



(21) 申請案號：111116003 (22) 申請日：中華民國 111 (2022) 年 04 月 27 日

(51) Int. Cl. : *A61K48/00 (2006.01)* *C12N15/63 (2006.01)*  
*C12N15/64 (2006.01)* *C12N15/86 (2006.01)*  
*C12N15/861 (2006.01)* *A61P1/16 (2006.01)*

(30) 優先權：2021/04/27 美國 63/180,603  
2021/09/09 美國 63/242,474  
2021/09/14 美國 63/244,205  
2022/01/21 美國 63/301,933  
2022/04/15 美國 63/331,385

(71) 申請人：賓州大學委員會 (美國) THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA (US)  
美國

(72) 發明人：威爾森 詹姆士 M WILSON, JAMES M. (US)；王 利利 WANG, LILI (US)；崔堤寇瓦 安娜 P TRETIAKOVA, ANNA P. (US)

(74) 代理人：王彥評；黃政誠

申請實體審查：無 申請專利範圍項數：52 項 圖式數：19 共 350 頁

## (54) 名稱

用於治療遺傳疾病的體內核酸酶媒介的基因靶向之組成物及方法

## (57) 摘要

提供一種用於治療遺傳疾病的雙重載體系統。該系統包括：(a) 包含表現匣之基因編輯載體，該表現匣包含編碼核酸酶之核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中之表現的調控序列；及 (b) 供體載體，其包含編碼用於從 PCSK9 基因座表現之外源產物的核酸序列，其中插入的核酸序列並不編碼 PCSK9，其中該系統進一步包含指導該核酸酶特異性地靶向天然 PCSK9 基因座的序列；且其中在以該雙重載體系統投藥後，該標靶細胞中的該天然 PCSK9 可選擇地消融或減少。

A dual vector system for treating a genetic disorder is provided. The system includes (a) a gene editing vector comprising an expression cassette comprising a nucleic acid sequence encoding a nuclease and regulatory sequences that direct expression of the nuclease in a target cell comprising a PCSK9 gene; and (b) a donor vector comprising a nucleic acid sequence encoding an exogenous product for expression from the PCSK9 locus, wherein the inserted nucleic acid sequence does not encode PCSK9, wherein the system further comprises sequences that direct the nuclease to specifically targets the native PCSK9 gene locus; and wherein the native PCSK9 in the target cell is optionally ablated or reduced post-dosing with the dual vector system.



## 【發明摘要】

### 【中文發明名稱】

用於治療遺傳疾病的體內核酸酶媒介的基因靶向之組成物及方法

### 【英文發明名稱】

COMPOSITIONS AND METHODS FOR IN VIVO NUCLEASE-MEDIATED GENE TARGETING FOR THE TREATMENT OF GENETIC DISORDERS

### 【中文】

提供一種用於治療遺傳疾病的雙重載體系統。該系統包括：(a)包含表現匣之基因編輯載體，該表現匣包含編碼核酸酶之核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中之表現的調控序列；及(b)供體載體，其包含編碼用於從 PCSK9 基因座表現之外源產物的核酸序列，其中插入的核酸序列並不編碼 PCSK9，其中該系統進一步包含指導該核酸酶特異性地靶向天然 PCSK9 基因座的序列；且其中在以該雙重載體系統投藥後，該標靶細胞中的該天然 PCSK9 可選擇地消融或減少。

**【 英文 】**

A dual vector system for treating a genetic disorder is provided. The system includes (a) a gene editing vector comprising an expression cassette comprising a nucleic acid sequence encoding a nuclease and regulatory sequences that direct expression of the nuclease in a target cell comprising a PCSK9 gene; and (b) a donor vector comprising a nucleic acid sequence encoding an exogenous product for expression from the PCSK9 locus, wherein the inserted nucleic acid sequence does not encode PCSK9, wherein the system further comprises sequences that direct the nuclease to specifically targets the native PCSK9 gene locus; and wherein the native PCSK9 in the target cell is optionally ablated or reduced post-dosing with the dual vector system.

