

公告本

發明專利說明書

(填寫本書件時請先行詳閱申請書後之申請須知，作※記號部分請勿填寫)

※ 申請案號： 91121625 ※IPC 分類： C07K14/16

※ 申請日期： 91-09-20

壹、發明名稱

(中文) 疫苗

(英文) VACCINES

貳、發明人 (共 6 人)

發明人 1 (如發明人超過一人，請填說明書發明人續頁)

姓名：(中文) 安德魯 畢藤

(英文) ANDREW BEATON

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

參、申請人 (共 1 人)

申請人 1 (如申請人超過一人，請填說明書申請人續頁)

姓名或名稱：(中文) 英商葛蘭素集團有限公司

(英文) GLAXO GROUP LIMITED

住居所或營業所地址：(中文) 英國米德賽克斯郡格林福德市柏克力大道葛蘭素大廈

(英文) GLAXO WELLCOME HOUSE, BERKELEY
AVENUE, GREENFORD, MIDDLESEX, UB6
ONN, UNITED KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

代表人：(中文) 彼得 約漢 吉第絲 (英文) PETER JOHN GIDDINGS

發明人 2

姓名：(中文) 彼德 法蘭茲 爾托

(英文) PETER FRANZ ERTL

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克
藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

發明人 3

姓名：(中文) 傑拉 韋恩 古

(英文) GERALD WAYNE GOUGH

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克
藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

發明人 4

姓名：(中文) 安德魯 李爾

(英文) ANDREW LEAR

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克
藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

發明人 5

姓名：(中文) 約翰 菲利 台特

(英文) JOHN PHILIP TITE

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克
藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

發明人 6

姓名：(中文) 凱瑟琳 安 梵 威利

(英文) CATHERINE ANN VAN WELY

住居所地址：(中文) 英國赫佛雪郡史帝芬艾吉市岡諾林路葛蘭素史密斯克
藍大廈

(英文) GLAXOSMITHKLINE, GUNNELS WOOD ROAD,
STEVENAGE, HERTFORDSHIRE, SG1 2NY, UNITED
KINGDOM

國籍：(中文) 英國

(英文) UNITED KINGDOM

捌、聲明事項

☐ 本案係符合專利法第二十條第一項 ☐ 第一款但書或 ☐ 第二款但書規定之期間，其日期為：_____

☒ 本案已向下列國家（地區）申請專利，申請日期及案號資料如下：

【格式請依：申請國家（地區）；申請日期；申請案號 順序註記】

1. 專利合作條約；2001 年 9 月 20 日；PCT/GB01/04207

2. 英國；2002 年 3 月 19 日；0206462.4

3. 英國；2001 年 12 月 11 日；0129604.5

☒ 主張專利法第二十四條第一項優先權：

【格式請依：受理國家（地區）；日期；案號 順序註記】

1. 英國；2002 年 3 月 19 日；0206462.4

2. 英國；2001 年 12 月 11 日；0129604.5

3. _____

4. _____

5. _____

6. _____

7. _____

8. _____

9. _____

10. _____

☐ 主張專利法第二十五條之一第一項優先權：

【格式請依：申請日；申請案號 順序註記】

1. _____

2. _____

3. _____

☐ 主張專利法第二十六條微生物：

☐ 國內微生物 【格式請依：寄存機構；日期；號碼 順序註記】

1. _____

2. _____

3. _____

☐ 國外微生物 【格式請依：寄存國名；機構；日期；號碼 順序註記】

1. _____

2. _____

3. _____

☐ 熟習該項技術者易於獲得，不須寄存。

(1)

玖、發明說明

(發明說明應敘明：發明所屬之技術領域、先前技術、內容、實施方式及圖式簡單說明)

發明領域

本發明係有關於核酸結構、含有該結構之宿主細胞、及其於核酸疫苗中之用途。本發明進一步有關於包含該結構之疫苗調配物及該調配物之醫療用途。特定言之，本發明係有關於可用於預防及治療HIV感染之DNA疫苗，更特定言之係經由顆粒遞送投予該疫苗。

發明背景

HIV-1係後天免疫不全症候(AIDS)之主要原因，其被視為世界重大健康問題之一。雖然全球已進行有關製作疫苗之大規模研究，但其努力至今仍未成功。

HIV-1之非-被膜蛋白質已經為人所描述過，其中包含如內結構蛋白質，如gag及pol基因產物與其它非結合蛋白質，如Rev、Nef、Vif及Tat(葛林(Green)等人，New England J. Med, 324, 5, 308頁以下(1991)及布萊恩(Bryant)等人(皮索(Pizzo)編)，Pediatr. Infect. Dis. J., 11, 5, 390頁以下(1992))。

Gag基因係由未刪節的RNA轉譯而產生前驅多蛋白質，其隨後分裂為3-5個衣殼(capsid)蛋白質、基質蛋白質、衣殼蛋白質及核酸接合蛋白質與蛋白酶。(1. Fundamental Virology, 菲爾茲(Fields) BN、奈普(Knipe)DM及郝雷(Howley) M 1996 2. Fields Virology, 第2卷, 1996)。

gag基因可產生55-千道耳吞(kD)Gag前驅蛋白質，亦稱為p55，其係由未接合病毒mRNA所表現。於轉譯時，p55

之N端會被十四醯基化，而觸發其結合細胞膜之細胞質方位。膜-關聯Gag多蛋白質召集二組病毒基因組RNA及其它病毒及細胞蛋白質，並引發病毒顆粒由受感染細胞表面出芽。於出芽形成後，在病毒成熟過程中，p55由病毒編碼蛋白酶(pol基因之產物)分裂為四種較小蛋白質，其為MA(基質[p17])、CA(衣殼[p24])、NC(病毒粒子[p9])、及p6。(4)。

除3種主要Gag蛋白質外，所有Gag前驅物皆包含數個其它部位，其可分裂出來以各種大小之肽型式存於病毒粒子中。這些蛋白質具有不同作用，如p2蛋白質可調節蛋白酶活性並可幫助矯正蛋白質水解步驟之時機。

MA多肽係源自N-端、p55之十四醯基化末端。大部分MA分子仍保持附著於病毒粒子雙層脂質之內面，以穩定顆粒。一次組之MA聚集於病毒粒子深層內側，成為可護送病毒DNA至細胞核(5)之複合物的一部分。這些MA分子可促進病毒基因組之細胞核運輸，因細胞核輸入機制可識別MA之親核訊號。此現象使HIV可感染未分裂細胞，這對反轉錄病毒來說並不尋常。

p24(CA)蛋白質構成病毒顆粒之圓椎形核心。環飛林(Cyclophilin) A已證明可與p55之p24部位交互作用並導致其併入HIV顆粒中。介於Gag及環飛林A間之交互作用相當重要，因瓦解環司波林(cyclosporine) A之交互作用可抑制病毒複製。

Gag之NC部位係特別負責識別所謂的HIV包裝訊號。包

裝訊號係由四個位於近病毒RNA 5'末端之主幹環(stem loop)結構所構成，其可供媒介異種RNA併入HIV-1病毒粒子中。經由兩個鋅-指狀中心(zinc-finger motifs)所媒介的交互作用，NC可與包裝訊號結合。NC亦可促進反轉錄作用。

p6多肽部位可媒介p55 Gag與輔助蛋白質Vpr間之交互作用，其可導致Vpr併入組裝病毒粒子。p6部位亦包含所謂的後(late)功能部位，其為由受感染細胞釋出出芽之病毒粒子時所必需。

Pol基因可編碼二種蛋白質，其作用活性為早期病毒感染所需，即將病毒DNA整合入細胞DNA所需之RT及接合酶蛋白質。Pol之主產物為病毒粒子蛋白酶裂解後可產生胺基端RT肽，其作用活性為DNA合成所需(經RNA及DNA指導之DNA聚合酶、核糖核酸酶H)及羧基端接合酶蛋白質。

HIV RT係由完整長度之RT(p66)及缺少羧基端核糖核酸酶接合酶功能部位之分裂產物(p51)所組成之異種二聚物。

RT為最高度保存性蛋白質之一，其係由反轉錄病毒基因組所編碼。RT二種主要活性為DNA Pol及核糖核酸酶H。RT之DNA Pol活性功能可替換地使用RNA及DNA做為模板，且如同所有已知之DNA聚合酶，如果缺少預存的分子充當引導子(RNA)就無法啟動DNA重新合成。

所有RT蛋白質中固有之核糖核酸酶H活性於複製早期

扮演重要的角色，其於DNA合成進行時移除RNA基因組。其可選擇性地自所有的RNA-DNA雜種分子中降解RNA。在結構上，聚合酶及ribo H各單獨占據不重疊的功能部位，其中Pol上三分之二覆蓋了胺基。

p66催化性次單元摺疊成為5個不同的次功能部位。這些次功能部位中之胺基端23具有RT活性部分。其羧基部份為核糖核酸酶H功能部位。

在感染宿主細胞後，反轉錄病毒之RNA基因組可藉存在於感染顆粒中之反轉錄酶複製為線狀雙股DNA。接合酶(請見史喀耳卡(Skalka) AM '99 Adv in Virus Res 52 271-273)可識別病毒DNA末端並將病毒DNA修剪、附加至宿主染色體位置，因此可催化整合作用。其可標定宿主DNA之許多部位以進行整合作用。雖然接合酶足以於生體外催化整合作用，活體內其並非唯一與病毒DNA有關之蛋白質-由受感染細胞分離出之大型蛋白質-病毒DNA複合物即為預先整合之複合物。其可幫助子代病毒基因組取得宿主細胞基因。

接合酶係由3個不同的功能部位所組成，即N端功能部位、催化核心及C端功能部位。催化核心功能部位包含所有聚核苷酸轉移所需之化學條件。

已知Nef蛋白質可使CD4(HIV受體)由細胞表面移除，但此作用在生物學上的重要性仍有待討論。此外，Nef可與T細胞之訊號通路交互作用以誘發活性狀態，其隨後可促進更多有效之基因表現。某些HIV分離株於該部位有突

變，其可導致無法編碼功能性蛋白質並嚴重危及其於活體內複製與致病的能力。

DNA疫苗通常由細菌質體載體構成，其中可嵌入一強力啟動基因、可編碼所欲抗原肽之基因及一組聚核苷酸化作用/轉錄終止序列。所欲基因應可編碼出完整蛋白質或僅編碼與病原、腫瘤或其它欲保護對抗的物質相關之抗原肽序列。對宿主投藥前，該質體可於細菌(如大腸桿菌(E.coli))中生長，隨後加以分離並依預期之投藥途徑以適當的培養基製備。於投藥後，質體由宿主細胞攝入並產生編碼過的肽。質體載體來源最好選擇不會在真核細胞中複製者以預防質體於哺乳動物宿主體內複製並與相關動物之染色體DNA整合。

相對於傳統疫苗科技而言，DNA疫苗接種有許多的優點。首先可預期的是，因為由DNA序列編碼之蛋白質係於宿主體內合成，故蛋白質之結構或構造將會類似與造成病狀相關之天然蛋白質。況且藉由產生可識別來自保存性蛋白質之抗原決定位的殺手T淋巴細胞反應，DNA疫苗接種可能提供保護對抗同一種病毒之不同病毒株。此外，因質體係被可產生抗原蛋白質之宿主細胞攝入，故可引發持久的免疫反應。本技藝亦提供結合多種抗原體成為單一製備物之可能性以利於同時免疫對抗多種病狀。

關於DNA疫苗接種之各類有益背景資訊係由東內力(Donnelly)等人所提供，"DNA Vaccines" Ann. Rev Immunol. 1997 15: 617-648，其完整內容在此以引用的方式併入本

文中。

發明簡述

本發明係提供可用於預防及治療 HIV 感染及 AIDS 之核酸疫苗中的新穎結構。

因此，本發明之第一種觀點係提供一種核酸分子，其包括可編碼 HIV gag 蛋白質或其片段之核苷酸序列，並連接至可編碼另一 HIV 抗原或其片段之核苷酸序列，且於實行時可連接至異種啟動基因。該核苷酸序列之片段可編碼 HIV 抗原決定位且典型上編碼至少 8 個胺基酸之肽。該核苷酸序列較佳為 DNA 序列且較佳包含於來自不同複製來源的質體中。此種核酸分子可以藥學上可接受賦形劑、載體、稀釋劑或佐劑調配以製造適於治療及 / 或預防 HIV 感染及 AIDS 之藥學組合物。

於一較佳具體實施例中，DNA 序列係調配至適於顆粒媒介藥物遞送之惰性顆粒或有孔小珠表面。有孔小珠較佳為金製。

於本發明一較佳具體實施例中提供一種 DNA 序列，其中可高度表現 gag 蛋白質之密碼已經最佳化以便類似哺乳動物細胞之基因所採用的密碼子。特定言之，gag 蛋白質已經最佳化以便類似可高度表現之相對人類基因。

DNA 密碼有 4 個字母 (A、T、C 及 G)，採用這些字母可拼出包含三個字母之 "密碼子"，其於生物基因中代表可編碼之胺基酸與蛋白質。沿著 DNA 分子之密碼子線狀序列可轉譯由那些基因所編碼之蛋白質的線狀胺基酸序列。密碼係

高度退化的，其中61個可編碼出20種天然胺基酸，而有3個代表"終止"訊號之密碼子。因此，大多數胺基酸之編碼可來自一個以上之密碼子-事實上有數種胺基酸可由四種或更多不同之密碼子所編碼而成。

雖然有一種以上的密碼子可供編碼出某一特定胺基酸，但所觀察到生物之密碼子利用模式卻相當的不隨機。不同生物品種之密碼子選擇各顯示不同之偏好，更甚者，即使同一物種在表現不同濃度基因時所使用的密碼子也大不相同。病毒、植物、細菌及哺乳動物細胞之偏好皆不同，且某些物種顯示出比他者更強烈偏離隨機選擇密碼子之偏好。如與某些細菌或病毒相較之下，人類及其它哺乳動物比較不會有強烈的偏好。由於這些原因，在大腸桿菌中表現之哺乳動物基因或在哺乳動物細胞表現之外來或重組基因極有可能產生不適當的密碼子分佈狀況，因而影響有效表現性。咸信若將少見於宿主體內的多組密碼子或大量密碼子之異種DNA序列於該宿主進行表現，則其可預期到低度之異種表現性。

於本發明一具體實施例中係提供一種可編碼胺基酸序列之gag多核苷酸序列，其中多核苷酸序列之密碼子利用模式類似高度表現哺乳動物基因之密碼子利用模式。較佳多核苷酸序列為DNA序列。多核苷酸序列之密碼子利用模式典型較佳為高度表現人類基因模式。

於本發明之多核苷酸中，密碼子利用模式由典型的人類免疫不全病毒模式改變為更接近代表標的生物之密碼子

偏好，如哺乳動物，特別為人類。"密碼子利用係數"係用來測量指定多核苷酸序列密碼子模式與標的物種密碼子模式之類似程度。許多物種之高度表現基因的密碼子頻率可得自文獻記載(見於如那卡目拉(Nakamura)等人，Nucleic Acid Research 1996, 24: 214-215)。將61個密碼子中各密碼子頻率(表示成為所選出基因等級中每1000個密碼子中之出現率)以20種天然胺基酸分類加以常態化，如此可將各胺基酸之最常用密碼子數值設定為1而較不常見之密碼子頻率則按比例排列於0與1間。如此，61個密碼子中每一個皆可依標的物種而指定一個數值為1或稍低於1之高度表現基因。為計算特定多核苷酸相對於該物種之高度表現基因之密碼子利用係數，記錄各特定多核苷酸之密碼子之分級數值並計算所有這些數值之幾何平均數(將這些數值之自然對數值總和除以所有的密碼子總數並取其反-對數值)。該係數值介於0至1間且係數越高，多核苷酸中含有越多經常使用之密碼子。若多核苷酸序列之密碼子利用係數為1，則所有密碼子皆為標的物種高度表現基因"最常用"之密碼子。

如本發明，多核苷酸之密碼子利用模式較佳排除RSCU值小於0.2之密碼子(相對於標的生物之高度表現基因)。或者，若一組密碼子僅提供不到10%表現特定胺基酸所應使用之密碼子，則應排除於密碼子利用模式之外。相對同義密碼子利用(RSCU)值為所觀測密碼子總數除以若該胺基酸之所有密碼子皆以同等頻率使用時之預期數目。本發

明多核苷酸相對於高度表現人類基因之密碼子利用係數(或RSCU)通常大於0.3,較佳大於0.4,最佳為大於0.5,人類之密碼子利用表亦可見於基因銀行(Genebank)。

對照之下,高度表現性肌動蛋白基因之RSCU為0.747。
人類之密碼子利用表如下:

密碼子利用表1:

現代人 [gbpri]: 27143 CDS's (12816923 密碼子)

表示法 fields: [三聯體][頻率:每千組]([數量])

UUU 17.0(217684)	UCU 14.8(189419)	UAU 12.1(155645)	UGU 10.0(127719)
UUC 20.5(262753)	UCC 17.5(224470)	UAC 15.8(202481)	UGC 12.3(157257)
UUA 7.3(93924)	UCA 11.9(152074)	UAA 0.7(9195)	UGA 1.3(16025)
UUG 12.5(159611)	UCG 4.5(57572)	UAG 0.5(6789)	UGG 12.9(165930)
CUU 12.8(163707)	CCU 17.3(222146)	CAU 10.5(134186)	CGU 4.6(59454)
CUC 19.3(247391)	CCC 20.0(256235)	CAC 14.9(190928)	CGC 10.8(137865)
CUA 7.0(89078)	CCA 16.7(214583)	CAA 12.0(153590)	CGA 6.3(80709)
CUG 39.7(509096)	CCG 7.0(89619)	CAG 34.5(441727)	CGG 11.6(148666)
AUU 15.8(202844)	ACU 12.9(165392)	AAU 17.0(218508)	AGU 12.0(154442)
AUC 21.6(277066)	ACC 19.3(247805)	AAC 19.8(253475)	AGC 19.3(247583)
AUA 7.2(92133)	ACA 14.9(191518)	AAA 24.0(308123)	AGA 11.5(147264)
AUG 22.3(285776)	ACG 6.3(80369)	AAG 32.6(418141)	AGG 11.3(145276)
GUU 10.9(139611)	GCU 18.5(236639)	GAU 22.4(286742)	GGU 10.8(138606)
GUC 14.6(187333)	GCC 28.3(362086)	GAC 26.1(334158)	GGC 22.7(290904)
GUA 7.0(89644)	GCA 15.9(203310)	GAA 29.1(373151)	GGA 16.4(210643)
GUG 28.8(369006)	GCG 7.5(96455)	GAG 40.2(515485)	GGG 16.4(209907)

密碼 GC 52.51% 第一字母 GC 56.04% 第二字母 GC 42.35%

第三字母 GC 59.13%

密碼子利用表2(較佳):

人類(高度表現)基因 1/24/91 (human_high.cod)之密碼子利用

(10)

發明說明書續頁

胺基酸	密碼子	數量	/1000	片段
Gly	GGG	905.00	18.76	0.24
Gly	GGA	525.00	10.88	0.14
Gly	GGT	441.00	9.14	0.12
Gly	GGC	1867.00	38.70	0.50
Glu	GAG	2420.00	50.16	0.75
Glu	GAA	792.00	16.42	0.25
Asp	GAT	592.00	12.27	0.25
Asp	GAC	1821.00	37.75	0.75
Val	GTG	1866.00	38.68	0.64
Val	GTA	134.00	2.78	0.05
Val	GTT	198.00	4.10	0.07
Val	GTC	728.00	15.09	0.25
Ala	GCG	652.00	13.51	0.17
Ala	GCA	488.00	10.12	0.13
Ala	GCT	654.00	13.56	0.17
Ala	GCC	2057.00	42.64	0.53
Arg	AGG	512.00	10.61	0.18
Arg	AGA	298.00	6.18	0.10
Ser	AGT	354.00	7.34	0.10
Ser	AGC	1171.00	24.27	0.34
Lys	AAG	2117.00	43.88	0.82
Lys	AAA	471.00	9.76	0.18
Asn	AAT	314.00	6.51	0.22
Asn	AAC	1120.00	23.22	0.78
Met	ATG	1077.00	22.32	1.00
Ile	ATA	88.00	1.82	0.05
Ile	ATT	315.00	6.53	0.18
Ile	ATC	1369.00	28.38	0.77
Thr	ACG	405.00	8.40	0.15

(11)

發明說明書頁

Thr	ACA	373.00	7.73	0.14
Thr	ACT	358.00	7.42	0.14
Thr	ACC	1502.00	31.13	0.57
Trp	TGG	652.00	13.51	1.00
End	TGA	109.00	2.26	0.55
Cys	TGT	325.00	6.74	0.32
Cys	TGC	706.00	14.63	0.68
End	TAG	42.00	0.87	0.21
End	TSS	46.00	0.95	0.23
Tyr	TAT	360.00	7.46	0.26
Tyr	TAC	1042.00	21.60	0.74
Leu	TTG	313.00	6.49	0.06
Leu	TTA	76.00	1.58	0.02
Phe	TTT	336.00	6.96	0.20
Phe	TTC	1377.00	28.54	0.80
Ser	TCG	325.00	6.74	0.09
Ser	TCA	165.00	3.42	0.05
Ser	TCT	450.00	9.33	0.13
Ser	TCC	958.00	19.86	0.28
Arg	CGG	611.00	12.67	0.21
Arg	CGA	183.00	3.79	0.06
Arg	CGT	210.00	4.35	0.07
Arg	CGC	1086.00	22.51	0.37
Gln	CAG	2020.00	41.87	0.88
Gln	CAA	283.00	5.87	0.12
His	CAT	234.00	4.85	0.21
His	CAC	870.00	18.03	0.79
Leu	CTG	2884.00	59.78	0.58
Leu	CTA	166.00	3.44	0.03
Leu	CTT	238.00	4.93	0.05
Leu	CTC	1276.00	26.45	0.26

Pro	CCG	482.00	9.99	0.17
Pro	CCA	456.00	9.45	0.16
Pro	CCT	568.00	11.77	0.19
Pro	CCC	1410.00	29.23	0.48

如本發明進一步之觀點係提供一種表現性載體，其包括如本發明第一種觀點之多核苷酸序列並可指導如本發明第一種觀點之多核苷酸序列的表現性，特定言之為包含了可典型高度表現哺乳動物基因(較佳為高度表現人類基因)之gag多核苷酸序列之密碼子利用模式。該載體適於驅使異種DNA於細菌、昆蟲或哺乳動物細胞(特別係人類細胞)中之表現。於一具體實施例中之表現性載體為p7313(見圖1)。

於第三個具體實施例中係提供一種受異種啟動基因控制之gag基因，其融合至一組可編碼NEF、其片段、或HIV反轉錄酶(RT)或其片段之DNA序列中。該基因之gag部分可位於融合基因之N端或C端部分。

於較佳具體實施例中，gag基因並不編碼gag p6肽。較佳為截短NEF基因、移除可編碼N端部位之序列，即移除30-85個，較佳為60-85個，典型地約81個，較佳為N端65個胺基酸。

於更進一步之具體實施例中，RT基因亦可最佳化使類似高度表現之人類基因。RT較佳可編碼一可實質上鈍化任何反轉錄酶活性之突變。較佳之鈍化突變牽涉到以W色胺酸229取代K(離胺酸)。

於更進一步之具體實施例中，RT基因亦可最佳化使類似高度表現之人類基因。RT較佳可編碼一可實質上鈍化任何反轉錄酶活性之突變。較佳之鈍化突變牽涉到以W色胺酸229取代K(離胺酸)。

如本發明進一步之觀點，其提供一宿主細胞，其中包含如本發明之多核苷酸序列，或如本發明之表現性載體。宿主細胞可為細菌(如大腸桿菌)、哺乳動物(如人類)、或可為昆蟲細胞。包含如本發明載體之哺乳動物細胞可於生體外培養進行細胞轉移感染或對哺乳動物投予載體於活體內進行轉移感染。

本發明進一步提供一種藥學組合物，其中包含如本發明之多核苷酸序列。組合物較佳包含DNA載體。於較佳具體實施例中，組合物包括許多顆粒，較佳為金製顆粒，表面塗覆之DNA包含可編碼本發明之多核苷酸序列之載體。該序列較佳可編碼HIV gag胺基酸序列，其中的多核苷酸序列之密碼子利用模式為典型高度表現之哺乳動物基因，特別係人類基因。於另一具體實施例中，該組合物包括藥學上可接受賦形劑及如本發明第二種觀點之DNA載體。該組合物亦可包含佐劑。

因此，本發明之一具體實施例中，本發明載體可與免疫興奮劑共同運用。免疫興奮劑較佳與本發明核酸載體同時給予，且於較佳之具體實施例中此二者可調配在一起。此種免疫興奮劑包含(但以下名單無法詳盡且並不排除其它藥劑)：合成異吡唑啉，如依米奎摩德(imiquimod)

[S-26308, R-837], (哈里森 (Harrison) 等人, 'Reduction of recurrent HSV disease using imiquimod alone or combined with a glycoprotein vaccine', Vaccine 19 : 1820-1826, (2001) 及 雷西奎摩德 (resiquimod) [S-28463, R-848] (華西拉寇司 (Vasilakos) 等人, 'Adjuvant activities of immune response modifier R-848 : Comparison with CpG ODN', Cellular immunology 204 : 64-74 (2000).), 固定表現於抗原表現細胞及 T-細胞表面之羧及胺之西福 (Schiff) 鹽基, 如吐卡雷颯 (tucaresol) 羅德斯 (Rhodes), J. 等人, 'Therapeutic potentiation of the immune system by costimulatory Schiff-base-forming drugs', Nature 377 : 71-75 (1995))、細胞動力素 (cytokine)、化學動力素 (chemokine) 及輔-興奮分子 (co-stimulatory), 如蛋白質或肽, 其可包含前-炎症性細胞動力素 (pro-inflammatory cytokines) (如 GM-CSF、IL-1 α 、IL-1 β 、TGF- α 及 TGF- β)、Th1 誘導劑 (如干擾素 γ 、IL-2、IL-12、IL-15 及 IL-18)、Th2 誘導劑 (如 IL-4、IL-5、IL-6、IL-10 及 IL-13) 與其它化學動力素及輔-興奮基因 (如 MCP-1、MIP-1 α 、MIP-1 β 、RANTES、TCA-3、CD80、CD86 及 CD40L)、其它免疫興奮標的配位體 (如 CTLA-4 及 L-什雷克汀 (selectin))、細胞凋亡 (apoptosis) 刺激蛋白質及肽, 如 Fas, (49)、合成脂基佐劑, 如華克斯菲可汀 (vaxfectin) (雷西斯 (Reyes) 等人, 'Vaxfectin enhances antigen specific antibody titeres and maintains Th1 type

immune responses to plasmid DNA immunization', Vaccine 19: 3778-3786) 鯊烯、 α -生育酚、多乙氧基醚 (polysorbate) 80、DOPC 及膽固醇、內毒素、[LPS]，布特樂 (Beutler) B. ('Endotoxin, Toll-like receptor 4, and the afferent limb of innate immunity', Current Opinion in Microbiology 3: 23-30 (2000)); CpG 寡-及二核苷酸，撒透 (Sato), Y. 等人，'Immunostimulatory DNA sequences necessary for effective intradermal gene immunization', Science 273 (5273): 352-354 (1996)。嘿米 (Hemmi), H. 等人，'A Toll-like receptor recognizes bacterial DNA', Nature 408: 740-745, (2000) 及其它可能引發內毒素受體 (Toll receptors) 以產生由 Th1 所誘導之細胞動力素之配位體，如合成成分枝菌脂蛋白質、分枝菌蛋白質 p19、肽聚糖 (peptidoglycan)、磷壁酸質及脂質 A。

某些可誘發明顯 Th1-類型反應包含之較佳佐劑包括如脂質 A 衍生物，如單磷酸脂質 A，或較佳為 3-去-O-醯化單磷酸基脂質 A。MPL[®] 佐劑係出品自克利莎公司 (Corixa Corporation) (Seattle, WA，見如美國專利號 4,436,727、4,877,611、4,866,034 及 4,912,094)。含 CpG 之寡核苷酸 (其中該 CpG 二核苷酸係去甲基化者) 亦可誘發明顯的 Th1 反應。此種寡核苷酸為眾所皆知並可見述於，如 WO 96/02555、WO 99/33488 及美國專利案號 6,008,200 與 5,856,462。免疫興奮 DNA 序列亦描述於如撒透 (Sato) 等人，Science 273: 352, 1996。另一較佳佐劑包括皂素，

如奎爾(Quil) A，或其衍生物，包含QS21及QS7(阿奎拉製藥(Aquila Biopharmaceuticals Inc.)、Framingham (MA)、依信(Escin)、滴極多寧(Digitonin)、或絲石竹(Gypsophila)或奎藜(Chenopodium quinoa)皂素。

亦提供如本發明多核苷酸或如本發明載體於治療或預防HIV感染之用途。

本發明亦提供治療或預防HIV感染、任何與此相關之症狀或疾病之方法，其包括投予有效量之如本發明之多核苷酸、載體或藥學組合物。投予藥學組合物時可採用一或多次單獨劑量型式，如以重覆劑量之相同DNA質體模式，或以"首次接觸-補強"治療接種攝生法。在某些情況下，"首次接觸"接種可經由顆粒媒介DNA遞送如本發明之多核苷酸，較佳為併入質體-來源載體，並藉由投予包含相同多核苷酸序列之重組病毒載體進行"補強"，或以存於佐劑中之蛋白質進行補強。反之，首次接觸亦可使用病毒載體或蛋白質調配物(通常為調配在佐劑中之蛋白質)，而以如本發明之DNA疫苗進行補強。首次接觸及/或補強皆可使用多次劑量。

本發明之具體實施例中考慮採取gag、nef或RT蛋白質片段。如，本發明之多核苷酸可編碼HIV gag片段、nef或RT蛋白質。本發明範圍所考慮之多核苷酸係可編碼至少8個，如8-10個胺基酸或多達20、50、60、70、80、100、150或200個胺基酸長度之片段者，且須能編碼出可顯示HIV抗原性之低聚肽或多肽。特定言之(但並不排除外

者)，本發明之此項觀點所包含之條件為該多核苷酸可編碼一完整HIV蛋白質序列片段及可代表一或多個該蛋白質之分離抗原決定位。該片段可為經最佳化之密碼子，使該片段得以類似高度表現哺乳動物基因之密碼子利用模式。

如本發明之較佳結構包含：

1. p17、p24、融合至截短之NEF(缺少可編碼末端胺基酸1-65之核苷酸)
2. p17、p24、RT、截短之NEF(缺少可編碼末端胺基酸1-65之核苷酸)
3. p17、p24(最佳化gag)、截短之NEF(缺少可編碼末端胺基酸1-65之核苷酸)
4. p17、p24(最佳化gag)、RT(最佳化)、截短之NEF(缺少可編碼末端胺基酸1-85之核苷酸)
5. p17、p24、RT(最佳化)、截短之NEF(缺少核苷酸可編碼末端胺基酸1-65)
6. 截短之NEF-(缺少核苷酸1-65)融合至最佳化p17、p24 gag。
7. 本發明特佳結構包含三聯融合體RT-NEF-Gag、及特定言之為RT-Gag-Nef。
8. 最佳化RT、截短之NEF及最佳化P17、p24(gag)(RNG)及
9. 最佳化RT、最佳化p17、24(gag)、截短之Nef(缺少胺基酸1-65)RGN

較佳 HIV 結構係源自 HIV 分化枝 B 或分化枝 C，特別係分化枝 B。

如上述，本發明包括含有本發明核苷酸序列之表現性載體。該表現性載體可例行地以分子生物學技藝建構並可使用如質體 DNA 及適當起始因子、啟動基因、增強物及其它元件，如可能需要聚腺苷酸化作用訊號，且其應置於正確方向以利於蛋白質表現。其它適當載體亦為熟諳此藝者所詳知。進一步實例可參考山姆布魯克 (Sambrook) 等人之 *Molecular Cloning: Laboratory Manual*. 第 2 版，CSH Laboratory Press. (1989)。

較佳地，本發明之多核苷酸，或可用於本發明載體之多核苷酸，可加以操作連接至控制序列上，後者可供宿主細胞表現編碼序列，即該載體成為一種表現性載體。"操作性連接"一詞係指一種並列方式，其中所述組分間的關係使得各成分皆可發揮其原預期功能。將調節序列(如啟動基因)"操作性連接"至編碼序列之安置方式使該編碼序列可於與調節序列相容之條件下進行表現。

載體可為，如質體，人造染色體(如 BAC、PAC、YAC)、病毒或噬菌體載體，其中包括複製來源、視需要選用之啟動基因(用來表現多核苷酸)及視需要選用(用來調節啟動基因)之調節基因。載體可包含一或多種選擇性標記基因，細菌質體可採用如胺苄青黴素或康黴素耐藥性基因，而噬菌體載體可採用耐藥性基因。載體可於生體外用於製造 DNA 或 RNA 或用來轉移感染或轉化宿主細胞(如哺乳動物

宿主細胞)，如藉由載體編碼以製造蛋白質。載體亦可適用於活體內，如利用DNA接種方法或基因療法。

可選擇啟動基因及其它表現性調節訊號以配合相容於設計來表現之特定宿主細胞。如哺乳動物啟動基因，包含可對重金屬(如鎘)反應誘發之金屬硫肽啟動基因，及 β -肌動蛋白啟動基因。病毒啟動基因，如SV40大型T抗原啟動基因，人類細胞巨大病毒(CMV)極早期(IE)啟動基因、勞司氏(rous)肉瘤病毒LTR啟動基因、腺病毒啟動基因、或HPV啟動基因，特定言之，HPV上游調節部位(URR)亦可使用。所有這些啟動基因皆有詳細敘述且在本技藝中隨時待用。

較佳之啟動基因元件為CMV極早期啟動基因，缺少插入序列A但包含表現序列1。啟動基因元件可為最小啟動基因元件或強化啟動基因，較佳者為強化啟動基因。依此可提供一種包含本發明多核苷酸之載體，其係由HCMV IE早期啟動基因所控制。

適當病毒載體實例包含單純皰疹(herpes simplex)病毒載體、牛痘或 α -病毒載體及反轉錄病毒，包含慢病毒、腺病毒及腺-相關病毒。利用這些病毒之基因轉移技術為熟諳此藝者所嫻知。如反轉錄病毒載體可用於將本發明多核苷酸穩定地併入宿主基因組中，然而這種重組並非較佳者。相較之下，複製-缺陷之腺病毒載體殘留有基因附體(episomal)並因而提供短暫表現性。能驅使昆蟲細胞(如桿狀病毒載體)、人類細胞、酵母或細菌表現的載體可應用

來產生大量由本發明多核苷酸編碼之HIV蛋白質，並可作為如次單位疫苗或免疫分析之用。

如本發明之多核苷酸可用以表現生產編碼蛋白質，此表現可發生在生體外、活體內或活體外。因此，核苷酸可利用重組蛋白質合成技術以便，如增加產量；或甚至可發現以其本身之條件、利用DNA接種技術做為治療劑。當採用本發明多核苷酸於生體外或活體外製備編碼蛋白質時，細胞(如於細胞培養中)可經修改包含欲表現之多核苷酸。此種細胞可包含短暫，或較佳為穩定之哺乳動物細胞株。藉由嵌入可編碼如本發明多肽之載體而修改之細胞特定實例包含哺乳動物HEK293T、CHO、HeLa、293及COS細胞。較佳地，所選擇之細胞株不僅穩定而且可提供成熟糖基化及多肽之細胞表面表現性。轉化卵細胞可滿足該表現性。多肽可由本發明多核苷酸表現於轉殖基因之非-人類之動物(較佳為鼠)細胞。本發明範圍包含可表現本發明多核苷酸之多肽的轉殖基因非-人類動物。

本發明進一步提供一種對哺乳動物對象注射疫苗之方法，其包括對其投予有效量之疫苗或疫苗組合物。使用於DNA疫苗、疫苗組合物及免疫療法之表現性載體最佳為質體載體。

DNA疫苗可以"單純裸DNA"型式投藥，如以液體調配物型式、使用注射器或高壓噴嘴投藥，或以脂質體或可刺激轉移感染之增強劑來調配DNA，或利用顆粒媒介DNA遞送法(PMDD)。所有這些遞送系統皆為本技藝中為吾人所熟

知之方法。載體可藉由如病毒載體遞送系統進入哺乳動物體內。

本發明組合物可藉多種途徑遞送，如肌肉注射、皮下注射、腹腔內注射、靜脈注射或經黏膜投藥。

於一較佳具體實施例中，組合物係以皮內注射方式遞送。特定言之，組合物係以基因槍(砲)方式遞送，特定言之為顆粒砲擊投藥技術，其包含將載體塗佈至有孔小珠(如金製)上，再接再利用高壓投藥進入表皮，如黑恩斯(Haynes)等人所描述者，J Biotechnology 44: 37-42 (1996)。

於一說明性實例中，可藉如包德傑克特製藥公司(Powderject Pharmaceuticals PLC; Oxford, UK)及包德傑克特疫苗公司(Powderject Vaccine Inc.; Madison, WI)所製造之裝置達成氣動(gas-driven)顆粒加速作用，其部分實例可見美國專利號5,846,796、6,010,478、5,865,796、5,584,807、及EP專利案號0500 799。此方法提供一種不需注射針之遞送方法，其中微小顆粒(如多核苷酸)之乾粉調配物藉由手持裝置所產生之氮氣流中提昇至高速，推動顆粒進入所欲之標的組織(典型為皮膚)。

顆粒較佳為0.4-4.0微米之金製有孔小珠，更佳為直徑0.6-2.0微米，將DNA共軛物(conjugate)塗覆於其上，並填入藥筒或藥匣以便裝入"基因槍"中。

於一相關具體實施例中，其它可用於本發明組合物之氣動無針頭注射的裝置及方法包含由拜歐傑克特公司

(Bioject, Inc.(Portland, OR))所提供者，若干實例描述於美國專利案號 4,790,824、5,064,413、5,312,335、5,383,851、5,399,163、5,520,639及5,993,412。

含有可編碼抗原肽之核苷酸序列的載體之投藥劑量應可有效預防或治療。顆粒-媒介性遞送投藥量範圍通常為每劑1微微克至1毫克，較佳為1微微克至10微克之核苷酸，而其它途徑為每劑100毫微克至1毫克，較佳為10微克至1毫克之核苷酸。確切用量可考量取決於需免疫病患之體重及投藥途徑而變化。

