較（四價疫苗：給予任何rAd26.mos1GagPol、rAd26.mos2GagPol、rAd26.mos1Env和rAd26.C4D7；或給予rAd26.mos1GagPol、rAd26.mos2GagPol、rAd26.mos1Env和rAd26.sC4）。該三價疫苗（缺少編碼本發明的合成的HIV包膜蛋白的任何載體）與該四價疫苗（包含編碼本發明的合成的HIV包膜蛋白的載體）的比較允許確定本發明的HIV包膜蛋白是否有助於保護的寬度。

【0187】 在疫苗方案中進行給予，其中該等Ad26載體在第0和6週作為雙重引發來給予，並且分枝C gp140蛋白（三價Env gp140

蛋白，其具有不帶有殘基1-29的信號肽序列的SEQ ID NO: 7，還參見WO 2010/042942）在第12和18週作為雙重加強來給予（參見例如，巴魯什等人，2015，科學（Science）349: 320-324）。表1描述了用於本研究的疫苗方案。空rAd26係指缺少編碼HIV抗原蛋白的序列的任何基因的對照載體。每個組包括六隻兔子。

表1：兔中的免疫原性研究中測試的疫苗方案

組	第一和第二免疫			第三和第四免疫			N=
	腺苷載體	劑量（vp）	總劑量（vp）	蛋白加強	劑量（ug）	佐劑	
1	rAd26.Mos1Env	2.5x10 ¹⁰	5x10 ¹⁰	GP140（分枝C）	10	AdjuPhos 250 µg	6
	rAd26.Mos1GagPol	1.25x10 ¹⁰					
	rAd26.Mos2Gagpol	1.25x10 ¹⁰					
2	rAd26.Mos1Env	1.25x10 ¹⁰	5x10 ¹⁰	GP140（分枝C）	10	AdjuPhos 250 µg	6
	rAd26.C4D7	1.25x10 ¹⁰					
	rAd26.Mos1GagPol	1.25x10 ¹⁰					
	rAd26.Mos2Gagpol	1.25x10 ¹⁰					
3	rAd26.Mos1Env	1.25x10 ¹⁰	5x10 ¹⁰	GP140（分枝C）	10	AdjuPhos 250 µg	6
	rAd26.sC4	1.25x10 ¹⁰					
	rAd26.Mos1GagPol	1.25x10 ¹⁰					
	rAd26.Mos2Gagpol	1.25x10 ¹⁰					
對照	空rAd26	5x10 ¹⁰	5x10 ¹⁰	NA	0	AdjuPhos 250 µg	6

【0188】 該三價Ad26疫苗（缺乏本發明的新穎Env抗原）與該四價Ad26疫苗（其包含新型sC4或C4D7 Env抗原）的比較允許測試新穎抗原是否有助於保護的寬度。使用建立的基於T2M-b1細

胞的中和測定[Montefiori DC. 分子生物學方法 (*Methods Mol Biol*) 2009,**485**:395-405; Sarzotti-Kelsoe M等人, 免疫學方法 (*J Immunol Methods*) 2014,**409**:131-146]來測量候選疫苗的中和活性。

【0189】 結果示於圖7中，並且藉由使用該三價疫苗（表1中的組1）作為對照組並與每種新穎四價疫苗（表1中的組2和組3）進行比較來統計分析該等結果。

【0190】 總的來說，在兔中進行兩次同源肌內免疫後，新穎C4-衍生的（即，編碼包含SEQ ID NO: 8的Env蛋白，係mos2Env的替代物）腺病毒構建體係免疫原性的。

【0191】 兔免疫血清針對Tier 1B假病毒的中和能力不存在（數據未顯示），這係不出乎意料的，因為已知此類病毒更難以中和。

【0192】 兔免疫血清針對分枝B Tier 1A病毒的假病毒中和能力不受添加新組分的影響（數據未顯示）。這證明新抗原不會不利地幹擾該疫苗中存在的現有分枝B抗原的免疫原性（儘管該等新組分針對分枝C，但是這種不期望的幹擾在其被測試前不能被先驗排除）。

【0193】 與三價（僅具有mos1Env）單獨的免疫（組1）相比，在含有腺病毒的該四價新穎C4D7（四價，組2）中兔免疫血清針對分枝C Tier 1A病毒的假病毒中和能力顯著增強（圖7，B圖）。此外，與三價（僅具有mos1Env）單獨的免疫（組1）相比，在含有腺病毒的該四價新穎sC4（四價，組3）中兔免疫血清針對分枝C Tier 1A

病毒的假病毒中和能力在第8週顯著增強（圖7，B圖）。

【0194】 總之，Ad26中編碼的C4D7和sC4構建體係免疫原性的並且其添加擴大了疫苗的結合能力和中和能力，該疫苗具有mos1Env（主要是分枝B）作為唯一的朝向分枝C菌株的Ad26編碼的Env組分（圖7B）。

實例6：包括編碼本發明的合成的HIV包膜蛋白的載體的疫苗方案之免疫原性

【0195】 一種另外的兔研究評估了該四價載體組合Ad26.Mos4.HIV（由四種腺病毒載體：Ad26.Mos1GagPol [編碼SEQ ID NO: 28]、Ad26.Mos2GagPol [編碼SEQ ID NO: 29]、Ad26.Mos1Env [編碼SEQ ID NO: 5]和Ad26.Mos2SEnv [如以上使用的名稱“C4D7”還被稱為“Mos2S”；該載體以1:1:1:1混合，以 5×10^9 vp的總劑量編碼根據本發明的新穎SEQ ID NO: 18]組成），在第0週和第6週肌內給予該四價載體組合作為雙重引發免疫，在第13週和第19週與重組的HIV-1 Env 蛋白加強進行組合，使用分枝C gp140 [具有SEQ ID NO: 7的胺基酸殘基30-708的序列]、花葉gp140 [具有SEQ ID NO: 36的胺基酸殘基30-724的序列]、或分枝C gp140和花葉gp140的組合。以10或50微克蛋白質組合免疫當天配製的250 mcg磷酸鋁佐劑的總劑量肌內施用該等蛋白質加強。

【0196】 結果表明，所有測試的方案在所有動物中是免疫原性的，這誘導高抗體滴度和針對Tier 1 Env假型病毒的中等中和活性。如果將花葉gp140單獨地或與分枝C gp140組合用作疫苗抗

原，則與單獨用分枝C gp140加強的參照組相比，花葉gp140特異性ELISA滴度和分枝B假病毒識別在第15週顯著增加。改善的總體效應大小係中度的，並且與單獨的花葉gp140相比，用二價分枝C gp140-花葉gp140組合而增強的組更大。在研究的第21週，失去該等差異，並且針對接受二價分枝C gp140-花葉gp140加強或單價分枝C gp140加強的群體而測量的免疫應答在統計學上不可區分。

【0197】 該二價蛋白質方案顯示分枝C ELISA滴度和假病毒識別與分枝C gp140單獨加強方案的可比較的誘導，這表明分枝B相關的免疫原花葉gp140的包含對分枝C抗原覆蓋沒有不利影響，同時在本研究的第15週顯著增強了分枝B覆蓋。

【0198】 數據證實，編碼根據本發明的合成的Env抗原的Ad26.Mos2SEnv載體可成功地用於疫苗方案中。

參考文獻

1. 巴魯什等人，自然醫學 (*Nat Med*) 2010，16: 319-323
2. WO 2010/059732
3. Schiernle等人，美國科學院院刊 (*PNAS*) 94: 8640-8645，
1997
4. 亞伯拉罕等人，病毒學雜誌 (*J Virol*) 79: 106-115，2005
5. US 20120076812
6. 巴魯什等人，細胞 (*Cell*) 155: 1-9，2013
7. 哈文格等人，2006，普通病毒學雜誌 (*J Gen Virol*) 87:
2135-43；
8. WO 03/104467
9. WO 2004/001032
10. WO 2007/104792
11. 阿賓克等人，(2007) 病毒學 (*Virol*) 81(9): 4654-63
12. 美國專利號7,270,811
13. 沃格爾(Vogels)等人，(2003) 病毒學雜誌(*J Virol*)77(15):
8263-71
14. WO 00/70071
15. WO2012/082918

16. 沃克LM、Phogat SK、陳慧 (Chan-Hui) PY、瓦格納 (Wagner) D、彭 (Phung) P、戈斯 (Goss) JL等人，來自非洲供體的廣泛並有效的中和性抗體揭示了新的HIV-1疫苗靶標 (Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target)，科學 (*Science*) 2009, **326**: 285-289。

17. 海恩斯BF、吉伯特 (Gilbert) PB、麥克拉特 (McElrath) MJ、左拉 (Zolla) -Pazner S、Tomaras GD、阿拉姆 (Alam) SM等人，HIV-1疫苗效力試驗的免疫相關分析 (Immune-correlates analysis of an HIV-1 vaccine efficacy trial)，新英格蘭醫學雜誌 (*N Engl J Med*) 2012, **366**: 1275-1286。

18. 巴魯什等人，(2015) 科學 (*Science*) 349: 320-324

19. 蒙特菲奧裡 (Montefiori) DC，在螢光素酶報導基因測定中測量HIV中和性 (Measuring HIV neutralization in a luciferase reporter gene assay)，分子生物學方法 (*Methods Mol Biol*) 2009, **485**: 395-405。

20. Sarzotti-凱爾索 (Kelsoe) M、貝勒 (Bailer) RT、Turk E、林 (Lin) CL、Bilska M、格林 (Greene) KM等人，針對HIV-1的中和抗體的標準化評估的TZM-bl測定的優化和驗證 (Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1)，免疫學方法雜誌 (*J Immunol Methods*) 2014, **409**: 131-146。

21. 愛德華茲等人，病毒學雜誌 (*J. Virology*)，2002, 76: 2683-2691。

【序列表】

<110> 荷蘭商傑森疫苗防護公司
<120> 人類免疫缺陷病毒抗原、載體、組成物、及其使用方法
<130> 0265 EP P01 PRI
<160> 37
<170> PatentIn 3.5版本
<210> 1
<211> 500
<212> PRT
<213> 人工序列
<220>
<223> mos1Gag 花葉抗原序列
<400> 1
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
1 5 10 15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys
20 25 30
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
35 40 45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
50 55 60
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
65 70 75 80
Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
85 90 95
Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
100 105 110
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val
115 120 125
Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His
130 135 140
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
145 150 155 160
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
165 170 175
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
180 185 190
Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
195 200 205
Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala
210 215 220
Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225 230 235 240
Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
245 250 255

Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
275 280 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
290 295 300

Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
305 310 315 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
325 330 335

Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
340 345 350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg
370 375 380

Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
385 390 395 400

Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
405 410 415

Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe
435 440 445

Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
450 455 460

Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
465 470 475 480

Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
485 490 495

Pro Ser Ser Gln
500

<210> 2
<211> 491
<212> PRT
<213> 人工序列

<220>
<223> mos2Gag 花葉抗原序列

<400> 2

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu
50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln
100 105 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn
115 120 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile
130 135 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala
145 150 155 160

Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala
165 170 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
180 185 190

Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu
195 200 205

Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala Pro Gly Gln
210 215 220

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Asn Leu
225 230 235 240

Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val Gly
245 250 255

Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
260 265 270

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
275 280 285

Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
290 295 300

Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
305 310 315 320

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
325 330 335

Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
340 345 350

Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
355 360 365

Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
370 375 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys
385 390 395 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
420 425 430

Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro
435 440 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
450 455 460

Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu
465 470 475 480

Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln
485 490

<210> 3
<211> 850
<212> PRT
<213> 人工序列

<220>
<223> mos1Pol 花葉抗原序列

<400> 3

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro
1 5 10 15

Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
20 25 30

Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu Lys Glu Gly Lys
35 40 45

Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala
50 55 60

Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg
65 70 75 80

Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile
85 90 95

Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala
100 105 110

Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Gly Phe Arg Lys
115 120 125

Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu Thr Pro Gly Ile
130 135 140

Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala
145 150 155 160

Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro Phe Arg Ala Lys
165 170 175

Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly
180 185 190

Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile Glu Glu Leu Arg
195 200 205

Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln
210 215 220

Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys
225 230 235 240

Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp Ser Trp Thr Val
245 250 255

- Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile
260 265 270

- Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Ala
275 280 285

Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu Ala Glu Leu Glu
290 295 300

Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr
305 310 315 320

Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly His
325 330 335

Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys
340 345 350

Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr Asn Asp Val Lys
355 360 365

Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu Ser Ile Val Ile
370 375 380

Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln Lys Glu Thr Trp
385 390 395 400

- Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp
405 410 415

-

- Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu
420 425 430

-

Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val Ala Gly Ala Ala
435 440 445

Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly
450 455 460

Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn Gln Lys Thr Ala
465 470 475 480

Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Ser Glu Val Asn
485 490 495

Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
500 505 510

Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile
515 520 525

Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro Ala His Lys Gly Ile
530 535 540

Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ser Gly Ile Arg Lys
545 550 555 560

Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys
565 570 575

Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro
580 585 590

Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Gln Cys Gln Leu Lys
595 600 605

Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln
610 615 620

Leu Ala Cys Thr His Leu Glu Gly Lys Ile Ile Leu Val Ala Val His
625 630 635 640

Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly
645 650 655

Gln Glu Thr Ala Tyr Phe Ile Leu Lys Leu Ala Gly Arg Trp Pro Val
660 665 670

Lys Val Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Ala Val
675 680 685

Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln Glu Phe Gly Ile Pro
690 695 700

Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu
705 710 715 720

Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
725 730 735

Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
740 745 750

Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile Asp Ile Ile Ala Thr
755 760 765

Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Ile Lys Ile Gln Asn
770 775 780

Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Ile Trp Lys Gly Pro
785 790 795 800

Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
805 810 815

Ser Asp Ile Lys Val Val Pro Arg Arg Lys Val Lys Ile Ile Lys Asp
820 825 830

Tyr Gly Lys Gln Met Ala Gly Ala Asp Cys Val Ala Gly Arg Gln Asp
835 840 845

Glu Asp
850

<210> 4
<211> 850
<212> PRT
<213> 人工序列

<220>
<223> mos2Pol 花葉抗原序列

<400> 4

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro
1 5 10 15

Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
20 25 30

Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys
35 40 45

Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe Ala
50 55 60

Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg
65 70 75 80

Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile
85 90 95

Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala
100 105 110

Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys
115 120 125

Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile
130 135 140

Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala
145 150 155 160

Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln
165 170 175

Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly
180 185 190

Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg
195 200 205

Gln His Leu Leu Arg Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln
210 215 220

Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys
225 230 235 240

Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val
245 250 255

Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile
260 265 270

Tyr Ala Gly Ile Lys Val Lys Gln Leu Cys Lys Leu Leu Arg Gly Thr
275 280 285

Lys Ala Leu Thr Glu Val Val Pro Leu Thr Glu Glu Ala Glu Leu Glu
290 295 300

Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr
305 310 315 320

Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln
325 330 335

Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys
340 345 350

Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys
355 360 365

Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Thr Glu Ser Ile Val Ile
370 375 380

Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp
385 390 395 400

Glu Ala Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp
405 410 415

Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu
420 425 430

Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala
435 440 445

Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly
450 455 460

Arg Gln Lys Val Val Ser Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala
465 470 475 480

Leu Gln Ala Ile His Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
485 490 495

Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
500 505 510

Asp Lys Ser Glu Ser Glu Leu Val Ser Gln Ile Ile Glu Gln Leu Ile
515 520 525

Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
530 535 540

Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Arg Gly Ile Arg Lys
545 550 555 560

Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys
565 570 575

Tyr His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu Pro Pro
580 585 590

Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys
595 600 605

Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln
610 615 620

Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His
625 630 635 640

Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly
645 650 655

Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val
660 665 670

Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Thr Val
675 680 685

Lys Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro
690 695 700

Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Ile Asn Lys Glu Leu
705 710 715 720

Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
725 730 735

Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
740 745 750

Ile Gly Glu Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Ser
755 760 765

Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn
770 775 780

Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Leu Trp Lys Gly Pro
785 790 795 800

Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
805 810 815

Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
820 825 830

Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp
835 840 845

Glu Asp
850

- <210> 5
- <211> 685
- <212> PRT
- <213> 人工序列

<220>
<223> mos1Env 花葉抗原序列

<400> 5

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Ala Gly Lys
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala Thr Asn Thr
130 135 140

Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile Lys Asn Cys
145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln Lys Gln Tyr
165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asp Ser Asn
180 185 190

Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln
195 200 205

Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
210 215 220

Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly
225 230 235 240

Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile
245 250 255

Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu
260 265 270

Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Ala Lys Thr
275 280 285

Ile Met Val Gln Leu Asn Val Ser Val Glu Ile Asn Cys Thr Arg Pro
290 295 300

Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe
305 310 315 320

Tyr Thr Ala Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn
325 330 335

Ile Ser Arg Ala Asn Trp Asn Asn Thr Leu Arg Gln Ile Val Glu Lys
340 345 350

Leu Gly Lys Gln Phe Gly Asn Asn Lys Thr Ile Val Phe Asn His Ser
355 360 365

Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly
370 375 380

Glu Phe Phe Tyr Cys Asn Ser Thr Lys Leu Phe Asn Ser Thr Trp Thr
385 390 395 400

Trp Asn Asn Ser Thr Trp Asn Asn Thr Lys Arg Ser Asn Asp Thr Glu
405 410 415

Glu His Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp
420 425 430

Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
435 440 445

Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
450 455 460

Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met
465 470 475 480

Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
485 490 495

Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln
500 505 510

Ser Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
515 520 525

Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val
530 535 540

Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
545 550 555 560

Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
565 570 575

Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
580 585 590

Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
595 600 605

Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu
610 615 620

Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile
625 630 635 640

Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn
645 650 655

Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
660 665 670

Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp
675 680 685

<210> 6
<211> 684
<212> PRT
<213> 人工序列

<220>

<223> mos2Env 花葉抗原序列

<400> 6

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe
355 360 365

Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly
385 390 395 400

Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr
405 410 415

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
420 425 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Ser
500 505 510

Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
565 570 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
660 665 670

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
675 680

<210> 7
<211> 708
<212> PRT
<213> 人工序列

<220>
<223> 穩定化的進化枝 C gp140三聚體: 具有切割突變的C97ZA012-gp140-foldon

<400> 7

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn
20 25 30

Leu Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val
35 40 45

Trp Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala
50 55 60

Tyr Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro
65 70 75 80

Thr Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn
85 90 95

Phe Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile
100 105 110

Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro
115 120 125

Leu Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val
130 135 140

Thr Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr
145 150 155 160

Thr Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg
165 170 175

Pro Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser
180 185 190

Glu Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys
195 200 205

Pro Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala
210 215 220

Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly
225 230 235 240

Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro
245 250 255

Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu
260 265 270

Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile
275 280 285

Val His Leu Asn Lys Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn
290 295 300

Asn Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala
305 310 315 320

Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser
325 330 335

Gly Ser Lys Trp Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Gln
340 345 350

Glu Asn Tyr Asn Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly
355 360 365

Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe
370 375 380

Phe Tyr Cys Asn Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asp
385 390 395 400

Glu Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp
405 410 415

Gln Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile
420 425 430

Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Val Arg Asp Gly Gly
435 440 445

Glu Asp Asn Lys Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met
450 455 460

Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Ile Glu Leu
465 470 475 480

Lys Pro Leu Gly Ile Ala Pro Thr Gly Ala Lys Glu Arg Val Val Glu
485 490 495

Arg Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
500 505 510

Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Leu Thr Leu Thr Val
515 520 525

Gln Ala Arg Gln Leu Leu Ser Ser Ile Val Gln Gln Gln Ser Asn Leu
530 535 540

Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp
545 550 555 560

Gly Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu
565 570 575

Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
580 585 590

Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln
595 600 605

Thr Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile
610 615 620

Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr
625 630 635 640

Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys
645 650 655

Asn Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys
660 665 670

Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro
675 680 685

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
690 695 700

Ser Thr Phe Leu
705

<210> 8
<211> 625
<212> PRT
<213> 人工序列

<220>
<223> C4片段：沒有信號肽和跨膜結構域的gp120-截短的 gp41

<400> 8

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr
100 105 110

Tyr Asn Ile Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser
115 120 125

Phe Asn Ala Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala
130 135 140

Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser
145 150 155 160

Glu Lys Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys
165 170 175

Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro
180 185 190

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys
195 200 205

Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr
210 215 220

Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu
225 230 235 240

Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn
245 250 255

Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val
260 265 270

Asn Ile Thr Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg
275 280 285

Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp
290 295 300

Ile Arg Gln Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr
305 310 315 320

Leu Gln Gly Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr
325 330 335

Ile Lys Phe Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His
340 345 350

Thr Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu
355 360 365

Phe Asn Glu Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro
370 375 380

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala
385 390 395 400

Ile Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile
405 410 415

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro
420 425 430

Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn
435 440 445

Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu
450 455 460

Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys
465 470 475 480

Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala
485 490 495

Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg
500 505 510

Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala
515 520 525

Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys
530 535 540

Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln
545 550 555 560

Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr
565 570 575

Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile
580 585 590

Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr
595 600 605

Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu
610 615 620

Lys
625

<210> 9
<211> 29
<212> PRT
<213> 人工序列

<220>
<223> 信號序列

<400> 9

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val
20 25

<210> 10
<211> 29
<212> PRT
<213> 人工序列

<220>
<223> 信號序列

<400> 10

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala
20 25

<210> 11
<211> 29
<212> PRT
<213> 人工序列

<220>
<223> 信號序列

<400> 11

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val
20 25

<210> 12
<211> 29
<212> PRT
<213> 人工序列

<220>
<223> 信號序列

<400> 12

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala
20 25

<210> 13
<211> 22
<212> PRT
<213> 人工序列

<220>
<223> 跨膜結構域

<400> 13

Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe
1 5 10 15

Ala Val Leu Ser Ile Val
20

<210> 14
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 截短的細胞質區域

<400> 14

Asn Arg Val Arg Gln Gly Tyr
1 5

<210> 15
<211> 32
<212> PRT
<213> 人工序列

<220>
<223> GCN4三聚結構域

<400> 15

Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr
1 5 10 15

His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val
20 25 30

<210> 16
<211> 30
<212> PRT
<213> 人工序列

<220>
<223> foldon三聚結構域

<400> 16

Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
1 5 10 15

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
20 25 30

<210> 17
<211> 862
<212> PRT
<213> 人工序列

<220>
<223> C4 序列

<400> 17

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe
355 360 365

- Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn
370 375 380

- Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu
385 390 395 400

Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile
405 410 415

Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala
420 425 430

Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu
435 440 445

Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr
450 455 460

Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser
465 470 475 480

Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala
485 490 495

Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val
500 505 510

. Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr
515 520 525
.

. Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
530 535 540
.

Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala
545 550 555 560

Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
565 570 575

Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu
580 585 590

Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro
595 600 605

Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn
610 615 620

Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu
625 630 635 640

Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu
645 650 655

Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn Trp Phe
660 665 670

Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val
675 680 685

Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val
690 695 700

Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Leu Gln Thr Leu Thr
705 710 715 720

Gln Asn Pro Gly Gly Leu Asp Arg Leu Gly Arg Ile Glu Glu Glu Gly
725 730 735

Gly Glu Gln Asp Lys Asp Arg Ser Ile Arg Leu Val Asn Gly Phe Phe
740 745 750

Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His
755 760 765

Arg Leu Arg Asp Phe Ile Leu Ile Val Ala Arg Ala Val Glu Leu Leu
770 775 780

Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly Trp Glu Ile Leu Lys
785 790 795 800

Tyr Leu Gly Ser Leu Leu Gln Tyr Trp Gly Leu Glu Leu Lys Lys Ser
805 810 815

Ala Ile Asn Leu Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr
820 825 830

Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys Arg Ala Ile Cys Asn
835 840 845

Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala Ala Leu Gln
850 855 860

<210> 18
<211> 711
<212> PRT
<213> 人工序列

<220>
<223> C4D7 序列

<400> 18

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe
355 360 365

Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu
385 390 395 400

Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile
405 410 415

Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala
420 425 430

Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu
435 440 445

Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr
450 455 460

Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser
465 470 475 480

Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala
485 490 495

Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val
500 505 510

Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr
515 520 525

Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
530 535 540

Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala
545 550 555 560

Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
565 570 575

Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu
580 585 590

Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro
595 600 605

Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn
610 615 620

Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu
625 630 635 640

Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu
645 650 655

Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn Trp Phe
660 665 670

Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val
675 680 685

Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val
690 695 700

Asn Arg Val Arg Gln Gly Tyr
705 710

<210> 19
<211> 704
<212> PRT
<213> 人工序列

<220>
<223> sC4 序列

<400> 19

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys
35 40 45

• Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

• His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

• Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

• Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe
355 360 365

Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu
385 390 395 400

Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile
405 410 415

Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala
420 425 430

Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu
435 440 445

Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr
450 455 460

Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser
465 470 475 480

Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala
485 490 495

Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Glu Arg Ala Val
500 505 510

Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr
515 520 525

Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
530 535 540

Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala
545 550 555 560

Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
565 570 575

Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu
580 585 590

Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro
595 600 605

Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn
610 615 620

Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu
625 630 635 640

Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Met Lys
645 650 655

Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr His Ile
660 665 670

Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val Gly Ser
675 680 685

Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
690 695 700

<210> 20
<211> 4050
<212> DNA
<213> 人工序列

<220>
<223> 編碼mos1GagPol的核苷酸序列

<400> 20
atgggagcca gagccagcgt gctgtccgga ggggagctgg accgctggga gaagatcagg 60
ctgaggcctg gagggaagaa gaagtacagg ctgaagcaca tcgtgtgggc cagcagagag 120
ctggaacggt ttgccgtgaa ccttggcctg ctggaaacca gcgagggtg taggcagatt 180
ctgggacagc tgcagcccag cctgcagaca ggcagcgagg aactgcggag cctgtacaac 240
accgtggcca cctgtactg cgtgcaccag cggatcgaga tcaaggacac caaagaagcc 300
ctggaaaaga tcgaggaaga gcagaacaag agcaagaaga aagcccagca ggctgccgct 360
gacacaggca acagcagcca ggtgtcccag aactaccca tcgtgcagaa catccaggga 420
cagatggtgc accaggccat cagccctcgg accctgaacg cctgggtgaa ggtggtggag 480
gaaaaggcct tcagccctga ggtgatcccc atgttctctg cctgagcga gggagccaca 540
ccccaggacc tgaacaccat gctgaacacc gtgggagggc accaggctgc catgcagatg 600
ctgaaagaga caatcaacga ggaagctgcc gagtgggaca ggggtccacc agtgcacgct 660
ggacctatcg ctcttgacca gatgagagag ccagaggca gcgatattgc tggcaccacc 720
tccacactgc aggaacagat cggtgggatg accaacaacc ctcccatccc tgtgggagag 780
atctacaagc ggtggatcat tctgggactg aacaagatcg tgcggatgta cagccctgtg 840
agcatcctgg acatcaggca gggacccaaa gagcccttca gggactacgt ggaccggttc 900
tacaagacc tgagagccga gcaggccagc caggacgtga agaactggat gaccgagaca 960
ctgctggtgc agaacgcaa cctgactgc aagaccatcc tgaagccct gggacctgct 1020
gccaccctgg aagagatgat gacagcctgc cagggagtgg gaggacctgg ccacaaggcc 1080
agggtgctgg ccgaggccat gagccagtg accaactctg ccaccatcat gatgcagaga 1140
ggcaacttcc ggaaccagag aaagaccgtg aagtgcttca actgtggcaa agaggacac 1200
attgccaaga actgcagggc tcccaggaag aaaggctgct ggaagtgcgg aaaagaaggc 1260
caccagatga aggactgcac cgagaggcag gccaaattcc tgggcaagat ctggcctagc 1320
aacaagggca ggcctggcaa ctctctgcag aacagaccgc agcccaccgc tctcccag 1380
gaaagcttcc ggtttgcgga ggaaccacc acccctagcc agaagcagga acccatcgac 1440
aaagagatgt accctctggc cagcctgaag agcctgttcg gcaacgaccc cagcagccag 1500
atggctccca tcagcccaat cgagacagt cctgtgaagc tgaagcctgg catggacgga 1560
cccagggtga agcagtggcc tctgaccgag gaaaagatca aagccctgac agccatctgc 1620
gaggaaatgg aaaaagagg caagatcacc aagatcggac ccgagaaccc ctacaacacc 1680
cctgtgttcg ccatcaagaa gaaagacagc accaagtga ggaactggt ggaactcaga 1740

gagctgaaca agcggacceca ggactttctgg gaggtgcagc tgggcatccc tcaccctgct	1800
ggcctgaaga aaaagaaaag cgtgaccgtg ctggctgttg gagatgccta cttcagcgtg	1860
cctctggagc agggcttccg gaagtacaca gccttcacca tccccagcac caacaacgag	1920
acacctggca tcagatacca gtacaacgtg ctgcctcagg gctggaaagg cagccctgcc	1980
atcttccagt gcagcatgac cagaatcctg gaacccttca gagccaagaa ccttgagatc	2040
gtgatctacc agtatatggc tgccctctac gtgggcagcg acctggaaat cggacagcac	2100
agagccaaaa tcgaagaact ccgcgagcac ctgctgaagt ggggattcac caccctgac	2160
aagaagcacc agaaagagcc tcccttccct tggatgggct acgagctgca ccctgacaag	2220
tggaccgtgc agcccatcca gctgccagag aaggactcct ggaccgtgaa cgacatccag	2280
aaactggtcg gcaagctgaa ctgggccagc cagatctacc ctggcatcaa agtcagacag	2340
ctgtgtaagc tgctgagggg agccaaagca ctgaccgaca tcgtgcctct gacagaagaa	2400
gccgagctgg aactggccga gaacagagag atcctgaaag aaccctgca cggagtgtac	2460
tacgaccct ccaaggacct gattgccgag atccagaaac agggacacga ccagtggacc	2520
taccagatct atcaggaacc tttcaagaac ctgaaaacag gcaagtacgc caagatgcgg	2580
acagcccaca ccaacgacgt gaagcagctg accgaagccg tgcagaaaat cgccatggaa	2640
agcatcgtga tctggggaaa gacaccaag ttcaggctgc ccatccagaa agagacatgg	2700
gaaacctggt ggaccgacta ctggcaggcc acctggattc ccgagtggga gttcgtgaac	2760
accccacccc tggatgaagct gtggtatcag ctggaaaagg accctatcgc tggcgtggag	2820
acattctacg tggctggagc tgccaacaga gagacaaagc tgggcaaggc tggctacgtg	2880
accgacagag gcagacagaa aatcgtgagc ctgaccgaaa ccaccaacca gaaaacagcc	2940
ctgcaggcca tctatctggc actgcaggac agcgggaagcg aggtgaacat cgtgacagcc	3000
agccagtatg ccttgggcat catccaggcc cagcctgaca agagcgagag cgagctggtg	3060
aaccagatca tcgagcagct gatcaagaaa gaacgggtgt acctgagctg ggtgccagcc	3120
cacaagggca tcggagggaa cgagcaggtg gacaagctgg tgtccagcgg aatccggaag	3180
gtgctgttcc tggacggcat cgataaagcc caggaagagc acgagaagta ccacagcaat	3240
tggagagcca tggccagcga cttcaacctg cctcccgtgg tggccaaaga aatcgtggcc	3300
agctgcgacc agtgccagct gaaaggcgag gccatgcacg gacaggtgga ctgctccct	3360
ggcatctggc agctggcatg ccccacctg gaaggcaaga tcattctggt ggccgtgcac	3420
gtggccagcg gatacatga agccgaagt atccctgccg agacagggca ggaacagcc	3480
tacttcatcc tgaagctggc tggcagatgg cctgtgaagg tgatccacac agccaacggc	3540
agcaacttca cctctgtgc cgtgaaggct gcctgttggg gggctggcat tcagcaggaa	3600
tttggcatcc cctacaatcc ccagtctcag ggagtgttgg ccagcatgaa caaagagctg	3660
aagaagatca tcggacaggt cagggatcag gccgagcacc tgaaaactgc cgtccagatg	3720
gccgtgttca tccacaactt caagcggaag ggagggatcg gaggtactc tgctggcgag	3780
cggatcatcg acatcattgc caccgatatc cagaccaaag agctgcagaa acagatcatc	3840
aagatccaga acttcagggt gtactacagg gacagcaggg accccatctg gaaggacct	3900
gccaagctgc tgtggaagg cgaaggagcc gtcgtcatcc aggacaacag cgacatcaag	3960
gtggtgccc gacggaaggt gaaaatcatc aaggactacg gcaaacagat ggctggagcc	4020
gactgtgtcg ctggcaggca ggacgaggac	4050

<210> 21
<211> 4023
<212> DNA
<213> 人工序列

<220>
<223> 編碼mos2GagPol的核苷酸序列

<400> 21	
atgggagcca gagccagcat cctgcgagga gggaaagctgg acaagtggga gaagatcagg	60
ctgaggcctg gagggaagaa acactacatg ctgaagcacc tggctctgggc cagcagagag	120
ctggaacggt ttgccctcaa tcctggcctg ctggaaacca gcgagggctg caagcagatc	180
atcaagcagc tgcagcctgc cctgcagaca ggcaccgagg aactgcggag cctgttcaac	240
accgtggcca ccctgtactg cgtgcatgcc gagatcgaag tgagggacac caaagaagcc	300
ctggacaaga tcgaggaaga gcagaacaag agccagcaga aaaccagca ggccaaagaa	360
gccgacggca aggtctccca gaactacccc atcgtgcaga acctgcaggg acagatggtg	420
caccagccca tcagccctcg gacactgaat gcctgggtga aggtgatcga ggaaaaggcc	480
ttcagccctg aggtgatccc catgttcaca gccctgagcg agggagccac accccaggac	540
ctgaacacca tgctgaacac cgtgggaggg caccaggctg ccatgcagat gctgaaggac	600
accatcaacg aggaagctgc cgagtgggac aggctgcacc ctgtgcacgc tggacctgtg	660
gtcctctggc agatgagaga gccagaggc agcgatatg ctggcaccac ctccaatctg	720
caggaacaga tcgcctggat gaccagcaac cctcccatcc ctgtgggaga catctacaag	780
cgggtggatca tcctgggact gaacaagatc gtgcggatgt acagccctac ctccatcctg	840
gacatcaagc agggacccaa agagcctttc agggactacg tggaccggtt cttcaagacc	900
ctgagagccg agcaggccac ccaggacgtg aagaactgga tgaccgacac cctgctggtg	960
cagaacgcca accctgactg caagaccatc ctgagagccc tgggacctgg agccaccctg	1020
gaagagatga tgacagcctg ccagggagtg ggaggaccct ctcacaaggc tagggtgctg	1080
gccgaggcca tgagccagac caacagcacc atcctgatgc agcggagcaa cttcaagggc	1140
agcaagcgga tcgtgaagtg cttcaactgt ggcaaagagg gacacattgc cagaaactgt	1200
agggcaccca ggaagaaagg ctgctggaag tgcggaaaag aaggccacca gatgaaggac	1260
tgcaccgaga ggcaggccaa cttcctgggc aagatctggc ctagccacaa gggcagacct	1320
ggcaacttcc tgcagagcag acccgagccc accgctcctc cagccgagag cttccggttc	1380
gaggaacca cccctgctcc caagcaggaa cctaaggaca gagagcctct gaccagcctg	1440
agaagcctgt tcggcagcga ccctctgagc cagatggctc ccatctcccc tatcgagaca	1500
gtgctgtga agctgaagcc tggcatggac ggaccaagtg tgaacagtg gcctctgacc	1560
gaggaaga tcaaagccct ggtggagatc tgtaccgaga tggaaaaaga gggcaagatc	1620
agcaagatcg gaccgagaa cccctacaac acccctatct tcgccatcaa gaagaaagac	1680
agcaccaagt ggaggaaact ggtggacttc agagagctga acaagcggac ccaggacttc	1740
tgggaggtgc agctgggcat ccctcaccct gctggcctga agaaaaagaa aagcgtgacc	1800
gtgctggccg tgggagatgc ctacttcagc gtgcctctgg acgaggactt cagaaagtac	1860
acagccttca ccatccccag catcaacaac gagacacctg gcacagata ccagtacaac	1920
gtgctgcctc agggatggaa gggtctcct gcaatcttcc agagcagcat gaccaagatc	1980
ctggaacctt tccggaagca gaacctgac atcgtgatct accagtacat ggcagccctg	2040
tacgtcggca gcgacctgga aatcggacag caccggacca agatcgaaga actcaggcag	2100
cacctgctgc ggtggggatt caccaccct gacaagaagc accagaaaga gcctcccttc	2160
ctgtggatgg gctacgagct gcaccagac aagtggaccg tgcagcccat cgtgctgcct	2220
gagaaggact cctggaccgt gaacgacatc cagaaactgg tcggcaagct gaactgggcc	2280
agccagatct acgtggcat caaagtgaag cagctgtgta agctcctgag aggcaccaaa	2340
gccctgaccg aggtggtgcc actgacagag gaagccgagc tggaactggc cgagaacaga	2400