**【 指定代表圖 】**

無。

**【 代表圖之符號簡單說明 】**

無。

**【 特徵化學式 】**

無。

## 【發明說明書】

### 【中文發明名稱】

用於治療遺傳疾病的體內核酸酶媒介的基因靶向之組成物及方法

### 【英文發明名稱】

COMPOSITIONS AND METHODS FOR IN VIVO NUCLEASE-MEDIATED GENE TARGETING FOR THE TREATMENT OF GENETIC DISORDERS

### 【技術領域】

【0001】本案係關於用於治療遺傳疾病的體內核酸酶媒介的基因靶向之組成物及方法。

### 【先前技術】

【0002】位點特異性核酸酶(諸如 CRISPR-Cas9 或巨型核酸酶(meganuclease))在染色體中產生雙股斷裂(DSB)，從而導致 DNA 修復。在存在供體 DNA 的情況下，會發生同源定向修復(homology directed repair, HDR)，並以來自供體基因的新資訊替換染色體中的遺傳資訊。

【0003】同源定向修復(HDR)是使用 DNA 模板藉由同源重組修復 DNA 雙鏈斷裂(DSB)的過程。此模板可在細胞週期的晚期 S 期或 G2 期來自細胞內，此時姐妹染色分體在有絲分裂完成之前可用。此外，外源的修復模板可被遞送至細胞中，最通常是以合成的單鏈 DNA 供體寡核苷酸或供體質體的形式，以在基因體中產生精確的變化。

【0004】安全港位點(Safe harbor sites, SHS)是基因或其它遺傳元件可安全插入和表現的基因體基因座。此等 SHS 對於有效的人類疾病基因治療；用於研究基因結構、功能和調控；及用於細胞標記和追蹤至關重要。

【0005】需要的是用於基因編輯的改進組成物及方法。

【發明內容】

【0006】本文提供用於基因編輯之組成物、方法、系統及套組，該基因編輯允許天然 PCSK9 基因的減弱或消融以及外源轉基因在 PCSK9 基因座中的插入及/或表現。

【0007】在第一態樣中，本文提供用於治療遺傳疾病之系統。該系統包括含有表現匣的基因編輯組分，該表現匣包含編碼靶向 PCSK9 基因之核酸酶的核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中的表現的調控序列。系統進一步包括含有轉基因匣之供體載體，該轉基因匣包含編碼轉基因之核酸序列及指導該轉基因在標靶細胞中的表現的調控序列，供體載體進一步包含對於轉基因匣之同源定向重組(HDR)臂 5'及 3'，其中該轉基因並不是 PCSK9。核酸酶靶向 PCSK9 基因。在一些具體實施例中，核酸酶靶向 PCSK9 外顯子 7。在一些具體實施例中，巨型核酸酶為 ARCUS 巨型核酸酶。

【0008】在一些具體實施例中，基因編輯組分包含編碼 Cas9 之序列。在某些具體實施例中，基因編輯載體進一步包含編碼含至少 20 個核苷酸種子區之 sgRNA

的序列，其中該 sgRNA 特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (protospacer-adjacent motif, PAM) 的 5' 處。

【0009】在其它具體實施例中，供體載體進一步包含編碼包含至少 20 個核苷酸種子區之 sgRNA 的序列，其中該 sgRNA 特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處。

【0010】在另一態樣中，本文提供一種治療遺傳疾病之系統。該系統包括基因編輯組分，該基因編輯組分包含編碼靶向 PCSK9 基因之核酸酶的核酸序列。該系統進一步包括包含轉基因匣之供體載體，該轉基因匣包含編碼轉基因之核酸序列及指導該轉基因在標靶細胞中之表現的調控序列，該供體載體進一步包含對於轉基因匣的同源定向重組 (HDR) 臂 5' 及 3'，其中該轉基因並非 PCSK9。該核酸酶靶向 PCSK9 基因。在某些具體實施例中，基因編輯組分係以脂質奈米顆粒形式提供。

【0011】在一些具體實施例中，基因編輯組分包含編碼 Cas9 之序列。在某些具體實施例中，基因編輯載體進一步包含編碼包含至少 20 個核苷酸種子區之 sgRNA 的序列，其中該 sgRNA 特異性結合在 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處。

【0012】在其它具體實施例中，供體載體進一步的組分包含編碼包含至少 20 個核苷酸種子區之 sgRNA 的序列，其中該 sgRNA 特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處。

【0013】在某些具體實施例中，轉基因與肝臟代謝性病症有關。在某些具體實施例中，轉基因為 OTC、PKU、CTLN1 或 LDLR。

【0014】在某些具體實施例中，載體為腺相關病毒(AAV)載體，且該載體包含 AAV 5'ITR 及 AAV 3'ITR。

【0015】在另一具體實施例中，治療遺傳疾病的雙重載體系統包括：基因編輯 AAV，該基因編輯 AAV 含有 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5'ITR、編碼巨型核酸酶之序列及 3'ITR，該巨型核酸酶在調控序列的控制下靶向 PCSK9，該調控序列指導該巨型核酸酶在包含 PCSK9 基因的標靶細胞中的表現；及包含 AAV 衣殼及第二載體基因體的供體 AAV 載體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3'HDR 臂、及 3'ITR，其中該轉基因並不編碼 PCSK9。