包含可編碼抗原肽之核苷酸序列的抗原體組分可以一次性投藥或重覆投藥，如1至7次，較佳1至4次，間隔介於約1天至約18個月。然而，該治療攝生法可能變化甚大，主要考量為病患體型、核苷酸序列投藥量、投藥途徑、及專業獸醫或開業醫師明顯考慮的其它因素。病患可接受一或多種其它抗HIV反轉錄病毒藥物作為其整體治療攝生法之一部分。此外，核酸免疫原可與佐劑共同投予。

在此指定之佐劑組分可同樣地經各種不同投藥途徑投予，如經口服、鼻腔投藥、肺部投藥、肌肉注射、皮下注射、皮內注射或局部途徑投藥。佐劑組分較佳係經由皮內注射或局部投藥。最佳為藉由局部途徑投藥。佐劑組分可於核苷酸序列投藥前約14日至投藥後約14日間進行投藥，較佳介於投予核苷酸序列前約第1日至投藥後約第3日。於一具體實施例中，實質上佐劑組分可與核苷酸序列

同時藉皮下注射投藥。"實質上同時"意指佐劑組分較佳與核苷酸序列同時投藥，或若未同一時間投藥，至少於核苷酸序列投藥前或投藥後數小時內投予。於最佳治療計劃中，佐劑組分投藥係與核苷酸序列投藥同時進行。顯然該計劃可視需要依上述提及之可變因素型態而變更。較佳佐劑為1H-異吡唑[4,5c]喹啉-4-胺衍生物，如伊米奎摩德(imiquimod)。通常伊米奎摩德可製成局部膏狀調配物並可以如上述計劃投藥。

如前，根據各可變因素，衍生物之投藥劑量亦可改變，但範圍介於如每公斤約0.1毫克至每公斤約100毫克，其中"每公斤"係指哺乳動物之體重。1H-異吡唑[4,5-c]喹啉-4-胺衍生物之投藥較佳為隨著核苷酸序列之後續或補強而重覆投藥。最佳投藥劑量介於每公斤約1毫克至約50毫克。以在此所述"首次接觸抗原-補強"之情況下，伊米奎摩德或其它1H-異吡唑[4,5-c]喹啉-4-胺衍生物可於首次接觸抗原或補強時給藥，或在首次接觸抗原及補強時都投藥。

若佐劑組分僅包含1H-異吡唑[4,5-c]喹啉-4-胺衍生物時，其可以原始化學狀態直接投藥，但較佳投藥型式為其藥學調配物。即佐劑組分較佳包括1H-異吡唑[4,5-c]喹啉-4-胺，並結合一或多種藥學可接受載劑，及其它選用性治療成分。"可接受"載劑意指其可與調配物中之其它成分相容，並對接受者無害。調配物特性可依預期之投藥途徑而加以變更，並可藉藥學技藝中為吾人所熟知之方法製

備。所有方法皆包含使 1H-異吡唑[4,5-c]喹啉-4-胺衍生物與適當載劑或載體結合之步驟。一般說來，調配物之製備係將該衍生物與液體載體或微細分割之固體載體、或前述二者均勻充分混合；隨後，若有必要時可將產物塑成所欲之調配物形狀。適合口服投藥之本發明調配物可製成分離單位(如包含預定份量活性成分的膠囊、扁囊或片劑)、粉末或顆粒、水性液體或非水性液體之溶液或懸浮液、或為水中油液體乳劑或油中水乳劑。活性成分亦可製成巨丸劑、藥糖劑或糊劑。

片劑可藉由壓縮或塑模製造，可視需要添加一或多種附加成分。製造壓縮片劑時可將自由流動型式之活性成分(如粉末或顆粒)於適當機器中壓縮而成，其可視需要與黏合劑、潤滑劑、惰性稀釋劑、潤滑劑(lubricating agent)、表面活性劑或分散劑混合。製造塑模片劑時可將以惰性液體稀釋劑濕潤之粉狀化合物的混合物於適當機器中塑模而成。

片劑可視需要塗覆或刻痕，並可調配成為能緩慢或控制釋出活性成分。

可經由如肌肉注射、腹腔內注射、或皮下注射途徑投藥之注射用調配物包含水性及非水性無菌注射溶液，其可包含抗氧化劑、緩衝溶液、抑菌劑及可使調配物與預期受藥者之血液等張之溶質，及水性與非水性無菌懸浮液，其中可包含懸浮劑及增稠劑。調配物可存放在於單位劑量或多劑量容器中，如密封安瓿及小藥瓶，並可貯存於冷凍乾燥

(凍乾)狀態下，僅需於使用前添加無菌液狀載劑，如注射用水。現場製作注射溶液及懸浮液可由前述各類無菌粉末、顆粒及片劑來製備。適合經由口腔或鼻腔進行肺部投藥之調配物可製成含活性成分之顆粒，其適合之直徑範圍為0.5至7微米，以便遞送至受藥者之支氣管樹中。該調配物可為研成細粉之粉末型式，其可方便地置於適用於吸入裝置之可刺穿膠囊(如明膠製)中，或者可製成自動噴射調配物，其中包含活性成分、適當液體推進劑及視需要添加之其它成分，如表面活性劑及/或固體稀釋劑。當活性成分係分佈成溶液或懸浮液中之小滴型式時，亦可應用成為自動噴射調配物。該自動噴射調配物類同於本技藝中為人所知者且可藉現有步驟來製備。其可適當地製成具有所需噴霧特性之手動或自動活門型式；後者可有利地設計成可計量型式以便每次操作時能遞送固定劑量，如50至100微升。

在更進一步的可能性中，佐劑組分可為能用於霧化器或噴霧器之溶液型式，藉由加速氣流或超音速震動可產生可吸入之微細小滴噴霧。

適於鼻內投藥之調配物通常包含類似上述肺部投藥之產品型式，但其中該調配物較佳之顆粒直徑範圍約10至約200微米，以使其可保留於鼻腔內。欲達此目的可使用適當顆粒大小之粉末，或選擇適當活門。其它適當調配物包含顆粒直徑範圍約20至約500微米之較粗粉劑，使用時將其容器靠近鼻子經鼻道快速吸入；以及含約0.2至5%重量

百分比活性成分之水性或油性溶液鼻滴劑。於本發明一具體實施例中，含有可編碼抗原肽之核苷酸序列的載體與1H-異吡唑[4,5-c]喹啉-4-胺衍生物可於同一調配物中一起投藥。因此於本具體實施例中，免疫原性及佐劑組分係見於同一調配物中。

於一具體實施例中，其中之佐劑組分係製備成為適於基因-槍投藥之型式，並經由該投藥途徑、實質上同時投予核苷酸序列。在製備適用於本方法之調配物時，1H-異吡唑[4,5-c]喹啉-4-胺衍生物需經過冷凍乾燥並依附至，如適於以基因-槍投藥之金製有孔小珠。

於可替代之具體實施例中，佐劑組分可經高壓氣體推進，而以乾粉型式投藥。

即使佐劑組分與核苷酸序列並未共同調配，亦可將佐劑組分投藥在與核苷酸序列相同投藥部位或其附近。

其它藥學製備物細節可見Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, 賓州(1985)，其中所揭露全部內容在此以引用的方式併入本文中。

可將裸多核苷酸或載體引渡至病患之適當技術亦包含採用適當載劑進行局部應用。核酸可局部投予至皮膚、或至黏膜表面，如藉由鼻內投藥、口服、陰道內投藥或直腸投藥。裸多核苷酸或載體可與藥學上可接受賦形劑(如磷酸鹽緩衝食鹽水(PBS))共存。採用助益劑(如布皮華凱(bupivacaine))可進一步促進DNA攝入，其可個別使用或

包含於DNA調配物中。其它可將核酸直接投予接受者之方法包含超音波、電刺激、電穿孔法及描述於美國專利第5,697,901號之微播種(microseeding)。

藉數種已知轉移感染技術可增加攝入核酸結構，如其中採用轉移感染劑者。此類試劑實例包含陽離子試劑，如磷酸鈣及DEAE-葡聚糖及脂肥克頓(lipofectants)，如脂肥克潭(lipofectam)及轉司肥克潭(transfectam)。核酸投藥劑量可改變。

本發明之核酸序列亦可藉常用於基因療法之特定遞送載體來投藥。基因療法之方法詳述見如佛梅(Verme)等人，Nature 1997, 389: 239-242。病毒性及非-病毒載體系統皆可使用。以病毒為主的系統包含反轉錄病毒、慢病毒(lentivirus)、腺病毒、腺-相關病毒、皰疹病毒、金絲雀痘病毒(Canarypox)及以牛痘病毒為主的系統。非病毒系統則包含直接投予核酸、中心體封裝技術(聚(乳酸-共-甘醇酸))及以脂質體為主的系統。必要時可結合病毒性及非病毒遞送系統以供初次接種後之補強注射，如最初的"首次接觸抗原"DNA接種使用非病毒載體(如質體)，隨後再使用病毒載體或非病毒系統進行一或多次之"補強"接種。同樣地，本發明亦考慮到含有本發明多核苷酸之首次接觸抗原啟動系統，隨後再以含蛋白質之佐劑進行補強；或反之亦可。

本發明核酸序列亦可藉轉化細胞投予。該類細胞包含由治療對象採集來之細胞。在生體外將本發明之裸多核苷酸

或載體引入該細胞中，隨後將該轉化細胞送回該治療對象。藉同源重組之結果可將本發明多核苷酸併入已存在於細胞之核酸中。若有必要，可於生體外培養轉化細胞並將一或多種所產生之細胞用於本發明。

藉已知之外科或顯微外科技術(如移植、顯微注射等)自病患適當部位取得細胞。

本發明藥學組合物可包含詳述如上之佐劑化合物，或其它可供增加藉由DNA編碼之蛋白質所誘發的免疫反應之物質。這些物質可藉由DNA編碼，無論是獨立或與抗原融合皆可，或成為調配物之非-DNA元件。可包含於本發明調配物之佐劑-類型物質實例包含泛素、溶酶體相關膜蛋白質(LAMP)、B型肝炎病毒核心抗原、FLT3-配位體(一種在產生專任(professional)抗原表現細胞(特定言之為齒狀(dendritic)細胞)時相當重要的細胞動力素)及其它細胞動力素，如IFN- γ 及GM-CSF。其它較佳佐劑包含伊米奎摩德及雷辛奎摩德(Resimquimod)及土卡拉索(Tucarasol)。伊米奎摩德為特佳者。

於本發明一較佳具體實施例中，本發明係提供以在此所述之核酸分子來治療或預防HIV感染的用途。核酸分子較佳係與伊米奎摩德一同投藥。伊米奎摩德較佳係局部投藥，而核酸分子較佳係藉顆粒媒介遞送投藥。

因此，本發明提供一種治療罹患或疑似罹患HIV感染病患之方法，其包括投予在此所述之核酸分子及伊米奎摩德。

本發明將以下列參考實例進行描述：

實例

實例 1：將 p55 gag(p17、p24、p13)最佳化以類似可供高度表現人類基因使用之密碼子

所欲基因

利用 PCR 自重疊寡核苷酸組合出可編碼 HIV-1 分化枝 B 菌株 HXB2 (GenBank entry k03455) 之 p55 gag 抗原的合成基因，使其最佳化而能於哺乳動物細胞中表現。

最佳化包含改變病毒基因之密碼子利用模式，使密碼子頻率接近高度表現人類基因。使用名為辛基因 (Syngene) (卡耳克基因 (Calcgene) 之更新版本，R.S. 哈爾 (Hale) 及 G. 湯普森 (Tompson) 所撰之 Protein Expression and Purification

，第 12 卷，185-188 頁，1998) 之統計學形象基礎 (Visual Basic) 程式來指定密碼子。

選殖：

以凝膠純化 1528 個鹼基對 (bp) 之 gag PCR 產物，以限制性核酸內切酶 NotI 及 Bam HI 切割並連接至經 NotI/Bam HI 切割之載體 WRG7077 上。其可將基因置於 CMV 啟動基因 / 插入序列 A 及牛生長激素聚腺苷酸化作用訊號之間。

將上述選殖株排序並檢查錯誤。沒有任何單一選殖株為 100% 正確。因此需將源自二種選殖株之正確序列部位加以結合，其係以重疊 PCR 反應、適當結合最佳化單體組而獲得完整長度密碼子之最佳化 gag 基因。隨後發現這個最

終的選植株中刪除了單一個核苷酸而造成移碼並導致提前終止轉譯。修補刪去部位時，先刪去包含不正確序列之基因並將來自另一選植株對等部位之正確序列選殖入本選植株中。如此可獲得最終最佳化密碼子之 p55 gag 選植株：Gagopt_{trpr2}(見圖 2)。

實例 2：p17/p24 截短之 Nef 融合基因的製備
所欲基因

將源自 HIV-1 分化枝 B 菌株 HXB2 之 p55 gag 基因的 p17 及 p24 部份經由質體 pHXB?Pr 以 PCR 反應放大 (B. 馬殊拉 (Maschera), E 佛芬 (Furfine) 及 E.D. 布萊爾 (Blair) 1995 J. Virol 69 5431-5436)。利用相同質體將源自 HXB2 nef 基因之 3' 末端的 pHXB?Pr. 426bp 加以放大。由於 HXB2 nef 基因包含提前終止密碼子，故以兩個重疊 PCR 修復該密碼子 (由 TGA [stop] 改為 TGG [Trp])。

以於 PCR 反應 (反向) 結合 p17/p24 連接子及 trNEF 連接子之 PCR 產物以形成 p17p24trNEF 融合基因 (圖 3)。

將 1542bp 之產物以凝膠純化，以限制性核酸內切酶 NotI 及 BamHI 切割並選殖入載體 WRG7077 之 NotI BamHI 部位。其可將該基因置於 CMV 啟動基因 / 插入序列 A 及牛生長激素聚腺苷酸化作用訊號之間。

實例 3：Gag p17/24opt/trNef1 ('Gagopt/Nef') 融合基因之製備

所欲基因

將源自 HIV-1 分化枝 B 菌株 HXB2 之 p55 gag 基因最佳化密

碼子 p17/p24 部分經由質體 pGagOPT_{Trp}2 以 PCR 反應放大。將其中提前終止密碼子已修補(由 TGA[stop]改為 TGG[Trp])之截短 HXB2 Nef 基因經由質體 7077trNef20 以 PCR 反應放大。二種 PCR 產物設計成為具有重疊末端以便二種基因可於第二種 PCR 中連結。

1544bp 產物以凝膠純化，以限制性核酸內切酶 NotI 及 BamHI 切割並選殖(見圖)至載體 WRG7077 之 NotI BamHI 部位。其可將該基因置於 CMV 啟動基因/插入序列 A 及牛生長激素聚腺苷酸化作用訊號之間。

實例 4：質體：p7077-RT3 選殖株 #A

所欲基因：

可編碼 HIV-1 分化枝 B 菌株 HXB2 之 pol 基因的 RT 部分之合成基因，經最佳化以便於哺乳動物細胞中表現，其係藉 PCR 由重疊寡核苷酸組合而成。經選殖之序列相當於 HXB2 重疊參照序列(GenBank entry K03455)之 2550-4222 位置。為保證其表現性，所選殖序列於 5' 末端具有二個並未存在於原始基因之附加密碼子 -AUG GGC(Met Gly)。

最佳化包含改變病毒基因之密碼子利用模式以獲得接近高度表現人類基因之密碼子頻率，但排除鮮少使用之密碼子。使用名為辛基因(卡耳克基因之更新版本，R.S. 哈爾及 G. 湯普森所撰之 Protein Expression and Purification, 第 12 卷, 185-188 頁, 1998)之統計學形象基礎程式來指定密碼子。

最終之選殖株係由二個中間選殖株 #16 及 #21 所構成。

選殖：

1.7kb之PCR產物以凝膠純化，以NotI及BamHI切割，並經PCR淨化，然後連接至經NotI/BamHI切割之pWRG7077。其可將該基因置於CMV啟動基因及牛生長激素聚腺苷酸化作用訊號之間。將選殖株排序。雖選殖株並非100%正確，但藉由以源自選殖株#21之正確KpnI-BamHI片段取代包含3個錯誤之403bp KpnI-BamHI片段，可校正選殖株#16。藉排序核對最終選殖株。(見於圖5)

實例5：最佳化RT

所欲基因

將可編碼HIV-1分化枝B菌株HXB2之pol基因的RT部分之合成基因(經最佳化以便表現於哺乳動物細胞)自質體p7077-RT3中切下成為1697bp之NotI/BamHI片段，以凝膠加以純化，並選殖入p7313-ie(源自pepC31)之NotI及BamHI部位，將該基因置於愛荷華段落(Iowa length)HCMV啟動基因+表現序列1之下游，及兔子血球蛋白聚腺苷酸化作用訊號之上游。(R7004 p27)(圖6)。

實例6

質體：7077trNef20

所欲基因：

嵌入基因包括源自HIV-1分化枝B菌株HXB2之部分Nef基因。由該基因之5'末端刪除195bp以便移除Nef之前65個胺基酸密碼子。此外將已公開HXB2 nef序列之提前終止密碼子修補(由TAG改為TGG[Trp])，步驟如質體

(33)

發明說明書

p17/24trNEF1 中所述。截短之 nef 序列經由質體 p17/24trNef1 以 PCR 反應放大。選殖之序列相當於 HXB2 參考序列 (GenBank entry K03455) 之 8992-9417 位置。為確保表現性，選殖序列於 5' 末端具有一不存在於原始基因之附加密碼子 -AUG (Met)。

引子：

StrNef(正 向) ATAAGAATGCGGCCGCCATGGTGGGTTTT
CCAGTCACACCTT [序列鑑別編號：1]

AStrNef (反向)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [序列鑑別編號：2]

PCR： 94°C 2分鐘，再接 25個循環： 94°C 30秒、50°C 30秒、72°C 2分鐘，終結 72°C 5分鐘

選殖：

以凝膠純化 455bp 之 RT 的 PCR 產物，以限制性內切酶 NotI 及 Bam HI 切割並連接進入以 NotI/BamHI 切割之載體 WRG7077。其可將該基因置於 CMV 啟動基因 / 插入序列 A 及牛生長激素聚腺苷酸化作用訊號之間。

實例 7

質體：7077RT 8

所欲基因

pol 基因之 RT 部分係源自 HIV-1 分化枝 B 菌株 HXB2。由質體 p7077Pol14 以 PCR 反應將其放大。

所選殖序列相當於 HXB2 參考序列 (GenBank entry K03455)

(34)



之 2550-4234 位置。為確保表現性，選殖序列於 5' 末端具有二個未存在於原始基因之額外密碼子 -AUG GGC(Met Gly)。

引子：

SRT(正 向)ATAAGAATGCGGCCGCCATGGGCCCCATTAGC
CCTATTGAGACT [序列鑑別編號：3]

ASRT(反 向)CGCGGATCCTTAATCTAAAAATAGTACTTTC
CTGATT [序列鑑別編號：4]

PCR：94℃ 2分鐘，再接 25個循環：94℃ 30秒、50℃ 30
秒、72℃ 4分鐘，終結 72℃ 5分鐘

選殖：

將 1720bp 之 RT PCR 產物以凝膠純化，以限制性內切酶 NotI 及 Bam HI 切割並連接入以 NotI/BamHI 切割之載體 WRG7077。其可將該基因置於 CMV 啟動基因 / 插入序列 A 及牛生長激素聚腺苷酸化作用訊號之間。

實例 8

p17/24opt/RT/trNef13 ('Gagopt/RT/Nef')

此構造包含可引起 R 至 H 胺基酸改變之 PCR。

所欲基因：

將源自 HIV-1 分化枝 B 菌株 HXB2 之最佳化 p55gag 基因密碼子之 p17/p24 部分由質體 pGagOPTrpr2 以 PCR 反應放大。RT 編碼序列係由質體 7077RT 8 以 PCR 反應放大。將其中提前終止密碼子已修補 (由 TGA[stop] 改為 TGG[Trp]) 之截短 HXB2 Nef 基因經由質體 7077trNef20 以 PCR 反應放

(35)



大。指定三種具有重疊末端之PCR產物以便三種基因可於第二種PCR中連結。

引子：

(P17/24)

Sp17p24opt(正向)

ATAAGAATGCGGCCGCCATGGGTGCCCCGAGCTTCGGT [序列鑑

別編號：5] ASp17p24optRT連接基因 (反向)

TGGGGCCCATCAACACTCTGGCTTTGTGTC [序列鑑別編

號：6]

PCR： 94℃ 1分鐘，再接20個循環：94℃ 30秒、50℃ 30秒、72℃ 2分鐘，終結72℃ 4分鐘

以凝膠純化1114bp之p17/24opt產物。

(RT)

Sp17p24optRT連接基因 (正向)

CAGAGTGTTGATGGGCCCCATTAGCCCTAT [序列鑑別編

號：7]

ASRTtrNef連接基因 (反向)

AACCCACCATATCTAAAAATAGTACTTTCC [序列鑑別編

號：8]

PCR：如上

以凝膠純化1711bp之RT PCR產物。

(5'截短之nef)

SRTtrNef 連接基因 (正向)

CTATTTTATAGATATGGTGGGTTTTCCAGTCAC [序列鑑別編

(36)

發明說明書

號：9]

AstrNef (反向)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [序列鑑別編號：10]

如上述進行PCR。

以凝膠純化448bp之產物。

三種PCR產物再接於第二次PCR中以引子Sp17/24opt及AstrNef聯結在一起。

PCR：94℃ 1分鐘，再接30個循環：94℃ 30秒、50℃ 30秒、72℃ 4分鐘，終結72℃ 4分鐘

以凝膠純化3253bp之產物，以限制性內切酶NotI及BamHI切割並選殖入載體WRG7077之NotI BamHI部位。此可將該基因置於CMV啟動基因/插入序列A及牛生長生長激素聚腺苷酸化作用訊號之間。

實例9

質體：pGRN#16 (p17/p24opt corr /RT/trNef.)

所欲基因：

由p17/24opt/RT/trNef13('Gagopt/RT/Nef')產生之多蛋白質中可注意到因p24中胺基酸270附近一簇不佳之密碼子而表現出的~30kDa之截短產物。這些密碼子可以藉PCR聯結致突變取代成最佳密碼子。以p17/24opt/RT/trNef13做為模板，以Sp17/p24opt及GTR-A引子放大突變物之Gag 5'部分，並以GTR-S及Asp17/p24optRT連接基因引子放大突變物之Gag 3'部分。使用Sp17/p24opt及Asp17/p24optRT

連接基因引子，將含有改變密碼子之產物重疊部分與經凝膠純化之產物聯結在一起。以 NotI 及 AgeI 切割產物並嵌入至同樣切割之 p17/24opt/RT/trNef13 中以產生 pGRN。證實選植株 #16 並加以發展利用。

引子：

5' PCR：

Sp17p24opt (正向)

ATAAGAATGCGGCCGCCATGGGTGCCCCGAGCTTCGGT [序列鑑別編號：11]

GTR-A(反向)

GCGCACGATCTTGTTTCAGGCCAGGATGATCCACCGTTTATAGATTTCTCC [序列鑑別編號：12]

3' PCR

正向：GTR-S(正向)

ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCC [序列鑑別編號：13]

ASp17p24optRT連接基因 (反向)

TGGGGCCCATCAACACTCTGGCTTTGTGTC [序列鑑別編號：14]

使用 PWO DNA 聚合酶 (羅氏 (Roche)) 時，適合個別產物之 PCR 及聯結的條件係：

95℃ 1分鐘，再接 20個循環：95℃ 30秒、55℃ 30秒、72℃ 180秒，終結 72℃ 120秒並保持於 4℃。

以凝膠純化 1114bp 之產物並以 NotI 及 AgeI 切割以釋出

(38)

明說明明

6647bp 片段，將其以凝膠純化並連接至經 NotI/AgeI 切割、以凝膠純化之 p17/24opt/RT/trNef13，而產生 pGRN#16。

實例 10：

質體： p73i-GRN2 選殖株 #19

(p17/p24(opt)/RT(opt)trNef)-已修補

所欲基因：

最佳化之 gag 密碼子的 p17/p24 部分、密碼子最佳化 RT 及截短之 Nef 基因；後者來自 HIV-1 分化枝 B 菌株 HXB2，其為愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

含有源自質體 p17/24trNef 之 trNef 基因的質體中包含一 PCR 錯誤，因而產生一 R 至 H 胺基酸改變，影響了 nef 末端 19 個胺基酸。其可藉 PCR 致突變加以矯正，經矯正之 nef PCR 可聯結至來自 p7077-RT3 之密碼子最佳化 RT，並以 ApaI 及 BamHI 切割該聯結片段，再選殖入經 ApaI/BamHI 切割之 p73i-GRN。

引子：

使用以下引子自 p7077-RT3 進行 coRT 之 PCR 反應：

(聚合酶 = PWO(羅氏)) 全程。

正向：U1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGT
GCCGGTGAAGCTGAAACCCGGGAT [序列鑑別編號：15]

AScoRT-Nef

(39)



GGTGTGACTGGAAAACCCACCATCAGCACCTTTCTAATCCCCGC

[序列鑑別編號：16]

循環： 95°C (30秒)，再接20個循環：95°C (30秒)、55°C (30秒)、 72°C (180秒)，再接72°C (120秒)並保持於4°C

以凝膠純化1.7kb之PCR產物。

使用以下引子自p17/24trNef1進行5' Nef之PCR反應：

正向：S-Nef

ATGGTGGGTTTTCCAGTCACACC [序列鑑別編號：17]

反向：ASNef-G：

GATGAAATGCTAGGCGGCTGTCAAACCTC [序列鑑別編號：18]

循環： 95°C (30秒)，再接15個循環：95°C (30秒)、55°C (30秒)、72°C (60秒)，再接72°C (120秒)並保持於4°C

使用以下引子自p17/24trNef1進行3' Nef之PCR反應：

正向：SNEF-G

GAGGTTTGACAGCCGCGCTAGCATTTCATC [序列鑑別編號：19]

反向：

AStrNef(反向)

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [序列鑑別編號：20]

循環： 95°C (30秒)，再接15個循環：95°C (30秒)、55°C (30秒)、72°C (60秒)，再接72°C (120秒)並保持於4°C

以凝膠純化PCR產物，並使用5'(S-Nef)及3'(AstrNef)引

(40)



子聯結最初二種 Nef 產物。

循環：95℃ (30 秒)，再接 15 個循環：95℃ (30 秒)、55℃ (30 秒)、72℃ (60 秒)，再接 72℃ (180 秒) 並保持於 4℃。

以 PCR 反應純淨 PCR 產物，並使用 U1 及 AstrNef 引子聯結至 RT 產物：

循環：95℃ (30 秒)，再接 20 個循環：95℃ (30 秒)、55℃ (30 秒)、72℃ (180 秒)，再接 72℃ (180 秒)，並保持於 4℃

以凝膠純化 2.1 kb 之產物，並以 ApaI 及 BamHI 切割。質體 p73I-GRN 亦使用 ApaI 及 BamHI 切割，再以凝膠純化並與 ApaI-Bam RT3trNef 連接以再生 p17/p24(opt)/RT(opt) trNef 基因。

實例 11

p73i-GN2 選植株 #2 (p17/p24opt/trNef)-已修補

所欲基因：

密碼子最佳化 gag 之 p17/p24 部分及截短之 Nef 基因；後者係來自 HIV-1 分化枝 B 菌株 HXB2，位於愛荷華段落 HCMV 啟動基因 + 表現序列 I 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

含源自質體 p17/24trNef1 之 trNef 基因的質體中包含一 PCR 錯誤，因而導致 R 至 H 之胺基酸改變，影響 Nef 末端之 19 個胺基酸。此改變可藉 PCR 致突變矯正，並以 BglII 及 BamHI 切割矯正後之片段，並選殖入經 BglII/BamHI 切割之 p73I-GN 中。(圖 12) 再生矯正後之 p17/p24opt/trNef 融合基因，其為愛荷華段落 HCMV 啟動基因 + 表現序列 I 之下

(41)

發明說明書

游，及兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

使用引子自 p17/24trNef1 進行 5' Nef 之 PCR：

聚合酶 = PWO (羅氏) 全程。

正向：S-Nef

ATGGTGGGTTTTCCAGTCACACC [序列鑑別編號：21]

反向：ASNef-G：

GATGAAATGCTAGGCGGCTGTCAAACCTC [序列鑑別編號：22]

循環：95°C (30秒)，再接 15 個循環：95°C (30秒)、55°C (30秒)、72°C (60秒)，再接 72°C (120秒)，並保持於 4°C

使用引子自 p17/24trNef1 進行 3' Nef 之 PCR：

正向：SNEF-G

GAGGTTTGACAGCCGCCTAGCATTTTCATC [序列鑑別編號：23]

反向：AStrNef

CGCGGATCCTCAGCAGTTCTTGAAGTACTCC [序列鑑別編號：24]

循環：95°C (30秒)，再接 15 個循環：95°C (30秒)、55°C (30秒)、72°C (60秒)，再接 72°C (120秒)，並保持於 4°C

以凝膠純化 PCR 產物，並使用 5'(S-Nef) 及 3'(AstrNef) 引子加以聯結。

循環：95°C (30秒)，再接 15 個循環：95°C (30秒)、55°C (30秒)、72°C (60秒)，再接 72°C (180秒) 並保持於 4°C

以 PCR 反應潔淨 PCR 產物，以 BglII/ BamHI 切割，並以

凝膠純化該367bp片段後選殖入經BglII/BamHI切割、凝膠純化之p73i-GN中。

實例 12

質體：p73I-RT w229k (不活化RT)

所欲基因：

產生一不活化RT基因，其位於愛荷華段落HCMV啟動基因+表現序列1之下游，及兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

由於考慮避免在治療性疫苗中採用活性HIV RT物種，故需將基因不活化。以PCR致突變將RT (源自P73I-GRN2) 胺基酸位置229自Trp改為Lys(R7271 pl-28)即可達此目的。

引子：

使用下列引子進行5' RT之PCR + 突變：

(聚合酶=PWO (羅氏) 全程)

正向：RT3-u：1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGT
GCCGGTGAAGCTGAAACCCGGGAT [序列鑑別編號：25]

反向：AScoRT-Trp229Lys

GGAGCTCGTAGCCCATCTTCAGGAATGGCGGCTCCTTCT [序列
鑑別編號：26]

循環：

1 x [94°C (30秒)]

15 x [94°C (30秒)/55°C (30秒)/72°C (60秒)]

(43)



1 x [72°C (180秒)]

PCR凝膠純化

使用引子進行 3' RT之 PCR+突變：

反向：RT3-1：1

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCA
GCTTGTCGACCTGCTCGTTGCCGC [序列鑑別編號
：27]

正向：ScoRT-Trp229Lys

CCTGAAGATGGGCTACGAGCTCCATG [序列鑑別編號：28]

循環：

1 x [94°C (30秒)]

15 x [94°C (30秒)/55°C (30秒)/72°C (60秒)]

1 x [72°C (180秒)]

PCR凝膠純化

PCR產物以凝膠純化且使用 5' (RT3-U1)及 3' (RT3-L1)引子聯結 RT之 5'及 3'末端。

循環：

1 x [94°C (30秒)]

15 x [94°C (30秒)/55°C (30秒)/72°C (120秒)]

1 x [72°C (180秒)]

PCR產物以凝膠純化，並選殖入 p7313ie 中，利用 NotI 及 BamHI 限制位置，以產生 p73I-RT w229k(見於圖 13)。

實例 13：

質體：p73i-Tgrn (#3)

所欲基因：

密碼子最佳化 gag 之 p17/p24 部分、RT 及截短之 Nef 基因；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其位於愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

包含 RT 活性型之三重融合結構可能因不能取得調節授權而無法供人類使用，故可藉將源自 p73i-RT w229k 之 *NheI* 及 *ApaI* 切割片段嵌入 *NheI/ApaI* 切割 p73i-GRN2#19 (圖 14)，而使 RT 不活化。其可導致 RT 之 229 位置的 W→K 變化。

實例 14

p73I-Tnrg (#16)

所欲基因：

截短之 Nef，不活化之密碼子最佳化 RT 及密碼子最佳化 gag 之 p17/p24 部分；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其位於愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

藉 PCR 反應及 PCR 聯結重新排列由 p73i-Tgrn 所編碼之多蛋白質的基因順序，並產生 p73I-Tnrg (圖 15)。各基因以 PCR 反應放大、以凝膠純化，在以 PCR 聯結基因以形成單一多蛋白質。以凝膠純化該產物，以 *NotI/BamHI* 消化並連接入由 *NotI/BamHI* 所切割之 p7313ie 中。

引子：

trNef PCR

(45)



S-Nef (Not I)

CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [序列鑑別
編號：29]

AS-Nef-coRT連接基因

GATGGGACTGATGGGGCCCATGCAGTTCTTGA ACTACTCCGG
[序列鑑別編號：30]

RTw229k PCR

S-coRT

ATGGGCCCCATCAGTCCCATCGAG [序列鑑別編號：31]

AS-coRT-p17p24 連接基因

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCG
C [序列鑑別編號：32]

p17p24opt PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [序列鑑別編號：33]

AS-p17p24opt (BamHI)

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [序列
鑑別編號：34]

個別產物及使用 VENT DNA聚合酶 (NEB) 聯結之 PCR 條件：

1 x [94°C (30秒)]

25 x [94°C (30秒)/55°C (30秒)/72°C (120秒 [p17p24或 RT]
或 60秒 [trNef])]

1 x [72°C (240秒)]

(46)



PCR 產物以凝膠純化並利用引子 S-trNef(NotI) 及 AS-p17p24opt(BamHI)以用於 PCR 聯結。

1 x [94°C (30秒)]

25 x [94°C (30秒)/55°C (30秒)/72°C (210秒)]

1 x [72°C (240秒)]

以凝膠純化 3000bp 之產物並以 NotI 及 BamHI 切割，以 PCR 法潔淨該產物並連接入至經 NotI/BamHI 消化、凝膠純化之 p7313ie 以產生 p73i-Tnrg。

實例 15：

1. 質體：P73i-Tnrg (#3)

所欲基因：

截短之 Nef 與密碼子最佳化 gag 之 p17/p24 部分及不活化密碼子最佳化 RT 基因；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其位於愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

藉 PCR 反應重新排列以 p73i-Tgrn 編碼之多蛋白質的基因順序以產生 p73I-Tnrg (圖 16)。將密碼子最佳化之 p17/p24 及 RT 製造成單一產物，並藉 PCR 反應聯結至已放大 trNef。產物經凝膠純化、NotI/BamHI 消化，並連接入經 NotI/BamHI 切割之 p7313ie 中。

引子：

P17/p24-RT 3'PCR：

Sp17p24opt (正向)

ATGGGTGCCCCGAGCTTCGGTACTG [序列鑑別編號：35]

(47)



RT3 l: 1 (反 向)

GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCA
GCTTGTCGACCTGCTCGTTGCCGC [序 列 鑑 別 編 號
: 36]

TrNef 5'PCR

S-Nef (NotI)

CATTAGAGCGGCCGCGATGGTGGGTTTTCCAC [序 列 鑑 別
編 號 : 37]

AS-Nef-p17p24

CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGA ACTACTCCG
G [序 列 鑑 別 編 號 : 38]

個 別 產 物 及 使 用 VENT DNA 聚 合 酶 (NEB) 進 行 聯 結 之
PCR 條 件 :

1 x [94°C (30秒)]

25 x [94 °C (30 秒)/55 °C (30 秒)/72 °C (180 秒
[p17p24+RT] 或 60秒 [trNef] 或 210秒 [聯 結])]

1 x [72°C (240秒)]

3000bp之產物經凝膠純化並以NotI及BamHI切割，再以
PCR潔淨並連接入經NotI/BamHI消化、凝膠純化誌之
p73l3ie以產生p73i-Tngr。

實 例 16 :

質 體 : p73I-Trgn (#6)

所 欲 基 因 :

不 活 化 之 密 碼 子 最 佳 化 RT、密 碼 子 最 佳 化 之 gag 的

p17/p24 部分及截短之 Nef 基因；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其位於愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

欲完成該構造內之基因順序，可藉由 PCR 反應分別自質體 p73I-GN2 及 p73I-RTw229k 進行 p17p24-trNef 與 RTw229k 之放大作用。進行 PCR 聯結並將產物以凝膠純化、以 NotI/BamHI 切割，並連接至經 NotI/BamHI 消化之 p7313ie。排序顯示 p17p24 未完全最佳化，因此接著以 AgeI/MunI 由編碼部位切割一 700bp 片段並以源自包含正確編碼序列之 p73i-Tgrn#3 的 MunI/AgeI 片段加以取代。(見圖 17)。

引子：

p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [序列鑑別編號：39]

AstrNef (BamHI)

RTw229k

RT3-U：1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGT

GCCGGTGAAGCTGAAACCCGGGAT [序列鑑別編號：40]

AS-coRT-p17p24opt 連接基因

CAGTACCGAAGCTCGGGCACCCATCAGCACCTTTCTAATCCCCG

C [序列鑑別編號：41]

個別產物及使用 VENT DNA 聚合酶 (NEB) 聯結時之 PCR

(49)

發明說明書

條件：

1 x [94°C (30秒)]

25 x [94°C (30秒)/55°C (30秒)/72°C (120秒 (PCR)
或 180秒 (聯結)]

1 x [72°C (240秒)]

以凝膠純化源自 PCR 反應聯結之 3000bp 產物並以 NotI 及 BamHI 切割，再經 PCR 潔淨並連接入經 NotI/BamHI 消化、以凝膠純化之 p7313ie 以產生 p73i-Tngr。序列分析顯示由 p73I-GN2 所獲得之 p17p24 序列並未完全的密碼子最佳化且此缺陷會被帶入新質體中。欲加以矯正，可自以 MunI 及 AgeI 切割之 p73i-Tngr 中切割一 700bp 片段，並替代連接一源自 p73i-Tgrn 之 700bp 的 MunI/AgeI 消化產物以產生構造 p73I-Tngr#6。

實例 17

質體：p73i-Trng (#11)

所欲基因：

不活化之密碼子最佳化 RT、截短之 Nef 及密碼子最佳化 gag 基因之 p17/p24 部分；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其為愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游，及兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

欲達成構造內之基因順序，可利用 PCR 法自 p73i-Tgrn 放大 RT-trNef 及 p17p24 基因。以 PCR 聯結二個 DNA 片段並以凝膠純化 3kb 之產物且以 NotI/BamHI 切割，再連接經 NotI/BamHI 消化之 p7313ie，可獲得 p73I Trng (#11)。

(50)