gagatcctga aagaacccgt gcacggagtg tactacgacc ccagcaagga cctgattgcc	2460
gagatccaga agcagggaca gggacagtgg acctaccaga tctaccagga acccttcaag	2520
aacctgaaaa caggcaagta cgccaggatg aggggagccc acaccaacga cgtcaaacag	2580
ctgaccgaag ccgctgcagaa gatcgccacc gagagcatcg tgatttgggg aaagacaccc	2640
aagttcaagc tgcccatcca gaaagagaca tgggaggcct ggtggaccga gtactggcag	2700
gccacctgga ttcccagatg ggagttcgtg aacaccccac ccctggtgaa gctgtggtat	2760
cagctggaaa aagaacccat cgtgggagcc gagacattct acgtggctgg agctgccaac	2820
agagagacaa agctgggcaa ggctggctac gtgaccgaca gaggcaggca gaaagtgggtg	2880
tccttgaccg ataccaccaa ccagaaaaca gccctgcagg ccatccacct ggctctgcag	2940
gactctggcc tggaagtgaa catcgtgaca gccagccagt atgccctggg catcattcag	3000
gcacagcctg acaagagcga gagcgagctg gtgtctcaga tcattgagca gctgatcaag	3060
aaagaaaagg tgtacctggc ctgggtgccca gccacaagg ggatcggagg gaacgagcag	3120
gtggacaagc tgggtgtccag gggcatccgg aaggtgctgt ttctggacgg catcgacaaa	3180
gcccaggaag agcacgagaa gtaccacagc aattggagag ccatggccag cgagttcaac	3240
ctgcctccca tcgtggccaa agaatcgtg gcctcttgcg acaagtgcc a gctgaaaggc	3300
gaggccattc acggacaggt ggactgcagc ccaggcatct ggcagctggc ctgcacccac	3360
ctggaaggca aggtgatect ggtggccgtg cacgtggcct ctggatacat cgaagccgaa	3420
gtgatccctg ccgagacagg ccaggaaaca gcctacttcc tgctgaagct ggctggcagg	3480
tggcctgtga aaaccatcca cacagccaac ggcagcaact tcacctctgc caccgtgaag	3540
gctgcctgtt ggtgggctgg cattaagcag gaatttggca tcccctacaa cctcagtt	3600
caggagtggtg tggcctccat caacaaagag ctgaagaaga tcatcggaca ggtcagggat	3660
caggccgagc atctgaaaac agccgtccag atggccgtgt tcattccaaa cttcaagcgg	3720
aaggagggga tcggagagta ctctgctggc gagaggatcg tggacattat cgccagcgat	3780
atccagacca aagaactgca gaagcagatc acaaagatcc agaacttcag ggtgtactac	3840
agggacagca gagatcccct gtggaaggga cctgccaaagc tgctgtggaa aggcgaagga	3900
gccgtcgtca tccaggacaa cagcgacatc aaggtgggtg ccagacggaa ggccaagatc	3960
atcagagact acggcaaaca gatggctggc gacgactgcg tcgcctctag gcaggacgag	4020
gac	4023

- <210> 22
- <211> 2055
- <212> DNA
- <213> 人工序列

- <220>
- <223> 編碼mos1Env的核苷酸序列

<400> 22	
atcggggtga ccggcatccg gaagaactac cagcacctgt ggcgggtgggg caccatgctg	60
ctgggcatcc tgatgatttg ctctgccgcc ggaaagctgt gggtgaccgt gtactacggc	120
gtgccctgtt ggaaagaggc caccaccacc ctgtttctgcg ccagcgacgc caaggcctac	180
gacaccgagg tgcacaacgt gtgggccacc cagcctgcg tgcccaccga cccaacccc	240
caggaagtgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag	360
ctgaccccc tgtagctgac cctgaactgc accgacgacg tgcggaacgt gaccaacaac	420
gccaccaaca ccaacagcag ctggggcgag cctatggaaa agggcgagat caagaactgc	480
agcttcaaca tcaccacctc catccggaac aaggtgcaga agcagtagc cctgtttctac	540

aagctggacg tgggtgccc	atcgacacgac agcaacaaca ccaactaccg gctgatcagc	600
tgcaacacca gcgatgatc	acaggcctgc cccaaggtgt ccttcgagcc catccccatc	660
cactactgcg cccctgccg	cttcgccatc ctgaagtgca acgacaagaa gttcaacggc	720
accggcccct gcaccaacg	gagcaccgtg cagtgacccc acggcatccg gcccggtgtg	780
tccaccacgc tgctgctga	acggcagcctg gccgaggaag aggtggtgat cagaagcgag	840
aatttcacca acaatgcc	aa gaccatcatg gtgcagctga acgtgagcgt ggagatcaac	900
tgccccggc ccaacaaca	acacccggaag agcatccaca tcggccctgg cagggccttc	960
tacacagccg gcgacatc	atcgcgacatc cggcaggccc actgcaacat cagccggggc	1020
aactggaaca acacctgc	gcgagatcgtg gagaagctgg gcaagcagtt cggcaacaac	1080
aagaccatcg tgttcaacc	acagcagcggc ggagaccccg agatcgtgat gcacagcttc	1140
aactgtggcg gcgagttc	tt ctactgcaac agcaccaagc tgttcaacag cacctggacc	1200
tggaacaact ccacctgg	aa taacaccaag cggagcaacg acaccgaaga gcacatcacc	1260
ctgccctgcc ggatcaag	ca gattatcaat atgtggcagg aggtcggcaa ggccatgtac	1320
gccccctcca tccggggc	ca gatccggtgc agcagcaaca tcaccggcct gctgctgacc	1380
cgggacggcg gcaacgata	cacgcggcacc gagatcttcc ggcctggcgg cgagatatg	1440
cgggacaact ggcgagcga	gctgtacaag tacaaggtgg tgaagatcga gccctgggc	1500
gtggctccca ccaaggcca	aa gcggcgggtg gtgcagagcg agaagagcgc cgtgggcac	1560
ggcgccgtgt ttctgggc	tt cctgggagcc gccggaagca ccatgggagc cgccagcatg	1620
accctgaccg tgcaggccc	g gctgctgctg tccggcatcg tgcagcagca gaacaacctg	1680
ctccgggcca tcgaggcca	ac gcagcacctg ctgcagctga ccgtgtgggg catcaagcag	1740
ctgcaggcca ggggtgctg	gc cgtggagaga tacctgaagg atcagcagct cctggggatc	1800
tggggctgca gcggcaag	ct gatctgcacc accaccgtgc cctggaacgc cagctggtcc	1860
aacaagagcc tggacaaga	t ctggaacaat atgacctgga tggaatggga gcgcgagatc	1920
aacaattaca ccagcctga	t ctacacctg atcgaggaaa gccagaacca gcaggaaaag	1980
aacgagcagg aactgctgg	a actggacaag tgggccagcc tgtggaactg gttcgacatc	2040
agcaactggc tgtgg		2055

- <210> 23
- <211> 2052
- <212> DNA
- <213> 人工序列
- <220>
- <223> 編碼mos2Env的核苷酸序列

<400> 23																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						</
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

cccatcccca tccactactg cgcccctgcc ggctacgcc tccctgaagtg caacaacaag 720
accttcaacg gcaccggccc ctgcaacaac gtgagcaccg tgcagtgcac ccacggcatc 780
aagcccgtgg tgtccacca gctgctgctg aacggcagcc tggccgagga agagatcatc 840
atccgggtccg agaacctgac caacaacgcc aagaccatca tcgtgcacct gaatgagacc 900
gtgaacatca cctgcacccg gcccacaac aacacccgga agagcatccg gatcggccct 960
ggccagacct tttagccac cggcgacatc atcggcgaca tccggcaggc ccactgcaac 1020
ctgagccggg acggctggaa caagacctg cagggcgta agaagaagct ggccgagcac 1080
ttccccaata agaccatcaa ctaccagc agcagcggcg gagacctgga aatcaccacc 1140
cacagcttca actgcagggg cgagttcttc tactgcaata cctccggcct gttcaatggc 1200
acctacatgc ccaacggcac caacagcaac agcagcagca acatcacctt gccctgccgg 1260
atcaagcaga tcatcaatat gtggcaggag gtcggcaggg ccatgtacgc cctcccatc 1320
gccggcaata tcacctgccg gtccaacatc accggcctgc tgcctgaccag ggacggcggc 1380
agcaacaacg gcgtgcctaa cgacaccgag accttccggc ctggcggcgg agatatgcgg 1440
aacaactggc ggagcgagct gtacaagtac aaggtggttg aggtgaagcc cctgggcgtg 1500
gctcctaccg aggccaagcg gcgggtggtg gagagcgaga agagcgccgt gggcatcggc 1560
gccgtgtttc tgggcattct gggagccgcc ggaagcacca tgggagccgc cagcatcacc 1620
ctgaccgtgc aggcccggca gctgctgtcc ggcctcgtgc agcagcagag caacctgctg 1680
agagccatcg aggcccagca gcacatgctg cagctgaccg tctggggcat caagcagctg 1740
cagacccggg tgctggccat cgagagatac ctgcaggatc agcagctcct gggcctgtgg 1800
ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacaccag ctggtccaac 1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agagatcggc 1920
aactacaccg gcgagatcta caggctgctg gaagagagcc agaaccagca ggaaaagaac 1980
gagaaggacc tgctggccct ggacagctgg aagaacctgt ggaactggtt cgacatcacc 2040
aactggctgt gg 2052

<210> 24
<211> 829
<212> DNA
<213> 人工序列

<220>
<223> 用於表現Ad26載體中的抗原的CMV啟動子

<400> 24
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
ttggccattg catacgttgt atccatatca taatatgtac atttatattg gctcatgtcc 120
aacattaccg ccatgttgac attgattatt gactagttat taatagtaat caattacggg 180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240
gcctggctga ccgcccacg acccccgcc attgacgtca ataatgacgt atgttcccat 300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaactgc 360
ccacttggca gtacatcaag tgtatcatat gccaaagtag cccctatttg acgtcaatga 420
cggtaaatgg ccgcctggc attatgccca gtacatgacc ttaatgggact ttctacttg 480
gcagtacatc tacgtattag tcacgctat taccatggtg atgcggtttt ggcagtacat 540
caatggggct ggatagcgtt ttgactcacg gggatttcca agtctccacc ccattgacgt 600
caatggggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc 660
cgccccattg acgcaaatgg gcggtaggcg tgtacggttg gaggtctata taagcagagc 720
tcgttttagtg aaccgtcaga tcgcctggag acgcatcca cgctgttttg acctccatag 780

aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg tgcattgga 829

- <210> 25
- <211> 2586
- <212> DNA
- <213> 人工序列

<220>
<223> 編碼C4的核苷酸序列

<400> 25
atgagagtgc ggggcatgct gagaaactgg cagcagtggt ggatctggtc cagcctgggc 60
ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120
gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac 180
gagaaagagg tgcacaacgt ctgggccacc cagcctgttg tggccaccga ccccaatccc 240
caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtc 300
gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgcgtgaag 360
ctgacccttc tgtgcgtgac cctgaactgc cggaacgtgc gcaacgtgtc cagcaacggc 420
acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480
accaccgtgg tcgaggaccg gaagcagaag gtgcaegccc tgtttetaccg gctggacatc 540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac 600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac 660
cctatcccca tccactactg cgccccctgc ggctacgcca tcctgaagtg caacaacaag 720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780
aagcccgtgg tgtctaccca gctgctgctg aacggcagcc tggccgagga agagatcatt 840
atcagaagcg agaacctgac caacaacgcc aaaacatca tcgtccacct gaacgaaacc 900
gtgaacatca cctgtacccg gcctaacaac aacacccgga agtccatccg gatcggccct 960
ggccagacct tttacgccac cgcgatatt atcggcgaca tccggcaggc ccaactgcaat 1020
ctgagccggg acggttgaa caagacactg cagggcgta agaagaagct ggccgaacac 1080
ttccctaaca agactatcaa gttcgccct cactctggcg gcgacctgga aatcaccacc 1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag 1200
agcaacatcg agcggaacga cagcatcatc acactgcctt gccggatcaa gcagattatc 1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca 1320
tgccggtcca atatcaccgg cctgtgtgtc accagagatg gcggtccaa caatggcgtg 1380
ccaaacgaca ccgagacatt cagacccggc ggaggcgaca tgcggaacaa ttggcggagc 1440
gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcgtggcccc taccgaggcc 1500
aagagaagag tggtcgaacg cgagaagcgg gccgtgggaa tcggagccgt gtttctggga 1560
atcctggggag ccgtggctc taccatgggc gctgcctcta tcacctgac agtgcaggcc 1620
agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgetgagagc cattgaggcc 1680
cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agtccagac acgggtgctg 1740
gccatcgaga gatactgca ggatcagcag ctcttgggcc tgtggggctg tagcggcaag 1800
ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac 1860
atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag 1920
atctatagac tgctggaaga gtcccagaac cagcaggaaa agaacgagaa ggacctgctg 1980
gccctggatt cttggaacaa tctgtggaac tggttcagca tctccaagtg gctgtgttac 2040
atcaagatct tcatcatgat cgtgggcggc ctgatcgcc tgcggatcat ctttgccgtg 2100
ctgagcatcg tgaaccgcgt gcggcaggga tacagccctc tgaacctgca gacctgact 2160

cagaaccctg gcggaactgga cagactgggc cggattgagg aagaaggcgg cgagcaggac 2220
aaggatcggg gcatcaggct gtcaacggc ttcttcgctc tgttttggga cgacctgcgg 2280
agcctgtgcc tgttcagcta ccacagactg cgggacttta tcctgattgt ggccagagcc 2340
gtcgaactgc tggggagaag ctctctgaga ggccctgcagc ggggctggga gattctgaag 2400
tacctgggct ccctgctgca gtactggggc ctggaactga agaagtctgc catcaatctg 2460
ctcgacacaa tcgctattgc cgtggccgaa ggcaccgata gaatcatcga gctgatccag 2520
cggatctgcc gggccatctg caacatcccc agacggatca gacagggtt cgaggccgct 2580
ctgcag 2586

<210> 26
<211> 2133
<212> DNA
<213> 人工序列

<220>
<223> 編碼C4D7的核苷酸序列

<400> 26
atgagagtgc ggggcatgct gagaaactgg cagcagtggt ggatctggtc cagcctgggc 60
ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120
gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac 180
gagaaagagg tgcacaacgt ctgggccacc cagcctgtg tgcccaccga cccaatccc 240
caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtc 300
gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgcgtgaag 360
ctgacccctc tgtgcgtgac cctgaactgc cggaacgtgc gcaacgtgc cagcaacggc 420
acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc 540
gtgcccctgg acgagaacaa cagcagcgag aagtccctcg agaacagctc cgagtactac 600
agactgatca actgcaacac cagcgccatc acccaggcct gccccaaagt gtccttcgac 660
cctatcccca tccactactg cgccccctgc ggctacgcca tcctgaagtg caacaacaag 720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780
aagcccgtgg tgtctaccca gctgctgctg aacggcagcc tggccgagga agagatcatt 840
atcagaagcg agaacctgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc 900
gtgaacatca cctgtacccg gcctaacaac aacacccgga agtccatccg gatcggccct 960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc cactgcaat 1020
ctgagccggg acggctggaa caagacactg cagggcgtca agaagaagct ggccgaacac 1080
ttccctaaca agactatcaa gttcgccct cactctggcg gcgacctgga aatcaccacc 1140
cacacctta actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag 1200
agcaacatcg agcggaacga cagcatcatc aactgcctt gccggatcaa gcagattatc 1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccg caacatcaca 1320
tgccggtcca atatcaccg cctgctgctc accagagatg gcggctccaa caatggcgtg 1380
ccaaacgaca ccgagacatt cagaccggc ggaggcgaca tgcggaacaa ttggcggagc 1440
gagctgtaca agtacaaggt ggtggaagtg aagccccctg gcgtggcccc taccgaggcc 1500
aagagaagag tggtcgaacg cgagaagcgg gccgtgggaa tcggagccgt gtttctggga 1560
atcctgggag ccgctggctc taccatgggc gctgcctcta tcaccctgac agtgcaggcc 1620
agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc 1680
cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agctccagac acgggtgctg 1740

gccatcgaga gataacctgca ggatcagcag ctctctgggcc tgtggggctg tagcggcaag 1800
ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac 1860
atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag 1920
atctatagac tgctggaaga gtcccagaac cagcaggaaa agaacgagaa ggacctgctg 1980
gccctggatt ctiggaacaa tctgtggaac tggttcagca tctccaagtg gctgtggtac 2040
atcaagatct tcatcatgat cgtgggcggc ctgatcggcc tgcggatcat ctttgccgtg 2100
ctgagcatcg tgaaccgcgt gcggcagggc tac 2133

<210> 27
<211> 2112
<212> DNA
<213> 人工序列

<220>
<223> 編碼sC4的核苷酸序列

<400> 27
atgagagtgc ggggcatgct gagaaactgg cagcagtggg ggatctggtc cagcctgggc 60
ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcacctg gtactacggc 120
gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac 180
gagaaagagg tgcacaacgt ctgggccacc cagcctgtg tggccaccga cccaatccc 240
caggaaatcg tcttgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtc 300
gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgctgaag 360
ctgacccttc tgtgctgac cctgaactgc cggaacgtgc gcaacgtgtc cagcaacggc 420
acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgt 480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc 540
gtgcccctgg acgagaacaa cagcagcgag aagtctctcg agaacagctc cgagtactac 600
agactgatca actgcaacac cagcgccatc acccaggcct gccccaaagt gtccttcgac 660
cctatcccca tccactactg cggccctgcc ggctacgcca tctgaagtg caacaacaag 720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatt 840
atcagaagcg agaacctgac caacaacgcc aaaaccatca tegtccacct gaacgaaacc 900
gtgaacatca cctgtaccog gcctaacaac aacacccgga agtccatccg gatcggccct 960
ggccagacct tttacgccac cggcgatatt atcggcgaca tccggcaggc cacttgaat 1020
ctgagccggg acggttgga caagacactg cagggcgta agaagaagct ggccgaacac 1080
ttccctaaca agactatcaa gtgcggccct cactctggcg gcgacctgga aatcaccacc 1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag 1200
agcaacatcg agcggaacga cagcatcatc aactgcctt gccggatcaa gcagattatc 1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca 1320
tgccggtcca atatcaccgg cctgctgctc accagagatg gcggctcaa caatggcgtg 1380
ccaaacgaca ccgagacatt cagaccggc ggaggcgaca tgcggaacaa ttggcggagc 1440
gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcgtggcccc taccgaggcc 1500
aagagaagag tggtcgaacg cgaggaacgg gccgtgggaa tcggagccgt gtttctggga 1560
atcctgggag ccgctggctc taccatgggc gctgcctcta tcacctgac agtgcaggcc 1620
agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc 1680
cagcagcaca tgctgcagct gaccgtgttg ggcattaagc agctccagac acgggtgctg 1740
gccatcgaga gataacctgca ggatcagcag ctctctgggcc tgtggggctg tagcggcaag 1800

ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac1860

atctggggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag1920

atctatagac tgctggaaga gtcccagaac cagcaggaaa agatgaagca gatcgaggac1980

aagatcgaag agattctgag caagatctac cacatcgaga acgagatcgc ccgcatcaag2040

aaactgatcg gcgaagtggg atccggcgct cccacaaagg ccaaaaagacg ggtggtgcag2100

cgcgagaaac gc2112

<210> 28

<211> 1350

<212> PRT

<213> 人工序列

<220>

<223> mos1GagPol 花葉抗原序列

<400> 28

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp11015

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys202530

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro354045

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu505560

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn65707580

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp859095

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys100105110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val115120125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His130135140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu145150155160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser165170175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly180185190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu195200205

Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala210215220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr225230235240

Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
245 250 255

Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
275 280 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
290 295 300

Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
305 310 315 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
325 330 335

Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
340 345 350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg
370 375 380

Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
385 390 395 400

Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
405 410 415

Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe
435 440 445

Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
450 455 460

Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
465 470 475 480

Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
485 490 495

Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val
500 505 510

Lys Leu Lys Pro Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu
515 520 525

Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu
530 535 540

Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr
545 550 555 560

Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu
565 570 575

Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val
580 585 590

Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val
595 600 605

Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu
610 615 620

Gly Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu
625 630 635 640

Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys
645 650 655

Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro
660 665 670

Phe Arg Ala Lys Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala
675 680 685

Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile
690 695 700

Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp
705 710 715 720

Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu
725 730 735

His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp
740 745 750

Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp
755 760 765

Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu
770 775 780

Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu
785 790 795 800

Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val
805 810 815

His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln
820 825 830

Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe
835 840 845

Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr
850 855 860

Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu
865 870 875 880

Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln
885 890 895

Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp
900 905 910

Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp
915 920 925

Tyr Gln Leu Glu Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val
930 935 940

Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val
945 950 955 960

Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn
965 970 975

Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly
980 985 990

Ser Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile
995 1000 1005

Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile
1010 1015 1020

Ile Glu Gln Leu Ile Lys Lys Glu Arg Val Tyr Leu Ser Trp Val
1025 1030 1035

Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu
1040 1045 1050

Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
1055 1060 1065

Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
1070 1075 1080

Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile
1085 1090 1095

Val Ala Ser Cys Asp Gln Cys Gln Leu Lys Gly Glu Ala Met His
1100 1105 1110

Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr
1115 1120 1125

His Leu Glu Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser
1130 1135 1140

Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu
1145 1150 1155

Thr Ala Tyr Phe Ile Leu Lys Leu Ala Gly Arg Trp Pro Val Lys
1160 1165 1170

Val Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Ala Val
1175 1180 1185

Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln Glu Phe Gly Ile
1190 1195 1200

Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys
1205 1210 1215

Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His
1220 1225 1230

Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys
1235 1240 1245

Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile
1250 1255 1260

Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln
1265 1270 1275

Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg
1280 1285 1290

Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
1295 1300 1305

Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
1310 1315 1320

Arg Arg Lys Val Lys Ile Ile Lys Asp Tyr Gly Lys Gln Met Ala
1325 1330 1335

Gly Ala Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
1340 1345 1350

<210> 29
<211> 1341
<212> PRT
<213> 人工序列

<220>
<223> mos2GagPol 花葉抗原序列

<400> 29

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu
50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln
100 105 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn
115 120 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile
130 135 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala
145 150 155 160

Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala
165 170 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
180 185 190

Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu
195 200 205

Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala Pro Gly Gln
210 215 220

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Asn Leu
225 230 235 240

Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val Gly
245 250 255

Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
260 265 270

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
275 280 285

Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
290 295 300

Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
305 310 315 320

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
325 330 335

Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
340 345 350

Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
355 360 365

Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
370 375 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys
385 390 395 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
420 425 430

Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro
435 440 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
450 455 460

Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu
465 470 475 480

Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln Met Ala Pro Ile Ser
485 490 495

Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro
500 505 510

Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val
515 520 525

Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly
530 535 540

Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp
545 550 555 560

Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg
565 570 575

Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly
580 585 590

Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr
595 600 605

Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr
610 615 620

Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn
625 630 635 640

Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser
645 650 655

Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val
660 665 670

Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile
675 680 685

Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg
690 695 700

Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe
705 710 715 720

Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro
725 730 735

Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys
740 745 750

Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys
755 760 765

Val Lys Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu
770 775 780

Val Val Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg
785 790 795 800

Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys
805 810 815

Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr
820 825 830

Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala
835 840 845

Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala
850 855 860

Val Gln Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro
865 870 875 880

Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr
885 890 895

Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
900 905 910

Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val
915 920 925

Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys
930 935 940

Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val
945 950 955 960

Ser Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile His
965 970 975

Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser
980 985 990

Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser
995 1000 1005

Glu Leu Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys
1010 1015 1020

Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn
1025 1030 1035

Glu Gln Val Asp Lys Leu Val Ser Arg Gly Ile Arg Lys Val Leu
1040 1045 1050

Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys Tyr
1055 1060 1065

His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu Pro Pro
1070 1075 1080

Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu
1085 1090 1095

Lys Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser Pro Gly Ile
1100 1105 1110

Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val
1115 1120 1125

Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro
1130 1135 1140

Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala
1145 1150 1155

Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn
1160 1165 1170

Phe Thr Ser Ala Thr Val Lys Ala Ala Cys Trp Trp Ala Gly Ile
1175 1180 1185

Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
1190 1200

Val Ala Ser Ile Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
1205 1210 1215

Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val
1220 1225 1230

Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Glu Tyr Ser
1235 1240 1245

Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Ser Asp Ile Gln Thr
1250 1255 1260

Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val
1265 1270 1275

Tyr Tyr Arg Asp Ser Arg Asp Pro Leu Trp Lys Gly Pro Ala Lys
1280 1285 1290

Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser
1295 1300 1305

Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
1310 1315 1320

Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln
1325 1330 1335

Asp Glu Asp
1340

<210> 30
<211> 709
<212> PRT
<213> 人工序列

<220>
<223> sC1 序列

<400> 30

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe
355 360 365

Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly
385 390 395 400

Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr
405 410 415

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
420 425 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
500 505 510

Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
565 570 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
645 650 655

Gln Glu Lys Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser
660 665 670

Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile
675 680 685

Gly Glu Val Gly Ser Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val
690 695 700

Gln Arg Glu Lys Arg
705

<210> 31
<211> 867
<212> PRT
<213> 人工序列

<220>
<223> C1 序列

<400> 31

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

- Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp
100 105 110

- Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
165 170 175

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser
180 185 190

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
195 200 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225 230 235 240

. Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
245 250 255

. Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
260 265 270

- Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe
355 360 365

Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly
385 390 395 400

Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr
405 410 415

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
420 425 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
500 505 510

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
565 570 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
660 665 670

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
675 680 685

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
500 505 510

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
565 570 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
660 665 670

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
675 680 685

Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala
690 695 700

Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr
705 710 715

<210> 33
<211> 2127
<212> DNA
<213> 人工序列

<220>
<223> 編碼 sC1的核苷酸序列

<400> 33
atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc 60
ttttgtagta tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120

gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac	180
gagaaagagg tgcacaacgt ctgggccacc cagcctgtg tgcccaccga cccaatccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgctgaag	360
ctgacccctc tgtgctgtac cctggaatgc cggaacgtgc gcaacgtgtc cagcaacggc	420
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgt	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtccctcg agaacagctc cgagtactac	600
agactgatca actgcaaac cagcgccatc acccaggcct gcccgaaggt gtccttcgac	660
cctatcccca tccactactg cgccttgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctaccca gctgctgtg aacggcagcc tggccgagga agagatcatc	840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtaccg gcctaacaac aacacccgga agtccatccg gatcgccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccactgcaat	1020
ctgagccggg acggctggaa caagacactg caggcggtca agaagaagct ggccgaacac	1080
ttccccaaca aaacatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc	1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga	1200
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccgg	1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc cctccaatc	1320
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc	1380
tctaacaatg gcgtgccaaa cgacaccgag acattcagac ccggcggagg cgacatgcgg	1440
aacaattggc ggagcgagct gtacaagtac aagggtgtgg aagtgaagcc cctgggctg	1500
gccctaccg aagccaagag aagagtggtc gaacgcgagg aacgggccgt gggcattgga	1560
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tgggcgctgc ctctatcaca	1620
ctgacagtgc aggccagaca gtcctgagc ggcacgtgc agcagcagag caacctgctg	1680
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc	1740
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg	1800
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac	1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc	1920
aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagatg	1980
aagcagatcg aggacaagat cgaagagatt ctgagcaaaa tctaccacat cgagaacgag	2040
atcgcccgca tcaagaaact gatcggcgaa gtgggatccg gcgctccac aaaggccaaa	2100
agacgggtgg tgcagcgca gaaacgc	2127

<210> 34
<211> 2601
<212> DNA
<213> 人工序列

<220>
<223> 編碼 C1的核苷酸序列

<400> 34	
atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc	60
ttttgtagta tcattatctg ccgctgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccggtg ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac	180

gagaaagagg tgcacaacgt ctggggccacc cacgcctgtg tgcccaccga cccaatccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgttgaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgctgaag	360
ctgaccctc tgtgcgtgac cctggaatgc cggaaactgc gcaactgtc cagcaacggc	420
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccctogac	660
cctatcccca tccactactg cggccctgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctacca getgtctgtg aacggcagcc tggccgagga agagatcatc	840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaacc	900
gtgaacatca cctgtaccog gcctaacaac aacacccgga agtccatccg gatcgccct	960
ggccagacct tttacgccac cgcgatatt atcggcgaca tccggcaggc cactgcaat	1020
ctgagccggg acggctggaa caagacactg cagggcgtca agaagaagct ggccgaacac	1080
ttccccaaca aaacctcaa cttaccagc tcctctggcg gcgacctgga aatcaccacc	1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga	1200
acctatagc ccaacgggac caacagcaac tccagcagca atatcacct gccttgccgg	1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacg ccctccaatc	1320
gcccgaaca tcacatgcag aagcaacatt accggcctgc tgctcaccg ggacggcggc	1380
tctaacaatg gcgtgcaaaa cgacaccgag acattcagac cggcgaggagg cgacatcgg	1440
aacaattggc ggagcgagct gtacaagtac aagggtgtgg aagtgaagcc cctgggcgtg	1500
gcccctaccg aagccaagag aagagtgtc gaacgcgaga agcgggccgt gggcattgga	1560
gccgtgtttc tgggaatcct gggagccgct ggcagacca tgggcgctgc ctctatcaca	1620
ctgacagtgc aggccagaca gtcctgagc ggcacgtgc agcagcagag caacctgtg	1680
agagccatcg aggcacagca gcacatgtg cagctgaccg tgtggggcat taagcagctc	1740
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg	1800
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac	1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc	1920
aactataccg gcgagatcta ccgactgtg gaagagtccc agaaccagca ggaaaagaac	1980
gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattggtt cgacatcacc	2040
aactggctgt ggtacatcaa gatcttcac atgatcgtgg gcggcctgat cggcctgcgg	2100
atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggatacag ccctctgagc	2160
ctgcagacc tgaccagaa tccaggcgga ctggatcggc tgggcgggat tgaggaagaa	2220
ggcggcgagc aggacaagga ccgcagcatc agactcgtga acggcttctt cgctctgttt	2280
tgggacgacc tcgggagcct gtgcctgttc tcctaccaca gactcgggga ctttatcctg	2340
attgtggcca gagccgtcga gctgctgggc agatcttctc tgagaggcct gcagcggggc	2400
tgggagattc tgaagtacct gggctccctg ctgcagtatt ggggcctgga actgaagaag	2460
tcgccatca atctgtcga cacaatcgct attgccgtgg ccgaaggcac cgacagaatc	2520
atcgagctga tccagcgat ctgccgggcc atctgcaaca tcccagacg gatcagacag	2580
ggctttgaag ccgcctcca g	2601

<210> 35
<211> 2148
<212> DNA
<213> 人工序列

<220>
<223> 編碼C1D7的核苷酸序列

<400> 35	
atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc	60
ttttggatga tcattatctg ccgctgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac	180
gagaaagagg tgcacaacgt ctgggccacc cacgcctgtg tgcccaccga cccaatccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgctgaag	360
ctgacccctc tgtgctgac cctggaatgc cggaacgtgc gcaacgtgtc cagcaacggc	420
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaaac cagcgccatc acccaggcct gcccgaaggt gtccttcgac	660
cctatcccca tccactactg cggccctgcc ggctacgcca tctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctaccca gctgctgctg aacggcagcc tggccgagga agagatcatc	840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtaccgg gcctaacaac aacacccgga agtccatccg gatcgccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccactgcaat	1020
ctgagccggg acggctggaa caagacactg cagggcgtca agaagaagct ggccgaacac	1080
ttcccaaca aaaccatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc	1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga	1200
acctacatgc ccaacgggac caacagcaac tccagcagca atatcacct gccttgccgg	1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc ccctccaatc	1320
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc	1380
tctaacaatg gcgtgcaaaa cgacaccgag acattcagac ccggcggagg cgacatgcgg	1440
aacaattggc ggagcgagct gtacaagtac aagggtggtg aagtgaagcc cctgggcgtg	1500
gcccctaccg aagccaagag aagagtggtc gaacgcgaga agcgggccgt gggcattgga	1560
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tgggcgtgc ctctatcaca	1620
ctgacagtgc aggccagaca gtcctgagc ggcacgtgc agcagcagag caacctgctg	1680
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc	1740
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg	1800
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac	1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc	1920
aactatacgg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagaac	1980
gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattggtt cgacatcacc	2040
aactggctgt ggtacatcaa gatcttcac atgatcgtgg gcggcctgat cggcctgcgg	2100
atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggctac	2148

<210> 36
<211> 724

<212> PRT
<213> 人工序列

<220>
<223> 花葉Env 三聚體序列

<400> 36
Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Gly Lys
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala Thr Asn Thr
130 135 140

Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile Lys Asn Cys
145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln Lys Gln Tyr
165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asp Ser Asn
180 185 190

Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln
195 200 205

Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
210 215 220

Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly
225 230 235 240

Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile
245 250 255

Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu
260 265 270

Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Ala Lys Thr
275 280 285

Ile Met Val Gln Leu Asn Val Ser Val Glu Ile Asn Cys Thr Arg Pro
290 295 300

Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe
305 310 315 320

Tyr Thr Ala Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn
325 330 335

Ile Ser Arg Ala Asn Trp Asn Asn Thr Leu Arg Gln Ile Val Glu Lys
340 345 350

Leu Gly Lys Gln Phe Gly Asn Asn Lys Thr Ile Val Phe Asn His Ser
355 360 365

Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly
370 375 380

Glu Phe Phe Tyr Cys Asn Ser Thr Lys Leu Phe Asn Ser Thr Trp Thr
385 390 395 400

Trp Asn Asn Ser Thr Trp Asn Asn Thr Lys Arg Ser Asn Asp Thr Glu
405 410 415

Glu His Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp
420 425 430

Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
435 440 445

Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
450 455 460

Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met
465 470 475 480

Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
485 490 495

Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Glu Arg Val Val Gln
500 505 510

Arg Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
515 520 525

Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val
530 535 540

Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
545 550 555 560

Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
565 570 575

Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
580 585 590

Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
595 600 605

Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu
610 615 620

Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile
625 630 635 640

Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn
645 650 655

Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
660 665 670

Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys
675 680 685

Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro
690 695 700

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
705 710 715 720

Ser Thr Phe Leu

<210> 37
<211> 28
<212> PRT
<213> 人工序列

<220>
<223> SEQ ID NO:18的殘基655-682

<400> 37

Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn
1 5 10 15

Trp Phe Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys
20 25

申請專利範圍

1. 一種編碼合成的HIV包膜蛋白之核酸，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及 (iii) EK479-480RRRR、I529P、A471C和T575C的組合。
2. 如申請專利範圍第1項所述之核酸，其中該合成的HIV包膜蛋白進一步包含一信號序列，例如包含選自由以下各項組成之群組的胺基酸序列之信號序列：SEQ ID NO: 9至SEQ ID NO: 12。
3. 如申請專利範圍第1或2項所述之核酸，其中該合成的HIV包膜蛋白進一步包含一跨膜結構域，例如包含SEQ ID NO: 13的跨膜結構域，較佳的是該合成的HIV包膜蛋白進一步包含融合到SEQ ID NO: 8的C-末端和該跨膜結構域的N-末端的SEQ ID NO: 37。
4. 如申請專利範圍第3項所述之核酸，其中該合成的HIV包膜蛋白進一步包含一細胞質結構域之片段，較佳的是包含該SEQ ID NO: 14的胺基酸序列或其殘基1-4的細胞質結構域之片段。
5. 如申請專利範圍第4項所述之核酸，其中該合成的HIV包膜蛋白包含該SEQ ID NO: 18的胺基酸序列。
6. 如申請專利範圍第1或2項所述之核酸，其中該合成的HIV包膜蛋白：(a) 進一步包含一個三聚結構域，例如具有SEQ ID NO: 15或SEQ ID NO: 16的胺基酸序列的三聚結構域，亦或 (b) 包含具有突變EK479-480RRRR、I529P、A471C和T575C的組合的SEQ ID NO: 8。
7. 如申請專利範圍第6項所述之核酸，其中該合成的HIV包膜蛋白包含SEQ