【0016】在另一具體實施例中，治療遺傳疾病之雙重載體系統包括：基因編輯 AAV，該基因編輯 AAV 包含 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5'ITR、5'核定位訊號(nuclear localization signal，

NLS)、編碼 Cas9 的序列及在包含 PCSK9 基因之標靶細胞中指導該 saCas9 表現的調控序列、3'NLS、及 3'ITR；及包含 AAV 衣殼及第二載體基因體之供體 AAV 載體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3'HDR 臂、U6 啟動子、包含至少 20 個核苷酸之 sgRNA(其特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處)、及 3'ITR，其中該轉基因並不編碼 PCSK9。

【0017】又一具體實施例中，治療遺傳疾病之雙重載體系統包括：基因編輯 AAV 載體，該基因編輯 AAV 載體包含 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5'ITR、U6 啟動子、包含至少 20 個核苷酸之 sgRNA(其特異性結合在 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處)、5'核定位訊號(NLS)、編碼 Cas9 的序列及指導該 Cas9 在包含 PCSK9 基因之標靶細胞中的表現的調控序列、3'NLS、及 3'ITR；及包含 AAV 衣殼及第二載體基因體的供體 AAV 載體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3'HDR 臂、及 3'ITR。

【0018】在一些具體實施例中，基因編輯 AAV 載體與供體 AAV 載體具有相同 AAV 衣殼。在其它具體實施

例中，基因編輯 AAV 載體與供體 AAV 載體具有不同 AAV 衣殼。在一些具體實施例中，AAV 衣殼選自 AAV8、AAV9、rh10、AAV6.2、AAV3B、hu37、rh79 及 rh64。

【0019】在另一態樣中，提供藉由共同投予如本文所述之雙重載體系統來治療人類病症的方法。

【0020】在另一態樣中，提供一種在受試者中治療肝臟代謝性病變之方法，該方法包括對該患有肝臟代謝性病變之受試者共同投予：基因編輯 AAV 載體，該基因編輯 AAV 載體包含編碼核酸酶之序列及指導該核酸酶在標靶細胞中之表現的調控序列，該標靶細胞包含 PCSK9 基因；及供體 AAV 載體，該供體 AAV 載體包含轉基因及指導該轉基因在標靶細胞中之表現的調控序列，供體載體進一步包含對於轉基因匣之同源定向重組 (HDR) 臂 5' 及 3'，其中該轉基因不為 PCSK9。在某些具體實施例中，肝臟代謝性病變為鳥胺酸胺甲醯基轉移酶 (ornithine transcarbamylase)。在其它具體實施例中，為家族性高膽固醇血症或苯丙酮尿症。在一具體實施例中，受試者為新生兒。

【0021】在另一態樣中，提供一種治療遺傳疾病之系統。該系統包括：脂質奈米顆粒 (LNP)，該脂質奈米顆粒包含編碼靶向 PCSK9 基因之核酸酶的 mRNA 序列；及供體 AAV 載體，該供體 AAV 載體包含轉基因及指導其在標靶細胞中之表現的調控序列，該供體載體進一步包含對於轉基因之同源定向重組 (HDR) 臂 5' 及 3'，其中

該轉基因不為 PCSK9。在一些具體實施例中，核酸酶靶向 PCSK9 外顯子 7。在一些具體實施例中，巨型核酸酶為 ARCUS 巨型核酸酶。

【0022】在其它具體實施例中，基因編輯載體編碼 Cas9。在某些具體實施例中，基因編輯載體進一步編碼包含至少 20 個核苷酸的 sgRNA，其特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處。在一些具體實施例中，其中該系統包括 LNP，該 LNP 包含 Cas9 編碼序列及 gRNA。

【0023】在其它具體實施例中，供體載體進一步編碼包含至少 20 個核苷酸種子區之 sgRNA，其中該 sgRNA 特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處。

【0024】又另一態樣中，提供一種治療遺傳疾病的雙重載體系統。該系統包括：基因編輯載體，其包含表現匣，該表現匣包含編碼核酸酶的核酸序列及指導該核酸酶在標靶細胞中表現的調控序列，該標靶細胞包含 PCSK9 基因；及供體載體，其包含編碼用於從 PCSK9 基因座表現外源產物的核酸序列，其中插入的核酸序列並不編碼 PCSK9，其中該系統進一步包含指導核酸酶特異性靶向天然 PCSK9 基因座的序列；且其中該標靶細胞中的天然 PCSK9 在以雙重載體系統給藥後可選擇地消融或減少。