發明說明書頁

引子：

RTw229k-trNef

RT3-u：1

GAATTCGCGGCCGCGATGGGCCCCATCAGTCCCATCGAGACCGT

GCCGGTGAAGCTGAAACCCGGGAT [序列鑑別編號：42]

AS-Nef-p17p24opt 連接基因

CAGTACCGAAGCTCGGGCACCCATGCAGTTCTTGA ACTACTCCG

G [序列鑑別編號：43]

P17p24

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [序列鑑別編號：44]

AS-p17p24opt(BamHI)

GATGGGGGATCCTCACAACACTCTGGCTTTGTGTCC [序列

鑑別編號：45]

個別產物及使用 VENT DNA 聚合酶 (NEB) 聯結時之 PCR
條件：

1 x [94°C (30秒)]

25 x [94°C (30秒)/55°C (30秒)/72°C (120秒 (PCR of
基因)或180秒 (聯結)]

1 x [72°C (240秒)]

源自 PCR 聯結之 3000bp 產物經凝膠純化並以 NotI 及
BamHI 切割，再經 PCR 潔淨並連接入經 NotI/BamHI 消
化、凝膠純化之 p7313ie 以產生 p73i-Tngr。

實例 18：

p73i-Tgnr (#f1)

所欲基因：

密碼子最佳化 gag 之 p17/p24 部分、截短之 Nef 及不活化之密碼子最佳化 RT 基因；後者係來自 HIV-1 分化枝 B 菌株 HXB2，其位於愛荷華段落 HCMV 啟動基因 + 表現序列 1 之下游、兔子 β -血球蛋白聚腺苷酸化作用訊號之上游。

欲達成構造內之基因順序，可藉 PCR 法將分別源自質體 p73I-GN2 及 p73I-RTw229k 之 p17p24-trNef 及 RTw229k 加以放大來完成。以 PCR 法進行聯結並將其產物以凝膠純化、以 NotI/BamHI 切割，並與經 NotI/BamHI 消化之 p7313ie 連接。於序列 (p17p24 及 RT) 中可發現二個序列之錯誤，其爾後可利用多蛋白質內之限制位置，取代以基因正確部份來加以修補。(見圖 19)

引子：

p17p24-trNef PCR

S-p17p24opt

ATGGGTGCCCCGAGCTTCGGTACTG [序列鑑別編號：46]

AS-Nef-coRT 連接基因

GATGGGACTGATGGGGCCCATGCAGTTCTTGA ACTACTCCGG

[序列鑑別編號：47]

RTw229k

S-coRT

ATGGGCCCCATCAGTCCCATCGAG [序列鑑別編號：48]

RT3-1：1

(52)



GAATTCGGATCCTTACAGCACCTTTCTAATCCCCGCACTCACCA
GCTTGTCGACCTGCTCGTTGCCGC [序列鑑別編號：49]

個別產物及使用 VENT DNA 聚合酶 (NEB) 聯結時之 PCR 條件：

1 x [94°C (30秒)]

25 x [94°C (30秒)/55°C (30秒)/72°C (120秒 (PCR) 或 180秒 (聯結)]

1 x [72°C (240秒)]

以凝膠純化 3000bp 產物並以 NotI 及 BamHI 切割，藉 PCR 反應潔淨並連接至以 NotI/BamHI 消化、凝膠純化之 p73l3ie 以產生 p73i-Tngr。排序可顯示 p17p24 並未徹底最佳化，故再以 AgeI/MunI 自編碼部位切割一 700bp 片段，並以源自 p73i-Tgrn#3 之包含正確編碼序列的 MunI/AgeI 片段加以取代。該多蛋白質亦包含單一點突變 (G2609A)，其導致多蛋白質之 RT 部分中之胺基酸 Thr 被取代成 Ala。藉 ApaI/BamHI 消化其構造並藉 PCR 反應清理以移除產生突變之序列可矯正該突變，其可經由聯結經 ApaI/BamHI 消化、源自 p73i-Tngr 之 RT 部分來加以取代。

實例 19：

製備可供 '基因槍' DNA 藥匣使用之質體 - 塗覆 '金漿 (gold slurry)'

將質體 DNA (約 1 微克 / 微升、如 100 微克) 及 2 微米金製顆粒 (如 50 毫克，包德傑克特 (PowderJect)) 懸浮於 0.05M 亞精

胺(如100微升)中(西格馬(Sigma))。添加1M之 CaCl_2 (如100微升(American Pharmaceutical Partners, Inc., USA))後，DNA沉澱至金製顆粒上。將DNA/金複合物於室溫培養10分鐘，以絕對酒精中沖洗3次，如3 x 1毫升，(預先於分子篩3A(BDH)中乾燥)。將樣本再懸浮於含0.05毫克/毫升之聚乙炔吡咯烷酮(PVP，西格馬)之絕對酒精中，並分為三等份裝入1.5毫升微量離心管(伊本多夫(Eppendorf))中。對該等份進行分析項目為：(a) '金漿'，(b) 自(a)洗脫之洗脫物-質體及(c) 製備供'基因槍'(見下列3組實例)使用之金製/質體塗覆鐵福傑(Tefzel)藥匣。為製備樣本(a)及(b)，將裝有含質體DNA/'金漿'之乙醇/PVP之離心管置於伊本多夫5418微量離心機中以最高速旋轉2分鐘，移除上清液並將'金漿'於室溫中乾燥10分鐘。將樣本(a)再懸浮於pH 8.0之含0.5-1.0微克/微升之質體DNA的TE中，假設約有50%塗佈率。為供洗脫之用，將樣本(b)再懸浮於pH 8.0之含0.5-1.0微克/微升質體DNA之TE中，並於37°C培養30分鐘，劇烈搖晃，接著於伊本多夫5418微量離心機中以最高速旋轉2分鐘並移除上清液、洗脫物並貯存於-20°C。洗脫之確切DNA濃度藉分光光度定量法使用基因定量(Genequant) II (法馬西亞生物科技(Pharmacia Biotech))測定。

實例 20：

DNA免疫藥匣之製備

如先前所述方法製備可供亞克協(Accell)基因轉移裝置

使用之藥匣(伊森布蘭(Eisenbraun)等人, DNA and Cell Biology, 1993, 第12卷 No 9 第791-797頁; 波特納(Pertner)等人)。簡單地說, 將質體DNA包覆於2微米金製顆粒上(迪高沙公司(DeGussa Corp.), South Plainfield, N.J., USA)並裝入鐵福傑管中, 再切割為1.27公分長度以作為藥匣並, 使用前皆於4°C脫水貯藏。在典型的接種時, 各藥匣包含0.5毫克的金, 其上塗覆以每藥匣總量為0.5微克之DNA。

實例 21:

使用基因槍接種DNA後對HIV抗原之免疫反應。

以由核酸編碼並置於二種載體中之抗原接種鼠(n=3/組)。P7077利用HCMV IE 啟動基因, 其包含插入序列A及表現序列1(fcmv啟動基因)。P73I雖遞送相同抗原, 但所含之HCMV IE啟動基因(iamv啟動基因)中缺少插入序列A, 僅包含表現序列1。

將質體遞送至已剃毛之F1(C3H x Balb/c)鼠腹部皮膚標的部位。於第0日對鼠投予首次免疫之2 x 0.5微克DNA, 於第35日以2 x 0.5微克DNA補強並於第40日使用IFN- γ 伊力司帕特(Elispot)來檢測細胞反應。

- P73I -空白載體
- P7077 -空白載體
- P7077 GRN-(f CMV啟動基因)Gag、RT、Nef
- P73I GRN-(i CMV啟動基因)Gag、RT、Nef
- P73I GR3N-(CMV啟動基因)最佳化Gag、最佳化 RT、Nef

(55)



- P7077 GN-(f CMV啟動基因)Gag、Nef
- P731 GN-(i CMV啟動基因)Gag、Nef

殺手T細胞反應

藉CD8+T細胞-限制性IFN- γ 伊力司帕特法分析5日後所收集之脾細胞以評估殺手T細胞之反應。鼠以頸椎脫臼法殺死並將脾臟採集至冰冷PBS中。將脾細胞理出至磷酸鹽緩衝食鹽水(PBS)中，再使紅血球細胞溶血(置於由155毫當量NH₄Cl、10毫當量KHCO₃、0.1毫當量EDTA所組成之緩衝溶液中1分鐘)。於PBS中沖洗二次以移除微粒狀物質後，將單次細胞懸浮液等分入伊力司帕特試盤中，試盤預先以捕獲IFN- γ 抗體塗覆並以CD8-限制性同源肽(Gag, Nef或RT)刺激。於培養隔夜後，藉應用抗-鼠科動物IFN- γ -生物素標記抗體(法名健(Pharmingen))，再用抗生蛋白鏈菌素-結合鹼性磷酸酶並使用圖像分析定量後可見IFN- γ 製造細胞。

實驗結果顯示於圖20、21及22。

實例22：

疫苗結構之免疫原性

1.細胞分析

細胞免疫反應包括殺手CD8細胞及輔助CD4細胞。可檢測特定CD8及CD4細胞之敏感方法為伊力司帕特分析法，其可定量能分泌干擾素- γ 或IL-2之細胞數量。伊力司帕特分析係仰賴捕捉由個別細胞分泌之細胞動力素。簡單地說，特定化之微滴定盤以經抗-細胞動力素抗體所塗

覆。由經免疫動物所分離之脾細胞與代表已知抗原決定位 (CD8) 或蛋白質 (CD4) 之特定肽一起下培養隔夜。若受刺激之細胞可釋出細胞動力素，其將與環繞各製造細胞周圍之試盤表面的抗體相結合。於細胞溶解並沖洗試盤後，餘留之細胞動力素仍依附於塗佈抗體上。該分析法之發展係類似 ELISA 分析法，其使用生物素/卵白素放大系統。斑點數量相當於細胞動力素製造細胞數量。

記錄在所有 6 種結構之中，CD8 對下列 K2^d-限制性鼠科動物抗原決定位之反應：Gag(AMQMLKETI)、Nef(MTYKAAVDL)

及 RT(YYDPSKDLI)，以及 CD4 對 Gag 及 RT 蛋白質之反應。將這些分析結果經統計分析並依其免疫原性將各結構評等。結果顯示於圖 23。

2. 體液分析

由二個試驗中於補強後第 7 日及第 14 日採集血液樣本供進行抗體分析。分離血清並冷凍貯存等待使用特定 ELISA 分析測量抗體力價。測試所有樣本中對 Gag、Nef 及 RT 之抗體。簡單地說，ELISA 試盤先以相關蛋白質塗覆。洗掉過量蛋白質後將稀釋血清樣本於孔槽中培養。沖掉血清樣本並添加結合適當標記之抗-鼠抗血清。培養試盤後利用試盤辨讀器進行辨認。結果顯示於圖 24。

3. 抗體數據

測試四個試驗中所有的六種抗體之力價。結構 p73i-GNR 對 Gag 始終不產生抗體反應並僅有限度地對 Nef 之抗體反

應。因自同一鼠所分離之脾細胞中可觀察到T-細胞反應，代表Gag蛋白質可表現於其體內，故此現象原因不明。

產生Gag特定抗體之評等如下：

RNG>GRN>NRG>RGN>NGR>GNR

分析細胞免疫學數據

目標為根據墨點計數資料自3項免疫學試驗排定6種結構之等級。共評估三組反應：

CD8於第7日對Gag、Nef及RT之反應(首次接種後第7日)，

CD4於第35日對Gag及RT之反應(補強後第7日)，

CD8於第35日對Gag、Nef及RT之反應(補強後第7日)。

使用SAS第8版之線性混效模型(linear mixed effect model)來塑造各反應(如CD8對Gag之反應)模型。該模型包含構造之固定效果，不論特定抗原(Gag、Nef或RT)及IL-2是否存在。此外，對CD8反應而言，每一隻鼠之資料皆可提供研究，故該模型中可包含研究對象成為隨機影響因素。該模型中所包含之交互作用可供研究各種抗原(存在/不存在)及IL-2(存在/不存在)之組合在各結構中之不同作用效果。

由該模型中，分別評估針對各種抗原(存在/不存在)及IL-2(存在/不存在)之組合而在各結構與p7313(控制組)間所產生經調整平均之反應差異，並以p-值指出該差異是否具有統計意義。在有抗原但缺少IL-2之情況下，根據差異與p-值來評等各結構，將具有最大差異之構造評等為6，差異次大者為5，依此類推，而0代表任何差異未達統計上

有意義之 5% 水準的結構。

該模型之假想為若變因固定的話，殘留物應正常分佈，故可採用圖解方法及敏感性分析來分析其真實性，其中須建立該反應之第一種對數及第二種平方根變換模型。結構之評等並未敏感到可偏離此模型假想。

在分別計算各實驗中各反應之評等後，計算所有 3 個實驗中 3 組反應之總評等。下表顯示 3 個實驗之總評等。

所有 3 個實驗中 3 組反應中各結構之總評等。

構造	第 7 日(首次接種後第 7 日)	第 35 日(補強後第 7 日)	
	CD8	CD4	CD8
GRN	5	18	3
GNR	17	24	28
RGN	28	23	33
RNG	25	27	37
NRG	25	19	0
NGR	4	14	10

RNG 在第 35 日之兩組反應中評等最高，而在第 7 日組中排名第二，僅次於 RGN。RGN 在第 35 日之兩組反應中亦得到很高的評等。

圖式簡單說明

- 圖 1 P7313-ie
- 圖 2 pGagOptrpr2 中 p55 gag 插入基因之序列
- 圖 3 p17/24trNEF1 中 p17/24trNEF 插入基因之序列
- 圖 4 p17/24opt/trNef1 中 p17/24opt/trNEF 插入基因之序列
- 圖 5 p7077-RT3 中 RT 插入基因之序列

- 圖 6 p73i-RT3 中 編 碼 插 入 基 因 之 序 列
- 圖 7 7077trNef20 中 Nef 插 入 基 因 之 序 列
- 圖 8 7077RT8 中 RT 插 入 基 因 之 序 列
- 圖 9 p17/24opt/RT/trNef13 中 p17/24opt/RT/trNef 插 入 基 因 之
序 列
- 圖 10 WRG7077 中 p17/p24opt(cor)/RT/trNef 編 碼 插 入 基 因 之
序 列
- 圖 11 p73i-GRN2 中 p17/p24(opt)/RT(opt)trNef 插 入 基 因 之 序 列
- 圖 12 p73i-GRN2 中 p17/p24opt/trNef 插 入 基 因 之 序 列
- 圖 13 RT w299k 之 序 列
- 圖 14 p73i-Tgrn
- 圖 15 Tnrg 之 序 列
- 圖 16 p73l3ie 中 Tngr 插 入 基 因 之 序 列
- 圖 17 插 入 基 因 Trgn#6 之 序 列
- 圖 18 Trng#11 插 入 基 因 之 序 列
- 圖 19 TgnR (F1) 之 序 列
- 圖 20 使 用 INF- γ 酵 素 聯 免 疫 斑 點 法 (ELIspot)， 測 量 補 強 5
日 後 對 Gag 肽 之 反 應
- 圖 21 使 用 INF- γ 酵 素 聯 免 疫 斑 點 法 (ELIspot)， 測 量 補 強 5
日 後 對 Nef 肽 之 反 應
- 圖 22 使 用 INF- γ 酵 素 聯 免 疫 斑 點 法 (ELIspot)， 測 量 補 強 5
日 後 對 Rt 肽 之 反 應
- 圖 23 本 實 例 係 為 老 鼠 模 式 中， 在 使 用 由 Gag、RT 及 Nef
以 三 聯 融 合 體 方 式 所 構 成 之 重 組 結 構 進 行 免 疫 後，

(60)



檢 測 其 CD8 免 疫 反 應 。

圖 24 使 用 ELISA 測 量 第 35 日 (補 強 7 日 後) 之 Gag 抗 體

肆、中文發明摘要

本發明係提供一種核苷酸序列，其可編碼 HIV-1 gag 蛋白質或其包含 gag 抗原決定位之片段及第二種 HIV 抗原或可編碼該第二種 HIV 抗原之抗原決定位的片段，其可操作連接至異種啟動基因。較佳之多核苷酸序列可進一步編碼 nef 或其片段及 RT 或其片段。

伍、英文發明摘要

The invention provides a nucleotide sequence that encodes an HIV-1 gag protein or fragment thereof containing a gag epitope and a second HIV antigen or a fragment encoding an epitope of said second HIV antigen, operably linked to a heterologous promoter. Preferred polynucleotide sequences further encodes nef or a fragment thereof and RT or a fragment thereof.

拾壹、圖式

公告本

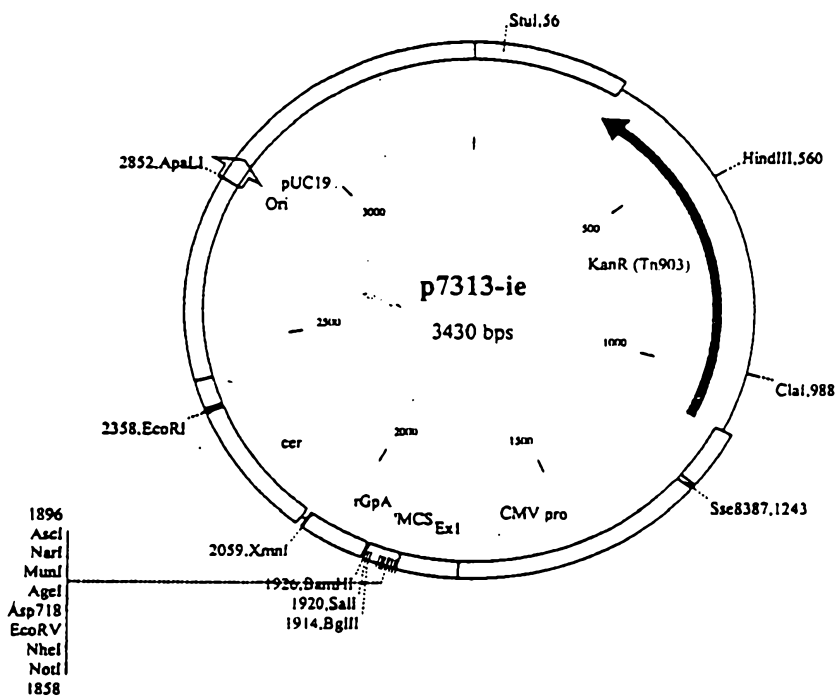


圖 1

pGagOptrpr2 中 p55 gag 插入基因之序列

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTGCAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGGCCGAAGCCATGAGCCAGGTGACGAACTCCGCAACCATCATGATGCAGAGA
 GGGAACTTCCGCAATCAGCGGAAGATCGTGAAGTGTTCATTTGCGGCAAGGAGGGT
 CATAACGCCCCGCAACTGTCGGGCCCCCTAGGAAGAAAGGGTGTGGAAGTGCGGCAAG
 GAGGGACACCAGATGAAAGACTGTACAGAACGACAGGCCAATTTTCTTGGAAGATT
 TGGCCGAGCTACAAGGGGAGACCTGGTAATTTCTGCAAAGCAGGCCCGAGCCCACC
 GCCCCCCTGAGGAATCCTTCAGGTCCGGAGTGGAGACCACAACGCCTCCCCAAAAA
 CAGGAACCAATCGACAAGGAGCTGTACCCTTTAACTTCTCTGCGTTCTCTCTTTGGC
 AACGACCCGTCGTCTCAATAA [序列鑑別編號: 50]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTMLNT VGGHQAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNPANPDC KTI LKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLAEAMSQV TNSATIMMQR GNFRNQRKIV KCFNCGKEGH
 TARNCRAPRK KGCWKCGKEG HQMKDCTERQ ANFLGKIWPS YKGRPGNFLQ
 SRPEPTAPPE ESFRSGVETT TPPQKQEPID KELYPLTSLR SLFGNDPSSQ

*

[序列鑑別編號: 51]

圖2

p17/24trNEF1 中 p17/24trNEF 插入基因之序列

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATT
CGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGC
AGGGAGCTAGAACGATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGT
AGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAAGCTTAGA
TCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAA
GACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAA
GCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTGAGCCAAAATTACCCT
ATAGTGCAGAACATCCAGGGGGCAAATGGTACATCAGGCCATATCACCTAGAAGCTTTA
AATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTTCAGCCCAGAAGTGATACCCATG
TTTTTCAGCATTATCAGAAGGAGCCACCCCAAGATTTAAACACCATGCTAAACACA
GTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCT
GCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATG
AGAGAACCAAGGGGAAGTGACATAGCAGGAAGTACTAGTACCCTTCAGGAACAAATA
GGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATTTATAAAAGATGGATA
ATCCTGGGATTAAATAAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATA
AGACAAGGACCAAAAGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTA
AGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTC
CAAATGCGAACCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACA
CTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGA
GTTTTTGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAG
GCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATT
CACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTAC
TTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTT
GGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAA
GGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAG
AGAGAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGA
GAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [序列鑑別編號: 52]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTMLNT VGGHQAMQM
LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGP EPFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
QGVGGPGHKA RVLVGFPVTP QVPLRPMTYK AAVDLSHFLK EKGGLEGLIH
SQRRQDILD DL WIYHTQGYFP DWQNYTPGPG VRYPLTFGWC YKLVPVEPDK
VEEANKGENT SLLHPVSLHG MDDPEREVLE WRFDSHLAFH HVARELHPEY
FKNC*

[序列鑑別編號: 53]

圖3

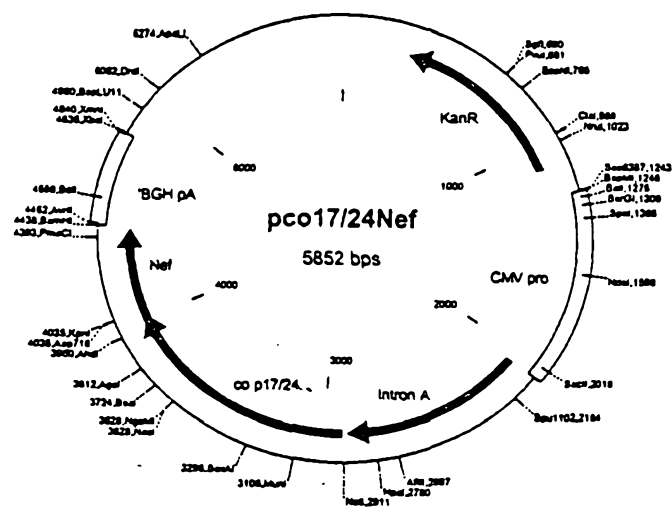
p17/24opt/trNef1 中 p17/24opt/trNef 插入基因之序列

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGFCGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
 AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTA
 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
 TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACC
 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
 AAAGGAGAGAACACCAGCTTGTTACACCTGTGAGCCTGCATGGGATGGATGACCCG
 GAGAGAGAAGTGTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCC
 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [序列鑑別編號: 54]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAISPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIIILGL NKIVRMYSP SILDIRQGP EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTIILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVL MVGF PVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
 HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYP LTFGW CYKLVPVEPD
 KVEEANKGEN TSL LHPVSLH GMDDPEREVL EWRFD SHLAF HHVARELHPE
 YFKNC*

[序列鑑別編號: 55]

圖4

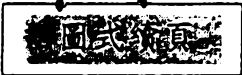


p7077-RT3 中 RT 插入基因之序列:

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTCGGAAGCAGAACCCEGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
 CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGCACGGCGTGTACTATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
 ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
 GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
 GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAAG
 GCCGGGTACGTACCAACCGGGGCCCGCAGAAGGTGCTCACCTGACCGACACCACC
 AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTACGGCCCAGCCGGAC
 CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
 GTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTGAC
 AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [序列鑑別編號: 56]

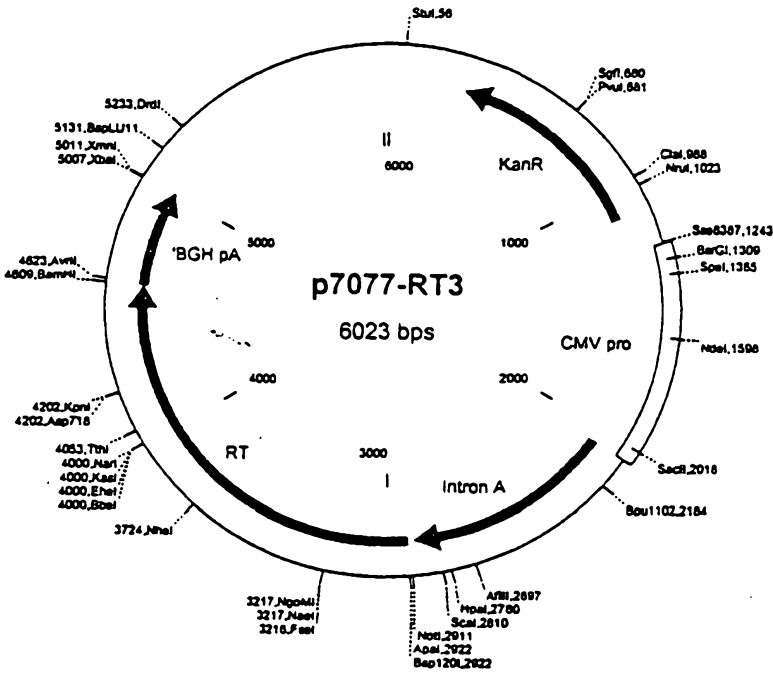
MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
 GLKKKKSVTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL YVGSLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
 ETKLGKAGYV TNRGRQKVV LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIIEQLIKK EKVYLAWVPA HKGIGGNEQV

圖 5



DKLVSAGIRK VL*

[序列鑑別編號: 57]





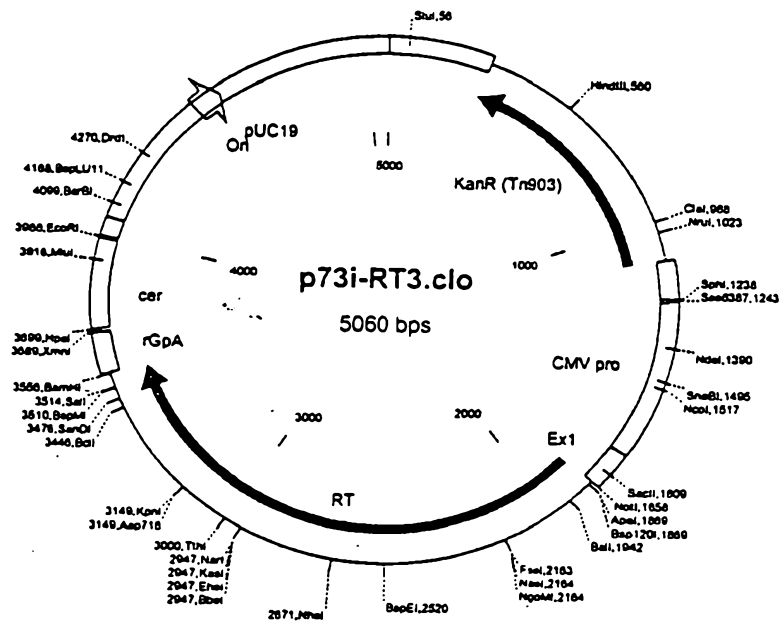
p73i-RT 中編碼插入基因之序列:

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
CCGTTTTCGGAAGCAGAACCCTGATATCGTCATCTACCAGTACATGGACGACCTGTAC
GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
TTCCTGTGGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTC
CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGCGCTGTACTATGACCCC
TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
GCCCCACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAA
GGCCGGGTACGTACCAACCGGGGCCGCGCAGAAGGTCTGCACCCTGACCGACACCAC
CAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGA
GGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTCAGGCCCAGCCGGA
CCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAA
GGTCTACCTCGCCTGGGTCCCGGCCCATAGGGCATTGGCGGCAACGAGCAGGTCTGA
CAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [序列鑑別編號: 58]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
GLKKKKSVMV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL YVSDLEIGQH
RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
ETKLKGAGYV TNRGRQKVV LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
SQYALGIIQA QPDQSESELV NQIIIEQLIKK EKVYLAWVPA HKGIGGNEQV

圖 6

DKLVSAGIRK VL*
[序列鑑別編號: 59]



7077trNef20中Nef插入基因之序列:

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA
GCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC
TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC
CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGA
TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGA
GAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA
GAAGTGTTAGAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCCGAGAG
CTGCATCCGGAGTACTTCAAGAACTGCTGA [序列鑑別編號: 60]

MVGFPVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDLWI
YHTQGYFPDW QNYTPGPGVR YPLTFGW CYK LVPVEPKVE EANKGENTSL
LHPVSLHGMD DPEREVLEWR FDSVLAFHHV ARELHPEYFK NC*
[序列鑑別編號: 61]

圖7

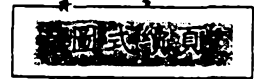
7077RT8中RT 插入基因之序列

ATGGGCCCCATTAGCCCTATTGAGACTGTGTTCAGTAAAATTAAAGCCAGGAATGGAT
 GGCCCAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAA
 ATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCA
 TACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTA
 GTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGA
 ATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGT
 GATGCATATTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCATTTACC
 ATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCA
 CAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAG
 CCTTTTAGAAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTAT
 GTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAA
 CATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCA
 TTCCTTTGGATGGGTTATGAACCTCATCTGTATAAATGGACAGTACAGCCTATAGTG
 CTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGGAAATTG
 AATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAACTCCTT
 AGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAA
 CTGGCAGAAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCA
 TCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAA
 ATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGT
 GCCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAA
 AGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAGGAAACA
 TGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGTAGTGGGAGTTT
 GTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCTAGTA
 GGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAA
 GCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACA
 AATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAA
 GTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGAT
 CAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAG
 GTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGAT
 AAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTTAGATTAA

[序列鑑別編號: 62]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
 GLKKKKS MTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL Y VGSDLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
 ETKLGKAGYV TNRGRQKVVT LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLA WVP A HKGIGGNEQV
 DKLVSAGIRK VLFLD*

圖 8



[序列鑑別編號: 63]

p17/24opv/RT/trNef13 中 p17/24opv/RT/trNef 插入基因之序列

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTGTGCGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAATTAAGCCAGGA
 ATGGATGGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAGCATTA
 GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
 AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGA
 AAATTAGTAGATTTTACAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
 TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
 GTGGGTGATGCATATTTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
 TTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTG
 CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATC
 TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGAT
 TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAATAATAGAGGAGCTG
 AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAA
 CCTCCATTCTTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
 ATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGG
 AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAA
 CTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAG
 CTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTAT
 GACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACA
 TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATG
 AGGGGTGCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACC
 ACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTGCCCATACAAAAG
 GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGG

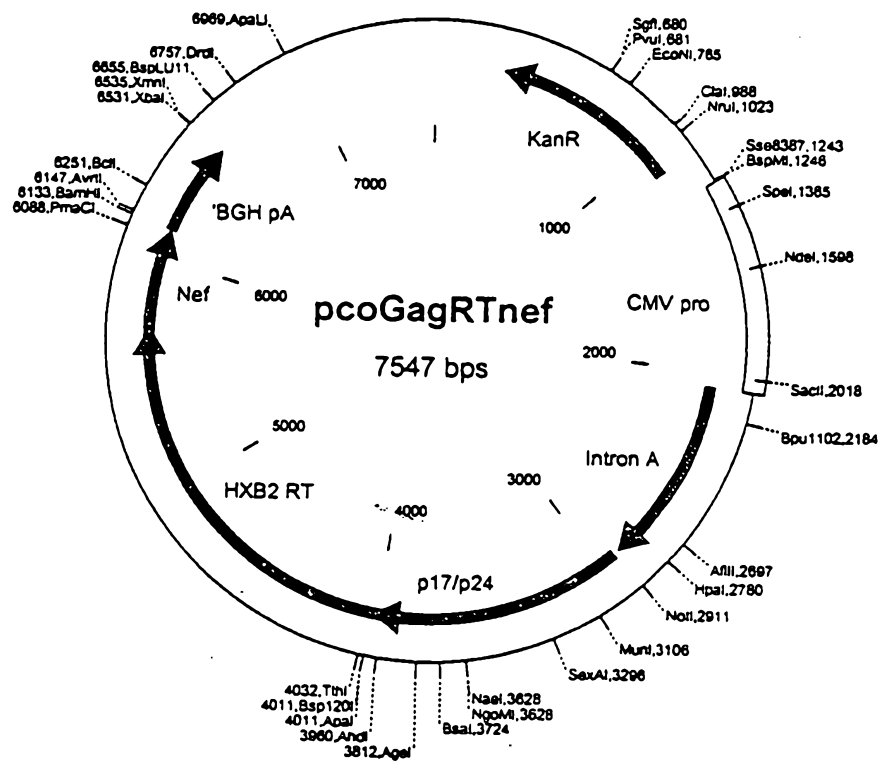
圖 9



GAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC
 ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGAC
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGA
 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAA
 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA
 GTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTAGATATGGTGGGT
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT
 CTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGA
 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG
 CAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTAC
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC
 AGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTA
 GAGTGGAGGTTTGACAGCCACCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCG
 GAGTACTTCAAGAACTGCTGA [序列鑑別編號: 64]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAISPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPKEPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNaNPDC KTIKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIIQYMD
 DLYVGSdleI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGQG
 QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWVK
 TPKFKLPIQK ETWETWWTEY WQATWIPEWE FVNTPLVLKL WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPQDSES ELVNQIIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLFLDMV GFPVTPQVPL RPMTYKAAVD
 LSHFLKEKGG LEGLIHSQRR QDILDLIWIYH TQGYFPDQWN YTPGPGVRY
 LTFGWCYKLV PVEPDKVEEA NKGENTSLLH PVSLHGMDDP EREVLEWRFD
 SHLAFHHVAR ELHPEYFKNC *

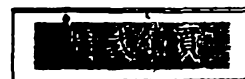
[序列鑑別編號: 65]



WRG7077 中 p17/24opt(cor)/RT/trNef 編碼插入基因之序列:

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGATATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCTCAAAACATTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATCCTGGGCCTGAACAAGATCGTGCGCATGTAATCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTGAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGGCCCCATTAGCCCTATTGAGACTGTGTGAGTAAATTAAGCCAGGA
 ATGGATGGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAGCATTAA
 GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAA
 AATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGA
 AAATTAGTAGATTTTCAAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAA
 TTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGAT
 GTGGGTGATGCATATTTTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAATATACTGCA
 TTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTG
 CTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATC
 TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGAT
 TTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTG
 AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAACATCAGAAAGAA
 CCTCCATTTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCT
 ATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGG
 AAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAA
 CTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATAACCACTAACAGAAGAAGCAGAG
 CTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTAT
 GACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACA
 TATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAATATGCAAGAATG
 AGGGGTGCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACC
 ACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTGCCCATACAAAAG
 GAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGG
 GAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCC

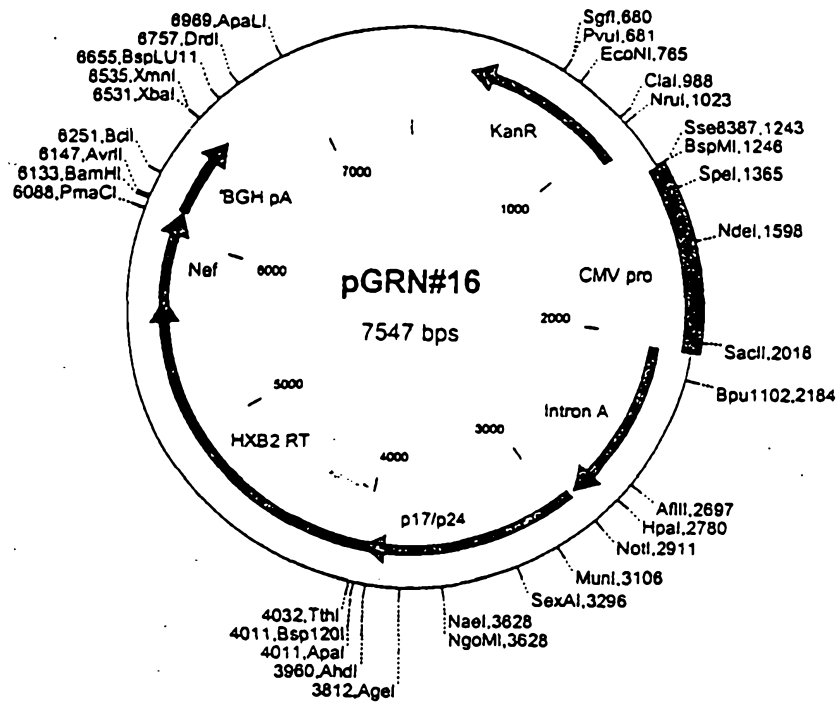
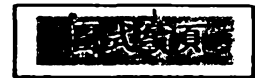
圖 10

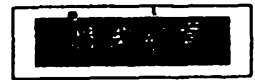


ATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTA
 GGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGAC
 ACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTTCGGGA
 TTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAA
 CCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAG
 GAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA
 GTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTATAGATATGGTGGGT
 TTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT
 CTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGA
 AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGG
 CAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTAC
 AAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACC
 AGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTA
 GAGTGGAGGTTTGACAGCCACCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCG
 GAGTACTTCAAGAACTGCTGA [序列鑑別編號: 66]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNMTMLNT VGGHQAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMR PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNaNPDC KTIKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVYQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYDPSKDL IAEIQKQGQG
 QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIVIWGK
 TPKFKLPIQK ETWETWTEY WQATWIPEWE FVNTPLVLK WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPQSES ELVNQIIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLFLDMV GFPVTPQVPL RPTYKAAVD
 LSHFLKEKGG LEGLIHSQRR QDILDWIYH TQGYFPDWQN YTPGPGVRY P
 LTFGWCYKLV PVEPDKVEEA NKGENTSLLH PVSLHGMDDP EREVLEWRFD
 SHLAFHHVAR ELHPEYFKNC *

[序列鑑別編號: 67]

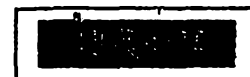




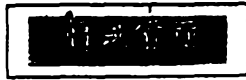
p73i-GRN2 中 p17/p24(opt)/RT(opt)trNef 插入基因之序列:

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGA
GAAAATTAGGCTGCGCCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCAT
ATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCT
GCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAGCCAT
CCCTCCAGACCGGGAGTGAAGAGCTGAGGTCCTTGTATAACACAGTGGCT
ACCCTCTACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGC
CTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAGGCCAG
CAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCC
TATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCC
GGACGCTCAATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCT
GAGGTTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGAC
CTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATGCAAAT
GTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCAT
CCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAGCCTCGCGG
CTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATCGGATGGA
TGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTT
GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTT
TTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGA
TGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATC
TTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTG
TCAGGGAGTAGGCGGACCCGGACACAAAGCCAGAGTGTTGATGGGCCCC
ATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGG
CCCCAAGGTCAAGCAGTGGCCACTACCGAGGAGAAGATCAAGGCCCTG
GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCG
GGCCTGAGAACCCATACAACACCCCCGTGTTTGCCATCAAGAAGAAGGAC
AGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGAC
CCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCCTGA
AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGC
GTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCT

圖 11



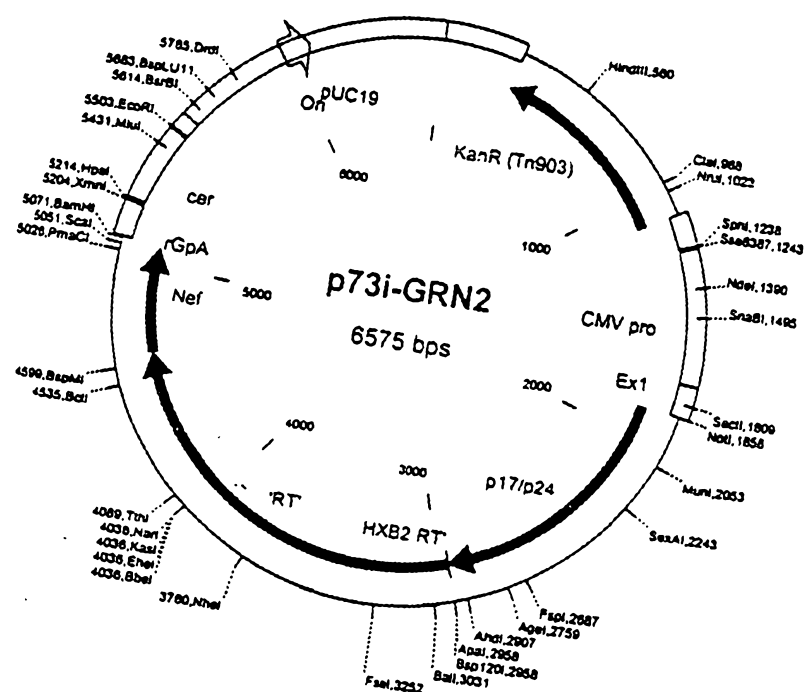
ATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCCA
GGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCC
TGGAGCCGTTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATG
GACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAA
GATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGG
ACAAGAAGCATCAGAAGGAGCCGCCATTCTGTGGATGGGCTACGAGCTC
CATCCCGACAAGTGGACCGTGCAGCCTATCGTCCTCCCGAGAAGGACAG
CTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTA
GCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGC
GGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCT
CGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGCACGGCGTGT
ACTATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAG
GGGCAGTGGACATAACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAGAC
CGGCAAGTACGCCCCGCATGAGGGGCGCCACACCAACGATGTCAAGCAG
CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGG
GAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGT
GGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTCGTG
AATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCAT
CGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACA
AAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCGCGCCAGAAGGTCG
TCACCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTAT
CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCA
GTACGCGCTGGGCATTATTCAGGCCAGCCGGACCGAGTCCGAGAGCGAAC
TGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTC
GCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAGGTCGACAA
GCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCA
CACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT
AGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATCACTCCCA
AAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACT
TCCCTGATTGGCAGAACTACACACCAGGGGCCAGGGGTCAGATATCCACTG
ACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGA
AGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGC

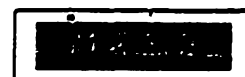


ATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAG
CCGCCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCA
AGAACTGCTGA [序列鑑別編號: 68]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSP T SILDIRQGPK EPFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
QGVGGPGHKA RVLMPISPI ETVSVKLKPG MDGPKVKQWP LTEEKIKALV
EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVIQYMD
DLVVGSDLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
PDKWTVQPIV LPEKDSWTVN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGGQ
QWTYQIYQEP FKNLKTGKYA RMRGAHTNDV KQLTEAVQKI TTESIWIWVK
TPKFKLPIQK ETWETWWTEY WQATWIPEWE FVNTPLVLKL WYQLEKEPIV
GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
LQDSGLEVNI VTDSQYALGI IQAQPQDSES ELVNQIIEQL IKKEKVYLAW
VPAHKGIGGN EQVDKLVSAG IRKVL MVGFP VTPQVPLRPM TYKAAVDLSH
FLKEKGGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
GWCYKLVPE PDKVEEANKG ENTSLHPVS LHGMDDPERE VLEWRFD SRL
AFHHVARELH PEYFKNC*

[序列鑑別編號: 69]



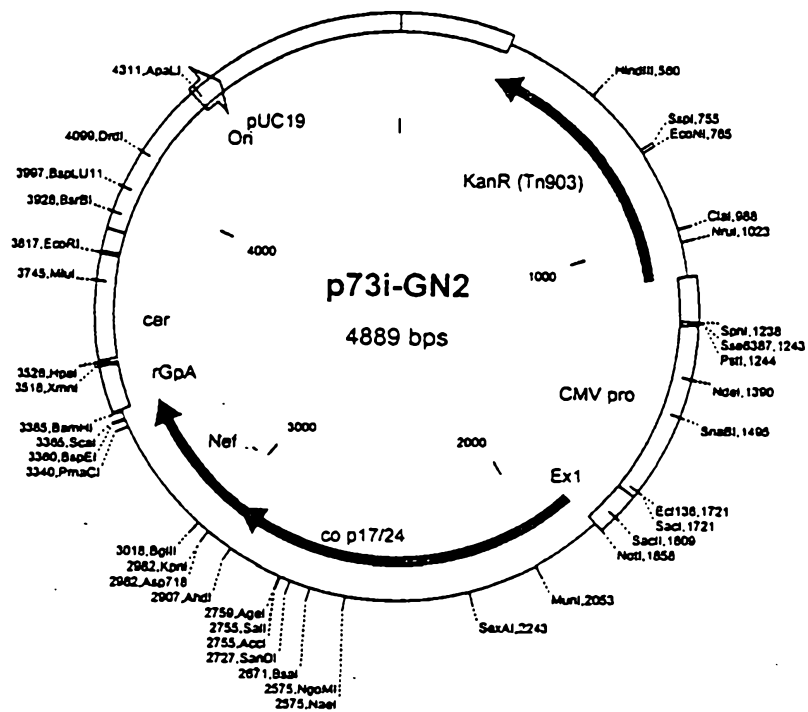
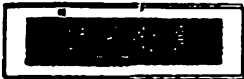


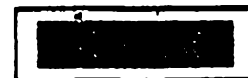
p73i-GN2 中 p17/p24opt/trNef 插入基因之序列:

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
 AGGCTGCGCCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
 AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
 CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
 TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
 GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
 GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
 ATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACGCTC
 AATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
 TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
 GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
 GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
 CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATC
 GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
 ATTCTCGGTCTCAATAAAATTGTTAGAATGTACTCTCCGACATCCATCCTTGACATT
 AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
 CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
 CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
 CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
 GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
 AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTA
 ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
 TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
 AAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCG
 GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCTAGCATTTTCATCACGTGGCC
 CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA [序列鑑別編號: 70]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSP SILDIRQGP EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTIKALGPA ATLEEMMTAC
 QGVGGPGHKA RVL MVGF PVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
 HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYP LTFGW CYKLVPVEPD
 KVEEANKGEN TSL LHPVSLH GMDDPEREVL EWRFD SRLAF HHVARELHPE
 YFKNC* [序列鑑別編號: 71]

圖 12

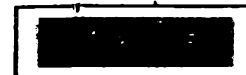




RT w229k 之序列:

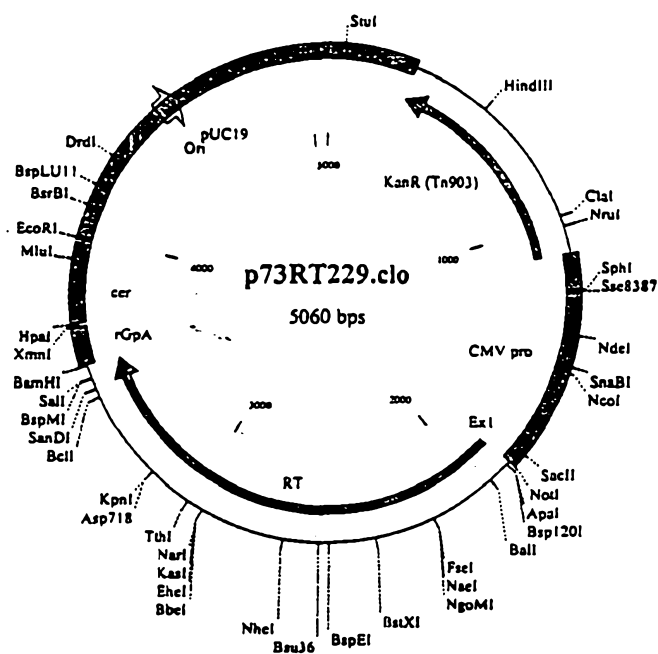
ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGGCCC
CAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCG
AGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCATACAACACCCCCGTG
TTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCGGGAGCTGAA
TAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCATCCGGCCGGCCTGAAGA
AGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAG
GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAG
ATATCAGTACAACGTCCTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCAGAGCTCCA
TGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATG
GACGACCTGTACGTGGGCTCTGACCTGGAATCGGGCAGCATCGCACGAAGATTGAGGAGCT
GAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGC
CATTCTG_{aa}GATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCTATCGTCCTC
CCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGC
TAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGG
CCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAG
ATCCTGAAGGAGCCCGTGACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAAT
CCAGAAGCAGGGCCAGGGGCAGTGGACATAACCAGATTTACCAGGAGCCTTCAAGAACCTCA
AGACCGGCAAGTACGCCCGCATGAGGGGCGCCACACCAACGATGTCAAGCAGCTGACCGAG
GCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCT
GCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATT
CCGAGTGGGAGTTTCGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAG
CCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGG
GAAGGCCGGGTACGTCACCAACCGGGGCGCCAGAAGGTGTCACCTGACCGACACCACCA
ACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAAC
ATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTCAGGCCCAGCCGACCAAGTCCGAGAG
CGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGG

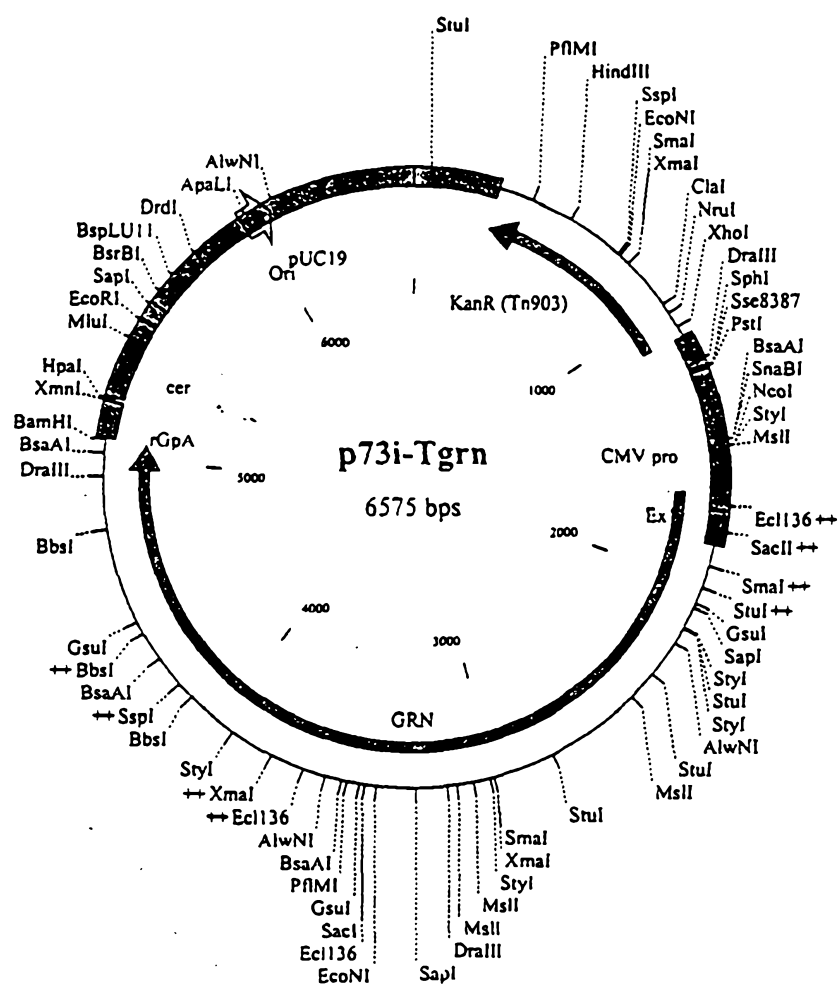
圖 13



TCCCGGCCCATAGGGCATTGGCGGCAACGAGCAGGTCGACAAGCTGGTGAGTGCGGGGATT

AGAAAGGTGCTGTAA [序列鑑別編號: 72]



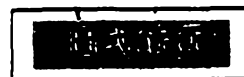


序列:

```

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
AGGCTGCGCCCGGGAGGCCAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCCGGGAGTGAAGAGCTGAGG
TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
ATTGTECAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
AATGCCTGGGTGAAGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACTCTACACTGCAAGAGCAAATC
GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
  
```

圖 14



ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAGACCCTG
CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAAGTGGATGACGGAGACACTCCTGGTA
CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCCGACCCGGACACAAAGCCAGA
GTGTTGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGG
ATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTG
GTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAG
AACCATAACAACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGC
AAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAG
CTGGGCATCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGAC
GTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCC
TTTACCATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTC
CTCCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATC
CTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGAC
CTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTG
AGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAG
CCGCCATTTCCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGACGCCT
ATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGC
AAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAG
CTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAG
CTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGCTACTAT
GACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACA
TACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCCGATG
AGGGGCGCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACG
ACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAG
GAGACCTGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGG
GAGTTCGTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCC
ATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTC
GGGAAGGCCGGGTACGTACCAACCGGGGGCCGAGAGGTGCTCACCCTGACCGAC
ACCACCAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGC
CTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTACAGGCCAG
CCGGACCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAA
GAGAAGGTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAG
GTCGACAAGCTGGTGAAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTC
ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC
TTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGAT
ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC
ACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA
CCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA
CACCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG
TTTGACAGCCGCTAGCATTTCATCACGTGGCCCCGAGAGCTGCATCCGGAGTACTTC
AAGAACTGCTGA [序列鑑別編號: 73]



MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
 LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
 LDKIEEEQNK SKKKAQQAAA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
 TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAAMQM
 LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
 TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGPK EPFRDYVDRF
 YKTLRAEQAS QEVKNWMTET LLVQNANPDC KTILKALGPA ATLEEMMTAC
 QGVGGPGHKA RVLMPISPI ETVSVKLGPG MDGPKVKQWP LTEEKIKALV
 EICTEMEKEG KISKIGPENP YNTPVFAIKK KDSTKWRKLV DFRELNKRTQ
 DFWEVQLGIP HPAGLKKKKS VTVLDVGDAY FSVPLDEDFR KYTAFTIPSI
 NNETPGIRYQ YNVLPQGWKG SPAIFQSSMT KILEPFRKQN PDIVYQYMD
 DLYVGSLEI GQHRTKIEEL RQHLLRWGLT TPDKKHQKEP PFLWMGYELH
 PDKWTVQPIV LPEKDSWTN DIQKLVGKLN WASQIYPGIK VRQLCKLLRG
 TKALTEVIPL TEEAELELAE NREILKEPVH GVYYDPSKDL IAEIQKQGQG
 QWTYQIYQEP FKNLKTGKYA RMGAHTNDV KOLTEAVQKI TTESIWIWGK
 TPKFKLPIQK ETWETWWTEY WQATWIPEWE FVNTPLVLKL WYQLEKEPIV
 GAETFYVDGA ANRETKLGKA GYVTNRGRQK VVTLTDTTNQ KTELQAIYLA
 LQDSGLEVNI VTDSQYALGI IQAQPQDSES ELVNQIIEQL IKKEKVYLAW
 VPAHKGIGGN EQVDKLVSAG IRKVLMSGFP VTPQVPLRPM TYKAAVDLSH
 FLKEKGGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF
 GWCYKLVPE PDKVEEANKG ENTSLLHPVS LHGMDDPERE VLEWRFDSRL
 AFHHVARELH PEYFKNC*

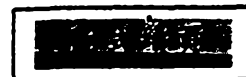
[序列鑑別編號: 74]



Tnrg 序列:

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCA
GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCAC
TCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC
CCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGA
TGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGA
GAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGA
GAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTTCATCACGTGGCCCCGAGAG
CTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAGACCGTG
CCGGTGAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACC
GAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAG
ATCAGCAAGATCGGGCCTGAGAACCCATACAAACCCCCCGTGTTTGCCATCAAGAAG
AAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACC
CAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCCTGAAGAAGAAG
AAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTGGACGAG
GACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACCCCTGGC
ATCAGATATCAGTACAACGTCCCTCCCCAGGGCTGGAAGGGCTCTCCCGCCATTTTC
CAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCCGATATCGTC
ATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGGGCAGCAT
CGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTCCG
GACAAGAAGCATCAGAAGGAGCCGCCATTCCTGaaGATGGGCTACGAGCTCCATCCC
GACAAGTGGACCGTGCAGCCTATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAAC
GACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATC
AAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATT
CCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAG
CCCGTGACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAATCCAGAAG
CAGGGCCAGGGGCAGTGGACATAACCAGATTTACCAGGAGCCTTTCAAGAACCTCAAG
ACCGGCAAGTACGCCCGCATGAGGGGCGCCACACCAACGATGTCAAGCAGCTGACC
GAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACACCCAAG
TTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATATTGGCAG
GCCACCTGGATTCCCGAGTGGGAGTTTCGTGAATACACCTCCTCTGGTGAAGCTGTGG
TACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGGCGCG
GCCAACC CGGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCCGCCAG
AAGGTCGTACCCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTAT
CTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCG
CTGGGCATTATTAGGCCCCAGCCGGACAGTCCGAGAGCGAACTGGTGAACCAGATT
ATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCCATTAAG
GGCATTGGCGGCAACGAGCAGGTGACAAAGCTGGTGAAGTGCAGGGGATTAGAAAGGTG
CTGATGGGTGCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAA
ATTAGGCTGCGCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCC
TCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGA
TGTCGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTG
AGGTCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATT

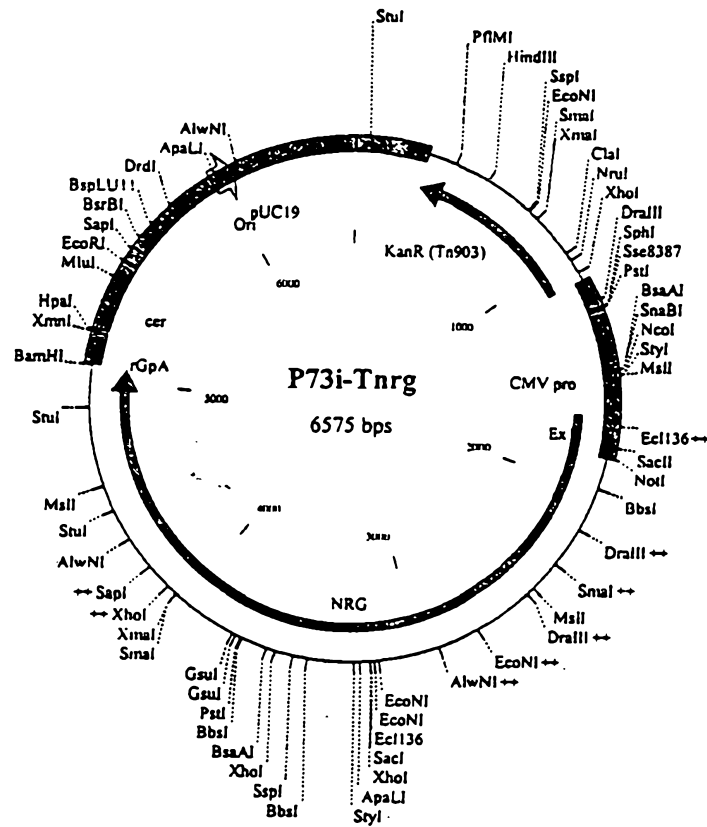
圖 15

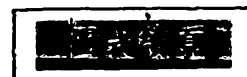


AAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAG
AAGGCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTAT
CCTATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGGACG
CTCAATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCC
ATGTTCTCCGCTTTGAGTGAGGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAAT
ACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAG
GCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAG
ATGCGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAA
ATCGGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGG
ATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGAC
ATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTTATAAGACC
CTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTG
GTACAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCC
ACCCTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCC
AGAGTGTGTGA [序列鑑別編號: 75]

MVGFPVTPQV	PLRPMTYKAA	VDLSHFLKEK	GGLEGLIHSQ	RRQDILDWLI
YHTQGYFPDW	QNYTPGPGVR	YPLTFGWCYK	LVPVEPDKVE	EANKGENTSL
LHPVSLHGMD	DPEREVLEWR	FDSRLAFHHV	ARELHPEYFK	NCMGPISPIE
TVSVKLKPGM	DGPKVKQWPL	TEEKIKALVE	ICTEMEKEGK	ISKIGPENPY
NTPVFAIKKK	DSTKWRKLVD	FRELNKRTQD	FWEVQLGIPH	PAGLKKKKSV
TVLDVGDAYF	SVPLDEDFRK	YTAFTIPSIN	NETPGIRYQY	NVLPQGWKGS
PAIFQSSMTK	ILEPFRKQNP	DIVIYQYMDD	LYVGSdleig	QHRTKIEELR
QHLLRWGLTT	PDKKHQKEPP	FLWMGYELHP	DKWTVQPIVL	PEKDSWTVND
IQKLVGKLNW	ASQIYPGIKV	RQLCKLLRGT	KALTEVIPLT	EEAELELAEN
REILKEPVHG	VYYDPSKDLI	AEIQKQGQGG	WTYQIYQEPF	KNLKTGKYAR
Mrgahtndvk	QLTEAVQKIT	TESIVIWGKT	PKFKLPIQKE	TWETWWTEYW
QATWIPEWEF	VNTPPLVKLW	YQLEKEPIVG	AETFYVDGAA	NRETKLGKAG
YVTNRGRQKV	VTLTDTTNQK	TELQAIYLAL	QDSGLEVNIV	TDSQYALGII
QAQPDQSESE	LVNQIIEQLI	KKEKVYLAWV	PAHKGIGGNE	QVDKLVSAGI
RKVLMGARAS	VLSSGELDRW	EKIRLRPGGK	KKYKLKHIVW	ASRELERFAV
NPGLLETSEG	CRQILGQLQP	SLQTGSEELR	SLYNTVATLY	CVHQRIEIKD
TKEALDKIEE	EQNKSKKKAQ	QAAADTGHNS	QVSQNYPIVQ	NIQQQMVHQA
ISPRTLNAWV	KVVEEKAFSP	EVIPMFSAIS	EGATPQDLNT	MLNTVGGHQA
AMQMLKETIN	EEAAEWDRVH	PVHAGPIAPG	QMREPRGSDI	AGTTSTLQEQ
IGWMTNPPPI	PVGEIYKRWI	ILGLNKIVRM	YSPTSILDIR	QGPKEPFRDY
VDRFYKTLRA	EQASQEVKNW	MTETLLVQNA	NPDCKTILKA	LGPAATLEEM

MTACQGVGGP GHKARVL* [序列鑑別編號: 76]

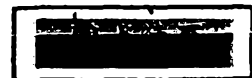




p7313ie中Tn_{gr} 插入基因之序列:

ATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGT
 AGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAA
 GACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAAC
 TACACACCAGGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACC
 AGTTGAGCCAGATAAGGTAGAAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
 TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTAGAGTGGAGGTTTGACAGCCGC
 CTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGTGC
 CCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATTAGGCTGCGCCCGG
 GAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTT
 GCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCA
 GCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGGTCCTTGTATAACACAGTGGCTACCCTCT
 ACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAG
 GAGCAAAACAAGAGCAAGAAGAAGGCCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCA
 GGTATCACAGAACTATCCTATTGTCCAAAACATTCAGGGCCAGATGGTTCATCAGGCCATCA
 GCCCCCGGACGCTCAATGCCTGGGTGAAGGTTGTCTGAAGAGAAGGCCTTTTCTCCTGAGGTT
 ATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAA
 TACCGTGCGCGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCAG
 CCGAGTGAGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAG
 CCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATCGGATGGATGAC
 CAACAATCCTCCCATCCAGTTGGAGAAATCTATAAACGGTGGATCATCCTGGGCCTGAACA
 AGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCT
 TTTAGGGATTACGTGACCGGTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGT
 CAAAAACTGGATGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAAACAATCT
 TGAAGGCACTAGGCCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTGTGAGGAGTAGGC
 GGACCCCGGACACAAAGCCAGAGTGTGATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGT
 GAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGA
 TCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGG
 CCTGAGAACCCATACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCG
 CAAGCTGGTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGG
 GCATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGAC
 GCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATC
 TATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCCAGGGCTGGAAGG
 GCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAAC
 CCCGATATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGG
 GCAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACCACTC
 CGGACAAGAAGCATCAGAAGGAGCCGCCATTCTGaaGATGGGCTACGAGCTCCATCCCGAC
 AAGTGACCGTGCAGCCTATCGTCCTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCA
 GAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGC
 TCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCC
 GAGCTCGAGCTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGACGGCGTGTACTATGA
 CCCCTCCAAGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAGA
 TTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGCGCCAC

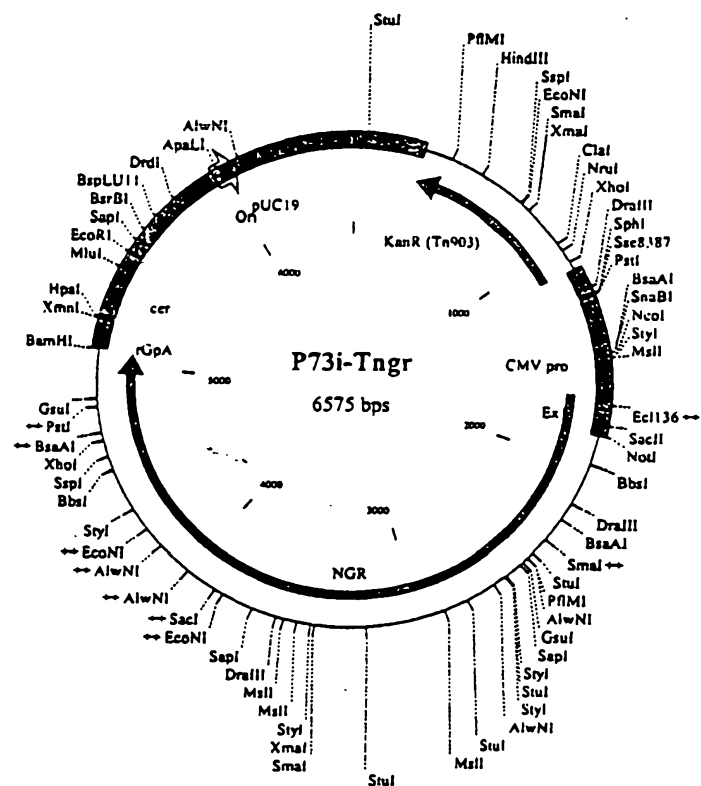
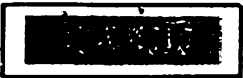
圖16

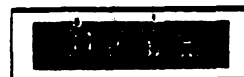


ACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGAT
 CTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGA
 CCGAATATTGGCAGGCCACCTGGATTCCCAGTGGGAGTTCTGTGAATACACCTCCTCTGGTG
 AAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGACGG
 CGCGGCCAACCGCGAAACAAAGCTCGGGAAGGCCGGGTACGTACCAACCGGGGCCGCCAGA
 AGGTCTGTCACCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCT
 CTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTAT
 TCAGGCCCAGCCGGACCAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCA
 AGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCCATAAGGGCATTGGCGGCAACGAGCAG
 GTCGACAAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGTAA [序列鑑別編號: 77]

MVGFPVTPQV PLRPMTYKAA VDLSHFLKEK GGLEGLIHSQ RRQDILDLWI
 YHTQGYFPDW QNYTPGPGVR YPLTFGWCYK LVPVEPDKE EANKGENTS
 LHPVSLHGMD DPEREVLEWR FDSRLAFHHV ARELHPEYFK NCMGARASVL
 SGGELDRWEK IRLRPGGKKK YLKHIVWAS RELERFAVNP GLLETSEGR
 QILGQLQPSL QTGSEELRSL YNTVATLYCV HQRIEIKDTK EALDKIEEEQ
 NKSKKKAQQA AADTGHSNQV SQNYPIVQNI QGQMVHQAI PRTLNAWVKV
 VEEKAFSPEV IPMFSALSEG ATPQDLNTML NTVGGHQAAM QMLKETINEE
 AAEWDRVHPV HAGPIAPGQM REPRGSDIAG TTSTLQEQIG WMTNNPPIPV
 GEIYKRWIL GLNKIVRMY PTSLDIRQG PKEPFRDYVD RFYKTLRAEQ
 ASQEVKNWMT ETLLVQNANP DCKTILKALG PAATLEEMMT ACQGVGGPGH
 KARVLMGPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK
 EGKISKIGPE NPYNTPVFAI KKKDSTKWRK LVDFRELNRK TQDFWEVQLG
 IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR
 YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDIVIYQY MDDLIVGSDL
 EIGQHRTKIE ELRQHLLRWG LTTPDKKHQK EPPFLWMGYE LHPDKWTVQP
 IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI
 PLTEEALEL AENREILKEP VHGVYDPSK DLIAEIQKQG QGQWYQIYQ
 EPFKNLKTGK YARMGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI
 QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD
 GAANRETKLG KAGYVTNRGR QKVVTLDTT NQKTELQAIY LALQDSGLEV
 NIVTDSQYAL GIIQAQPDQS ESELVNQIIE QLIKKEKVYL AWPVPAHKIG
 GNEQVDKLVS AGIRKVL*

[序列鑑別編號: 78]

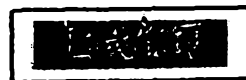




插入基因Tran#6之序列

ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
 GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
 ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCTGAGAACCCA
 TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
 GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
 ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
 GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
 ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCTCTCCC
 CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
 CCGTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
 GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
 CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
 TTCTTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
 CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
 AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
 CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
 CTGGCTGAGAACCGGGAGATCCTGAAGGAGCCCGTGCACGGCGTGTACTATGACCCC
 TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
 ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCGCATGAGGGGC
 GCCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
 TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
 TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
 GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
 GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAAG
 GCCGGGTACGTACCAACCGGGGCCGCCAGAAGTTCGTACCCCTGACCGACACCACC
 AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
 GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTACGGCCCAGCCGGAC
 CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
 GTCTACCTCGCCTGGGTCCCGGCCATAAAGGGCATTGGCGGCAACGAGCAGGTGAC
 AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGGTGCCCGAGCTTTCGGTACTG
 TCTGGTGGAGAGCTGGACAGATGGGAGAAAATTAGGCTGCGCCCGGGAGGCAAAAAG
 AAATACAAGCTCAAGCATATCGTGTGGGCCTCGAGGGAGCTTGAACGGTTTGCCGTG
 AACCCAGGCCTGCTGGAAACATCTGAGGGATGTCGCCAGATCCTGGGGCAATTGCAG
 CCATCCCTCCAGACCGGGAGTGAAGAGCTGAGGTCTTGTATAACACAGTGGCTACC
 CTCTACTGCGTACACCAGAGGATCGAGATTAAGGATACCAAGGAGGCCTTGGACAAA
 ATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAGGCCCAGCAGGCAGCTGCTGACACT
 GGGCATAGCAACCAGGTATCACAGAACTATCCTATTGTCCAAAACATTACAGGGCCAG
 ATGGTTCATCAGGCCATCAGCCCCCGACGCTCAATGCCTGGGTGAAGGTTGTGCAA
 GAGAAGGCCTTTTCTCCTGAGGTTATCCCCATGTTCTCCGCTTTGAGTGAGGGGGCC
 ACTCCTCAGGACCTCAATACAATGCTTAATACCGTGGGCGGCCATCAGGCCGCCATG

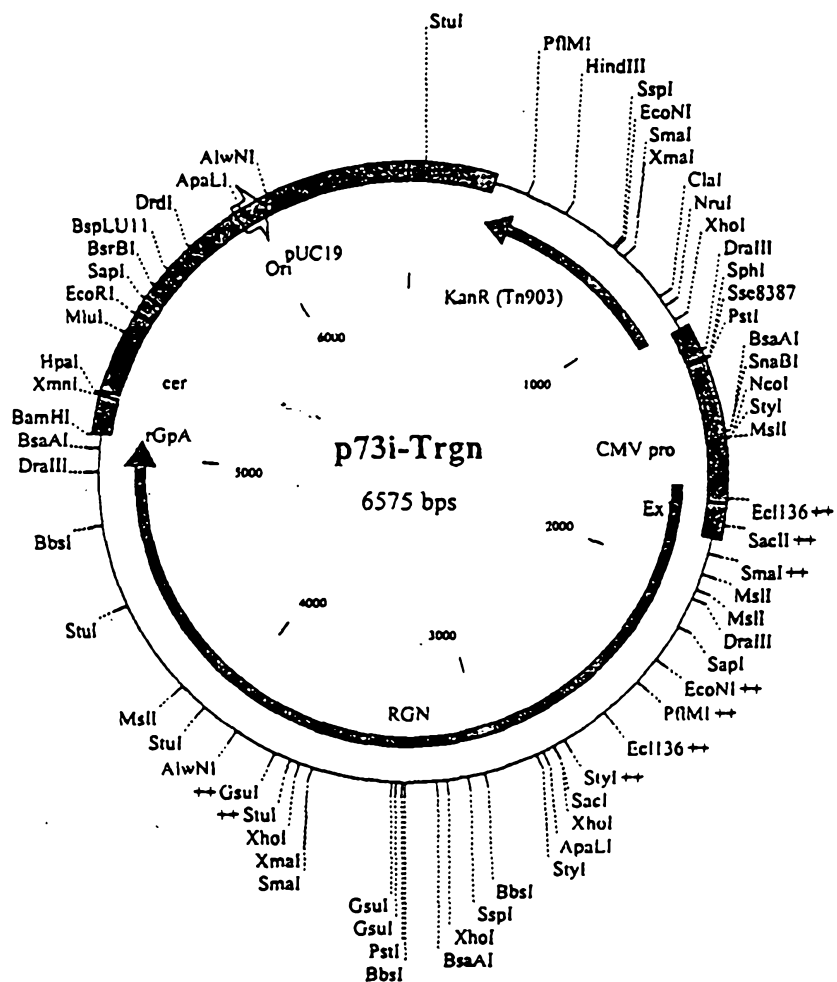
圖 17

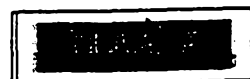


CAAATGTTGAAGGAGACTATCAACGAGGAGGCAGCCGAGTGGGACAGAGTGCATCCC
 GTCCACGCTGGCCCAATCGCGCCCGGACAGATGCGGGAGCCTCGCGGCTCTGACATT
 GCCGGCACCACCTCTACACTGCAAGAGCAAATCGGATGGATGACCAACAATCCTCCC
 ATCCCAGTTGGAGAAATCTATAAACGGTGGATCATCCTGGGCCTGAACAAGATCGTG
 CGCATGTACTCTCCGACATCCATCCTTGACATTAGACAGGGACCCAAAGAGCCTTTT
 AGGGATTACGTCGACCGGTTTTATAAGACCCTGCGAGCAGAGCAGGCCTCTCAGGAG
 GTCAAAAACCTGGATGACGGAGACACTCCTGGTACAGAACGCTAACCCCGACTGCAAA
 ACAATCTTGAAGGCACTAGGCCCGGCTGCCACCCTGGAAGAGATGATGACCGCCTGT
 CAGGGAGTAGGCGGACCCGGACACAAAGCCAGAGTGTGATGGTGGGTTTTCCAGTC
 ACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCAC
 TTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGAT
 ATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTAC
 ACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA
 CCAGTTGAGCCAGATAAGGTAAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTA
 CACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGG
 TTTGACAGCCGCCTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTC
 AAGAACTGCTGA [序列鑑別編號: 79]

MGPISPIETV SVKLKPGMDG PKVKQWPLTE EKIKALVEIC TEMEKEGKIS
 KIGPENPYNT PVFAIKKKDS TKWRKLVDFR ELNKRTQDFW EVQLGIPHPA
 GLKKKKS MTV LDVGDAYFSV PLDEDFRKYT AFTIPSINNE TPGIRYQYNV
 LPQGWKGSPA IFQSSMTKIL EPFRKQNPDI VIYQYMDDL VGS DLEIGQH
 RTKIEELRQH LLRWGLTTPD KKHQKEPPFL WMGYELHPDK WTVQPIVLPE
 KDSWTVNDIQ KLVGKLNWAS QIYPGIKVRQ LCKLLRGTKA LTEVIPLTEE
 AELELAENRE ILKEPVHGVY YDPSKDLIAE IQKQGQGWY YQIYQEPFKN
 LKTGKYARMR GAHTNDVKQL TEAVQKITTE SIVIWGKTPK FKLPIQKETW
 ETWTEYWQA TWIPEWEFVN TPPLVKLWYQ LEKEPIVGAE TFYVDGAANR
 ETKLGKAGYV TNRGRQKVVT LTDTTNQKTE LQAIYLALQD SGLEVNIVTD
 SQYALGIIQA QPDQSESELV NQIIEQLIKK EKVYLA WPA HKGIGGNEQV
 DKLV SAGIRK VLMGARASVL SGGELDRWEK IRLRPGGKKK YKLKHIVWAS
 RELERFAVNP GLLETSEGCR QILGQLQPSL QTGSEELRSL YNTVATLYCV
 HQRIEIKDTK EALDKIEEEQ NKS KKAQQA AADTGHSNQV SQNYPIVQNI
 QGQMVHQ AIS PRTLNAWVKV VEEKAFSPEV IPMFSALSEG ATPQDLN TML
 NTVGGHQAAM QMLKETINEE AA EWDRVHPV HAGPIAPGQM REPRGSDIAG
 TTSTLQEQIG WMTNPNPIP V GEIYKRWIIL GLNKIVRMYS PTSILDIRQG
 PKEPFRDYVD RFYKTLRAEQ ASQEVKNWMT ETLLVQANP DCKTILKALG
 PAATLEEMMT ACQGVGGPGH KARVLMVGFP VTPQVPLRPM TYKAAVDLSH
 FLKEKGGLEG LIHSQRRQDI LDLWIYHTQG YFPDWQNYTP GPGVRYPLTF