ID NO: 19的殘基1-686之胺基酸序列。

8. 一種包含如申請專利範圍第1-7項中任一項所述之核酸之載體，其中該核酸可操作地連接至一啟動子序列。
9. 如申請專利範圍第8項所述之載體，其係一病毒載體，較佳的是腺病毒載體。
10. 如申請專利範圍第9項所述之載體，其中該腺病毒載體係人類腺病毒血清型26（Ad26）載體。
11. 一種分離的細胞，該分離的細胞包含如申請專利範圍第8-10項中任一項所述之載體。
12. 一種組成物，該組成物包含免疫原性有效量的如申請專利範圍第8-10項中任一項所述之載體、以及運載體。
13. 一種疫苗組成物，該疫苗組成物包含：
 - (i) 第一組成物，該第一組成物包含免疫原性有效量的腺病毒載體，較佳的是腺病毒26載體，該腺病毒載體編碼具有SEQ ID NO: 8的、較佳的是SEQ ID NO: 18的胺基酸序列之合成的HIV包膜蛋白；
 - (ii) 第二組成物，該第二組成物包含免疫原性有效量的第二腺病毒載體，較佳的是第二腺病毒26載體，該第二腺病毒載體編碼包含SEQ ID NO: 5的胺基酸序列之HIV抗原性多肽；以及視情況
 - (iii) 至少一種另外的組成物，該至少一種另外的組成物包含免疫有效量的選自由以下各項組成之群組的至少一種
 - (iiia) 載體，該載體編碼具有選自由SEQ ID NO: 1-4、28和29組成之群組的胺基酸序列之抗原性多肽，以及

(iiib) 包含免疫原性有效量的分離的HIV抗原性多肽之多肽，該分離的HIV抗原性多肽具有SEQ ID NO: 7的胺基酸序列的殘基30-708或SEQ ID NO: 36的殘基30-724，

其中該第一組成物、第二組成物和另外的組成物存在於同一組成物或一種或多種不同組成物中。

14. 一種如申請專利範圍第12項所述之組成物或如申請專利範圍第13項所述之疫苗組合的用途，其係用於製備誘導針對人類免疫缺陷病毒（HIV）之免疫應答的藥物。
15. 一種合成的HIV包膜蛋白，該合成的HIV包膜蛋白包含SEQ ID NO: 8的、或具有一種或多種選自下組的突變的SEQ ID NO: 8的胺基酸序列，該組由以下各項組成：(i) I529P，(ii) K480E，以及 (iii) EK479-480RRRR、I529P、A471C和T575C的組合，較佳的是，該合成的HIV包膜蛋白包含(i) 具有突變EK479-480RRRR、I529P、A471C和T575C的組合的SEQ ID No:8；(ii) SEQ ID NO: 18的胺基酸殘基30-704，或 (iii) SEQ ID NO: 19的胺基酸殘基30-686。