【0025】又另一態樣中，提供一種使用本文所述的系統治療病患的方法，其中該病患的天然 PCSK9 表現水平被降低，且其中該病患表現外源的產物。

【0026】又另一態樣中，提供用於鳥胺酸胺甲醯基轉移酶之工程化編碼序列。亦包括載體、表現匣及包含其等之重組病毒。

【0027】本發明的其它態樣和優點將從以下本發明的詳細說明中顯而易見。

### 【圖式簡單說明】

#### 【0028】

圖 1 顯示 rhPCSK9 基因座之示意圖，其顯示外顯子 7 內的供體剪接位點，以及包含感興趣的供體模板的 HDR 供體載體，例如 hFIX、hOTC。

圖 2 顯示一項先期研究的時間線，該研究包含 ARCUS2 或 SaCas9 在新生 NHP 中的 PCSK9 基因座中的 hFIX 袖珍基因敲入。

圖 3A 至 3C 顯示用於 SaCas9 或 ARCUS 媒介的基因校正的雙 AAV 載體系統之示意圖。圖 3A 顯示用於 ARCUS2 媒介的基因校正的雙 AAVhu37 載體系統的示意圖，其中該 AAVhu37-供體載體包含 hOTC 供體模板序列。圖 3B 顯示用於 Sa-Cas9 媒介的基因校正(反式；AAVhu37-SaCas9)的雙 AAVhu37 載體系統的示意圖，其中該用於 SaCas9 及 sgRNA 之表現匣位於二個分別的載體，且 AAVhu37.sgRNA-供體載體包含 hOTC 供體模板序列及 U6.sgRNA 匣。圖 3C 顯示用於 Sa-Cas9 媒介的基

因校正的雙 AAVhu37 載體系統(順式；AAVhu37.PCSK9-sgRN.SaCas9)的示意圖，其中該用於 SaCas9 及 sgRNA 之表現匣位於相同載體，且 hOTC 供體載體位於分別的載體。

圖 4A 至 4H 顯示在新生 NHP 中核酸酶媒介的基因靶向的體內測試。向動物投予  $1 \times 10^{13}$  GC/kg 之 AAVhu37.ARCUS2.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR，或  $1 \times 10^{13}$  GC/kg 之 AAVhu37.SaCas9.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR.U6.sgR，或  $1 \times 10^{13}$  GC/kg 之 AAVhu37.GFP.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR.U6.sgR，如圖 4A、4B 及 5G 中所示。圖 4C 顯示在新生 NHP 中於指定時間點的 hFIX 水平(繪製為 ng/mL)。圖 4D 顯示在新生 NHP 中於指定時間點的 PCSK9 水平(繪製為第 0 天的基線百分比)。圖 4E 顯示在新生 NHP 中於指定時間點的 ALT(丙胺酸轉胺酶)水平(繪製為 U/L)。圖 4F 顯示在新生 NHP 中於指定時間點的抗 FIX IgG 水平(繪製為稀釋因子，1/稀釋度)。圖 4G 顯示在新生 NHP 中於指定時間點的 PCSK9 水平(繪製為 ng/mL)。圖 4H 顯示在新生 NHP 中量測的重量(繪製為 g)。

圖 5A 至 5H 顯示投予至 3 月齡嬰兒 NHP 的針對圖 4 所述之體內試驗的結果。圖 5A 顯示在嬰兒 NHP 中於指定時間點的 hFIX 水平(繪製為 ng/mL)。圖 5B 顯示在嬰兒 NHP 中於指定時間點的 PCSK9 水平(繪製為第 0 天的基線百分比)。圖 5C 顯示在嬰兒 NHP 中於指定時間點的 ALT(丙胺酸轉胺酶)水平(繪製為 U/L)。圖 5D 顯示在

嬰兒 NHP 中於指定時間點的抗 FIX IgG 水平(繪製為稀釋因子，1/稀釋度)。圖 5E 顯示在嬰兒 NHP 中於指定時間點的 PCSK9 水平(繪製為 ng/mL)。圖 5F 顯示在嬰兒 NHP 中於指定時間點所量測的重量(繪製為 g)。圖 5G 為總表，顯示自圖 4A-5G 中描述的實驗的數據。圖 5H 顯示測試的新生兒和嬰兒 NHP 之間的各种數據比較。