GWCYKLVPE PDKVEEANKG ENTSLHPVS LHGMDDPERE VLEWRFDSRL
AFHHVARELH PEYFKNC * [序列鑑別編號: 80]

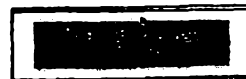




Tmg#11 插入基因之序列

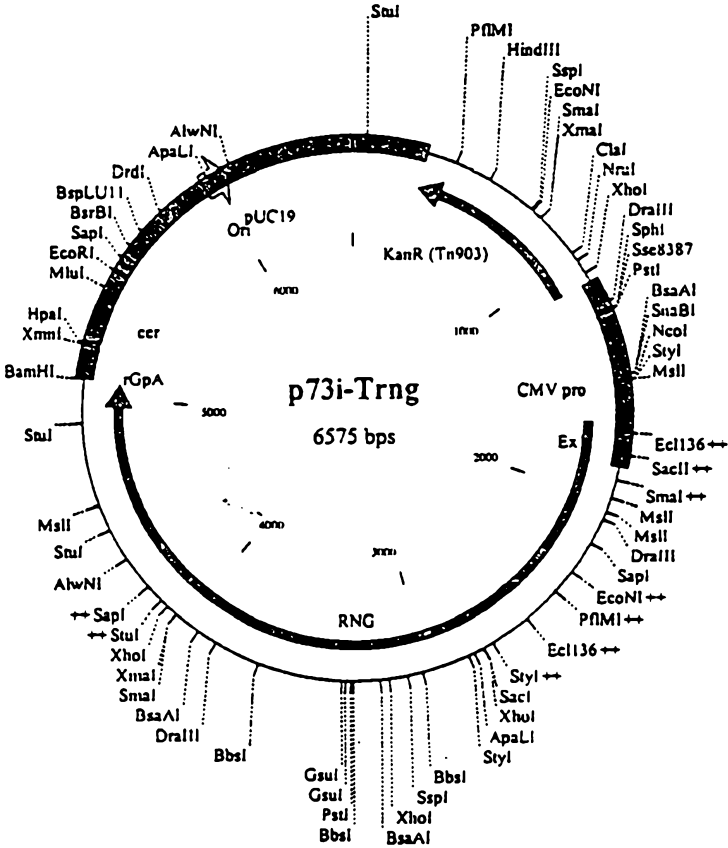
ATGGGCCCCATCAGTCCCATCGAGACCGTGCCGGTGAAGCTGAAACCCGGGATGGAC
GGCCCCAAGGTCAAGCAGTGGCCACTCACCGAGGAGAAGATCAAGGCCCTGGTGGAG
ATCTGCACCGAGATGGAGAAAGAGGGCAAGATCAGCAAGATCGGGCCGGAGAACCCA
TACAACACCCCCGTGTTTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG
GTGGATTTCCGGGAGCTGAATAAGCGGACCCAGGATTTCTGGGAGGTCCAGCTGGGC
ATCCCCCATCCGGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGC
GACGCTTACTTCAGCGTCCCTCTGGACGAGGACTTTAGAAAGTACACCGCCTTTACC
ATCCCATCTATCAACAACGAGACCCCTGGCATCAGATATCAGTACAACGTCCTCCCC
CAGGGCTGGAAGGGCTCTCCCGCCATTTTCCAGAGCTCCATGACCAAGATCCTGGAG
CCGTTTTCGGAAGCAGAACCCCGATATCGTCATCTACCAGTACATGGACGACCTGTAC
GTGGGCTCTGACCTGGAAATCGGGCAGCATCGCACGAAGATTGAGGAGCTGAGGCAG
CATCTGCTGAGATGGGGCCTGACCACTCCGGACAAGAAGCATCAGAAGGAGCCGCCA
TTCCTgaaGATGGGCTACGAGCTCCATCCCGACAAGTGGACCGTGCAGCCTATCGTC
CTCCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTC
AACTGGGCTAGCCAGATCTATCCCGGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTG
CGCGGCACCAAGGCCCTGACCGAGGTGATTCCCCTCACGGAGGAAGCCGAGCTCGAG
CTGGCTGAGAACC GGGAGATCCTGAAGGAGCCCGTGCACGGCGTGTACTATGACCCC
TCCAAGGACCTGATCGCCGAAATCCAGAAGCAGGGCCAGGGGCAGTGGACATACCAG
ATTTACCAGGAGCCTTTCAAGAACCTCAAGACCGGCAAGTACGCCCCGATGAGGGGC
GCCACACCAACGATGTCAAGCAGCTGACCGAGGCCGTCCAGAAGATCACGACCGAG
TCCATCGTGATCTGGGGGAAGACACCCAAGTTCAAGCTGCCTATCCAGAAGGAGACC
TGGGAGACGTGGTGGACCGAATATTGGCAGGCCACCTGGATTCCCGAGTGGGAGTTC
GTGAATACACCTCCTCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCGTG
GGCGCGGAGACATTCTACGTGGACGGCGCGGCCAACCGCGAAACAAAGCTCGGGAAG
GCCGGGTACGTCACCAACCGGGGCCCGCCAGAAGGTCGTCACCCTGACCGACACCACC
AACCAGAAGACGGAGCTGCAGGCCATCTATCTCGCTCTCCAGGACTCCGGCCTGGAG
GTGAACATCGTGACGGACAGCCAGTACGCGCTGGGCATTATTACAGGCCAGCCGGAC
CAGTCCGAGAGCGAACTGGTGAACCAGATTATCGAGCAGCTGATCAAGAAAGAGAAG
GTCTACCTCGCCTGGGTCCCGGCCCATTAAGGGCATTGGCGGCAACGAGCAGGTGAC
AAGCTGGTGAGTGCGGGGATTAGAAAGGTGCTGATGGTGGGTTTTCCAGTCACACCT
CAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTTA
AAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGATATCCTT
GATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCA
GGGCCAGGGGTGAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTT
GAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCT
GTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGAC
AGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAAC
TGCTGAATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAG
AAAATTAGGCTGCGCCCCGGGAGGCAAAAAGAAATACAAGCTCAAGCATATCGTGTGG
GCCTCGAGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAG
GGATGTCGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAG
CTGAGGTCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAG

圖 18



ATTAAGGATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAG
 AAGAAGGCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAAC
 TATCCTATTGTCCAAAACATTGAGGGCCAGATGGTTCATCAGGCCATCAGCCCCCGG
 ACGCTCAATGCCTGGGTGAAGGTTGTCGAAGAGAAGGCCTTTTCTCCTGAGGTTATC
 CCCATGTTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTT
 AATACCGTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAG
 GAGGCAGCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGA
 CAGATGCGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAG
 CAAATCGGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGG
 TGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTT
 GACATTAGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTCGACCGGTTTTATAAG
 ACCCTGCGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACTC
 CTGGTACAGAACGCTAACCCCGACTGCAAACAATCTTGAAGGCACTAGGCCCGGCT
 GCCACCCTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAA
 GCCAGAGTGTTG [序列鑑別編號: 81]

MGPISPIETV	SVKLKPGMDG	PKVKQWPLTE	EKIKALVEIC	TEMEKEGKIS
KIGPENPYNT	PVFAIKKKDS	TKWRKLVDFR	ELNKRTQDFW	EVQLGIPHPA
GLKKKKSVTV	LDVGDAYFSV	PLDEDFRKYT	AFTIPSINNE	TPGIRYQYNV
LPQGWKGSPA	IFQSSMTKIL	EPFRKQNPDI	VIYQYMDDL	VGSDLEIGQH
RTKIEELRQH	LLRWGLTTPD	KKHQKEPFL	WMGYELHPDK	WTVQPIVLPE
KDSWTVNDIQ	KLVGKLNWAS	QIYPGIKVRQ	LCKLLRGTKA	LTEVIPLTEE
AELELAENRE	ILKEPVHGVY	YDPSKDLIAE	IQKQGQGW	YQIYQEPFKN
LKTGKYARMR	GAHTNDVKQL	TEAVQKITTE	SIVIWGKTPK	FKLPIQKETW
ETWTEYWQA	TWIPEWEFVN	TPPLVKLWYQ	LEKEPIVGAE	TFYVDGAANR
ETKLKGAGYV	TNRGRQKVVT	LTDTTNQKTE	LQAIYLALQD	SGLEVNIVTD
SQYALGIIQA	QPDQSESELV	NQIIIEQLIKK	EKVYLAWVPA	HKGIGGNEQV
DKLVSAGIRK	VLMVGFPVTP	QVPLRPMTYK	AAVDLSHFLK	EKGGLEGLIH
SQRRQDILDL	WIYHTQGYFP	DWQNYTPGPG	VRYPITFGWC	YKLVPVEPDK
VEEANKGENT	SLLHPVSLHG	MDDPEREVLE	WRFDSRLAFH	HVARELHPEY
FKNCMGARAS	VLGGELDRW	EKIRLRPGGK	KKYKLKHIVW	ASRELERFAV
NPGLLETSEG	CRQILGQLQP	SLQTGSEELR	SLYNTVATLY	CVHQRIEIKD
TKEALDKIEE	EQNKSKKKAQ	QAAADTGHSN	QVSQNYPIVQ	NIQGQMVHQA
ISPRTLNAWV	KVVEEKAFSP	EVIPMFSA	EGATPQDLNT	MLNTVGGHQA
AMQMLKETIN	EEAAEWDRVH	PVHAGPIAPG	QMREPRGSDI	AGTTSTLQEQ
IGWMTNPNPI	PVGEIYKRWI	ILGLNKIVRM	YSPTSILDIR	QGPKEPFRDY
VDRFYKTLRA	EQASQEVKNW	MTETLLVQNA	NPDCKTILKA	LGPAATLEEM
MTACQGVGGP	GHKARVL	[序列鑑別編號: 82]		

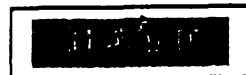




TgnR(F1) 序列:

ATGGGTGCCCCGAGCTTCGGTACTGTCTGGTGGAGAGCTGGACAGATGGGAGAAAATT
AGGCTGCGCCCGGGAGGCCAAAAAGAAATACAAGCTCAAGCATATCGTGTGGGCCTCG
AGGGAGCTTGAACGGTTTGCCGTGAACCCAGGCCTGCTGGAAACATCTGAGGGATGT
CGCCAGATCCTGGGGCAATTGCAGCCATCCCTCCAGACCGGGAGTGAAGAGCTGAGG
TCCTTGTATAACACAGTGGCTACCCTCTACTGCGTACACCAGAGGATCGAGATTAAG
GATACCAAGGAGGCCTTGGACAAAATTGAGGAGGAGCAAAACAAGAGCAAGAAGAAG
GCCCAGCAGGCAGCTGCTGACACTGGGCATAGCAACCAGGTATCACAGAACTATCCT
ATTGTCCAAAACATTTCAGGGCCAGATGGTTTCATCAGGCCATCAGCCCCCGGACGCTC
AATGCCTGGGTGAAGGTTGTGGAAGAGAAGGCCTTTTCTCCTGAGGTTATCCCCATG
TTCTCCGCTTTGAGTGAGGGGGCCACTCCTCAGGACCTCAATACAATGCTTAATACC
GTGGGCGGCCATCAGGCCGCCATGCAAATGTTGAAGGAGACTATCAACGAGGAGGCA
GCCGAGTGGGACAGAGTGCATCCCGTCCACGCTGGCCCAATCGCGCCCGGACAGATG
CGGGAGCCTCGCGGCTCTGACATTGCCGGCACCACCTCTACACTGCAAGAGCAAATC
GGATGGATGACCAACAATCCTCCCATCCCAGTTGGAGAAATCTATAAACGGTGGATC
ATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCTCCGACATCCATCCTTGACATT
AGACAGGGACCCAAAGAGCCTTTTAGGGATTACGTGACCGGTTTTATAAGACCCTG
CGAGCAGAGCAGGCCTCTCAGGAGGTCAAAAACCTGGATGACGGAGACACTCCTGGTA
CAGAACGCTAACCCCGACTGCAAAACAATCTTGAAGGCACTAGGCCCGGCTGCCACC
CTGGAAGAGATGATGACCGCCTGTCAGGGAGTAGGCGGACCCGGACACAAAGCCAGA
GTGTTGATGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTAC
AAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTA
ATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC
TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC
TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAAT
AAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCG
GAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCTAGCATTTTCATCACGTGGCC
CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCATGGGCCCCATCAGTCCCATCGAG
ACCGTGCCGGTGAAGCTGAAACCCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCA
CTCACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAAGAG
GGCAAGATCAGCAAGATCGGGCCTGAGAACCATAACAACCCCCCGTGTGTTGCCATC
AAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGATTTCCGGGAGCTGAATAAG
CGGACCCAGGATTTCTGGGAGGTCCAGCTGGGCATCCCCCATCCGGCCGGCCTGAAG
AAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTTACTTCAGCGTCCCTCTG
GACGAGGACTTTAGAAAGTACACCGCCTTTACCATCCCATCTATCAACAACGAGACC
CCTGGCATCAGATATCAGTACAACGTCCTCCCCCAGGGCTGGAAGGGCTCTCCCGCC
ATTTTTCAGAGCTCCATGACCAAGATCCTGGAGCCGTTTCGGAAGCAGAACCCCGAT
ATCGTCATCTACCAGTACATGGACGACCTGTACGTGGGCTCTGACCTGGAAATCGGG
CAGCATCGCACGAAGATTGAGGAGCTGAGGCAGCATCTGCTGAGATGGGGCCTGACC
ACTCCGGACAAGAAGCATCAGAAGGAGCCGCCATTCTTgaaGATGGGCTACGAGCTC
CATCCCGACAAGTGGACCGTGCAGCCTATCGTCTCCCCGAGAAGGACAGCTGGACC
GTGAACGACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCTAGCCAGATCTATCCC
GGGATCAAGGTGCGCCAGCTCTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAG

圖 19

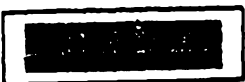


GTGATTCCCCTCACGGAGGAAGCCGAGCTCGAGCTGGCTGAGAACCGGGAGATCCTG
AAGGAGCCCCGTGCACGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAAATC
CAGAAGCAGGGCCAGGGGCAGTGGACATACCAGATTTACCAGGAGCCTTTCAAGAAC
CTCAAGACCGGCAAGTACGCCCCGATGAGGGGGCGCCACACCAACGATGTCAAGCAG
CTGACCGAGGCCGTCCAGAAGATCACGACCGAGTCCATCGTGATCTGGGGGAAGACA
CCCAAGTTCAAGCTGCCTATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAATAT
TGGCAGGCCACCTGGATTCCCAGTGGGAGTTCGTGAATACACCTCCTCTGGTGAAG
CTGTGGTACCAGCTCGAGAAGGAGCCCATCGTGGGCGCGGAGACATTCTACGTGGAC
GGCGCGGCCAACC CGCAAACAAAGCTCGGGAAGGCCGGGTACGTCACCAACCGGGGC
CGCCAGAAGGTCGTCACCCTGACCGACACCACCAACCAGAAGACGGAGCTGCAGGCC
ATCTATCTCGCTCTCCAGGACTCCGGCCTGGAGGTGAACATCGTGACGGACAGCCAG
TACGCGCTGGGCATTATTTCAGGCCCAGCCGGACCACTCCGAGAGCGAACTGGTGAAC
CAGATTATCGAGCAGCTGATCAAGAAAGAGAAGGTCTACCTCGCCTGGGTCCCGGCC
CATAAGGGCATTGGCGGCAACGAGCAGGTGACAAGCTGGTGAGTGCGGGGATTAGA
AAGGTGCTGTAA [序列鑑別編號: 83]

MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL
LETSEGCRI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA
LDKIEEEQNK SKKKAQQA AA DTGHSNQVSQ NYPIVQNIQG QMVHQAI SPR
TLNAWVKVVE EKAFSPEVIP MFSALSEGAT PQDLNTMLNT VGGHQAMQM
LKETINEEAA EWDRVHPVHA GPIAPGQMRE PRGSDIAGTT STLQEQIGWM
TNNPPIPVGE IYKRWIILGL NKIVRMYSPT SILDIRQGP KEFRDYVDRF
YKTLRAEQAS QEVKNWMTET LLVQANPDC KTILKALGPA ATLEEMMTAC
QGVGGPGHKA RVL MVGF PVT PQVPLRPMTY KAAVDLSHFL KEKGGLEGLI
HSQRRQDILD LWIYHTQGYF PDWQNYTPGP GVRYP LTFGW CYKLVPVEPD
KVEEANKGEN TSLLHPVSLH GMDDPEREVL EWRFD SRLAF HHVARELHPE
YFKNCMPIS PIETVSVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK
EGKISKIGPE NPYNTPVFAI KKKDSTKWRK LVDFRELNR TQDFWEVQLG
IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR
YQYNVLPQGW KGSPAIFQSS MTKILEPFRK QNPDI VIYQY MDDL YVGS DL
EIGQHRTKIE ELRQHLLRWG LTPDKKHQK EPPFLWMGYE LHPDKWTVQP
IVLPEKDSWT VNDIQKLVGK LN WASQIYPG IKVRQLCKLL RGTKALTEVI
PLTEEALEL AENREILKEP VHGVYYDPSK DLIAEIQKQG QGQWTYQIYQ
EPFKNLKTGK YARMGAHTN DVKQLTEAVQ KITTESI VIW GKTPKFKLPI
QKETWETWWT EYWQATWIPE WEFVNTPLV KLWYQLEKEP IVGAETFYVD
GAANRETKLG KAGYVTNRGR QKVVT LTDTT NQKTELQAIY LALQDSGLEV
NIVTDSQYAL GIIQAQPDQS ESELVNQIIE QLIKKEKVYL AWP AHKGIG
GNEQVDKLVS AGIRKVL*

[序列鑑別編號: 84]





使用IFN- γ 伊力司帕特法(ELISpot)測量補強5日後對Gag肽之反應

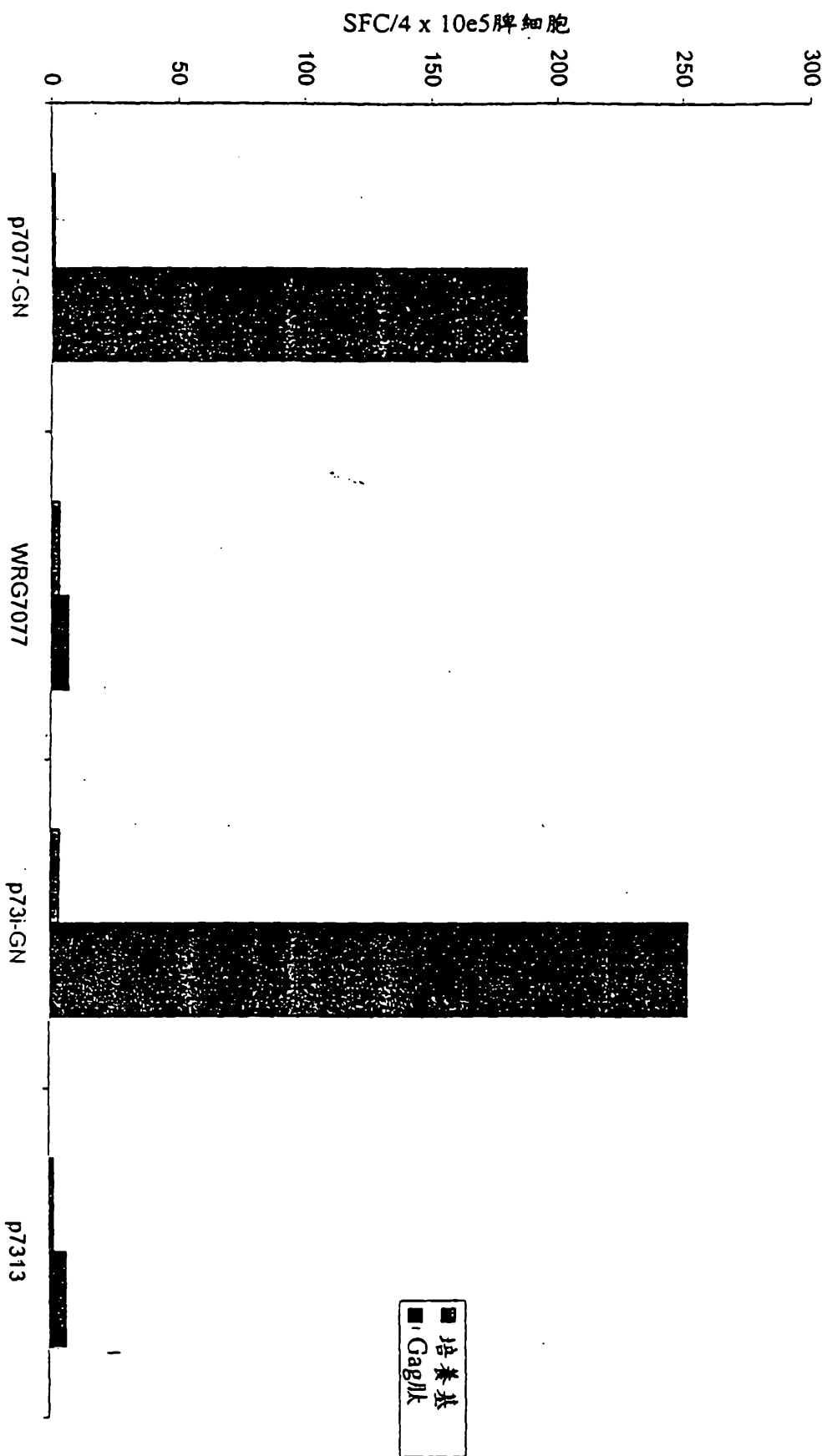


圖 20

11月8日

使用IFN- γ 伊力司帕特法(ELIspt)測量補強5日後對NeDk之反應

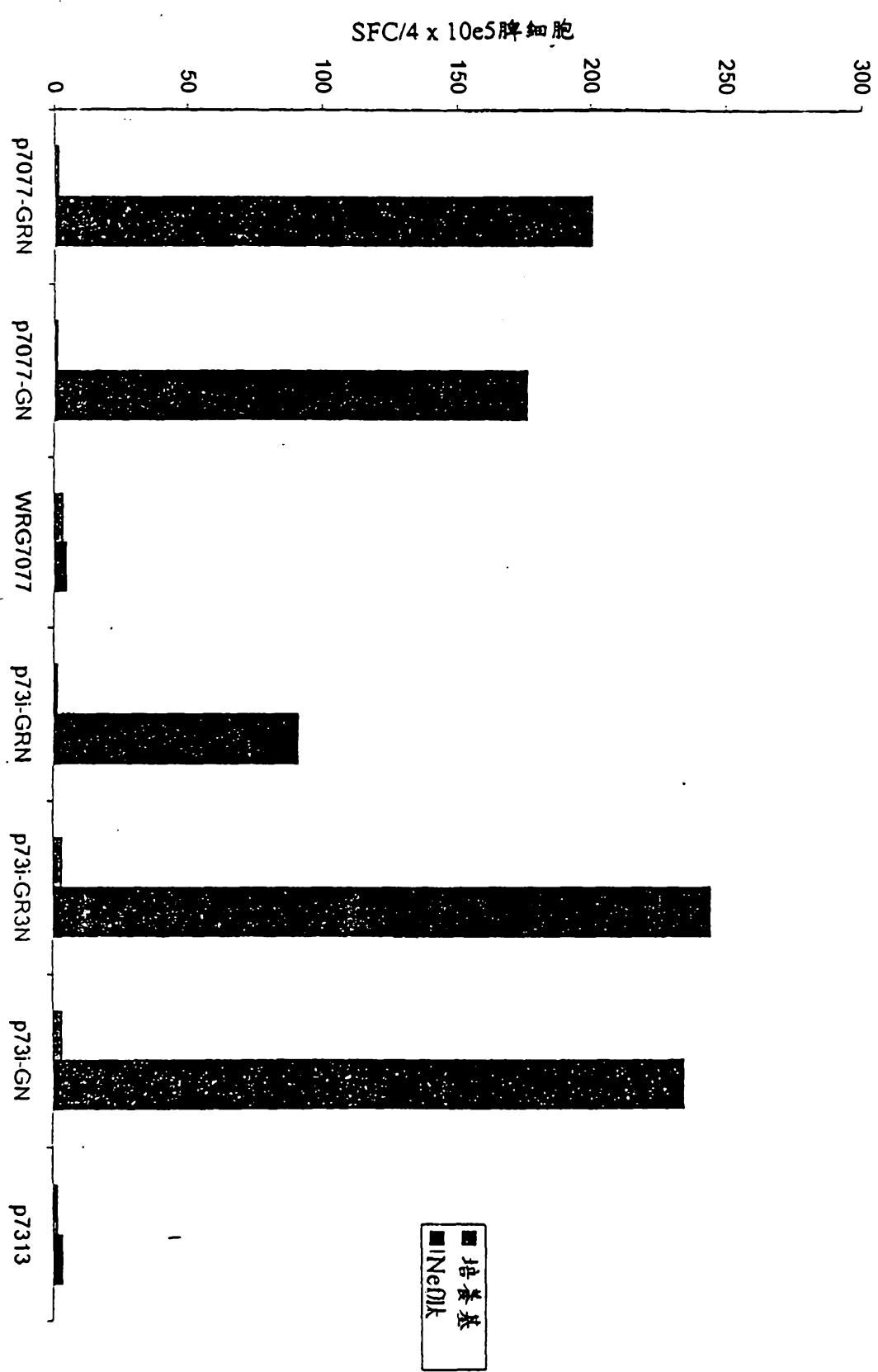


圖 21



使用IFN- γ 伊力司帕特法(ELIspot)測量補強5日後對R11A之反應

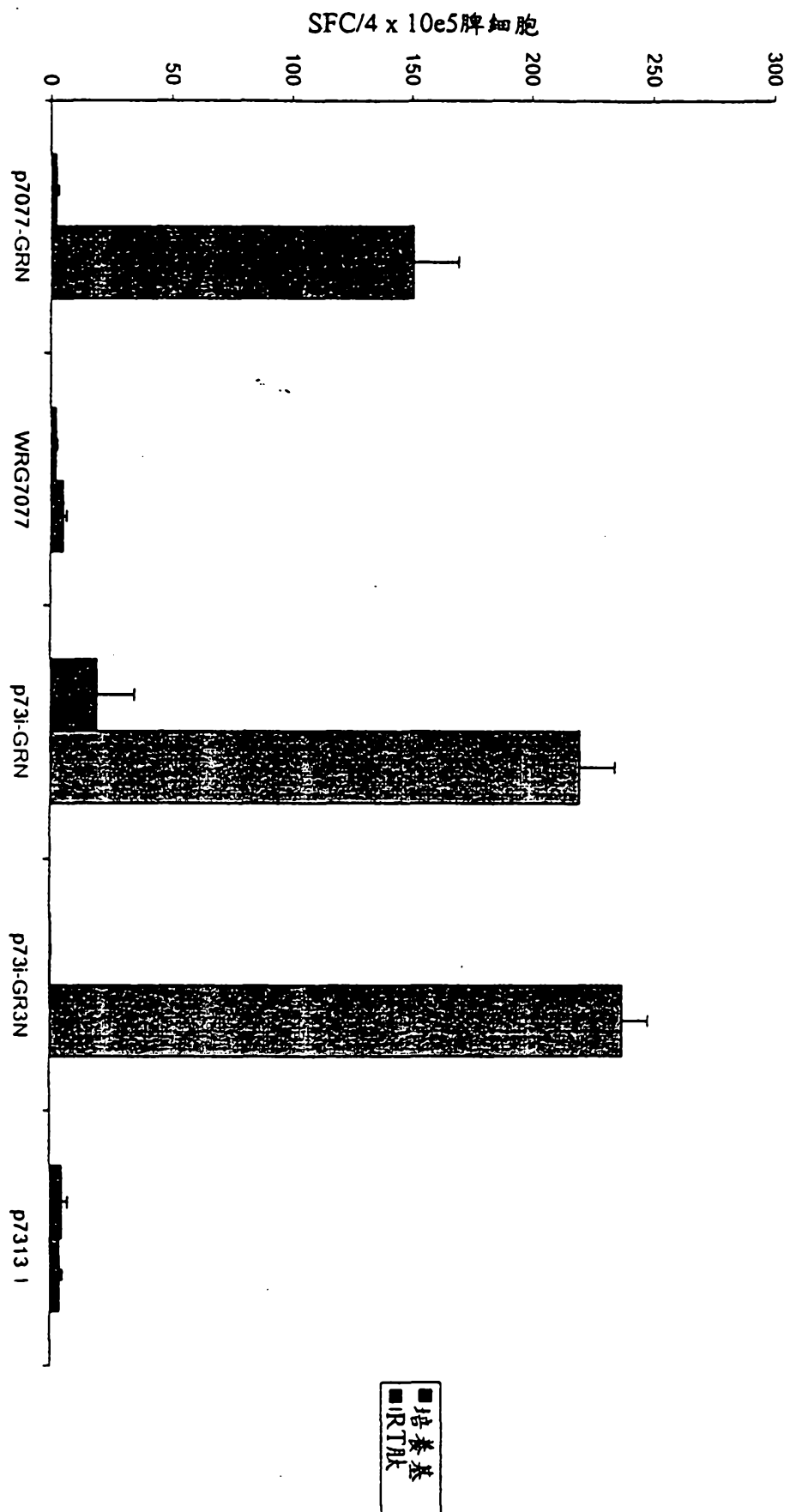
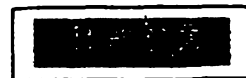
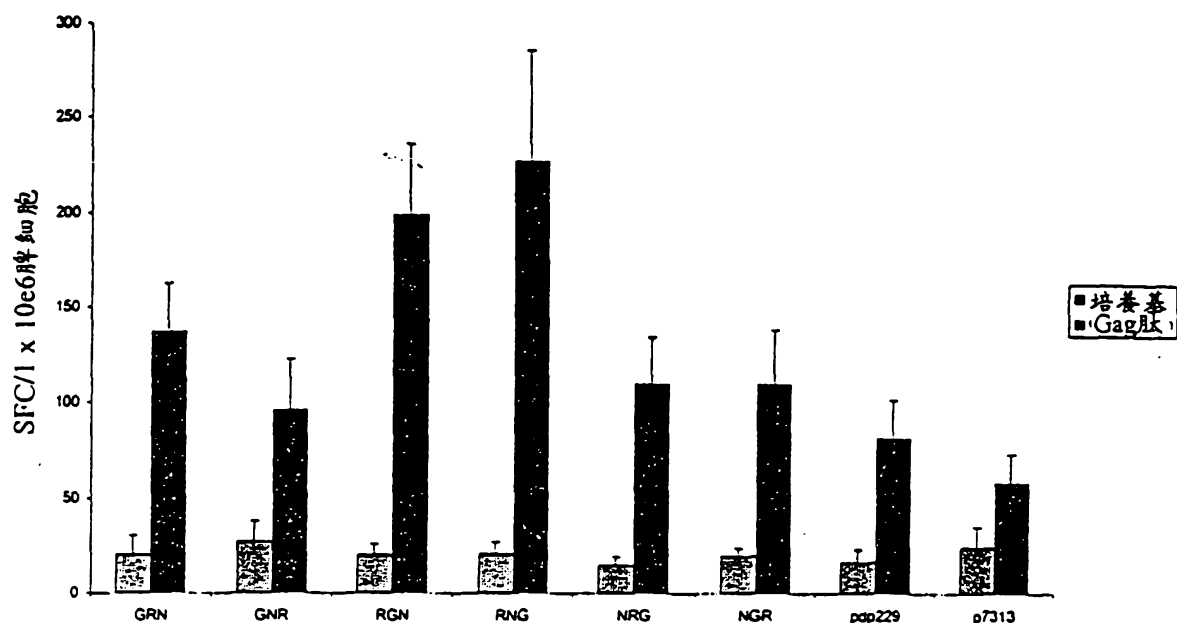


圖 22



本實例係以重組結構進行免疫後，檢測鼠模式之CD8
免疫反應。該結構係由Gag、RT及Nef以三聯融合體
方式所構成。G為Gag、R為RT、且N為Nef。

使用IFN- γ 伊力司帕特法(ELIsport)測量第35日(補強7日後)對Gag肽之反應



使用IFN- γ 伊力司帕特法(ELIsport)測量第35日(補強7日後)對Nef肽之反應

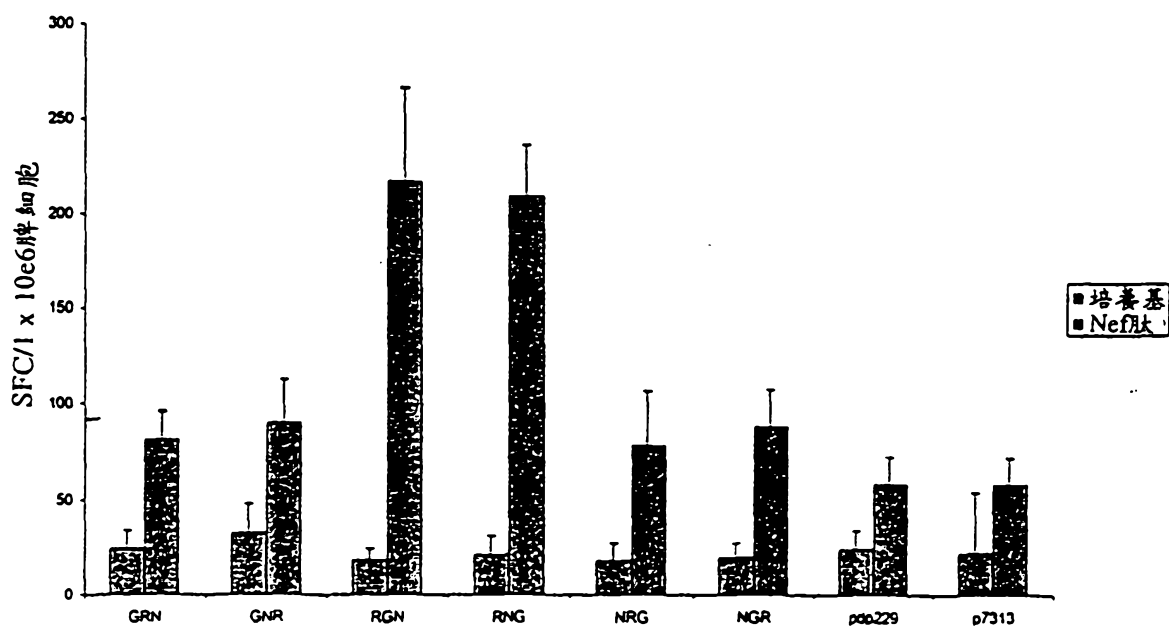
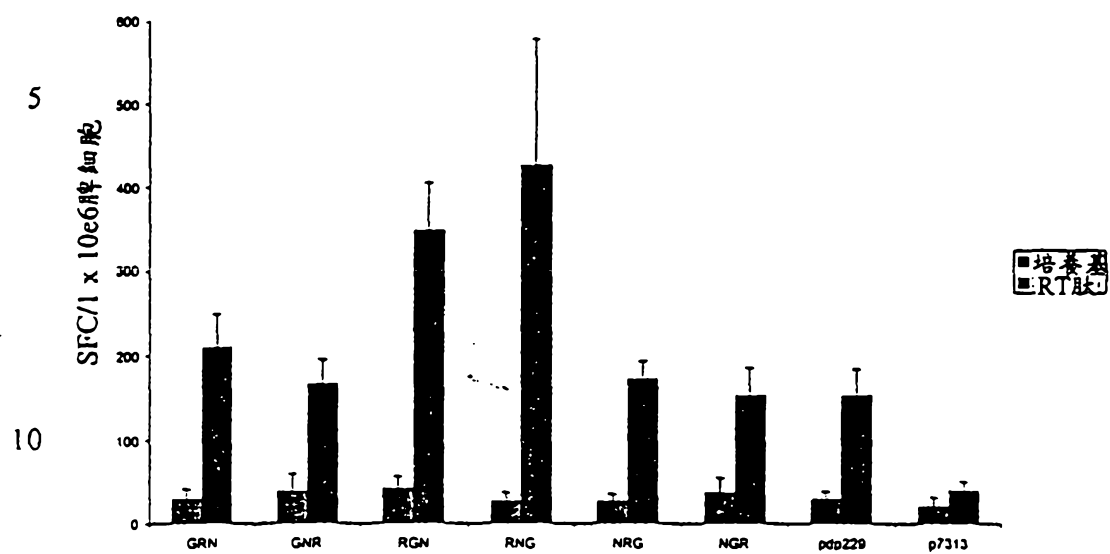


圖23



使用IFN- γ 伊力司帕特法(ELIspot)測量第35日(補強7日後)對RT肽之反應



15

20

25

以ELISA測量第35天(補強後第7天)之Gag抗體

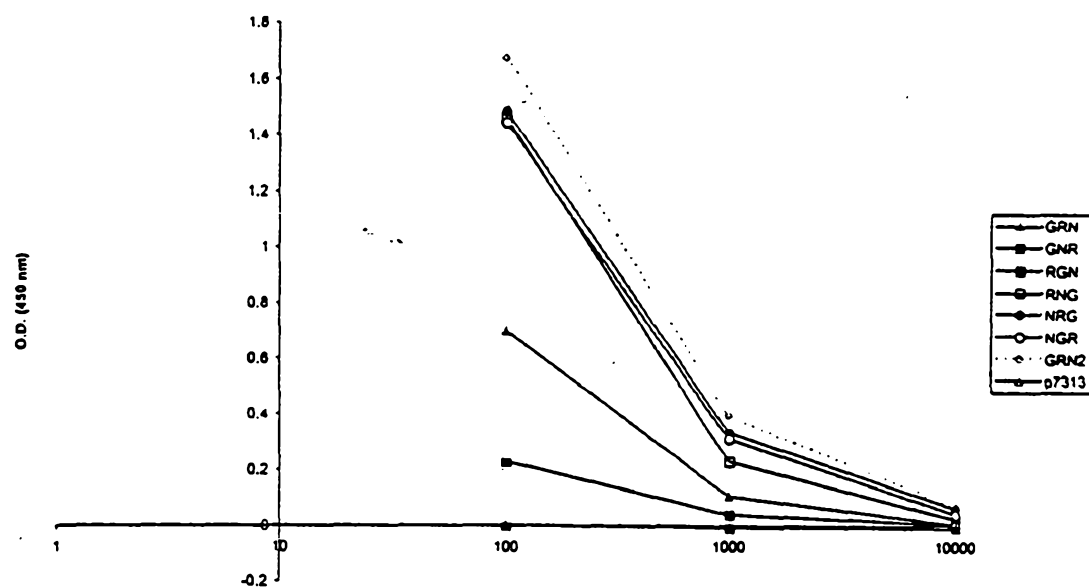


圖24

陸、(一)、本案指定代表圖為：第1圖

(二)、本代表圖之元件代表符號簡單說明：

無

柒、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

中文補充序列表

<110> 英商葛蘭素集團有限公司(Glaxo group limited)