圖式

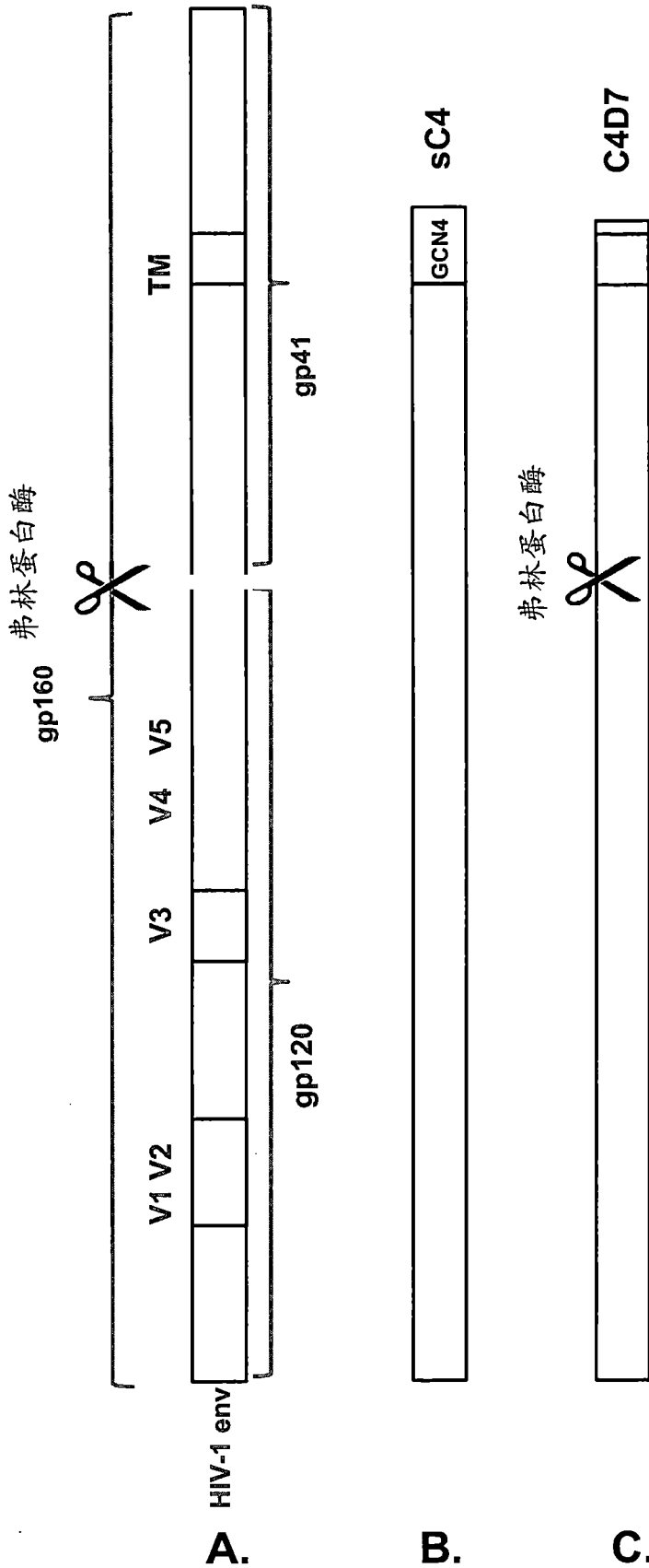


圖 1

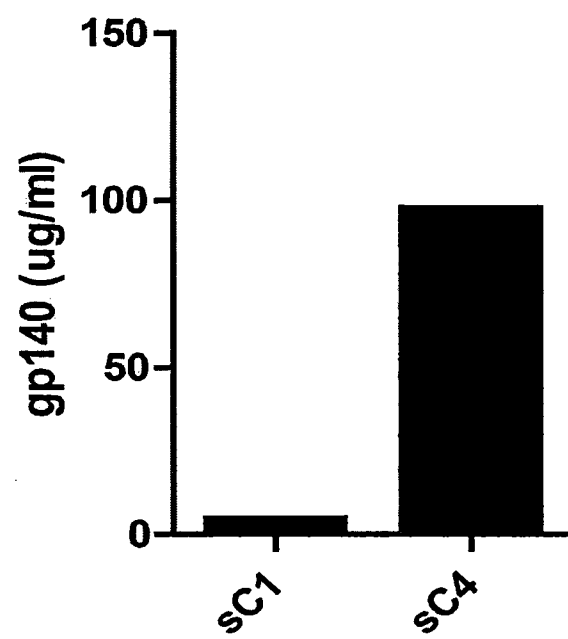


圖 2

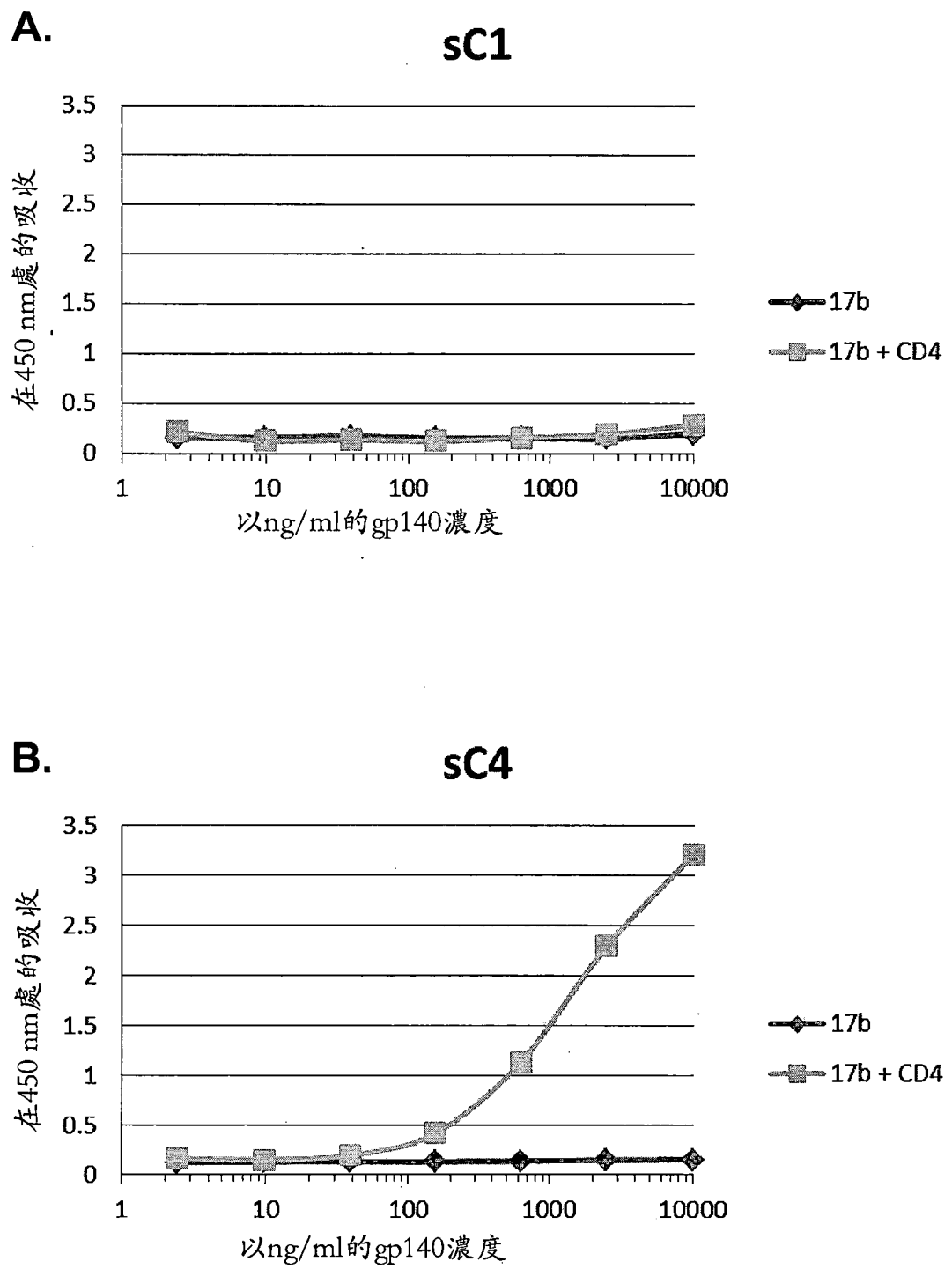


圖 3

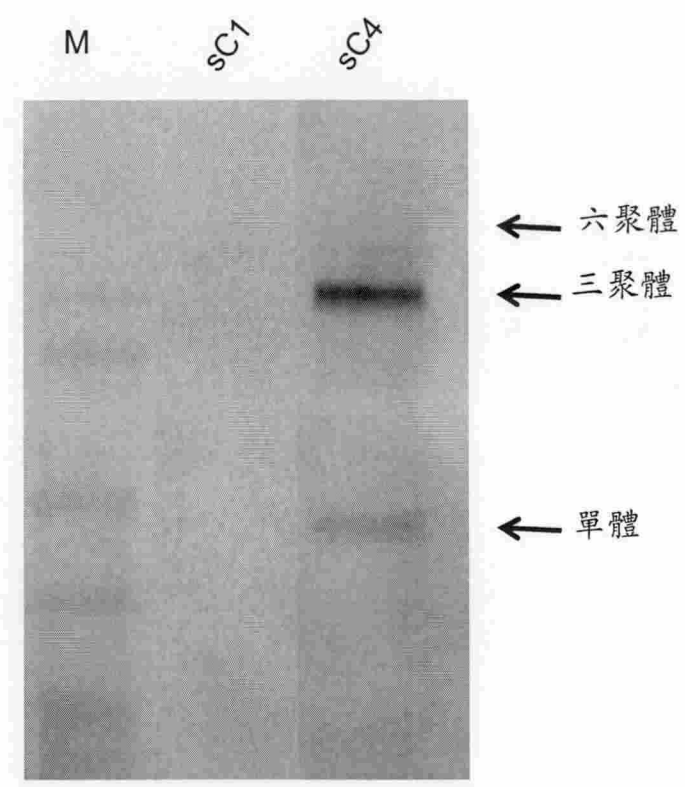


圖 4

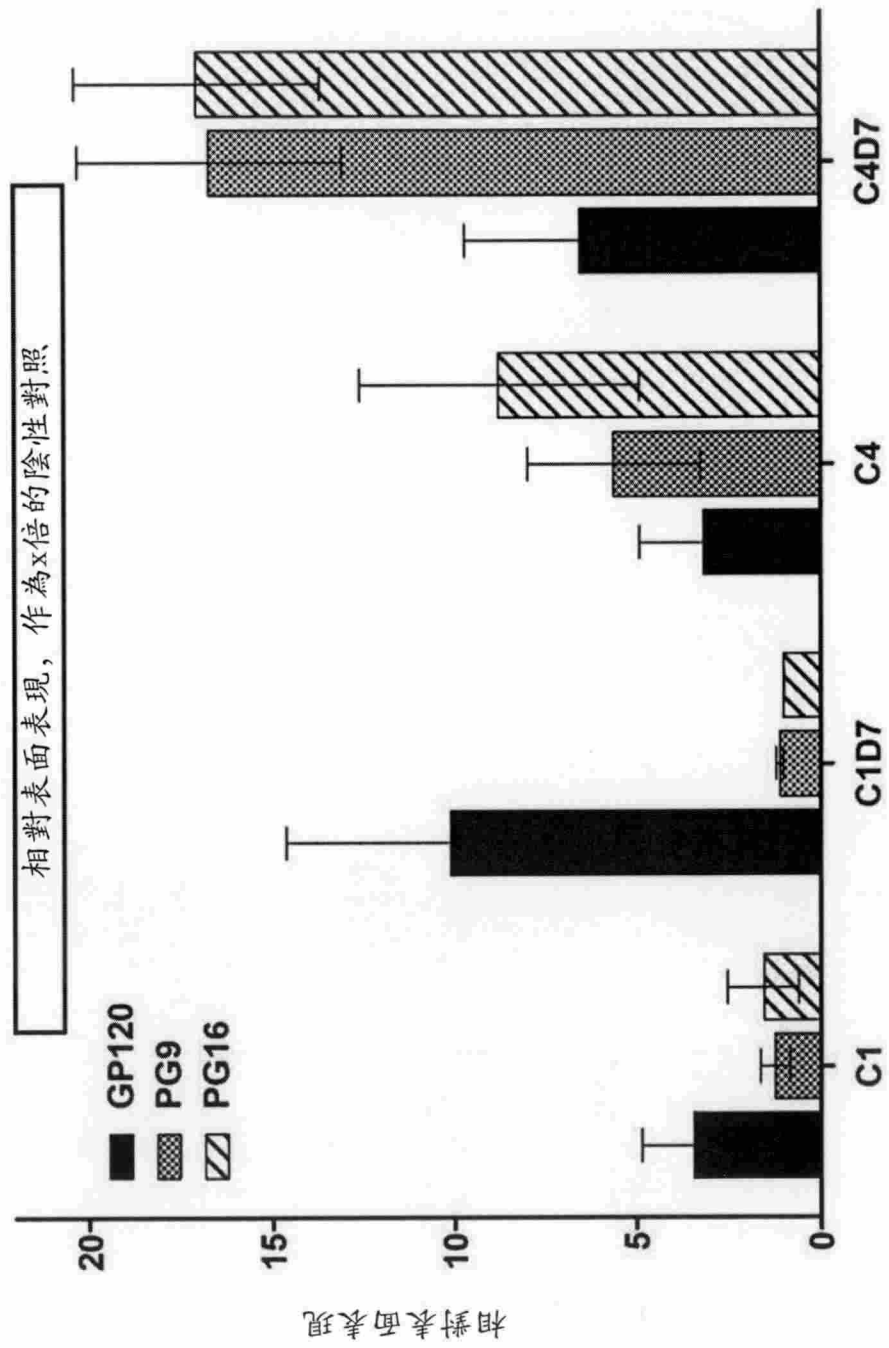


圖 5

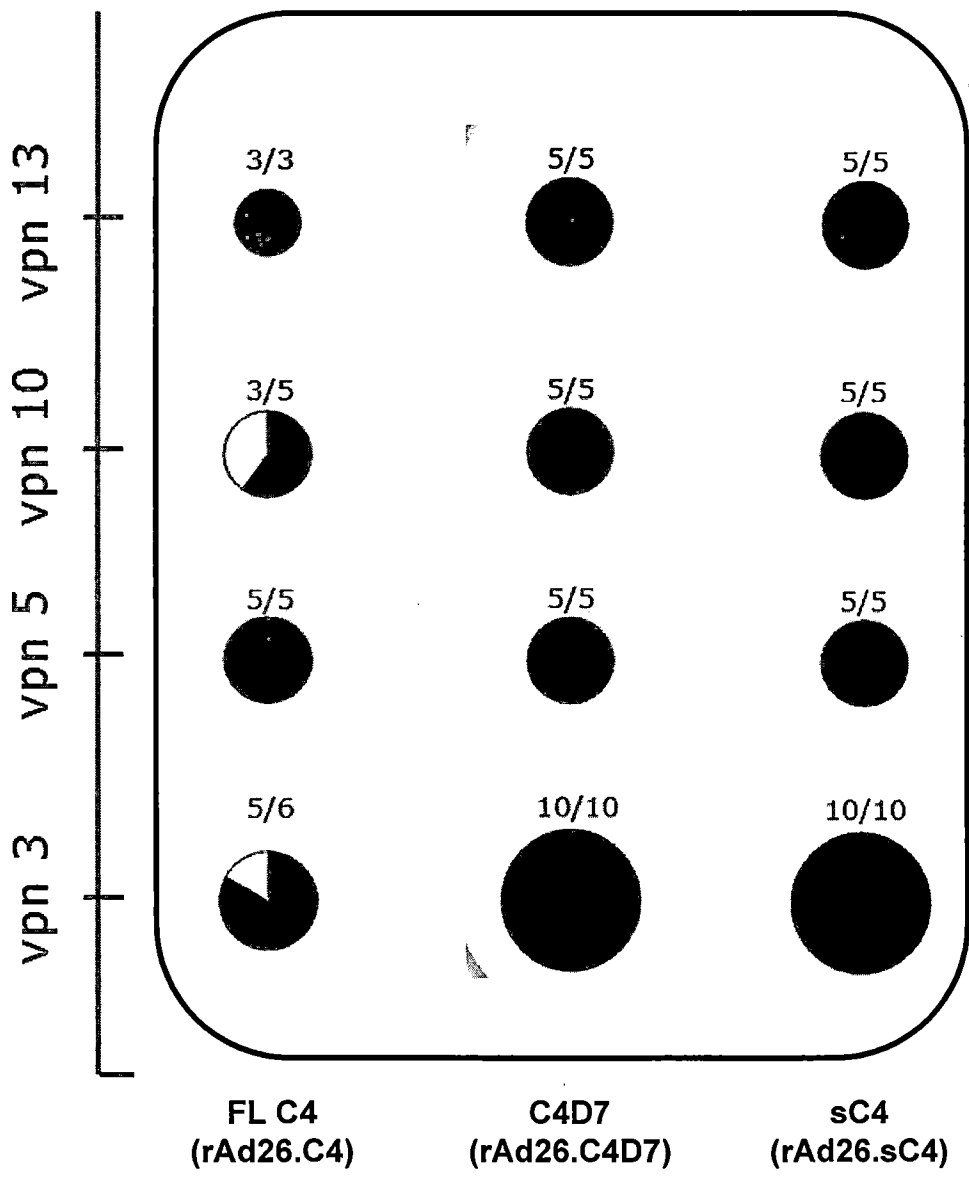


圖 6

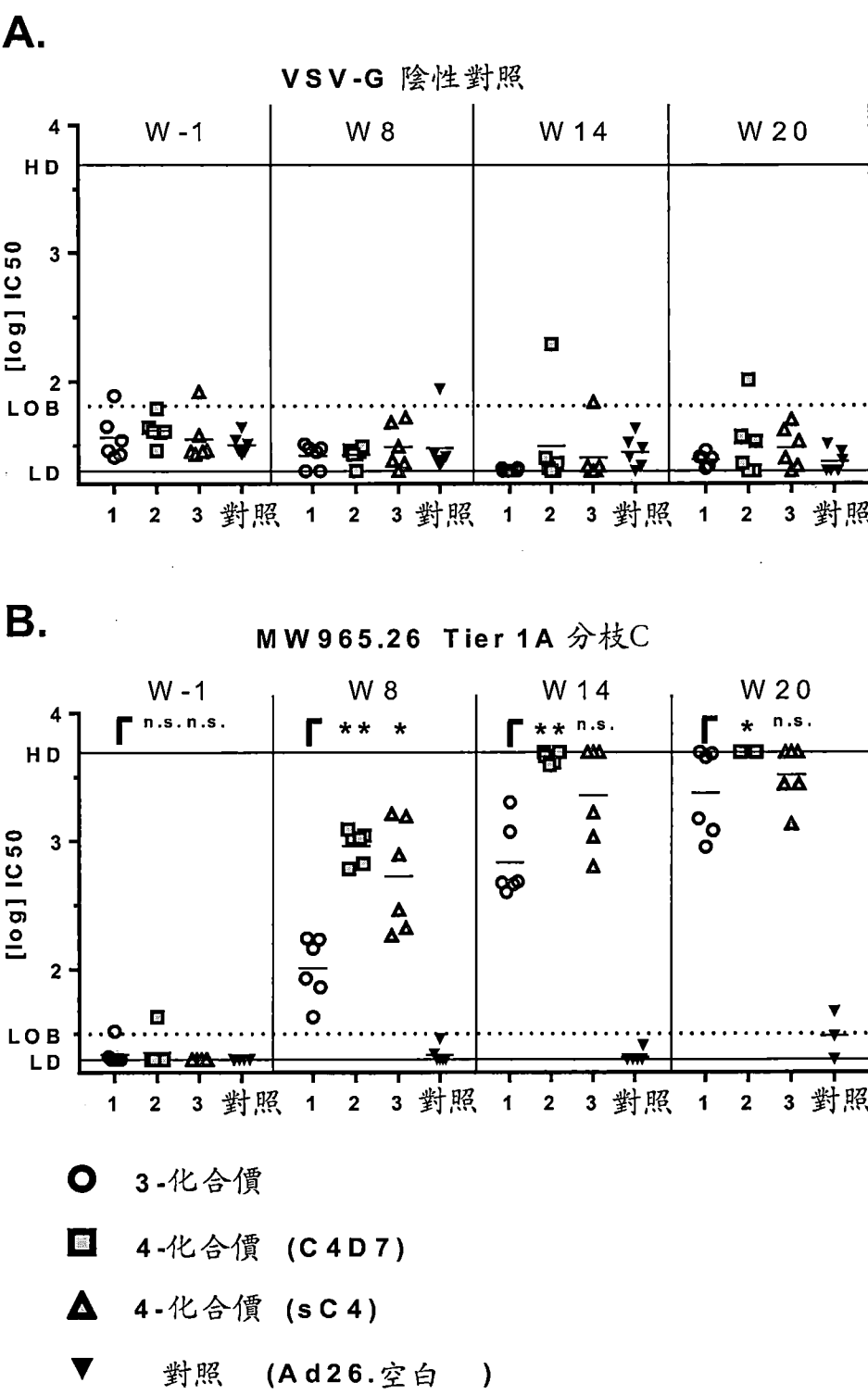


圖 7