圖 6A 至 6E 顯示在如圖 4A-4H 所述治療之 NHP 中在治療後於不同天收集的肝臟生檢樣品中的載體轉導(GC)和轉基因表現。圖 6A 顯示肝臟生檢樣品中的載體轉導水平，繪製為每二倍體細胞(diploid cell)中之 AAV 基因體拷貝數(GC)。圖 6B 顯示肝臟生檢樣品中轉基因 RNA 的相對表現。圖 6C 顯示使用特異性探針檢測肝臟生檢中的 FIX 和 ARCUS 的雙重原位雜交(ISH)。圖 6D 顯示用於轉導百分比量化的數位 ISH 圖像。圖 6E 顯示藉由 ISH 量化的 FIX 轉基因的轉導效率，並繪製為轉導百分比。

圖 7A 至 7L 顯示使用特異性探針檢 NHP 治療後 84 天收集的肝臟生檢中的 FIX 和 ARCUS 的雙重原位雜交(ISH)；顯示在各種放大視圖(以 AAVhu37.ARCUS2 及 AAVhu37.供體-HDR-hFIX 治療的 NHP)。圖 7A 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS。圖 7B 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 7C 顯示放大 4 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7D 顯示放大 4 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI(細胞核染色)的疊加圖

像。圖 7E 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS。圖 7F 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 7G 顯示放大 10 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7H 顯示放大 10 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。圖 7I 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS 表現。圖 7J 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 7K 顯示放大 20 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7L 顯示放大 20 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。

圖 8A 至 8M 顯示使用特異性探針檢測 NHP 治療後 84 天收集的肝臟生檢中的 FIX 和 ARCUS 之雙重原位雜交 (ISH)；顯示在各種放大視圖 (以 AAVhu37.EGFP 及 AAVhu37.Donor-HDR-hFIX.U6.sgR 治療的 NHP)。圖 8A 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE。圖 8B 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8C 顯示放大 4 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8D 顯示放大 4 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。圖 8E 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE。圖 8F 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8G 顯示放大 10 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8H 顯示放大 10 倍觀察到的 ISH 檢測的 GFP-

WPRE 和 hFIX 作為具有 DAPI (細胞核染色)的疊加圖像。圖 8I 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE 表現。圖 8J 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8K 顯示放大 20 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8L 顯示放大 20 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色)的疊加圖像。圖 8M 顯示在未治療的對照中放大 20 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色)的疊加圖像。

圖 9 顯示以 AAVhu37.ARCUS2 和 AAVhu37.供體-HDR-hFIX 治療的 NHP 中 ARCUS 媒介的靶向編輯。在治療後 84 天，收集肝臟生檢樣品，並根據擴增子序列 (amplicon-seq)計算存在的標靶區域中總插入缺失的百分比。

圖 10A 及 10B 顯示 PCSK9-hE7-KI 小鼠模型的示意圖。圖 10A 顯示以人類 *pcsk9* 外顯子 7 (hE7 包含 ARCUS 靶向序列)置換的小鼠 *pcsk9* 外顯子 7 的示意圖。人類 PCSK9 外顯子 7 序列顯示於 SEQ ID NO: 44。圖 10B 顯示將 PCSK9-hE7-KI 小鼠模型與其它疾病小鼠模型(諸如 OTC *spf<sup>ash</sup>*、KI- *spf<sup>ash</sup>* 模型)雜交的示意圖。*PCSK9-hE7-KI* 敲入小鼠模型首先是藉由以包含外顯子 7 的人類 *PCSK9* 基因區域置換包含鼠類 *Pcsk9* 基因之外顯子 7 的區域來生成的。然後將 *PCSK9-hE7-KI* 小鼠與 *sparse fur ash (spf<sup>ash</sup>)*小鼠雜交，由於在 *Otc* 基因外顯子 4 末端處的剪接供體位點具有 G 到 A 的點突變，其在 OTC 表現

上呈現 20 倍的減少。來自此雜交的小鼠被稱為 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠並如本文所述使用。縮寫：bp，鹼基對；E6，外顯子 6；E7：外顯子 7；E8，外顯子 8；HDR，同源定向重組；*PCSK9*，前蛋白轉化酶枯草溶菌素/kexin 9 型(基因，人類)；*Pcsk9*，前蛋白轉化酶枯草溶菌素/kexin 9 型(基因，小鼠)。