<120> 疫苗

<130> PG4653

<140> 091121625

<141> 2002-09-20

<160> 84

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 42

<212> DNA

<213> 人工序列

<220>

<223> Nef引子

<400> 1

ataagaatgc ggccgccatg gtgggttttc cagtcacacc tt

42

<210> 2

<211> 31

<212> DNA

<213> 人工序列

<400> 2

cgcggaatcct cagcagttct tgaagtactc c

31

<210> 3

<211> 44

<212> DNA

<213> 人工序列

<220>

<223> srt引子

<400> 3

ataagaatgc ggccgccatg ggccccatta gccctattga gact

44

<210> 4

<211> 44

<212> DNA

<213> 人工序列

<220>

<223> srt引子

<400> 4

ataagaatgc ggccgccatg ggccccatta gccctattga gact

44

<210> 5

<211> 37

<212> DNA

<213> 人工序列

<220>

<223> sp17p24引子

<400> 5

ataagaatgc ggccgccatg ggtgcccagag cticggt

37

<210> 6

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> spl7p24引子	
<400> 6 tggggcccat caacactctg gccttctgtc	30
<210> 7 <211> 30 <212> DNA <213> 人工序列	
<220> <223> 連接子	
<400> 7 cagagtgttg atgggcccga ttagccctat	30
<210> 8 <211> 30 <212> DNA <213> 人工序列	
<220> <223> 連接子	
<400> 8 aaccaccat atctaaaaat agtactttcc	30
<210> 9 <211> 32 <212> DNA <213> 人工序列	
<220> <223> 連接子	
<400> 9 ctatttttag atatgggggg ttttcagtc ac	32
<210> 10 <211> 31 <212> DNA <213> 人工序列	
<220> <223> 連接子	
<400> 10 cgcggaacct cagcagttct tgaagtactc c	31
<210> 11 <211> 37 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 11 ataagaatgc ggccgcatg ggtgcccagc ctccggt	47
<210> 12 <211> 51 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 12 gcgcacgata ttgttcagge ccaggatgat ccaccgttta tagatttttc c	51
<210> 13 <211> 49 <212> DNA <213> 人工序列	

<220>	
<223> PCR引子	
<400> 13	
atcctggggcc tgaacaagat cgtgcgcatg tactctccga catccatcc	49
<210> 14	
<211> 30	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 14	
lggggcccac caacactctg gctttgtgtc	30
<210> 15	
<211> 68	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 15	
gaattcgagg ccgcgatggg ccccatcagt cccatcgaga ccgigccggt gaagctgaaa	60
cccgaggat	68
<210> 16	
<211> 44	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 16	
gggtgtactg gaaaacccac catcagcacc ttctaatcc ccgc	44
<210> 17	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 17	
atgggtgggtt ttccagtcac acc	23
<210> 18	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 18	
gatgaaatgc taggcggctg tcaaacctc	29
<210> 19	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 19	
gaggtttgac agccgcctag catttcac	29
<210> 20	

<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 20	
cgcggaiccct cagcagttct tgaagttac c	31
<210> 21	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 21	
atggagggtt ttccagttac acc	23
<210> 22	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 22	
gatgaaatgc taggcggctg tcaaacctc	29
<210> 23	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 23	
gaggtttgac agccgcttag catttcacc	29
<210> 24	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 24	
cgcggaiccct cagcagttct tgaagttac c	31
<210> 25	
<211> 68	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 25	
gaattcgccg ccgcgatggg cccatcagtt cccatcgaga ccgtgcgggt gaagctgaa 60	
cccgggat	68
<210> 26	
<211> 39	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 26	

ggagctcgta gcccatcttc aggaatggcg gctccttct	39
<210> 27	
<211> 68	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 27	
gaattcggat ccttacagca cctttctaata ccccgcactc accagcttgt cgacctgctc	60
gttgccgc	68
<210> 28	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 28	
cctgaagatg ggctacgagc tccaatg	26
<210> 29	
<211> 32	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 29	
cattagagcg gccgcgatgg tgggttttcc ac	32
<210> 30	
<211> 42	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 30	
gatgggactg atggggccca tgcagttctt gaactactcc gg	42
<210> 31	
<211> 24	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 31	
atgggccccca tcagtcccat cgag	24
<210> 32	
<211> 45	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR引子	
<400> 32	
cagtaccgaa gctcgggcac ccatcagcac ctittctaate cccgc	45
<210> 33	
<211> 24	
<212> DNA	
<213> 人工序列	
<220>	

<223> PCR引子	
<400> 33 atgggtgccc gagcttcggt actg	24
<210> 34 <211> 36 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 34 gatgggggat cctcacaaca ctctggcttt gigtcc	36
<210> 35 <211> 24 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 35 atgggtgccc gagcttcggt actg	24
<210> 36 <211> 68 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 36 gaattcggat ccttacagca cctttctaat ccccgacac accagcttgt cgacctgctc gttgccgc	60 68
<210> 37 <211> 32 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 37 cattagagcg gccgcgatgg tgggttttcc ac	32
<210> 38 <211> 45 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 38 cagtaaccgaa gcctggggac ccatgcagtt ctggaactac tccgg	45
<210> 39 <211> 24 <212> DNA <213> 人工序列	
<220> <223> PCR引子	
<400> 39 atgggtgccc gagcttcggt actg	24
<210> 40 <211> 68 <212> DNA	

<213> 人工序列

<220>

<223> PCR引子

<400> 40
gaattcgcg cgcgatggg ccccatcagt cccatcgaga ccgtgccgt gaagctgaaa 60
cccgggat 68

<210> 41
<211> 45
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 41
cagtaccgaa gctcgggcac ccatcagcac ctctctaac cccgc 45

<210> 42
<211> 68
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 42
gaattcgcg cgcgatggg ccccatcagt cccatcgaga ccgtgccgt gaagctgaaa 60
cccgggat 68

<210> 43
<211> 45
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 43
cagtaccgaa gctcgggcac ccatgcagtt ctggaactac tccgg 45

<210> 44
<211> 24
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 44
atgggtgcc gagcttcgt actg 24

<210> 45
<211> 36
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 45
gatgggggat cctcacaaca ctctggcttt gtgtcc 36

<210> 46
<211> 24
<212> DNA
<213> 人工序列

<220>

<223> PCR引子

<400> 46
atgggtgcc gagcttcgt actg 24

<210> 47
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> PCR引子

<400> 47
gatgggactg atggggccca tgcagttcct gaactacicc gg 42

<210> 48
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> PCR引子

<400> 48
atgggcccac tcagtcctat cgag 24

<210> 49
<211> 68
<212> DNA
<213> 人工序列

<220>
<223> PCR引子

<400> 49
gaattcggat ccttacagca ccttctaat ccccgacac accagcttct cgacctgctc 60
gttgcgcg 68

<210> 50
<211> 1503
<212> DNA
<213> HIV

<400> 50
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaatttagg 60
ctgcgcccg gaggcacaaa gaaatacaag ctcaagcata tegtgtgggc ctgagggag 120
cttgaacgggt ttgccgtgaa cccaggectg ctggaaacat ctgagggatg tgcacagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc ctgtatatac 240
acagltggcta cctctactg cgtacaccag aggatcgaga ttatggatgc caaggaggtc 300
ttggacacaaa ttgaggagga gcaaaacaag agcaagaaga agcccacga ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaa catcaggtc 420
cagatgggtc atcagcccat cagccccgg agctcaatg cctgggtgaa ggttgtgaa 480
gagaaggcct ttctctctga ggttatcccc atgttctccg ctlttgatga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtggcgccg atcaggtcgc catgcaatg 600
ttgaaggaga ctatcaacga ggagccagcc gattgggaca gattgcatcc cgtccagct 660
ggcccaatcg cgcgccgaca gatggggag cctcggggt ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctccatccc agttggagaa 780
atctataaac ggttgatcat tctcgtctc aataaattg ttgaatgta ctctccgaca 840
tccatccttg acattagaca gggaccacaa gaccttita gggatlaagt cgaaccgttt 900
tataagaccc tgcagacaga gcaggtctct caggaggtca aaactggat gacggagaca 960
ctcttggtac agaacgttaa ccccgactgc aaaaatct ttgaagcaat aggcctggct 1020
gccacctgg aagagatgat gaccgctgt caggagtag gggaccggg acacaagcc 1080
agagtggtgg ccgaagccat gaccaggtg acgaactcc caaccatcat gatgcagaga 1140
gggaacttcc gcaatcagcg gaagatctgt aagtgttca attgggcaa ggagggtcat 1200
accgcccgca actgtcgggc ccttaggaag aaagggtgtt ggaagtgcg caaggagaga 1260
caccagatga aagactgac agaaccagag gccaatltt ttggaagat ttggccgagc 1320
tacaagggga gacctgtaa ttcttgcna agcaggtccc agcccaccg cccccctgag 1380
gaatcttca ggtccggagt ggagaacaca acgtctccc aaaaacagga accaatcgac 1440
aaggagctgt accctttaac ttctctggt tctctctttg gaaagaccc gtctgtctaa 1500
taa 1503

<210> 51
<211> 500
<212> PRT
<213> HIV

<400> 51
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	Lys 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 Asp Lys Ile	Glu Glu Gln Asn	Lys Ser Lys
100	105	110	
Lys Lys Ala Gln	Gln Ala Ala Ala	Asp Thr Gly His	Ser Asn 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 Thr	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 Glu	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	Ile Val Lys Cys	Phe Asn Cys Gly	Lys Glu Gly His
385	390	395	400
Thr Ala Arg 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	Tyr Lys Gly Arg	Pro Gly Asn Phe
435	440	445	
Leu Gln Ser Arg	Pro Glu Pro Thr	Ala Pro Pro Glu	Glu Glu Ser Phe Arg
450	455	460	
Ser Gly Val Glu	Thr Thr Pro Pro	Gln Lys Gln Glu	Pro Ile Asp
465	470	475	480
Lys Glu Leu Tyr	Pro Leu Thr Ser	Leu Arg Ser Leu	Phe Gly Asn Asp
485	490	495	
Pro Ser Ser Gln			
500			

<210> 52
 <211> 1515
 <212> DNA
 <213> HIV

<400> 52
 atgggtgcga gagcgtcagt attaaagcggg ggagaattag atcgatggga aaaaattcgg 60
 ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcaggag 120
 ctagaacgat tcgcagittaa tcctggcctg ttagaaacat cagaaggctg tagacaaata 180
 ctgggacagc tacaaccatc ccctcagaca ggatcagaag aacttagatc attatataat 240

```

acagtagcaa cccctctattg tggcatcaa aggatagaga taaaagacac caaggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagacaatca ggtagcaca aattacccta tagtgacaga calccagggg 420
caaattgtac atcaggccat atcacctaga actttaaatg catgggtaaa agtagtagaa 480
gagaaggctt tcagcccaga agtgatcccc atgttttcag cattatcaga aggagccacc 540
ccacaagatt taaacaccaa gctaaacaca gtggggggac atcaagcagc catgcaaatg 600
ttaaagaga ccatcaatga ggaagctgca gaatgggata gagtgcattc agtgcatgca 660
gggccttatg caccaggcca gatgagagaa ccaaggggaa glgacatagc aggaactact 720
agtacccctc aggaacaaa aggatggatg acaataaatc cacctatccc agtaggagaa 780
atttataaaa gatggataal cctgggatta aataaaaatg taagaatgta tagccctacc 840
agcattcttg acataagaca aggaccaaa gaacccttta gagactatgt agaccggttc 900
tataaaactc taagagccga gcaagcttca caggaggtaa aaaattggat gacagaaacc 960
ttgttggtcc aaaatgcgaa cccagattgt aagactatit taaaagcatt gggaccagcg 1020
gtacacttag aagaaatgat gacagcatgt caggagatag gaggaccgga ccataaggca 1080
agagtttttg tgggttttcc agtcacacct caggtaacct taagaccaat gacttacaag 1140
gcagctgtag atcttagcca ctttttataa gaaaaggggg gactggaagg gctaattcac 1200
tcccaagaaa gacaagatat ccttgatctg tggatctacc acacacaagg ctacttccct 1260
gattggcaga actacacacc agggccaggg gtcagataic cactgacctt tggatgggtc 1320
tacaagctag taccagttga gccagataag gtagaagagg ccaataaagg agagaacacc 1380
agcttggtac accctgtgag cctgcatggg atggatgacc cggagagaga agtgttagag 1440
tggaggtttg acagccacct agcatttcat cactggccc gagagctgca tccggagtac 1500
ttcaagaact gctga 1515

```

<210> 53
 <211> 504
 <212> PRT
 <213> HIV

<400> 53

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	Lys	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	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys
		100						105					110		
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	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	Thr	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	Glu	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	Val	Gly	Phe	Pro	Val
	355					360							365		

Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala Val Asp
370 375 380
Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His
385 390 395 400
Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr Gln
405 410 415
Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val Arg
420 425 430
Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu Pro
435 440 445
Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu His
450 455 460
Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu Glu
465 470 475 480
Trp Arg Phe Asp Ser His Leu Ala Phe His His Val Ala Arg Glu Leu
485 490 495
His Pro Glu Tyr Phe Lys Asn Cys
500

<210> 54
<211> 1518
<212> DNA
<213> HIV

<400> 54
atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgaggagcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctgaggggag 120
cttgaacgggt ttgccgtgaa cccaggcctg ctggaacat ctgagggatg tcgccagatc 180
ctgggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc ctgtataac 240
acagtggcta cctctactg cgtacaccag aggatcgaga ttaaggatac caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga agggccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaa cattcagggc 420
cagatgggtt atcaggccat cagcccccg acgctcaatg cctgggtgaa ggttgcgaa 480
gagaaggcct ttctctctga ggttatcccc atgttctccg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cgcccgaca gatcggggag cctcgcggt ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat tctcggtctc aataaaattg ttagaatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagacc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctgggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggt 1020
gccaccctgg aagagatgat gaccgcctgt caggagtag gcggaccgg acacaaagcc 1080
agagtgttga tgggtgggttt tccagtcaca cctcaggtac cttaagacc aatgacttac 1140
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 1200
cactcccaaa gaagacaaga tatcttgat ctgtggatct accacacaca aggcctacttc 1260
cctgattggc agaactacac accagggcca ggggtcagat atccactgac ctttggatgg 1320
tgctacaagc tagtaccagt tgagccagat aaggtagaag agggcaataa aggagagaac 1380
accagcttgt tacaccctgt gagcctgcat gggatggatg acccgagag agaagtgtta 1440
gagtggaggt ttgacagcca cctagcattt catcacgtgg cccgagagct gcatccggag 1500
tacttcaaga actgctga 1518

<210> 55
<211> 505
<212> PRT
<213> HIV

<400> 55
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 Lys 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 Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
100 105 110
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn 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	Thr	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	Glu	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	Met	Val	Gly	Phe	Pro
		355					360					365			
Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala	Val
370					375						380				
Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu	Ile
385				390						395					400
His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His	Thr
				405					410					415	
Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly	Val
		420						425					430		
Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val	Glu
		435					440					445			
Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	Leu
	450					455					460				
His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu	Val	Leu
465				470						475					480
Glu	Trp	Arg	Phe</												

<210>	56
<211>	1689
<212>	DNA
<213>	HIV

<400> 56

atgggcceca	tcagtcceat	cgagacegtg	ccpgtgaag	tgaacecegg	gatggacggc	60
cccaaggtca	agcagtggec	actcecegg	gagaagata	aggeccctgt	ggagatctgc	120
accgagatgg	agaaagaggg	caagatcagc	aagatcgggc	ctgagaaccc	atacaacacc	180
cccgtgtttg	ccatcaagaa	gaaggaacgc	accagttggc	gcagcttggt	ggatttccgg	240
gagctgaata	agcggaccac	ggatttcttg	gaggtccagc	tgggcatccc	ccatccgggc	300
ggcctgaaga	agaaagattg	cgttgacgtg	cctggacgtg	ggcagcttta	cttccagctc	360
ccctctggacg	aggacttttag	aaggtacacc	gcctttacca	tccactctat	catacaagag	420
acccctggca	tcagataata	gtacaaagtc	ctcccccagg	gctgggaagg	ctctcccgcc	480
attttccaga	gtcccatgac	caagatccct	gagccgtttc	gggaagcaga	cccagatata	540
gtcatctacc	agttacattg	cgacctgtac	gtgggtcttg	acctggaat	cgggcagcat	600
cgcacgaaga	ttagggagtc	gaggcagcat	ctgctgagat	ggggcttgac	cactccggac	660
aagaagcaat	agaaaggagc	gccattcttg	tggatgggtt	acgagctcca	ctccgacaa	720
ttggaccgtgc	agcctatctg	ctctcccgag	aaggacagct	ggaacctgaa	tcgacatcag	780
aagctggttg	gcaagctcaa	ctgggtctagc	cagatctata	ccgggataaa	ggttgcgccag	840
ctctgcaagc	tgtctgcggg	caccaaaggcc	ctgaccgagg	tgttctccct	caegggaggaa	900
gccgagctcg	agctggctga	gaaccggggg	atcttgaagg	agccccttga	cggcgtgtac	960
tatgacccct	ccaaggacct	gacgcgggaa	atccagaagc	aggggacagg	gcagctggaca	1020
laccagattt	accaggagcc	tttcaagaac	ctcagaagcg	gcgaagtacg	cgcgatggag	1080
ggcgcgccaca	ccaacgattg	caagcagctg	accgaagccg	tcagaaagat	cagaccggag	1140

```
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
acacctctc tggtaagct gtgtaccag ctcgagaagg agcccatcgt gggcgcgagg 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtacg cgtgggcat tattcaggcc cagccggacc agtccgagag cgaactggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaaggtct acctcgctg ggtcccgcc 1620
cataagggca ttggcggaac cgagcaggtc gacaagctgg tgagtgcggg gattagaaag 1680
gtgctgttaa 1689
```

<210> 57
<211> 562
<212> PRT
<213> HIV

<400> 57

```
Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser 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 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 Ser Val Thr Val Leu Asp
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 Asp Asp 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 Leu 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 Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
275 280 285
Lys Ala Leu Thr Glu Val Ile 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 Thr 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 Thr 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 Asp Gly Ala Ala
435 440 445
Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
450 455 460
```

Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
 465 470 475 480
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 485 490 495
 Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 500 505 510
 Asp Gln Ser Glu Ser Glu Leu Val Asn 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 Ala Gly Ile Arg Lys
 545 550 555 560
 Val Leu

<210> 58
 <211> 1689
 <212> DNA
 <213> HIV

<400> 58
 atgggcccac tcagtccecat cgagaccgtg ccggigaagc tgaaaccecg gatggacggc 60
 cccaaggcca agcagtgccc actcaccgag gagaagatca aggccctggg ggagatctgc 120
 accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
 cccgtgtttg ccatcaagaa gaaggacagc accaagtggc gcaagctggg ggatttccgg 240
 gagctgaata agcggaccca ggatttctgg gaggtccagc tgggcatccc ccatccggcc 300
 ggccigaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgcata ctccagcgtc 360
 cctctggagc aggaatttag aaagtacacc gcccttacc aaccatctat caacaacgag 420
 acccctggca tcagatatac gtacaacgtc ctccccagg gctggaaggg ctctcccgcc 480
 atttccaga gctccatgac caagatccg gagecgttgc ggaagcagaa ccccgataac 540
 gtatctacc agtacttgga cgactgttac gtgggtctcg acctggaaat cgggcagcat 600
 cgcacgaaga ttgaggagct gaggcagcat ctgtctgagat ggggcctgac cactccggac 660
 aagaagcacc agaaggagcc gccattccgt tggatgggct acgagctcca tcccgacaag 720
 tggaccgtgc agcctatcgt ctccccgag aaggacagct ggaccgtgaa cgacatccag 780
 aagctgggtgg gcaagctcaa ctgggttagc cagatctatc cggggatcaa ggtgcgccag 840
 ctctgcaagc tgcctgcggc caaccaaggcc ctgaccgagg tgattcccc caccgaggaa 900
 gccgagctcg agctggctga gaaccgggag atcttgaagg agcccgctga cggcgtgtac 960
 tatgacccct ccaaggacct gatcgccgaa atccagaagg agggccaggg gcagtggaac 1020
 taccagattt accaggagcc ttcaagaac ctcaagacc gcaagtacgc ccgatgagg 1080
 ggcgccacac caacgatgt caagcagctg accgagcgcc tccagaagat cagcaccgag 1140
 tccatcgtga tctgggggaa gacaccaag ttcaagctgc ctatccagaa ggagacctgg 1200
 gagacgttgt ggaccgaata ttgcagggcc acctggattc ccgagtgga gttctgaaat 1260
 acacctcttc tggtagagct gtgttaccag ctgagaagg agccatcgtt gggcgccgag 1320
 acattctacg tggacggcgc ggcacaccgc gaaacaaagc tgggaagcc cgggtacgtc 1380
 accaacgggg gccgcagaa ggtgttacc ctgaccgaca caaccaacca gaagacggag 1440
 ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
 agccagtagc cgttgggcat tatccagccc cagccggacc agtccgagag cgaactggtg 1560
 aaccagatta tggagcagct gataaagaaa gagaaggtct acctcgcctg ggtcccgccc 1620
 cataaggcca ttggcggaac cgagcaggtc gacaagctgg tgagtgcggg gattagaag 1680
 gtgtgttaa 1689

<210> 59
 <211> 562
 <212> PRT
 <213> HIV

<400> 59
 Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser 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 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 Ser Val Thr Val Leu Asp
 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
Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln		160
	165	170
Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly		175
	180	185
Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg		190
	195	200
Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln		205
	210	215
Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys		220
225	230	235
Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val		240
	245	250
Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile		255
	260	265
Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr		270
	275	280
Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu		285
	290	295
Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr		300
305	310	315
Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln		320
	325	330
Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys		335
	340	345
Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys		350
	355	360
Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile		365
	370	375
Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp		380
385	390	395
Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp		400
	405	410
Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu		415
	420	425
Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala		430
	435	440
Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly		445
	450	455
Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu		460
465	470	475
Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn		480
	485	490
Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro		495
	500	505
Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile		510
	515	520
Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile		525
	530	535
Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys		540
545	550	555
Val Leu		560

<210> 60
 <211> 429
 <212> DNA
 <213> HIV

<400> 60
 atggtgggtt ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct 60
 gtatgacttta gccacttttt aaaagaaaaag gggggacttg aagggtctaat tcaactccaa 120
 agaagacaag atatccttga tctgtggatc taccacacac aaggctactt ccctgatitgg 180
 cagaactaca caccagggcc aggggtcaga tatccactga cctttggatg gtgctacaag 240
 ctagtaccag ttgagccaga taaggtagaa gaggccaata aaggagagaa caccagcttg 300
 ttacaccctg tgagcctgca tgggatggat gaccgcgaga gagaagtgtt agagtggagg 360
 ttigacagcc acctagcatt tcatcacgtg gcccagagac tgcatccgga gtacttcaag 420
 aactgctga 429

<210> 61
 <211> 142
 <212> PRT

<213> HIV

<400> 61

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

<210> 62

<211> 1698

<212> DNA

<213> hiv

<400> 62

atgggccc	atagccctat	tgagactgtg	tcagtaaaat	taaagccagg	aatggatg	60
ccaaaagtta	aacaatggcc	atlgacagaa	gaaaaaataa	aagcattagt	agaaatttgt	120
acagagatgg	aaaaggaagg	gaaaatttca	aaaattgggc	ctgaaaatcc	atacaatact	180
ccagtatttg	ccataaagaa	aaaagacagt	actaaatgga	gaaaattagt	agatttcaga	240
gaacttaata	agagaactca	agacttcttg	gaagttcaat	taggaatacc	acateccgea	300
gggttaaaaa	agaaaaaatc	agtaacagta	ctggatgttg	gtgatgcata	tttttcagtt	360
cccttagatg	aagacttcag	gaataatact	gcatttacca	tacctagtat	aaacaatgag	420
acaccaggga	ttagatatac	gtacaatgtg	cttccacagg	gatggaaagg	atcaccagea	480
atatccaaa	gtagcatgac	aaaaatctta	gagcctttta	gaaaacaaaa	tecagacata	540
gtatctatc	aatacatgga	tgatttgtat	gtaggatctg	acttagaaat	agggcagcat	600
agaacaaaaa	tagaggagct	gagacaacat	ctgttgaggt	ggggacttac	cacaccagac	660
aaaaaacatc	agaaagaacc	tccttctctt	tggatggggt	atgaactcca	tcctgataaa	720
tggacaglac	agcctatagt	gtgtccagaa	aaagacagct	ggactgtcaa	tgacatacag	780
aagttagtag	ggaaattgaa	ttgggcaggt	cagatttacc	cagggaattaa	agtaaggcaa	840
ttatgtaaac	tccttagagg	aaacaaagca	ctaacagaag	taataccact	aacagaagaa	900
gcagagctag	aactggcaga	aaacagagag	attctaaaag	aaccagtaca	tggagtgtat	960
tatgacccat	caaaagacit	aatagcagaa	atacagaage	aggggcaggg	ccaatggaca	1020
tatcaaatit	atcaagagcc	atttataaat	ctgaaaacag	gaaaataatc	aagaatgagg	1080
ggtgccccaca	ctaatgatgt	aaaacaattt	acagaggeag	tgcataaaat	aaccacagaa	1140
agcatagtaa	tatggggaaa	gaactctaaa	tttaaaetgc	ccatacaaaa	ggaacatgg	1200
gaaacatggt	ggacagagta	ttggcagccc	acctggatgc	ctgagtggga	gtttgttaat	1260
acccctccct	tagtgaattt	atgttaccag	ttagagaagag	aaccctatgt	aggagcagaa	1320
accttctatg	tagatggggc	agactaaat	taggaanage	aggatatgtt		1380
actaatagag	gaagacaaaa	agttgtccac	ctaaetgaca	caacaataca	gaagactgag	1440
ttacaagcaa	tttatctage	tttgcaggat	tegggattag	aagtaaacat	agtaacagac	1500
tcacaatatg	cattaggaat	cattcaagca	caaccagatc	aaagtgaate	agagttagtc	1560
aatcaataaa	tagagcagtt	ataaaaaaag	gaaaagggtc	atctggcatg	ggtaccagca	1620
cacaaaggaa	tggaggaaa	tgaaacagta	gataatttag	tcagtgtctg	aatcaggaaa	1680
gtactatit	tagatttaa					1698

<210> 63

<211> 565

<212> PRT

<213> HIV

<400> 63

Met	Gly	Pro	Ile	Ser	Pro	Ile	Glu	Thr	Val	Ser	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	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 Ser Val Thr Val Leu Asp
 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 Asp Asp 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 Leu 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 Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
 275 280 285
 Lys Ala Leu Thr Glu Val Ile 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 Thr 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 Thr 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 Asp Gly Ala Ala
 435 440 445
 Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
 450 455 460
 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
 465 470 475 480
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 485 490 495
 Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 500 505 510
 Asp Gln Ser Glu Ser Glu Leu Val Asn 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 Ala Gly Ile Arg Lys
 545 550 555 560
 Val Leu Phe Leu Asp
 565

<210> 64
 <211> 3213
 <212> DNA
 <213> HIV

<400> 64
 atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
 ctgcgcccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgagggag 120
 ctitgaacggt ttgccgtgaa cccaggcctg ctggaaacat ctgagggatg tcgccagatc 180
 ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc cttgtataac 240
 acagtggcta ccttctactg cgtaaccag aggatcgaga ttaaggatag caaggaggcc 300

```

ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga aggcccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aactatccia ttgtccaaaa catlcagggc 420
cagatgglic atcaggccat cagcccccg acgtcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct ttctcctga ggttatcccc atgttctccg ctttgagtga gggggccact 540
cctcaggacc tcaatacaat gcllaalacc gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctalcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
ggcccaatcg cggccggaca gatgcgggag cctcggcggt ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatgggat accaacaatc ctcccatccc agltggagaa 780
atctataaac ggtlgaalca tctcggcttc aataaaattg ttagaatgla cctcccgaca 840
tccatcctlg acalltagaca gggacccaaa gagccltita gggattacgt cgaccggltt 900
tataagaccc tgcgagcaga caggagcctc caggaggica aaaactggat gacggagaca 960
ctcctgglac agaacgclaa ccccgactgc aaaacaatc tgaaggcact agggccggct 1020
gccaccctgg aagagatgat gaccgcctgt caggggagtag gcggaccggg acacaaagcc 1080
agagtgtlga tgggccccat tagccclatt gagacttgtt cagtaaaatt aaagccagga 1140
atggatggcc caaaagttaa acaatggcca ttgacagaag aaaaaataaa agcattagta 1200
gaaatttlla cagagatgga aaaggaaggg aaaaattcaa aaattgggct tgaaaaatcca 1260
tacaatactc cagttatllc calaaagaaa aaagacagta claaatggag aaaattagta 1320
gatttcagag aacttaataa gagaactcaa gacttctggg aagttaaat aggaatacca 1380
catcccgag ggtlaaaaaa gaaaaatca glaacaglac tggatgtggg tgatgcata 1440
ttttcagtic ccttagatga agacttcagg aaatatactg cattaccat acctagata 1500
aacaatgaga caccagggat tagatatcag tacaatgtgc tccacaggg atggaaagga 1560
tcaccagcaa tatccaag tagcatgaca aaaactctag agccctllag aaaacaaat 1620
ccagacatag ttatctatca atacatggat gatttgtatg taggatctga cttagaata 1680
ggcgagcala gaacaaaaat agaggagctg agacaacatc tgttgaggig gggacttacc 1740
acaccagaca aaaaacatca gaagaaacct ccattccctt ggatggglla tgaactccat 1800
cctgataaat ggacagtaaa gccatagatg ctgccagaaa aagacagctg gactgtcaat 1860
gacatacaga agtttagtgg gaaatllgat tgggcaaglc agatttacc agggatlaaa 1920
gtaaggcaat tatgtlaaact ccttagagga accaaagcac taacagaagi aataccacta 1980
acagaagaag cagagctaga actggcagaa aacagagaga ttctaaaaga accagtacat 2040
ggagtgtatt atgacccatc aaaagactta atagcagaaa tacagaagca ggggcaaggc 2100
caatggacat atcaaatlta tcaagagcca ttlaaaaaatc tgaaaacagg aaaatatgca 2160
agaatgaggg gtgcccacac taatgatgla aaacaattaa cagaggcagt gcaaaaaata 2220
accacagaaa gcatagtaat atggggaag actcctaaat ttaacttgc cacaacaaag 2280
gaaacatggg aaacatggg gacagatlat tggcaagcca cctggattcc tgagtgggag 2340
ttgttlaala cccctccctt agtgaattta tggtaaccgt tagagaaaga acccatagta 2400
ggagcagaaa ccttctatgt agatgggcca gctaacaggg agactaaat agggaaagca 2460
ggatatgtta ctaatagagg aagacaaaaa gttgtcacc taactgacac aacaaatcag 2520
aagactgagt tacaagcaat ttatctagct ttgcaggat cgggallaga agtaaacata 2580
glaacagact cacaatatgc attaggaatc attcaagcac aaccagatca aagtgaata 2640
gagtttagta atcaaatat agagcaglla ataaaaaag aaaaggleta tctggcatgg 2700
gtaccagcac acaaaagaa ttggaggaat gaacaagtag ataaattagt cagtgtgga 2760
atcaggaaag tactatlttt agatatgggt ggltttccag tcaacacctc ggtacctta 2820
agaccaatga cttacaagge agctgtatgt cttagccact ttltaaaaga aaggggggga 2880
ctggaagggc taattcact ccaagaaga caagataacc ttgatctgtg gatctaccac 2940
acacaaggct acttccctga ttgcagaae tacacaccag ggcaggggt cagataacca 3000
ctgacctttg gatgggtgta caagtagta ccagttgagc agataaaggt agaagaggee 3060
aataaaggag agaaccacag ctgtttacac cctgtgagcc tgcattggat ggtgacccg 3120
gagagagaag tgttagagtg gagglttgac agccacctag catttcaata cgtggcccca 3180
gagctgcatc cggagtaact caagaactgc tga 3213

```

<210> 65
 <211> 1070
 <212> PRT
 <213> HIV

<400> 65
 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 Lys 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 Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Asp Thr Gly His Ser Asn 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 Thr 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 Glu 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 Met Gly	Pro Ile Ser			
	355	360	365			
Pro Ile Glu Thr Val	Ser Val Lys Leu	Lys Pro Gly Met	Asp Gly Pro			
	370	375	380			
Lys Val Lys Gln Trp	Pro Leu Thr Glu	Glu Lys Ile Lys	Ala Leu Val			
385	390	395	400			
Glu Ile Cys Thr Glu	Met Glu Lys Glu	Gly Lys Ile Ser	Lys Ile Gly			
	405	410	415			
Pro Glu Asn Pro Tyr	Asn Thr Pro Val	Phe Ala Ile Lys	Lys Lys Asp			
	420	425	430			
Ser Thr Lys Trp Arg	Lys Leu Val Asp	Phe Arg Glu Leu	Asn Lys Arg			
	435	440	445			
Thr Gln Asp Phe Trp	Glu Val Gln Leu	Gly Ile Pro His	Pro Ala Gly			
	450	455	460			
Leu Lys Lys Lys Lys	Ser Val Thr Val	Leu Asp Val Gly	Asp Ala Tyr			
465	470	475	480			
Phe Ser Val Pro Leu	Asp Glu Asp Phe	Arg Lys Tyr Thr	Ala Phe Thr			
	485	490	495			
Ile Pro Ser Ile Asn	Asn Glu Thr Pro	Gly Ile Arg Tyr	Gln Tyr Asn			
	500	505	510			
Val Leu Pro Gln Gly	Trp Lys Gly Ser	Pro Ala Ile Phe	Gln Ser Ser			
	515	520	525			
Met Thr Lys Ile Leu	Glu Pro Phe Arg	Lys Gln Asn Pro	Asp Ile Val			
	530	535	540			
Ile Tyr Gln Tyr Met	Asp Asp Leu Tyr	Val Gly Ser Asp	Leu Glu Ile			
545	550	555	560			
Gly Gln His Arg Thr	Lys Ile Glu Glu	Leu Arg Gln His	Leu Leu Arg			
	565	570	575			
Trp Gly Leu Thr Thr	Pro Asp Lys Lys	His Gln Lys Glu	Pro Pro Phe			
	580	585	590			
Leu Trp Met Gly Tyr	Glu Leu His Pro	Asp Lys Trp Thr	Val Gln Pro			
	595	600	605			
Ile Val Leu Pro Glu	Lys Asp Ser Trp	Thr Val Asn Asp	Ile Gln Lys			
	610	615	620			
Leu Val Gly Lys Leu	Asn Trp Ala Ser	Gln Ile Tyr Pro	Gly Ile Lys			
625	630	635	640			
Val Arg Gln Leu Cys	Lys Leu Leu Arg	Gly Thr Lys Ala	Leu Thr Glu			
	645	650	655			
Val Ile Pro Leu Thr	Glu Glu Ala Glu	Leu Glu Leu Ala	Glu Asn Arg			
	660	665	670			
Glu Ile Leu Lys Glu	Pro Val His Gly	Val Tyr Tyr Asp	Pro Ser Lys			
	675	680	685			
Asp Leu Ile Ala Glu	Ile Gln Lys Gln	Gly Gln Gly Gln	Trp Thr Tyr			
	690	695	700			
Gln Ile Tyr Gln Glu	Pro Phe Lys Asn	Leu Lys Thr Gly	Lys Tyr Ala			
705	710	715	720			
Arg Met Arg Gly Ala	His Thr Asn Asp	Val Lys Gln Leu	Thr Glu Ala			
	725	730	735			
Val Gln Lys Ile Thr	Thr Glu Ser Ile	Val Ile Trp Gly	Lys Thr Pro			
	740	745	750			

Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr
755 760 765
Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
770 775 780
Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val
785 790 795 800
Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys
805 810 815
Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val
820 825 830
Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr
835 840 845
Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser
850 855 860
Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser
865 870 875 880
Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val
885 890 895
Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln
900 905 910
Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp
915 920 925
Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
930 935 940
Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
945 950 955 960
Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
965 970 975
Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
980 985 990
Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
995 1000 1005
Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
1010 1015 1020
Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
1025 1030 1035 1040
Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser His Leu Ala Phe His
1045 1050 1055
His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
1060 1065 1070