圖 11A-11I 顯示針對如圖 11I 中所示之載體的新生 NHP 中核酸酶媒介的基因靶向的體內測試。圖 11A 是顯示如實施例 3 中所述載體在新生 NHP 中核酸酶媒介之基因靶向的體內測試的實驗設計之圖表。動物 21-111、21-122 和 21-113 在給藥前是 AAV 結合抗體(BAb)陽性。在載體給藥後收集第 0 天的 21-178 樣品，這會干擾 Bab 測定(Bab assay)。°：列出在獨立 ITRseq 分析中鑑定的 OT 位點的數量。圖 11B 顯示如圖 11A 中所示組別的 PCSK9 水平，顯示為 ng/mL (頂行)或第 0 天的百分比(底行)。圖 11C 顯示組別的顯示為 U/L 的 ALT 水平(頂行)或顯示為 U/L 的 AST (底行)，如圖 11A 所示。圖 11D 顯示藉由 ISH 或 IF 量化之 OTC 轉基因的轉導效率，並繪製為轉導的肝細胞百分比。圖 11E 顯示小鼠的體重。圖 11F 顯示在第 84 天藉由定量 PCR 分析在肝臟中的載體 GC。圖 11G 顯示在第 84 天獼猴肝臟中 hOTC 及核酸酶的表現，其藉由定量 PCR 對從肝臟生檢樣品中分離的總 RNA 隨後進行逆轉錄測量，並表示為藉由 GAPDH 水平標準化的相對表現水平。圖 11H 顯示藉由擴增子序列對 *rhPCSK9* 靶向的基因座進行的插入或缺失分析。圖 11I

為新生 NHP 中核酸酶媒介的基因靶向的體內測試時間線示意圖，該新生 NHP 包括用於實施例 3 所述實驗測試的載體。

圖 12 顯示代表人類 PCSK9 序列的 PCSK9-hE7 敲入等位基因、小鼠 PCSK9 (mPCSK9) 及恆河猴 PCSK9 (rhPCSK9) 的 265 bp 序列的序列比對。縮寫：GAPDH，甘油醛-3-磷酸去氫酶 (glyceraldehyde-3-phosphate dehydrogenase)；GC，基因體拷貝；hOTC，人類鳥胺酰胺甲醯基轉移酶；OT，脫靶；PCR，聚合酶連鎖反應；*rhPCSK9*，前蛋白轉化酶枯草溶菌素/kexin 9 型(恆河猴基因)；RNA，核糖核酸。

圖 13 顯示用於 ARCUS2 媒介的基因校正的雙 AAV 載體系統的供體構建體的示意圖，其中該 AAV-供體載體包含 hOTC 供體模板序列。顯示了構建體中的 HDR 臂與敲入小鼠模型(圖 10A-10B)、NHP 及人類標靶區域的同源性。

圖 14A 顯示在 PCSK9-hE7-KI.spf-ash PUPS(部分 OTC 缺乏症模型)中進行的一項包含藉由 ARCUS2 的 hOTC 袖珍基因敲入 PCSK9 基因座的研究的時間線，如實施例 5 中所述。圖 14B 顯示每組將接受用於圖 14A 的研究的載體和劑量。

圖 14C-14I 顯示以如圖 7 所示載體治療、或未治療(KI WT)的小鼠並餵飼高蛋白質(HP)飲食 10 天的研究結果。圖 14C 顯示存活機率。圖 14D 顯示體重佔導入 HP 飲食之前體重的百分比。圖 14E 顯示在 HP 飲食的第 10

天的血漿  $\text{NH}_3$  水平。圖 14F 顯示在第 48 天的 mPCSK9 蛋白水平。圖 14G 顯示在第 59 天藉由擴增子序列量測的插入或缺失 %。圖 14H 顯示在第 59 天量測的肝臟生檢樣品中的載體轉導水平，繪製為每二倍體細胞的 AAV 基因體拷貝 (GC)。圖 14I 顯示 8 週的 OTC IF。

圖 15 為在 PCSK9-hE7-KI.1dlr-/1dlr-.apobec-/apobec-Pups (hoFH 模型) 中，於實施例 10 所述藉由 SaCas9 在 PCSK9 基因座中產生 hLDLR 袖珍基因敲入的實驗設計之示意圖。