<210> 66
<211> 3213
<212> DNA
<213> HIV

<400> 66

atgggtgccc gagcttgggt actgtctggt ggagagctgg acagatggga gaaattagg 60
ctgcgcccg gagcagaaa gaatacaag ctcaagcata tctgttggtg ctcagggag 120
cttgaacgggt ttgcctgtaa cccaggtctg ctggaaacat ctgaggatg tgcagagatc 180
ctggggcaat tgcagcatt cctcagacc gggagtgaag agctgaggtc ctgttatatc 240
acagttgcta cctctactg cgtacacag aggatcgaga ttaaggatac caaggagacc 300
ttggacaaaa ttgaggagga gaaaacaag agcaagaaga agcccagca ggcagctgt 360
gacactgggc atagcaacca ggtatcacg aactatccta ttgtccaaa catcagggc 420
cagatgggtc atcaggecat cagccccgg acgtctaatg cctgggtgaa ggttgtgaa 480
gagaaggcct ttctctctg ggttateccc atgttctcgg ctlttgatga ggggtccact 540
cctcaggacc tcaatacaat gcttaatacc gtggggggcc atcagggcgc catgcaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gattgggaca gattgcatcc cgtccacgt 660
ggcccaatcg cgcctggaca gatgcgggag cctgcgggt ctgacattgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctccatccc agttggagaa 780
atctataaac ggttgatcat cctgggcttg acaagatcg tgcctatga ctctccgaca 840
tccatccttg acattagaca ggaacccaaa gacctttta ggaattagct cgaagggttt 900
tataagagacc tgcagcaga gaggcctct caggaggta aaaaatggat gaaggagaca 960
ctcctggtac agaacctaa ccccgactgc aaaaatct tgaaggact agggccgggt 1020
gccaccctgg aagagatgat gaccgctgt caggagtag ggggcccgg acacaagcc 1080
agagtgttga tgggcccct tagcctatt gagactgtgt cagtataatt aaagccagga 1140
atggatggcc caaaagttaa acaatggcca ttgacagaag aaaaaataaa agcatlagta 1200
gaaatttgta cagagatgga aaaggagggg aaaaattcaa aaattgggac tgaataatca 1260
tacaatactc cagtatttc catanaaaa aaagacagta ctaaatggag aaatttagta 1320
gatttcagag aacttaataa gagaactcaa gaattctggg aattcaatt aggaatacca 1380
catccgcag ggttaaaaaa gaaaaatca gtacagtac tggatgtggg tgaigcatat 1440
tttccagttc cttagatga agacttcagg aaatatactg catctaccat acctagata 1500
aacaatgaga caccagggat tagatatcag tacaatgtgc ttccacaggg atggaaagga 1560
tcaccagcaa tatccaaag tagatgaca aaaaatttag agcttttag aaaaacaaat 1620

```

ccagacatag ttatctatca atacatggat gatttgtatg taggatctga cttagaata 1680
gggcagcata gaacaaaaat agaggagctg agacaacatc tgttgagggtg gggacttacc 1740
acaccagaca aaaacatca gaaagaacct ccattccitt ggatgggtta tgaactccat 1800
cctgataaat ggacagtaca gcctatagtg ctgccagaaa aagacagctg gactgtcaat 1860
gacatacaga agttagtggg gaaatigaat tgggcaagtc agatttacc agggaataaa 1920
gtaaggcaat tatgtaaat ccttagagga accaaagcac taacagaagt aataccacta 1980
acagaagaag cagagctaga actggcagaa aacagagaga ttctaaaaga accagtacat 2040
ggagtgtatt atgacccatc aaaagactta atagcagaaa tacagaagca ggggcaaggc 2100
caatggacat atcaaattta tcaagagcca tttaaaaatc tgaaacaggg aaaatatgca 2160
agaatgaggg gtgcccacac taatgatgta aaacaattaa cagaggcagt gcaaaaaata 2220
accacagaaa gcatagtaat atggggaaag actcctaata ttaactgcc catacaaaag 2280
gaaacatggg aaacatgggtg gacagagtat tggcaagcca cctggattcc tgagtgggag 2340
tttgttaata ccctccctt agtgaatta tggtagcagt tagagaaaga acccatagta 2400
ggagcagaaa ccttctatgt agatggggca gctaacaggg agactaaatt agggaaaagca 2460
ggatatgtta ctaatagagg aagacaaaaa gtgtgaccc taactgacac aacaaatcag 2520
aagactgagt tacaagcaat ttatctagct ttgcaggatt cgggattaga agtaaacata 2580
gtaacagact cacaatatgc attaggaatc attcaagcac aaccagatca aagtgaatca 2640
gagttagtca atcaataat agagcagtta ataaaaaagg aaaaggctca tctggcatgg 2700
gtaccagcac acaaaaggaaat tggaggaaat gaacaagtag ataaattagt cagtgtgga 2760
atcaggaaag tactatitit agatagggtg ggttttccag tcacacctca ggtaccttta 2820
agaccaatga cttaacaggc agctgtagat cttagccact ttttaaaaga aaagggggga 2880
ctggaagggc taattcactc ccaagaaga caagatatcc ttgatctgtg gatctaccac 2940
acacaaggct acttccctga ttggcagaac tacacaccag ggccaggggg cagatatcca 3000
ctgacctttg gatggtgcta caagctagta ccagttgagc cagataaggt agaagaggcc 3060
aataaaggag agaaccaccag cttgttacac cctgtgagcc tgcattgggt ggatgaccgc 3120
gagagagaag tgttagagtg gaggtttgac agccacctag catttcatca cgtggcccga 3180
gagctgcac cggagtactt caagaactgc tga 3213

```

<210> 67
 <211> 1070
 <212> PRT
 <213> HIV

<400> 67

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

Leu Gly Pro	Ala Ala Thr	Leu Glu Glu	Met Met Thr	Ala Cys Gln Gly
Val Gly Gly	Pro Gly His Lys	Ala Arg Val	Leu Met Gly	Pro Ile Ser
Pro Ile Glu	Thr Val Ser	Val Lys Leu	Lys Pro Gly	Met Asp Gly Pro
Lys Val Lys	Gln Trp Pro	Leu Thr Glu	Glu Lys Ile	Lys Ala Leu Val
Glu Ile Cys	Thr Glu Met	Glu Lys Glu	Gly Lys Ile	Ser Lys Ile Gly
Pro Glu Asn	Pro Tyr Asn Thr	Pro Val Phe	Ala Ile Lys	Lys Lys Asp
Ser Thr Lys	Trp Arg Lys Leu	Val Asp Phe	Arg Glu Leu	Asn Lys Arg
Thr Gln Asp	Phe Trp Glu	Val Gln Leu	Gly Ile Pro	His Pro Ala Gly
Leu Lys Lys	Lys Lys Ser	Val Thr Val	Leu Asp Val	Gly Asp Ala Tyr
Phe Ser Val	Pro Leu Asp	Glu Asp Phe	Arg Lys Tyr	Thr Ala Phe Thr
Ile Pro Ser	Ile Asn Asn	Glu Thr Pro	Gly Ile Arg	Tyr Gln Tyr Asn
Val Leu Pro	Gln Gly Trp Lys	Gly Ser Pro	Ala Ile Phe	Gln Ser Ser
Met Thr Lys	Ile Leu Glu	Pro Phe Arg	Lys Gln Asn	Pro Asp Ile Val
Ile Tyr Gln	Tyr Met Asp	Leu Tyr Val	Gly Ser Asp	Leu Glu Ile
Gly Gln His	Arg Thr Lys	Ile Glu Glu	Leu Arg Gln	His Leu Leu Arg
Trp Gly Leu	Thr Thr Pro	Asp Lys Lys	His Gln Lys	Glu Pro Pro Phe
Leu Trp Met	Gly Tyr Glu	Leu His Pro	Asp Lys Trp	Thr Val Gln Pro
Ile Val Leu	Pro Glu Lys	Asp Ser Trp	Thr Val Asn	Asp Ile Gln Lys
Leu Val Gly	Lys Leu Asn	Trp Ala Ser	Gln Ile Tyr	Pro Gly Ile Lys
Val Arg Gln	Leu Cys Lys	Leu Leu Arg	Gly Thr Lys	Ala Leu Thr Glu
Val Ile Pro	Leu Thr Glu	Glu Ala Glu	Leu Glu Leu	Ala Glu Asn Arg
Glu Ile Leu	Lys Glu Pro	Val His Gly	Val Tyr Tyr	Asp Pro Ser Lys
Asp Leu Ile	Ala Glu Ile	Gln Lys Gln	Gly Gln Gly	Gln Tip Thr Tyr
Gln Ile Tyr	Gln Glu Pro	Phe Lys Asn	Leu Lys Thr	Gly Lys Tyr Ala
Arg Met Arg	Gly Ala His	Thr Asn Asp	Val Lys Gln	Leu Thr Glu Ala
Val Gln Lys	Ile Thr Thr	Glu Ser Ile	Val Ile Tip	Gly Lys Thr Pro
Lys Phe Lys	Leu Pro Ile	Gln Lys Glu	Thr Tip Glu	Thr Tip Tip Thr
Glu Tyr Trp	Gln Ala Thr	Trp Ile Pro	Glu Tip Glu	Phe Val Asn Thr
Pro Pro Leu	Val Lys Leu	Trp Tyr Gln	Leu Glu Lys	Glu Pro Ile Val
Gly Ala Glu	Thr Phe Tyr	Val Asp Gly	Ala Ala Asn	Arg Glu Thr Lys
Leu Gly Lys	Ala Gly Tyr	Val Thr Asn	Arg Gly Arg	Gln Lys Val Val
Thr Leu Thr	Asp Thr Thr	Asn Gln Lys	Thr Glu Leu	Gln Ala Ile Tyr
Leu Ala Leu	Gln Asp Ser	Gly Leu Glu	Val Asn Ile	Val Thr Asp Ser
Gln Tyr Ala	Leu Gly Ile	Ile Gln Ala	Gln Pro Asp	Gln Ser Glu Ser
Glu Leu Val	Asn Gln Ile	Ile Glu Gln	Leu Ile Lys	Lys Glu Lys Val
Tyr Leu Ala	Trp Val Pro	Ala His Lys	Gly Ile Gly	Gly Asn Glu Gln
Val Asp Lys	Leu Val Ser	Ala Gly Ile	Arg Lys Val	Leu Phe Leu Asp

Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 930 935 940
 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 945 950 955 960
 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 965 970 975
 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 980 985 990
 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 995 1000 1005
 Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 1010 1015 1020
 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 1025 1030 1035 1040
 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser His Leu Ala Phe His
 1045 1050 1055
 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
 1060 1065 1070

<210> 68
 <211> 3204
 <212> DNA
 <213> Hiv

<400> 68
 atgggtgccc gagcttcggg actgtctggt ggagagctgg acagatggga gaaaattagg 60
 ctgcgcccgaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggagg 120
 ctltgaacgggt ttgccgtgaa cccaggccctg ctggaacat ctgagggatg tcgccagatc 180
 ctggggcaat tgcagccatc cctccagacc ggagtgaaag agctgaggtc ctgtataac 240
 acagtggcta cctctactg cgtacaccag aggatcgaga ttaaggatac caaggaggcc 300
 ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga agggccagca ggcagctgct 360
 gacactgggc atagcaacca ggtatcacag aactatccta ttgtccaaaa cattcagggc 420
 cagatgggtc atcaggccat cagcccccg acgctcaatg cctgggtgaa ggttgtcgaa 480
 gagaaggcct ttctctctga ggttatcccc atgttctccg ctttgagtga gggggccact 540
 cctcaggacc tcaatacaat gcttaatacc gtggcgggcc atcaggccgc catgcaaatg 600
 ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcattc cgtccacgct 660
 ggcccaatcg cgcgggaca gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
 tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
 atctataaac ggtggatcat cctgggcctg aacaagatcg tgcgcatgta ctctccgaca 840
 tccatccctt acattagaca gggacccaaa gagccttita gggattacgt cgaccggttt 900
 tataagaccc tgcgagcaga caggccctct caggagggtca aaaactggat gacggagaca 960
 ctcttggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggt 1020
 gccacccctg aagagatgat gaccgcctgt caggagtag gcggaccggg acacaaagcc 1080
 agagtgttga tgggccccat cagtcctatc gagaccgtgc cggtgaaagt gaaacccggg 1140
 atggacggcc ccaaggctaa gcagtgcca ctaccagagg agaagatcaa ggccctggtg 1200
 gaga tctgca ccgagatgga gaaagagggc aagatcagca agatcgggcc tgagaacca 1260
 tacaacaccc ccgtgttgc catcaagaag aaggacagca ccaagtggcg caagctggtg 1320
 gatttccggg agctgaataa gcggaccag gatttctggg aggtccagct gggcatcccc 1380
 catccggccg gcttgaagaa gaagaagagc gtgaccgtgc tggacgtggg cgacgcttac 1440
 ttacgctcc ctctggacga ggactttaga agttacacc cctttaccat cccatctatc 1500
 aacaacagga cccctggcat catgatcag tacaacgtcc tccccaggg ctggaagggc 1560
 tctcccgcca ttctcagag ctccatgacc aagatcctgg agccgtttcg gaagcagaac 1620
 cccgatatcg tcatctacca gtacatggac gacctgtacg tgggctctga cctggaaatc 1680
 gggcagcatc gcacgaagat tgaggagctg aggcagcatc tgcctgagatg gggcctgacc 1740
 actccggaca agaagcatca gaaggagccg ccattctgtt ggatgggcta cgagctccat 1800
 cccgacaagt ggaccgtgca gcctatcgtc ctccccgaga aggcagactg gaccgtgaac 1860
 gacatccaga agctgggtgg caagctcaac tgggctagcc agatctatcc cgggatcaag 1920
 gtgcgccagc tctgcaaggt cgtgcggggc accaaggccc tgaccgaggt gattccccctc 1980
 acggaggaag ccgagctcga gctggctgag aaccgggaga tccctgaagg gcccgtgcac 2040
 ggctgtact atgaccctc caaggacctg atcgccgaaa tccagaagca gggccagggg 2100
 cagtggacat accagattta ccaggagcct ttcaagaacc tcaagaccgg caagtacgcc 2160
 cgcatgaggg gcgcccacac caacgatgtc aagcagctga ccgaggccgt ccagaagatc 2220
 acgaccgagt ccatctgagt ctgggggaag acaccaagt tcaagctgcc tatccagaag 2280
 gagacctggg agacgtgggt gaccgaatat tggcaggcca cctggattcc cgagtgggag 2340
 ttctgtaata caccctctct ggtgaagctg tggtagcagc tcgagaagga gcccatctg 2400
 ggcgcggaga catctacgt ggacggcgcg gccaacgcg aaacaaagct cgggaaggcc 2460
 gggtagctca ccaaccgggg ccgccagaag gtcgtcacc tgaccgacac cacaaccag 2520
 aagacggagc tgcaggccat ctatctcgtc ctccaggact cggcctgga ggtgaacatc 2580
 gtgacggaca gccagatcgc cgtgggcatt attcaggccc agccggacca gtccgagagc 2640
 gacttggtga accagattat ctagcagctg atcaagaag agaaggtcta cctcgcttg 2700
 gtccggggcc ataaggcat tggcggaac gagcagctg acaagctgt gagtgcgggg 2760
 attagaaagg tgcgtatggt gggttttcca gtcacacct aggtacctt aagaccaatg 2820
 acttacaagg cagctgtaga tcttagccac tttttaaaag aaaagggggg actggaaggg 2880
 ctaattcact cccaaagaag acaagatatc ctgtatctgt ggtactacca cacacaaggc 2940

tacttccctg attggcagaa ctacacacca gggccagggg tcagatatcc actgaccttt 3000
 ggatggtgct acaagctagc accagltgag ccagataagg tagaagaggc caataaagga 3060
 gagaacacca gcctgttaca cccgtgtgagc ctgcatggga tggatgaccc ggagagagaa 3120
 gtgttagagc ggaggttga cagccgccta gcatttcac acgtggcccg agagctgcat 3180
 ccggagtact tcaagaactg ciga 3204

<210> 69
 <211> 1067
 <212> PRT
 <213> HIV

<400> 69
 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 Lys 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 Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn 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 Thr 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 Glu 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 Met Gly Pro Ile Ser
 355 360 365
 Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro
 370 375 380
 Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val
 385 390 395 400
 Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly
 405 410 415
 Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp
 420 425 430
 Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg
 435 440 445
 Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly
 450 455 460
 Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr
 465 470 475 480
 Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr
 485 490 495
 Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn

Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser
Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val
Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile
Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg
Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe
Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro
Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys
Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys
Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu
Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg
Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys
Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr
Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala
Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala
Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro
Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr
Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr
Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val
Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln	Lys	Val	Val
Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	Tyr
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln	Ser	Glu	Ser
Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val
Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln
Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Met	Val	Gly
Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala
Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly
Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr
His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro
Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro
Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser
Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu
Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala
Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys					

<210> 70
<211> 1518

<212> DNA
<213> HIV

<400> 70

```

atgggtgccc gagcttcggt acgtctcgtt ggagagctgg acagatggga gaaaattagg 60
ctgcgcccgg gaggcacaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggagg 120
cttgaacggt ttgccgtlga cccaggccct ctggaaacat ctgagggatg tcgccagatc 180
ctgggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggtc ctgtatatac 240
acagtggtcta cctcttactg cgtacaccag aggatcgaga ttaaggatag caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga agggccagca ggcagctgct 360
gacactgggc atagcaacca ggtatcacag aaatatacct ttgtccaaaa catlcagggc 420
cagatgggttc atcaggccat cagccccggg acgctcaatg cctgggtlga gggtgtcgaa 480
gagaaggcct ttctctccta gggtatcccc atgttctcgg ctltgagtga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtggcgggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gatlgtatcc cgtccacgct 660
ggcccaatcg cgcccggaac gatlgtggag cctcggggct ctgacatlgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agtltggaga 780
atctataaac ggttggatcat tctcgggttc aataaaatgt ttagaatgta ctctccgaca 840
tccatctctg acattagaca gggaaccaaa gagccttita gggatllact cgaccggltt 900
talaagaccc tgcgagcaga gcaggcctct caggagggtc aaaactggat gacggagaca 960
ctcctgggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
gccacccctg aagagatgat gaccgctctg cagggagtag gcggaccctg acacaaagcc 1080
agagtgltga tgggtgggtt tccagtcaca cctcagggtc ctlttaagacc aatgacttac 1140
aaggcagctg tagatcttag ccacttltta aaagaaaagg ggggactlga agggctlaat 1200
cactcccaaa gaagacaaga tatccttgat ctgtggatct accacacaca aggtacttcc 1260
cctgattggc agaactacac accagggtcca ggggtcagat atccactgac ctltggatgg 1320
tgttacaagc tagtaccagt tgaagcagat aaggtagaag agggcaataa aggagagatc 1380
accagctgtt tacacctgtt gaccctgcat gggatggatg acccgagag agaatgtlta 1440
gagtgagggt ttgacagccg cctagcattt catcacgtgg cccgagagct gcatccggag 1500
tacttcaaga actgctga

```

<210> 71
<211> 505
<212> PRT
<213> HIV

<400> 71

```

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 Tyr Lys 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 Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
100          105          110
Lys Lys Ala Gln Gln Ala Ala Asp Thr Gly His Ser Asn 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 Thr 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 Glu 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 Met Val Gly Phe Pro
 355 360 365
 Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala Val
 370 375 380
 Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile
 385 390 395 400
 His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr
 405 410 415
 Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val
 420 425 430
 Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu
 435 440 445
 Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu
 450 455 460
 His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu
 465 470 475 480
 Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg Glu
 485 490 495
 Leu His Pro Glu Tyr Phe Lys Asn Cys
 500 505

<210> 72
 <211> 1689
 <212> DNA
 <213> HIV

<400> 72
 atgggcccc tcagtcctcat cgagaccgtg ccggtgaagc tgaaccgccg gatggacggc 60
 cccaagggtca agcagtggtc actcaccgag gagaagatca aggccttggg ggagatctgc 120
 accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
 cccgtgtttt ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
 gagctgaata agcggaccca ggatttcttg gaggtccagc tgggcatccc ccatccggcc 300
 ggccctgaaga agaagaagag cgtgaccgtg ctggacgtgg gcgacgctta cttcagcgtc 360
 cctctggacg aggactttag aaagtacacc gcctttacca tcccatctat caacaacgag 420
 acccttggca tcagatatca gtacaacgtc ctccccagg gctggaaggg ctctcccgcc 480
 attttccaga gctccatgac caagatcctg gagccgtttc ggaagcagaa ccccgatc 540
 glcatctacc agtacctgga cgacctgtac gtgggctctg acctggaat cgggcagcat 600
 cgcacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcctgac cactccggac 660
 aagaagcatc agaaggagcc gccattcctg aagatgggct acgagctcca tcccagacaag 720
 tggaccgtgc agcctatctg cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
 aagcttggtg gcaagctcaa ctgggctagc cagatctatc ccgggatcaa ggtgcccagg 840
 ctctgcaagc tgcctgcggg caccaaggcc ctgaccgagg tgattccctt cacggaggaa 900
 gccgagctcg agctggctga gaaccgggag atcctgaagg agcccgtgca cggcgtgtac 960
 tatgaccctt ccaaggacct gatcgccgaa atccagaagc agggccaggg gcagtggaca 1020
 taccagattt accaggagcc tttaagaac ctcaagacc gcaagtacgc ccgcatgagg 1080
 tgcgcccaca ccaacgatgt caagcagctg accgagggcc tccagaagat caccagcgag 1140
 tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
 gagacgtggt ggaccgaata ttggcaggcc acctggattc ccgagtggga gttcgtgaat 1260
 acacctcttc tggtagagct gtgtgtaccg ctcgagaagg agcccatctg gggcgcgag 1320
 acattctacg tggacggcgc ggccaaccgc gaaacaaagc tgggaaggc cgggtacgtc 1380
 accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
 ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
 agccagtagc cgttgggcat tattcaggcc cagccggacc agtccgagag cgaactgggt 1560
 aaccagatta tgcagcagct gatcaagaaa gagaaggctt acctcgcttg ggtcccgccc 1620
 cataagggca ttggcggaac cgagcagggt gacaagctgg tgagtgcggg gattagaaag 1680
 gtgctgtaa 1689

<210> 73
 <211> 3204
 <212> DNA
 <213> HIV

<400> 73
 atgggtgcc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
 ctgcgcccg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggag 120
 ctigaacggt ttgccgtgaa cccaggcctg ctggaacat ctgagggatg tcgccagatc 180
 ctggggcaat tgcagccatc cctccagacc gggagtgaag agctgaggct ctgtataac 240
 acagtggcta cctctactg cgtacaccag aggatcgaga ttaaggatc caaggaggcc 300

```

ttggacaaaa ttgaggaggga gcaaaacaag agcaagaaga agggccagca ggcagctgct 360
gacactgggc atagcaacca ggatcaccag aactatccta ttgtccaaaa caticagggc 420
cagatgggtc atcaggccat cagccccggg acgctcaatg ccttgggtgaa ggttgtcgaa 480
gagaaggcct ttctctctga ggttatcccc atgttctccg ctltgagtga gggggccact 540
cctcaggacc tcaatacaal gctlaalacc gtgggcggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gagtgcacc cgtccacgct 660
ggcccaatcg cgcccgagca gatgcgggag cctcgcggct ctgacattgc cggcaccacc 720
tctacactgc aagagcaaal cggatggatg accaacaatc ctccccatcc agtlggagaa 780
atctataaac ggtggatcat cctgggcccg aacaagatcg tgcgcattga ctctccgaca 840
tccatcctlg acattlagaca gggaccctaa gagcctltta gggattacgt cgaccggltt 900
tataagaccc tgcgagcaga gcaggccctc caggagggtca aaaactggat gacggagaca 960
ctccgtgtac agaacgtctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
gccacccctgg aagagatgat gaccgctgtt caggagtagt gcggaccggg acacaaagcc 1080
agagtgltga tgggccccat cagtcctatc gagaccgtgc cggltgaagct gaaacccggg 1140
atggacggcc ccaagggtcaa gcagtgccca ctaccggagg agaagatcaa ggccctgggt 1200
gagatctgca ccgagatgga gaaagagggc aagatcagca agatcgggcc tgagaacca 1260
tacaacaccc ccgtgtllgc calcaagaag aaggacagca ccaagtggcg caagctgggt 1320
gatttccggg ccgtgaalaa gcgggaccag gatttctggg aggtccagct gggcatcccc 1380
catccggcgg gccigaagaa gaagaagagc glgaccgtgc tggacgtggg cgacgtttac 1440
ttcagcgtcc ctctggacga ggactllaga aaglacaccg cctltaccat cccatctatc 1500
aacaacgaga cccctggcat cagatctcag tacaacgtcc tccccagggg ctggaagggc 1560
tctcccccca llttccagag ctccatgacc aagatccctg agccgtttcg gaagcagaac 1620
cccgatctac tcatctacca glcatggag gacctgtacg tgggtcttga cctggaaatc 1680
gggcagcatc gcacgaagat lgaggagctg aggcagcatc tgcctgagat gggccctgacc 1740
actccggaca agaagcatca gaaggagccg ccatlccctga agatgggcta cgagctccat 1800
cccgacaagt ggaccgttga gcctatctgc ctccccgaga aggcagctgc gaccgtgaac 1860
gacatccaga agctgggtgg caagctcaac tgggtctagc agatctatcc cgggatcaag 1920
gtgcgccagc tctgcaagct gctgcgcggc accaaggccc tgaccgaggt gatccccctc 1980
acggagggaag ccgagcttga gctggctgag aaccgggaga tctgaagga gccctgtcac 2040
ggcgtgtact atgacccctc caaggacctg atgcgccgaa tccagaagca gggccagggg 2100
cagtggaat accagatlla ccaggagcct tcaagaacc tcaagaccgg caagtaecce 2160
cgcatgaggg gcgccacac caacgatgtc aagcagctga ccgaggccgt ccagaagatc 2220
acgaccgagt ccctcgtgat ctgggggaag acaccaagt tcaagctgc tatccagaag 2280
gagacctggg agacgtgggt gaccgaalat lggcaggcca cctggattcc cgagtgggag 2340
ttcgtgaata cactctctct ggltgaagctg tggtaaccag tggagaagga gccctctgtg 2400
ggcgcgagga cattctactg ggacggcgcg gccaacggcg aaacaaagt cgggaaggcc 2460
gggtacgtca ccaaccgggg ccgccagaag gtctgtaccc tgaccgaac caccacccag 2520
aagacggagc tgcaggccat ctatctctgt ctccaggact ccggccctga ggtgaacatc 2580
glgacggaca gccagtaecg gctgggcatl attcaggccc agccggacca gtccagagag 2640
gaactgttga accagatlat cgagcagctg atcaagaag agaaggtcta cctcgcctgg 2700
gtcccgcccc ataaggccat tggcgccaac gacaggttcg acaagctgtt gactgcgggg 2760
attagaaagg tgcgtatgtt ggglltcca gtcacacctc aggtaccttt aagaccaatg 2820
acttacaagg cagctgtaga tcttgcacac tttttaaag aaaagggggg actggaaggg 2880
ctaattcact cccaaagaag acagatate ctltgatctgt ggatctacca caccacagge 2940
tacttccctg atggcgagaa ctacacacca gggccagggg tcatatate actgaccttt 3000
ggatggtgtc acaagctagt acaagttgag ccagataagg tagaagagge caataaagga 3060
gagaacacca gcttgttaca cctctgagc ctgcattggg tggatgccc ggagagagaa 3120
gtgttagagt ggaggttga cagccgcta gcatltctat acttgcccg agagctgcat 3180
ccggagtact tcaagaactg ctga 3204

```

<210> 74
 <211> 1067
 <212> PRT
 <213> HIV

<400> 74

Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Ile
1				5					10				15		
Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys	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	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys
	100					105					110				
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	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 Thr	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 Glu	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 Met	Gly Pro Ile Ser			
	355	360	365			
Pro Ile Glu Thr	Val Ser Val Lys	Leu Lys Pro Gly	Met Asp Gly Pro			
	370	375	380			
Lys Val Lys Gln	Trp Pro Leu Thr	Glu Glu Lys Ile	Lys Ala Leu Val			
	385	390	395			400
Glu Ile Cys Thr	Glu Met Glu Lys	Glu Gly Lys Ile	Ser Lys Ile Gly			
	405	410	415			
Pro Glu Asn Pro	Tyr Asn Thr Pro	Val Phe Ala Ile	Lys Lys Lys Asp			
	420	425	430			
Ser Thr Lys Trp	Arg Lys Leu Val	Asp Phe Arg Glu	Leu Asn Lys Arg			
	435	440	445			
Thr Gln Asp Phe	Trp Glu Val Gln	Leu Gly Ile Pro	His Pro Ala Gly			
	450	455	460			
Leu Lys Lys Lys	Lys Ser Val Thr	Val Leu Asp Val	Gly Asp Ala Tyr			
	465	470	475			480
Phe Ser Val Pro	Leu Asp Glu Asp	Phe Arg Lys Tyr	Thr Ala Phe Thr			
	485	490	495			
Ile Pro Ser Ile	Asn Asn Glu Thr	Pro Gly Ile Arg	Tyr Gln Tyr Asn			
	500	505	510			
Val Leu Pro Gln	Gly Trp Lys Gly	Ser Pro Ala Ile	Phe Gln Ser Ser			
	515	520	525			
Met Thr Lys Ile	Leu Glu Pro Phe	Arg Lys Gln Asn	Pro Asp Ile Val			
	530	535	540			
Ile Tyr Gln Tyr	Met Asp Asp Leu	Tyr Val Gly Ser	Asp Leu Glu Ile			
	545	550	555			560
Gly Gln His Arg	Thr Lys Ile Glu	Glu Glu Leu Arg	Gln His Leu Leu Arg			
	565	570	575			
Trp Gly Leu Thr	Pro Asp Lys Lys	His Gln Lys Glu	Pro Pro Phe			
	580	585	590			
Leu Trp Met Gly	Tyr Glu Leu His	Pro Asp Lys Trp	Thr Val Gln Pro			
	595	600	605			
Ile Val Leu Pro	Glu Lys Asp Ser	Trp Thr Val Asn	Asp Ile Gln Lys			
	610	615	620			
Leu Val Gly Lys	Leu Asn Trp Ala	Ser Gln Ile Tyr	Pro Gly Ile Lys			
	625	630	635			640
Val Arg Gln Leu	Cys Lys Leu Leu	Arg Gly Thr Lys	Ala Leu Thr Glu			
	645	650	655			
Val Ile Pro Leu	Thr Glu Glu Ala	Glu Leu Glu Leu	Ala Glu Asn Arg			
	660	665	670			
Glu Ile Leu Lys	Glu Pro Val His	Gly Val Tyr Tyr	Asp Pro Ser Lys			
	675	680	685			
Asp Leu Ile Ala	Glu Ile Gln Lys	Gln Gly Gln Gly	Gln Trp Thr Tyr			
	690	695	700			
Gln Ile Tyr Gln	Glu Pro Phe Lys	Asn Leu Lys Thr	Gly Lys Tyr Ala			
	705	710	715			720
Arg Met Arg Gly	Ala His Thr Asn	Asp Val Lys Gln	Leu Thr Glu Ala			
	725	730	735			
Val Gln Lys Ile	Thr Thr Glu Ser	Ile Val Ile Trp	Gly Lys Thr Pro			
	740	745	750			

Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr
755 760 765
Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
770 775 780
Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val
785 790 795 800
Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys
805 810 815
Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val
820 825 830
Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr
835 840 845
Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser
850 855 860
Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser
865 870 875 880
Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val
885 890 895
Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln
900 905 910
Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Met Val Gly
915 920 925
Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala
930 935 940
Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly
945 950 955 960
Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr
965 970 975
His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro
980 985 990
Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro
995 1000 1005
Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser
1010 1015 1020
Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu
1025 1030 1035 1040
Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala
1045 1050 1055
Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
1060 1065

<210> 75
<211> 3204
<212> DNA
<213> HIV

<400> 75
atgggtgggtt ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct 60
gtagatctta gccacttttt aaagaanaag gggggacttg aaggctaat teacteccaa 120
agaagacaag atatccttga ttgttggtat taacacacac aaggctactt ccttgattgg 180
cagaactaca caccagggtc aggggtcaga tateacttga cctttggatg gtgtacaaag 240
ctaglacagg ttgagccaga taaggtagaa gaggcacata aaggagagaa caccagcttg 300
ttacaccttg tgagccttga ttggatggat gaccgggaga gagaagtgtt agagtggagg 360
tttgacagcc gcttagcatt tcttaccgtg gcccgagagc tgcatacggg gtaettcaag 420
aactgcatgg gccccatcag tcccatcgag acctgcccg tgaactttaa acctcgggatg 480
gacggcccca aggtcaagca gtggccactc acctgaggga agataagge cctgggtggag 540
atctgcaccg agatggagaa agagggcag atcagaaga tgggacctga gaacctaac 600
aacacccccg tgtttgcat caagaagaag gacagacaa agtggcgcaa gtgtgtgat 660
ttccgggagc tgaataagcg gaccaggat ttctgggagg tccagctggg cctcccccat 720
ccggccggcc tgaagaagaa gaagagcgtg acctgtcttg acctggcgga cgttacttc 780
agcgtccctc tggacagga ctttgaagaa taacctgct ttacctccc atctataac 840
aacgagaccc ctggcatcag atatacgtac aactgcttcc cccagggttg gaagggtct 900
cccgcattt tccagagctc catgaacaa atcctggagc cgtttcggaa gcaaacccc 960
galatcttca tctaccagta catggacgac ctgtacttgg gctctgactt ggaatccgg 1020
cagcatcgca cgaagattga ggagctgagg cagcatctgc ttgatatggg cctgacct 1080
ccggacaaga agcatcagaa ggagccgcaa ttctgaaga ttgggtacga gctccatccc 1140
gacaagtggc cgttgcagcc tatgtctctc cccgagaagg acagctggac cgtgaacgac 1200
atccagaagc tgggtgggca gtctacttgg gctagccaga tctatccgg galcaagggt 1260
cgccagctct gcaagctgtt ggcggcacc aaggccctga cggaggtgat tccccatcg 1320
gagggaagcg agctcagctt ggtcagaaac cgggagatcc tgaaggagcc cgtgcacggc 1380
gtgtactatg accctccaa ggacctgctc gcggaatcc agaagcaggg ccagggggcag 1440
tggacatacc agatttacc ggagcttctc aagaacctca agaccggcaa gtacgcccgc 1500
atgaggggcg cccacaccaa cgaatgcaag cagctgaccc aggccttcaa gaagatcacg 1560
accgagtcga tctgtacttg ggggaagaca cccnagttca agctgacctat ccagaaggag 1620

```

acctgggaga cgtgggtggac cgaatatattgg caggccacct ggattcccga gtgggagttc 1680
gtgaatacac ctctcttggg gaagctgtgg taccagctcg agaaggagcc catcgtgggc 1740
gcggagacat tctacgtgga cggcgcgggc aaccgcgaaa caaagctcgg gaaggccggg 1800
tacgtcacca accggggccg ccagaaggtc gtcacctga ccgacaccac caaccagaag 1860
acggagctgc aggccatcta tctcgtctc caggactccg ccctggaggt gaacatcgtg 1920
acggacagcc agtacgcgt gggcattatt caggcccagc cggaccagtc cgagagcgaa 1980
ctgggtgaac agattatcga gcagctgatc aagaaagaga aggtctacct cgcctgggtc 2040
ccggcccata agggcatlgg ccgcaacgag caggctcgaca agctggtagg tgcggggatt 2100
agaaagggtc tgaatgggtc ccgagcttcg gtactgtctg gtggagagct ggacagatgg 2160
gagaaaaatta ggctgcgccc gggaggcaaa aagaaatata agctcaagca tatcgtgtgg 2220
gcctcgaggg agcttgaacg gtttgccgtg aaccagggcc tgctggaaac atctgagggg 2280
tgtcgccaga tcttggggca attgcagcca tccctccaga ccgggagtga agagctgagg 2340
tcttgtata acacagtggc tacctcttac tgcgtacacc agaggatcga gattaaaggat 2400
accaaggagg ccttggacaa aattgaggag gagcaaaaca agagcaagaa gaaggcccag 2460
caggcagctg ctgacactgg gcatagcaac caggatcac agaactatcc taitgtccaa 2520
aacattcagg gccagatggt tcatcaggcc atcagcccc ggacgctcaa tgcctgggtg 2580
aagggtgtcg aagagaaggc cttttctcct gaggttatcc ccatgttctc cgttttgagt 2640
gagggggcca ctctcagga cctcaataca atgcttaata ccgtgggagg ccatcaggcc 2700
gccatgcaaa tgttgaagga gactatcaac gaggaggcag ccgagtgga cagagtgcgt 2760
ccgtccacg ctggcccaat cgcgcccga cagatgcggg agcctcgcg ccttgacatt 2820
gccggcacca ctcttacct gcaagagcaa atcggaatga tgaccaacaa tcttccatc 2880
ccagtggag aaatctataa acggtggatc atcctgggcc tgaacaagat cgtgcgcatg 2940
tactctccga catccatcct tgacattaga caggaccca aagagcctt tagggattac 3000
gtcgaccggt ttataagac cctgcgagca gaggaggcct ctgaggaggt caaaaactgg 3060
atgacggaga cactctggt acagaacgct aaccccgact gcaaaacaat cttgaaggca 3120
ctaggcccgg ctgccaccgt ggaagagatg atgaccgct gtcaggaggt aggcggaccc 3180
ggacacaaag ccagagtgtt gtga 3204

```

<210> 76
 <211> 1067
 <212> PRT
 <213> HIV

<400> 76

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

Leu	Glu	Ile	Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His
			340					345						350	
Leu	Leu	Arg	Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu
		355					360						365		
Pro	Pro	Phe	Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr
	370					375					380				
Val	Gln	Pro	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp
385					390					395					400
Ile	Gln	Lys	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro
			405					410						415	
Gly	Ile	Lys	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala
			420					425					430		
Leu	Thr	Glu	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala
		435				440						445			
Glu	Asn	Arg	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp
	450					455					460				
Pro	Ser	Lys	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln
465				470						475					480
Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly
			485						490					495	
Lys	Tyr	Ala	Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu
			500					505					510		
Thr	Glu	Ala	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly
			515				520					525			
Lys	Thr	Pro	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr
	530					535					540				
Trp	Trp	Thr	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe
545					550					555					560
Val	Asn	Thr	Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu
					565				570					575	
Pro	Ile	Val	Gly	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg
			580					585					590		
Glu	Thr	Lys	Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asn	Arg	Gly	Arg	Gln
		595				600						605			
Lys	Val	Val	Thr	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln
	610					615					620				
Ala	Ile	Tyr	Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val
625					630					635					640
Thr	Asp	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Gln
				645					650					655	
Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys
			660					665					670		
Glu	Lys	Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly
		675				680					685				
Asn	Glu	Gln	Val	Asp	Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu
	690					695					700				
Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Trp
705					710				715						720
Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys	Leu	Lys
			725					730					735		
His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro
			740					745					750		
Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu
			755				760					765			
Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn
	770					775					780				
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp
785					790				795						800
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys
				805					810					815	
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val
			820					825					830		
Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His
		835					840					845			
Gln	Ala	Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu
	850					855					860				
Glu	Lys	Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser
865					870				875						880
Glu	Gly	Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly
			885					890						895	
Gly	His	Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu
			900					905					910		
Ala	Ala	Glu	Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala
		915				920						925			

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 930 935 940
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 945 950 955 960
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 965 970 975
 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 980 985 990
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 995 1000 1005
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 1010 1015 1020
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 1025 1030 1035 1040
 Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 1045 1050 1055
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu
 1060 1065

<210> 77
 <211> 3204
 <212> DNA
 <213> HIV

<400> 77
 atggtgggtt ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct 60
 gtagatctta gccacttttt aaaagaaaag gggggactgg aagggtctaat tcactcccaa 120
 agaagacaag atatccttga tctgtggatc taccacacac aaggctactt ccttgattgg 180
 cagaactaca caccagggcc aggggtcaga tatccactga cctttggatg gtgctacaag 240
 ctagtaccag ttgagccaga taaggtagaa gaggccaata aaggagagaa caccagcttg 300
 ttacacccctg tgagcctgca tgggatggat gaccgggaga gagaagtgtt agagtggagg 360
 tttagacagcc gccttagcatt tcatcacgtg gcccagagagc tgcattccgga gtacttcaag 420
 aactgcatgg gtccccgagc ttccgttactg tctgggtggag agctggacag atggggagaaa 480
 attaggctgc gcccgggagg caaaaagaaa tacaagctca agcatatcgt gtgggccttcg 540
 agggagccttg aacgggttgc cgtgaaccga ggccctgtctgg aaacatctga gggatgtcgc 600
 cagatccttg ggcaattgca gccatccctc cagaccggga gtgaagagct gaggtccttg 660
 tataacacag tggctacctt ctactgcgta caccagagga tgcagattaa ggataccaag 720
 gaggccttgg acaaaatlga ggaggagcaa aacaagagca agaagaaggc ccagcaggca 780
 gctgttgaca ctgggcatag caaccaggta tcacagaact atcctattgt ccaaaacatt 840
 caggggccaga tggttcatca ggccatcagc ccccgagcgc tcaatgcctg ggtgaagggt 900
 gtcgaagaga aggccttttc tcctgagggt atccccatgt tctccgcttt gagttagggg 960
 gccactcctc aggacctcaa tacaatgctt aataccgttg gcggccatca ggccgcatg 1020
 caaatgttga aggagactat caacgaggag gcagccgagt gggacagagt gcattccgctc 1080
 cacgctggcc caatcgcgcc cggacagatg cgggagcctc gcggtcttga cattgccggc 1140
 accacctcta cactgcaaga gcaaatcgga tggatgacca acaatcctcc catcccagtt 1200
 ggagaaatct ataaacgggt gatcatcctg ggccctgaaca agatcgtgcg catgtactct 1260
 ccgacatcca tcccttgacat tagacaggga cccaaagagc cttttaggga ttacgtcgac 1320
 cggttttata agacccttgc agcagagcag gcctctcagg aggtcaaaaa ctggatgacg 1380
 gagacactcc tggtagagaa cgctaaccac gactgcaaaa caatcttgaa ggcactaggc 1440
 ccggctgcca ccttgggaaga gatgatgacc gcctgtcagg gaggtaggcg acccgagcac 1500
 aaagccagag tgttgatggg ccccatcagt cccatcgaga ccgtgccgtt gaagctgaaa 1560
 cccgggatgg acggccccaa ggtcaagcag tggccactca ccgaggagaa gatcaaggcc 1620
 ctggtagaga tctgcaccga gatggagaaa gagggcaaga tcagcaagat cgggcctgag 1680
 aaccataca acacccccgt gtttgccatc aagaagaagg acagcaccaa gtggcgcaag 1740
 ctggtggatt tccgggagct gaataagcgg acccaggatt tctgggaggt ccagctgggc 1800
 atccccatc cggccggcct gaagaagaag aagagcgtga ccgtgtgga cgtgggcgac 1860
 gcttacttca gcgtccctct ggacgaggac tttagaaagt acaccgcctt taccatccca 1920
 tctatcaaca acgagacccc tggcatcaga tatcagtlaca acgtcctccc ccaggcttg 1980
 aagggtcttc ccgcaatttt ccagagctcc atgaccaaga tccctggagcc gtttcggaag 2040
 cagaaccccg atatcgtcat ctaccagtac atggacgacc tgtacgtggg ctctgacctg 2100
 gaaatcgggc agcatcgac gaagattgag gagctgaggc agcatctgct gagatggggc 2160
 ctgaccactc cggacaagaa gcatcagaag gagccgcat tccctgaagat gggctacgag 2220
 ctccatcccc acaagtggac cgtgcagcct atcgtcctcc ccgagaagga cagctggacc 2280
 gtgaacgaca tccagaagct ggtgggcaag ctcaactggg ctagccagat ctatccggg 2340
 atcaagggtgc gccagctctg caagctgtct gcgggcacca aggcctgac cgagggtatt 2400
 cccctcacgg aggaagccga gctcgagctg gctgagaacc gggagatcct gaaggagccc 2460
 gtgcacggcg tgaatctatg cccctccaag gacctgatcg ccgaaatcca gaagcagggc 2520
 caggggcagt ggacatacca gatttaccag gagcctttca agaacctcaa gaccggcaag 2580
 tacgcccgc tggggggcgc ccacaccaac gatgtcaagc agctgaccga ggccgtccag 2640
 aagatcacga ccgagtcctt cgtgatcttg gggaagacac ccaagttcaa gctgcctatc 2700
 cagaaggaga cctgggagac ttgggtggacc gaatatlggc aggccacctg gattcccgag 2760
 tgggagtctg tgaatacacc tccctctggt aagctgtgtt accagctcga gaaggagccc 2820
 atcgtgggcg cggagacatt ctactggac ggcgcgccca accgcgaac aaagctcggg 2880
 aaggccgggt acgtcaccaa ccggggccgc cagaaggctg tcacctgac cgacaccacc 2940

aaccagaaga cggagcigca ggccaictat ctgcgtctcc aggactccgg cctggagggtg 3000
 aacatcgiga cggacagcca giacgcgctg ggcattattc agggccagcc ggaccagttc 3060
 gagagcgaac tggigaacca gattatcgag cagctgatca agaaagagaa ggtctacctc 3120
 gcctgggtcc cggcccalaa gggcatlggc ggcaacgagc aggtcgacaa gctggtgagt 3180
 gcggggatta gaaagggtgt gtaa 3204

<210> 78
 <211> 1067
 <212> PRT
 <213> HIV

<400> 78

Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 1 5 10 15
 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 20 25 30
 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 35 40 45
 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 50 55 60
 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 65 70 75 80
 Leu Val Pro Val Glu Pro Asp Lys Val Glu Ala Asn Lys Gly Glu
 85 90 95
 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 100 105 110
 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
 115 120 125
 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Met Gly
 130 135 140
 Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp Glu Lys
 145 150 155 160
 Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile
 165 170 175
 Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu
 180 185 190
 Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro
 195 200 205
 Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val
 210 215 220
 Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp Thr Lys
 225 230 235 240
 Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Lys
 245 250 255
 Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val Ser Gln
 260 265 270
 Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala
 275 280 285
 Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys
 290 295 300
 Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly
 305 310 315 320
 Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His
 325 330 335
 Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala
 340 345 350
 Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly
 355 360 365
 Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr
 370 375 380
 Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val
 385 390 395 400
 Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val
 405 410 415
 Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys
 420 425 430
 Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala
 435 440 445
 Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu
 450 455 460
 Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly
 465 470 475 480
 Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly
 485 490 495
 Gly Pro Gly His Lys Ala Arg Val Leu Met Gly Pro Ile Ser Pro Ile

500				505				510			
Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val											
515				520				525			
Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile											
530				535				540			
Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu											
545				550				555			
Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr											
565				570				575			
Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln											
580				585				590			
Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys											
595				600				605			
Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser											
610				615				620			
Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro											
625				630				635			
Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu											
645				650				655			
Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr											
660				665				670			
Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr											
675				680				685			
Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln											
690				695				700			
His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly											
705				710				715			
Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp											
725				730				735			
Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val											
740				745				750			
Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val											
755				760				765			
Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg											
770				775				780			
Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile											
785				790				795			
Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile											
805				810				815			
Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu											
820				825				830			
Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile											
835				840				845			
Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met											
850				855				860			
Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln											
865				870				875			
Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe											
885				890				895			
Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr											
900				905				910			
Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro											
915				920				925			
Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala											
930				935				940			
Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly											
945				950				955			
Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu											
965				970				975			
Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala											
980				985				990			
Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr											
995				1000				1005			
Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu											
1010				1015				1020			
Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu											
1025				1030				1035			
Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp											
1045				1050				1055			
Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu											
1060				1065							

<210> 79
<211> 3204

<212> DNA
<213> HIV

<400> 79

```

atgggcccac tcagtcacac cgagaccgtg ccggtgaagc tgaaccaccg gatggacggc 60
cccaagggtc agcagtggtc actaccgagc gagaagatca aggccttggt ggagatctgc 120
accgagatgg agaaagaggg caagatcagc aagatcgggc ctgagaaccc atacaacacc 180
cccgtgttgc ccatcaagaa gaaggacagc accaagtggc gcaagctggt ggatttccgg 240
gagctgaata agcggaccca ggatttctgg gaggtccagc tgggcatccc ccatccggcc 300
ggcctgaaga agaagaagag cgtgaccgtg ctggacgttg gcgacgcita cticagcgic 360
cccttgagcg aggaclttag aaagtacacc gccittacca tcccatctat caacaacgag 420
acccctggca tcagatatca glacaacgtc ctccccagg gctggaaggc ctccccgcc 480
attttccaga gctccatgac caagatcctg gagccgtttc ggaaagcagaa ccccgataic 540
gtcatctacc agtacctgga cgacctgac gtgggtcttg acctggaaat cgggcagcat 600
cgacacgaaga ttgaggagct gaggcagcat ctgctgagat ggggcttgac cactccggac 660
aagaagcatc agaaggagcc gccattcctg aagatgggct acgagctcca tcccgacaag 720
tggaccgtgc agcctatcgt cctccccgag aaggacagct ggaccgtgaa cgacatccag 780
aagctgggtg gcaagctcaa ctgggttagc cagatctatc ccgggatcaa ggtgcgccag 840
ctctgcaagc tgcctgcggc caacaaggcc ctgaccgagg tgattcccci caccgaggaa 900
ggcgagctcg agctggctga gaaccgggag atcttgaagg agcccggtgc cggcgtgtac 960
tatgacccct ccaaggacct gatcgccgaa atccagaagc agggccaggc gcagtgagca 1020
taccagattt accaggagcc ttcaagaac ctcaagaccg gcaagtagcc ccgatgagg 1080
ggcgccacca ccaacgatgt caagcagctg accgagggcg tccagaagat cagcaccgag 1140
tccatcttga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgttgt ggaccgaata ttggcaggcc acctggaltc ccgagtgagg gttcgtgaat 1260
acacctcttc tggtagagct gtgttaccag ctcgagaagg agcccatcgt gggcgccggg 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tggggaaggc cgggtacgtc 1380
accaaccggg gccgcccaga ggctgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggccttg aggtgaacat cgtacgggac 1500
agccagtagc agctggggcal taltcaggcc cagccggacc agtccgagag ggaactgggt 1560
aaccagatta tccagcagct gatcaagaaa gagaaggctc acctcgcctg ggctccggcc 1620
cataagggca ttggcggcaa cgaacaggtc gacaagcttg tgatgcggg gattagaaag 1680
gtctgtatgg gtgcccgagc ttccgtactg tctgggtggg agctggacag atgggagaaa 1740
attaggctgc gcccgggagg caaaaagaaa tacaagctca agcatatcgt gtggcctctg 1800
agggagcttg aacgglttgc cgtgaaccca ggctgtcttg aaacatctga gggatgtctg 1860
cagatcttgg ggcaatttgc gccatccctc cagaccggga gtgaagagct gaggctcttg 1920
tataacacag tggctacctt ctactgctga caccagagga tccagattaa ggataccaag 1980
gagggccttg acaaaattga ggaggagcaa aacaagagca agaagaagcc ccagcagcaa 2040
gtctgtgaca ctgggcctag caaccaggta tccagaaact atctatctgt ccaaaacatt 2100
cagggccaga tgggttcatc ggccatcagc ccccgagcgc tcaatgcctg ggtgaaggtt 2160
gtcgaagaga aggccttttc tcttgaggtt atccccatgt tctcgtcttt gattgagggg 2220
gccactcttc aggaactcaa tacaatgctt aataccgttg gggccatca ggcgcctatg 2280
caaatgttga aggagactat caacgaggag gcagccgagt gggacagagt gcatcccttc 2340
cacgttgccc caatcgccgc cggacagatg cgggagcctc gggcctctga catctccgce 2400
accacctctc cactgcaaga gcaaatcgga tggatgacca acaatctctc cctccagttt 2460
ggagaaatct ataaacgggt gataatcttg ggcctgaaca agatcgttgc catgtactct 2520
ccgacatcca tcttgcact tagacaggga cccaaagagc ctltttagga ttactctgac 2580
cgglttllata agacccttgc agcagagcag gctctctagg aggtcaaaaa ctggatgacg 2640
gagacactcc tggtagagaa cgttaacccc gactgcaaaa caatctttaa ggcactagge 2700
ccggttgcca ccttggagaa gatgatgacc gctgtctagg gattaggggg acccggacac 2760
aaagccagag tgttgaatgt gggttttcca gtacaccttc aggtaccttt aagaaccaatg 2820
acttacaagg cagctgtaga tcttagccac tttttaaag aaaaaggggg acttgaaggg 2880
ctaatcaact cccaaagaga caaagatate cttagatctg ggaatctaca caacaaagge 2940
tacttccctg attggcagaa ctacacacca ggccagggg tccagatctc actgaacttt 3000
ggatgggtgt acaagctagt accagttagc ccagataagg tagaagagge caataaagga 3060
gagaacacca gcttgttaca cctgttagc ctgcatggga tggatgaccc ggagagagaa 3120
gtgttagagt ggaggtttag cagcgcctta gcaatttate acctggcccc agagctgcat 3180
ccggagtagt tcaagaactg ctga

```

<210> 80
<211> 1067
<212> PRT
<213> HIV

<400> 80

```

Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser 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 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 Ser Val Thr Val Leu Asp
 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 Asp Asp 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 Leu 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 Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
 275 280 285
 Lys Ala Leu Thr Glu Val Ile 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 Thr 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 Thr 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 Asp Gly Ala Ala
 435 440 445
 Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
 450 455 460
 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
 465 470 475 480
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 485 490 495
 Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 500 505 510
 Asp Gln Ser Glu Ser Glu Leu Val Asn 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 Ala Gly Ile Arg Lys
 545 550 555 560
 Val Leu Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp
 565 570 575
 Arg Trp Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys
 580 585 590
 Leu Lys His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val
 595 600 605
 Asn Pro Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly
 610 615 620
 Gln Leu Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu
 625 630 635 640
 Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile
 645 650 655
 Lys Asp Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys
 660 665 670
 Ser Lys Lys Lys Ala Gln Gln Ala Ala Asp Thr Gly His Ser Asn

675	680	685
Gln Val Ser Gln Asn Tyr	Pro Ile Val Gln Asn	Ile Gln Gly Gln Met
690	695	700
Val His Gln Ala Ile Ser	Pro Arg Thr Leu Asn	Ala Trp Val Lys Val
705	710	715
Val Glu Glu Lys Ala Phe	Ser Pro Glu Val Ile	Pro Met Phe Ser Ala
725	730	735
Leu Ser Glu Gly Ala Thr	Pro Gln Asp Leu Asn	Thr Met Leu Asn Thr
740	745	750
Val Gly Gly His Gln Ala	Ala Met Gln Met Leu	Lys Glu Thr Ile Asn
755	760	765
Glu Glu Ala Ala Glu Trp	Asp Arg Val His Pro	Val His Ala Gly Pro
770	775	780
Ile Ala Pro Gly Gln Met	Arg Glu Pro Arg Gly	Ser Asp Ile Ala Gly
785	790	795
Thr Thr Ser Thr Leu Gln	Glu Gln Ile Gly Trp	Met Thr Asn Asn Pro
805	810	815
Pro Ile Pro Val Gly Glu	Ile Tyr Lys Arg Trp	Ile Ile Leu Gly Leu
820	825	830
Asn Lys Ile Val Arg Met	Tyr Ser Pro Thr Ser	Ile Leu Asp Ile Arg
835	840	845
Gln Gly Pro Lys Glu Pro	Phe Arg Asp Tyr Val	Asp Arg Phe Tyr Lys
850	855	860
Thr Leu Arg Ala Glu Gln	Ala Ser Gln Glu Val	Lys Asn Trp Met Thr
865	870	875
Glu Thr Leu Leu Val Gln	Asn Ala Asn Pro Asp	Cys Lys Thr Ile Leu
885	890	895
Lys Ala Leu Gly Pro Ala	Ala Thr Leu Glu Glu	Met Met Thr Ala Cys
900	905	910
Gln Gly Val Gly Gly Pro	Gly His Lys Ala Arg	Val Leu Met Val Gly
915	920	925
Phe Pro Val Thr Pro Gln	Val Pro Leu Arg Pro	Met Thr Tyr Lys Ala
930	935	940
Ala Val Asp Leu Ser His	Phe Leu Lys Glu Lys	Gly Gly Leu Glu Gly
945	950	955
Leu Ile His Ser Gln Arg	Arg Gln Asp Ile Leu	Asp Leu Trp Ile Tyr
965	970	975
His Thr Gln Gly Tyr Phe	Pro Asp Trp Gln Asn	Tyr Thr Pro Gly Pro
980	985	990
Gly Val Arg Tyr Pro Leu	Thr Phe Gly Trp Cys	Tyr Lys Leu Val Pro
995	1000	1005
Val Glu Pro Asp Lys Val	Glu Glu Ala Asn Lys	Gly Glu Asn Thr Ser
1010	1015	1020
Leu Leu His Pro Val Ser	Leu His Gly Met Asp	Asp Pro Glu Arg Glu
1025	1030	1035
Val Leu Glu Trp Arg Phe	Asp Ser Arg Leu Ala	Phe His His Val Ala
1045	1050	1055
Arg Glu Leu His Pro Glu	Tyr Phe Lys Asn Cys	
1060	1065	

<210> 81
 <211> 3204
 <212> DNA
 <213> HIV

<400> 81
 atgggcccac teagteccat egagaccgtg ceggtagaage tgaatcccgk gatggacggc 60
 cccaaggtea agcagtaggc actcaccgag gagaagatac aggccttggf ggagatctgc 120
 accgagatgg agaagagagg caagatcage aagatcgggc cggagaaacc atacaacacc 180
 cccgltgttg ccatcaagaa gaaggacage accaagtggc gcaagcttggf ggatttccgg 240
 gagctgaata ageggaccca ggatttctgg gaggteccage tgggcatecc ccatccggcc 300
 ggccitgaaga agaagaagag cgtgaccgtg ctggacgttg ggcacgttta ctteagcgtc 360
 cctctggacg aggcatttag aaagtaaccc gcctttacca tcccatctat caacaacgag 420
 acccctggca teagatatac gtacaacgtc ctccccagg gcttgaaggk ctctcccgcc 480
 attttccaga gctccatgac caagatccct gagecgttite ggaagcagaa ccccgatac 540
 gtcatctacc agtacatgga cgaactgtac gtgggtcttg acctggaat cgggcagcat 600
 cgcacgaaga ttgaggagct gaggcagcat ctgttgagat ggggcctgac cactccggac 660
 aagaagcacc agaaggagcc gccattccct aagatgggct acgagctcca tcccgacaag 720
 tggaccgtgc agcctatcgt ctctcccgag aaggacagct ggaacgtgaa cgaatccag 780
 aagctggtag gcaagctcaa ctgggtcage cagatctata cggggtacca ggttgcggag 840
 ctctgcaagg tgcctgcggc caccagggcc ctgaccgagg tgaattccct caccgaggaa 900
 gccagctctg agcttggtga gaaccgggag atcttgaagg agcccggtga cggcggtgac 960
 tatgaccctt ccaaggacct gatcgccgaa atccagaagg agggccaggk gcagtggaac 1020
 taccagattt accaggagcc tttaagaac cttaagacc gcaagtaacc cccatgagg 1080

ggcgcccaca ccaacgatgt caagcagctg accgaggccg tccagaagat caccgaccgag 1140
tccatcgtga tctgggggaa gacacccaag ttcaagctgc ctatccagaa ggagacctgg 1200
gagacgttgt ggaccgaata ttggcaggcc acctggattc ccgagtggga gticgtgaat 1260
acacctcctc tggatgaagct gtgttaccag ctcgagaagg agcccatcgt gggcgcggag 1320
acattctacg tggacggcgc ggccaaccgc gaaacaaagc tcgggaaggc cgggtacgtc 1380
accaaccggg gccgccagaa ggtcgtcacc ctgaccgaca ccaccaacca gaagacggag 1440
ctgcaggcca tctatctcgc tctccaggac tccggcctgg aggtgaacat cgtgacggac 1500
agccagtacg cgttgggcat tattcaggcc cagccggacc agtccgagag cgaactggtg 1560
aaccagatta tcgagcagct gatcaagaaa gagaaggctt acctcgctg ggtcccggcc 1620
cataagggca ttggcggcaa cgagcaggtc gacaagctgg tgagtgcggg gattagaaag 1680
gtgctgatgg tgggttttcc agtcacacct caggtacctt taagaccaat gacttacaag 1740
tcgactgtag atcttagcca ctttttaaaa gaaaaggggg gactggaagg gctaattcac 1800
tcccaaagaa gacaagatat ccttgatctg tggatctacc acacacaagg ctacttccct 1860
gattggcaga actacacacc agggccaggg gtcagatata cactgacctt tggatggtg 1920
lacaagctag taccagtga gccagataag gtagaaggag ccaataaagg agagaacacc 1980
agcttgttac acctgtgtg cctgcatggg atggatgacc cggagagaga agtgttagag 2040
tggaggtttg acagccgcct agcatttcat cacgtggccc gagagctgca tccggagtac 2100
ttcaagaact gctgaatggg tgcggagct tgggtactgt ctggggaga gctggacaga 2160
tgggagaaaa ttaggctgcg cccgggaggc aaaaagaaat acaagctcaa gcatactgt 2220
tgggctcga gggagctiga acggtttgcc gtgaaccag gcctgctgga aacatctgag 2280
ggatgtcgcc agatcctggg gcaattgcag ccattccctc agaccgggag tgaagagctg 2340
aggctcctgt ataacacagt ggctaccctc tactgcgtac accagaggat cgagattaag 2400
gataccaagg aggccttggg caaaattgag gaggagcaaa acaagagcaa gaagaaggcc 2460
cagcaggcag ctgctgacac tgggcatagc aaccaggtat cacagaacta tcctattgtc 2520
caaaaacatt agggccagat ggttcatcag gccatcagcc cccggacgct caatgcctgg 2580
gtgaaggttg tcgaagagaa ggccctttct cctgaggtta tcccatgtt ctccgctttg 2640
agtgaggggg ccactcctca ggacctcaat acaatgctta ataccgtggg cggccatcag 2700
gccgccatgc aaatgtlga ggagactatc aacgaggagg cagccgagtg ggacagagtg 2760
catcccgtcc acgttggccc aatcgcgccc ggacagatgc gggagcctcg cggctctgac 2820
attccgggca ccacctctac actgcaagag caaatcggat ggatgaccaa caatcctccc 2880
atccagttg gagaaatcta taaacggtgg atcatcctgg gcctgaacaa gatcgtgcgc 2940
atgtactctc cgacatccat ccttgacatt agacagggac ccaaagagcc ttttagggat 3000
tacgtcgacc ggttttataa gacctgcga gcagagcagg cctctcagga ggtcaaaaac 3060
tggatgacgg agacactcct ggtacagaac gctaaccctg actgcaaac aatcttgaag 3120
gcactaggcc cggctgccac cctggaagag atgatgaccg cctgtcaggg agtaggcgga 3180
cccggacaca aagccagagt gttg 3204

<210> 82
<211> 1067
<212> PRT
<213> HIV

<400> 82
Met Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro
1 5 10 15
Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr 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 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 Asp
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 Asp Asp 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 Leu 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 Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
 275 280 285
 Lys Ala Leu Thr Glu Val Ile 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 Thr 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 Thr 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 Asp Gly Ala Ala
 435 440 445
 Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
 450 455 460
 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
 465 470 475 480
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 485 490 495
 Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 500 505 510
 Asp Gln Ser Glu Ser Glu Leu Val Asn 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 Ala Gly Ile Arg Lys
 545 550 555 560
 Val Leu Met Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro
 565 570 575
 Met Thr Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys
 580 585 590
 Gly Gly Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu
 595 600 605
 Asp Leu Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn
 610 615 620
 Tyr Thr Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys
 625 630 635 640
 Tyr Lys Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys
 645 650 655
 Gly Glu Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp
 660 665 670
 Asp Pro Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala
 675 680 685
 Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
 690 695 700
 Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 705 710 715 720
 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 725 730 735
 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 740 745 750
 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 755 760 765
 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 770 775 780
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 785 790 795 800
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Gln Asn Lys Ser Lys
 805 810 815
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 820 825 830
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His
 835 840 845
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

850	855	860
Glu Lys Ala Phe Ser Pro	Glu Val Ile Pro Met	Phe Ser Ala Leu Ser
865	870	875
Glu Gly Ala Thr Pro	Gln Asp Leu Asn Thr	Met Leu Asn Thr Val Gly
885	890	895
Gly His Gln Ala Ala Met	Gln Met Leu Lys Glu Thr	Ile Asn Glu Glu
900	905	910
Ala Ala Glu Trp Asp Arg Val	His Pro Val His Ala Gly Pro	Ile Ala
915	920	925
Pro Gly Gln Met Arg Glu Pro	Arg Gly Ser Asp Ile Ala Gly Thr Thr	
930	935	940
Ser Thr Leu Gln Glu Gln Ile	Gly Trp Met Thr Asn Asn Pro Pro Ile	
945	950	955
Pro Val Gly Glu Ile Tyr Lys	Arg Trp Ile Ile Leu Gly Leu Asn Lys	
965	970	975
Ile Val Arg Met Tyr Ser Pro	Thr Ser Ile Leu Asp Ile Arg Gln Gly	
980	985	990
Pro Lys Glu Pro Phe Arg Asp	Tyr Val Asp Arg Phe Tyr Lys Thr Leu	
995	1000	1005
Arg Ala Glu Gln Ala Ser Gln	Glu Val Lys Asn Trp Met Thr Glu Thr	
1010	1015	1020
Leu Leu Val Gln Asn Ala Asn	Pro Asp Cys Lys Thr Ile Leu Lys Ala	
1025	1030	1035
Leu Gly Pro Ala Ala Thr Leu	Glu Glu Met Met Thr Ala Cys Gln Gly	
1045	1050	1055
Val Gly Gly Pro Gly His Lys	Ala Arg Val Leu	
1060	1065	

<210> 83
 <211> 3204
 <212> DNA
 <213> HIV

<400> 83

```

atgggtgccc gagcttcggt actgtctggt ggagagctgg acagatggga gaaaattagg 60
ctcgccccgg gaggcaaaaa gaaatacaag ctcaagcata tcgtgtgggc ctcgaggag 120
cttgaacggt ttgccgtgaa cccaggcctg ctggaacat ctgagggatg tcgccagatc 180
ctggggcaat tgcagccatc cctccagacc gggagtgaa agctgaggtc cttgtataac 240
acagtggcta cctctactg cgtacaccag aggatcgaga ttaaggatc caaggaggcc 300
ttggacaaaa ttgaggagga gcaaaacaag agcaagaaga agggccagca ggcagctgct 360
gacactgggc atagcaacca ggtaacacag aactatccta ttgtccaaaa cattcagggc 420
cagatggttc atcaggccat cagccccggg acgtcaatg cctgggtgaa ggttgtcgaa 480
gagaaggcct ttctctctga ggtaatcccc atgttctccg ctttgatga gggggccact 540
cctcaggacc tcaatacaat gcttaatacc gtggcgggcc atcaggccgc catgcaaatg 600
ttgaaggaga ctatcaacga ggaggcagcc gagtgggaca gattgcatcc cgtccacgt 660
ggcccaatcg cgccgggaca gatgcgggag cctcgcggtc ctgacatgc cggcaccacc 720
tctacactgc aagagcaaat cggatggatg accaacaatc ctcccatccc agttggagaa 780
atctataaac ggtggatcat cctgggcttg aacaagatcg tgcgcatgta ctctccgaca 840
tccatccttg acattagaca gggacccaaa gagcctttta gggattacgt cgaccggttt 900
tataagacc tgcgagcaga gcaggcctct caggaggtca aaaactggat gacggagaca 960
ctcctggtac agaacgctaa ccccgactgc aaaacaatct tgaaggcact agggccggct 1020
ggcaccctgg aagagatgat gaccgcctgt caggagtag gcggaccggg acacaaagcc 1080
agagtgttga tgggtgggtt tccagtaca cctcagggtac ctttaagacc aatgacttac 1140
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 1200
cactcccaa gaagacaaga tatccttgat ctgtggatct accacacaca aggtacttc 1260
cctgattggc agaactacac accagggccca ggggtcagat atccactgac ctttggatgg 1320
tgctacaagc tagtaccagt tgagccagat aaggtagaag agggcaataa aggagagaac 1380
accagctgtg tacaccctgt gagcttgcat gggatggatg acccgagag agaaagtgtta 1440
gagtgagggt ttgacagccg cctagcatgt catcacgtgg cccgagagct gcatccggag 1500
tacttcaaga actgcatggg ccccatcagt cccatcgaga ccgtgccgtg gaagctgaaa 1560
cccgggatgg acggccccaa ggtcaagcag tggccactca ccgaggagaa gatcaaggcc 1620
ctgggtggaga tctgcaccga gatggagaaa gagggcaaga tcagcaagat cgggcctgag 1680
aaccataca acacccccgt gtttgccatc aagaagaagg acagcaccaa gtggcgcaag 1740
ctgggtggat ttccggagct gaataagcgg acccaggatt tctgggaggt ccagctgggc 1800
atccccatc cgccgggctt gaagaagaag aagagcgtga ccgtgtgga cgtggcgac 1860
gcttacttca gcgtccctct ggacgaggac tttagaaagt acaccgcctt taccatccca 1920
tctatcaaca acgagacccc tggcatcaga tatcagtaca acgtcctccc ccagggtgg 1980
aagggtcttc ccgcatttt ccagagctcc atgaccaaga tcttggagcc gtttcggaag 2040
cagaaccccc atatctgcat ctaccagtac atggacgacc tgtactggg ctctgacctg 2100
gaaatcgggc agcatcgac gaagatlgag gagctgaggc agcatctgtg gagatggggc 2160
ctgaccactc cggacaagaa gcatcagaag gagccgcat tctgaagat gggctacgag 2220
ctccatcccg acaagtggac cgtgcagcct atcgtcctcc ccgagaagga cagctggacc 2280
gtgaacgaca tccagaagct ggtgggcaag ctcaactggg ctaggcagat ctatcccggg 2340
atcaaggtgc gccagctctg caagctgtct gcgggcacca agggcctgac cgagggtgatt 2400
  
```

```

ccccacagg aggaagccga gctcgagctg gctgagaacc gggagatcct gaaggagccc 2460
gtgcacggcg tgtactatga cccctccaag gacctgatcg ccgaatcca gaagcagggc 2520
caggggcagt ggacalacca gatttaccag gagcctttca agaacctcaa gaccggcaag 2580
tacgcccga tgagggggcg ccacaccaac gatgtcaagc agctgaccga ggccgtccag 2640
aagatcacga ccgagtcctat cgtgatctgg gggaagacac ccaagttcaa gctgcctatc 2700
cagaaggaga cctgggagac gtgggtggacc gaataitggc aggccacctg gattcccgag 2760
tgggagttcg tgaalacacc tctcttggtg aagctgtggt accagctcga gaaggagccc 2820
atcgtggcg cggagacatt ctacgtggac ggcgcggcca accgcgaaac aaagctcggg 2880
aaggccgggt acgtcaccaa ccggggccgc cagaaggctg tcacctgac cgacaccacc 2940
aaccagaaga cggagctgca ggccatctat ctgcctctcc aggactccgg cctggagggtg 3000
aacatcgtga cggacagcca gtacgcgtg ggcatlatic agggccagcc ggaccagtcc 3060
gagagcgaaac tggtagacca gattatcgag cagctgatca agaaagagaa ggcttacctc 3120
gcctgggtcc cggcccaataa ggcatlgtgc ggcaacgagc aggtcgacaa gctggtaggt 3180
cgggggaita gaaaggltgt glaa 3204

```

<210> 84
 <211> 1067
 <212> PRT
 <213> HIV

```

<400> 84
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 Lys 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 Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
100          105          110
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn 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 Thr 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 Thr 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 Glu 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 Met Val Gly Phe Pro
355          360          365
Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala Val
370          375          380
Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile
385          390          395          400
His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr
405          410          415
Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val
420          425          430

```

Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu
 435 440 445
 Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu
 450 455 460
 His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu
 465 470 475 480
 Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg Glu
 485 490 495
 Leu His Pro Glu Tyr Phe Lys Asn Cys Met Gly Pro Ile Ser Pro Ile
 500 505 510
 Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 515 520 525
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 530 535 540
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 545 550 555 560
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 565 570 575
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 580 585 590
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 595 600 605
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 610 615 620
 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 625 630 635 640
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 645 650 655
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
 660 665 670
 Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
 675 680 685
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 690 695 700
 His Arg Thr Lys Ile Glu Leu Arg Gln His Leu Leu Arg Trp Gly
 705 710 715 720
 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 725 730 735
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
 740 745 750
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 755 760 765
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
 770 775 780
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
 785 790 795 800
 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 805 810 815
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 820 825 830
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 835 840 845
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 850 855 860
 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 865 870 875 880
 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 885 890 895
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 900 905 910
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 915 920 925
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 930 935 940
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 945 950 955 960
 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
 965 970 975
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala
 980 985 990
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 995 1000 1005
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 1010 1015 1020
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu

1025 1030 1035 1040
Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 1045 1050 1055
Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu
 1060 1065

拾、申請專利範圍

1. 一種核苷酸序列，其可編碼 HIV-1 gag 蛋白質或 gag 抗原決定位及 HIV-1 Nef 蛋白質或 Nef 抗原決定位，其可操作連接至異種啟動基因。
2. 如申請專利範圍第 1 項之核苷酸序列，其中 gag 蛋白質包括 p17。
3. 如申請專利範圍第 2 項之核苷酸序列，其中 gag 蛋白質另包括 p24。
4. 如申請專利範圍第 1 至 3 項中任一項之核苷酸序列，其中 gag 序列經過密碼子最佳化成為類似高度表現人類基因所採用之密碼子，其具相對同義密碼子利用 (RSCU) 值為 0.5。
5. 如申請專利範圍第 1 至 3 項中任一項之核苷酸序列，其中該序列進一步可編碼 RT 蛋白質或 RT 抗原決定位。
6. 如申請專利範圍第 5 項之核苷酸序列，其中該序列之順序為 RT、gag、Nef 或 RT、Nef、gag。
7. 如申請專利範圍第 5 項之核苷酸序列，其中 RT 序列經過密碼子最佳化以類似高度表現人類基因。
8. 一種選自由下列所組成之群的核苷酸序列，
 - Gag(p17,p24)、截短的 Nef
 - Gag(p17,p24)(密碼子最佳化)、Nef(截短的)
 - Gag(p17,p24)、RT、Nef(截短的)
 - Gag(p17,p24)密碼子最佳化、RT、Nef(截短的)

Gag(p17,p24)密碼子最佳化、RT密碼子最佳化、截短之Nef

RT(密碼子最佳化)、Gag (p17,p24)密碼子最佳化、截短之Nef

RT(密碼子最佳化)，截短之Nef、gag p17、p24密碼子最佳化，

其中截短之Nef包括移除N端胺基酸，且其中該核苷酸序列可操作連接至異種啟動基因。

9. 如申請專利範圍第8項之核苷酸序列，其中截短之Nef包括移除65個N端胺基酸。
10. 如申請專利範圍第1至3及8至9項中任一項之核苷酸序列，其中異種啟動基因係來自HCMV IE基因之啟動基因。
11. 如申請專利範圍第10項之核苷酸序列，其中啟動基因之5'端包括表現序列1。
12. 如申請專利範圍第5項之核苷酸序列，其中RT係編碼一突變，以實質上去活化任何反轉錄酶活性。
13. 如申請專利範圍第12項之核苷酸序列，其中RT係經取代色胺酸229為離胺酸而突變之。
14. 一種包括如申請專利範圍第1至3及8至9項之核苷酸序列的載體。
15. 如申請專利範圍第14項之載體，其係病毒載體。
16. 如申請專利範圍第15項之載體，其係複製缺陷的腺病

毒。

17. 如申請專利範圍第14項之載體，其係雙股DNA質體。
18. 一種由如申請專利範圍第1至3及8至9項中任一項之核苷酸序列所編碼之蛋白質。
19. 一種醫藥組合物，其包括如申請專利範圍第1至3及8至9項之核苷酸序列及藥學上可接受賦形劑、稀釋劑、載劑或佐劑。
20. 一種醫藥組合物，其包括如申請專利範圍第14項之載體及藥學上可接受賦形劑、稀釋劑、載劑或佐劑。
21. 如申請專利範圍第19項之醫藥組合物，其適於肌肉注射或皮內遞送。
22. 如申請專利範圍第20項之醫藥組合物，其適於肌肉注射或皮內遞送。
23. 如申請專利範圍第19項之醫藥組合物，其中之載劑為金製有孔小珠。
24. 如申請專利範圍第20項之醫藥組合物，其中之載劑為金製有孔小珠。
25. 一種皮內遞送裝置，其包括如申請專利範圍第19項之醫藥組合物。
26. 一種皮內遞送裝置，其包括如申請專利範圍第20項之醫藥組合物。
27. 如申請專利範圍第19項之醫藥組合物，其可用於治療罹患或易感受疾病之病患。

28. 如申請專利範圍第20項之醫藥組合物，其可用於治療罹患或易感受疾病之病患。
29. 如申請專利範圍第1至3及8至9項中任一項之核苷酸序列，其可運用於醫學。
30. 如申請專利範圍第14項之載體，其可運用於醫學。
31. 如申請專利範圍第19項之醫藥組合物，其可運用於醫學。
32. 如申請專利範圍第20項之醫藥組合物，其可運用於醫學。
33. 一種如申請專利範圍第1至3及8至9項中任一項之核苷酸序列之用途，其係用於製備可治療疾病之藥劑。
34. 一種如申請專利範圍第14項之載體，其係用於製備可治療疾病之藥劑。
35. 一種製備如申請專利範圍第1至3及8至9項中任一項之核苷酸的方法，其包括經操作將可編碼HIV-1 gag蛋白質及第二種HIV-1 Nef蛋白質之核苷酸序列連接至異種啟動基因序列。