圖 16 為顯示實施例 10 使用的載體的示意圖。

圖 17 顯示實施例 10 的實驗設計。

圖 18A-18D 顯示實施例 10 之實驗結果。圖 18A 顯示對於 shHDR + saCas9、mhHDR + saCas9、僅 shHDR 及未治療的小鼠的血清 LDL 水平。圖 18B 顯示對於 shHDR + saCas9、mhHDR + saCas9、僅 shHDR 治療的小鼠的插入或缺失百分比。圖 18C 顯示在第 63 天於肝臟量測的每二倍體基因體的 hLDLR 基因體拷貝。圖 18D 顯示在第 63 天，對於 shHDR + saCas9、mhHDR + saCas9、僅 shHDR 及未治療的小鼠的血清 LDL 水平。

圖 19 顯示對於實施例 10 之小鼠在第 63 天取得之肝臟樣品的免疫組織化學數據。

### 【實施方式】

【0029】本文提供的是組成物、套組及方法，其對患有某些遺傳疾病(包括肝臟代謝性病症)的病患提供穩定、長期的治療效果。該組成物、套組及方法利用靶向

標靶細胞之 PCSK9 基因座的核酸酶，且供體載體提供包括用於整合到 PCSK9 基因座中並從其表現外源產物的模板，其中該插入的核酸序列並不編碼 PCSK9，且內源性 PCSK9 的表現被破壞，表現水平降低。

**【 0030 】**

### **[PCSK9]**

前蛋白轉化酶枯草溶菌素 kexin 9 型 (proprotein convertase subtilisin kexin 9, PCSK9) 是一種絲胺酸蛋白酶，可降低肝臟和肝外低密度脂蛋白 (LDL) 受體 (LDLR; 606945) 水平並增加血漿 LDL 膽固醇。PCSK9 在調節血漿膽固醇穩態中至關重要。PCSK9 與低密度脂質受體家族成員低密度脂蛋白受體 (LDLR)、極低密度脂蛋白受體 (VLDLR)、脂蛋白元 E 受體 (LRP1/APOER) 和脂蛋白元受體 2 (LRP8/APOER2) 結合，並促進它們在細胞內酸性腔室中的降解。人類 PCSK9 具有蛋白質序列 NP\_777596.2，如 SEQ ID NO: 23 所示，具有 SEQ ID NO: 22 中所示之編碼序列。

**【 0031 】** 雖然 PCSK9 基因已被靶向用於治療膽固醇相關疾病，但本文證實 PCSK9 基因座是用於插入其它非 PCSK9 轉基因的基因靶向的安全港 (safe harbor)。因此，本文所提供之組成物、套組及方法利用靶向 PCSK9 基因座之核酸酶，並使用供體模板將治療性轉基因插入至標靶 PCSK9 基因座中。

**【 0032 】** 本文所提供之組成物、套組及方法包括：基因編輯組分 (在一些具體實施例中，載體)，及供體載體，其提供在宿主細胞中表現的治療性轉基因。

## 【 0033 】

## [基因編輯組分]

本文提供的組成物、套組及方法包括基因編輯組分，其包含核酸酶(或其之編碼序列)及指導該核酸酶特異性靶向染色體 1 上天然 PCSK9 基因座的序列。如本文所使用，「標靶 PCSK9 基因座」或「PCSK9 基因座」為 PCSK9 編碼區域中需要插入異源轉基因的任何位點。在某些具體實施例中，標靶 PCSK9 基因座位於 PCSK9 編碼序列之外顯子 7 中。圖 12 提供人類(h)、恆河猴(rh)及小鼠(m) PCSK9 外顯子 7 剪接位點的比對，本文使用 SaCas9 和靶向 PCSK9 的巨型核酸酶(稱為 ARCUS)舉例說明。

【 0034 】 本文描述了組成物，特別是核酸酶，其可用於靶向用於插入轉基因的基因，例如，對 PCSK9 特異的核酸酶。在某些具體實施例中，核酸酶為天然存在的。在其它具體實施例中，核酸酶為非天然存在的，即，在 DNA-結合域及/或切割域工程化。例如，可改變天然存在的核酸酶的 DNA-結合域以結合選定的標靶位點(例如，已被工程化為與同源結合位點不同之位點結合的巨型核酸酶)。在其它具體實施例中，核酸酶包含異源 DNA-結合域及切割域(例如，鋅指核酸酶(zinc finger nuclease)；TAL-效應核酸酶；具有異源切割域的巨型核酸酶 DNA-結合域)。

【 0035 】 在某些具體實施例中，核酸酶為靶向 PCSK9 之巨型核酸酶。巨型核酸酶為內切去氧核糖核酸

酶，特徵在於大的識別位點(12 至 40 個鹼基對雙股的 DNA 序列)，例如，I-SceI。當與核酸酶結合時，可在特定位置切割 DNA。可將限制酶導入細胞中，用於基因編輯或基因體原位編輯。在某些具體實施例中，核酸酶是歸巢核酸內切酶(homing endonuclease)之 LAGLIDADG (SEQ ID NO: 31)家族的成員。在某些具體實施例中，核酸酶是歸巢核酸內切酶之 I-CreI 家族的成員，其識別並切割 22 個鹼基對識別序列 SEQ ID NO: 32-CAAACGTCGTGAGACAGTTTG。參見，例如，WO 2009/059195。描述了用於合理設計單 LAGLIDADG (SEQ ID NO: 32)歸巢核酸內切酶的方法，該方法能夠全面重新設計 I-CreI 及其它歸巢核酸內切酶以靶向廣泛不同的 DNA 位點，包括位於哺乳動物、酵母、植物、細菌及病毒基因體中的位點(WO 2007/047859)。在一具體實施例中，核酸酶經由 SEQ ID NO: 19 所示序列(nt 330 至 1424)或與其共享至少 95%、98%或 99%同一性的序列所編碼。在一具體實施例中，核酸酶蛋白質序列為 SEQ ID NO: 20 所示序列、或與其共享至少 95%、98%或 99%同一性的序列。此類核酸酶在本文中有時稱為 ARCUS 核酸酶。術語「歸巢核酸內切酶」與術語「巨型核酸酶」同義，參見，WO 2018/195449，描述了某些 PCSK9 巨型核酸酶，其全文併入本文。

**【0036】** 鋅指核酸酶(ZFN)是藉由將鋅指 DNA 結合域與 DNA 切割域融合而產生的人工限制酶。鋅指域可設計成靶向特定的所需的 DNA 序列，這使得鋅指核酸

酶能夠靶向複雜基因體中的獨特序列。藉由利用內源性 DNA 修復機制，這些試劑可用於精確改變高等生物的基因體，並作為基因體編輯領域的重要工具。轉錄活化因子樣效應子核酸酶 (Transcription activator-like effector nuclease, TALEN) 是限制性內切酶，可被工程化以切割特定的 DNA 序列。它們是藉由將 TAL 效應子 DNA 結合域與 DNA 切割域 (一種切割 DNA 股的核酸酶) 融合而成的。在另一具體實施例中，編碼序列編碼鋅指核酸酶或轉錄活化因子樣 (TAL) 效應子核酸酶 (TALEN)。

【0037】在某些具體實施例中，核酸酶為 CRISPR-相關核酸酶 (Cas)，可選擇地為 Cas9。「Cas9」(CRISPR 相關蛋白質 9) 係指 RNA 引導的 DNA 核酸內切酶家族，其特徵在於兩個標籤核酸酶域，即 RuvC (切割非編碼股) 和 HNH (編碼股)。Cas9 之合適的細菌來源包括金黃色葡萄球菌 (*Staphylococcus aureus*) (SaCas9)、釀膿鏈球菌 (*Streptococcus pyogenes*) (SpCas9) 和腦膜炎雙球菌 (*Neisseria meningitides*) [KM Estelt *et al*, Nat Meth, 10: 1116-1121 (2013)]。野生型編碼序列可用於本文所述的構建體中。或者，細菌密碼子被優化用於在人類中表現，例如，使用多種已知的人類密碼子優化算法中的任何一種。或者，這些序列可全部或部分合成產生。可選擇地替換具有相似性質的其它核酸內切酶。參見，例如，可在 <http://crispr.u-psud.fr/crispr> 可得的公共 CRISPR 資料庫 (db)。

【0038】在某些具體實施例中，關於組成物、套組及方法，核酸酶編碼序列包含於基因編輯載體中。基因編輯載體包括包含編碼核酸酶之核酸序列及指導該核酸酶在標靶細胞中表現之調控序列的表現匣，該標靶細胞包含 PCSK9 基因。

【0039】如本文所使用的「載體」是包含核酸序列的生物或化學部分，該核酸序列可被導入合適的宿主細胞中以複製或表現該核酸序列。常見的載體包括非病毒載體及病毒載體。如本文所使用，非病毒系統可選自奈米顆粒、電穿孔系統和新穎生物材料、裸 DNA、噬菌體、轉位子、質體、黏接質體 (Phillip McClean, [www.ndsu.edu/pubweb/~mcclean/-plsc731/cloning/cloning4.htm](http://www.ndsu.edu/pubweb/~mcclean/-plsc731/cloning/cloning4.htm)) 及人工染色體 (Gong, Shiaoqing, et al. "A gene expression atlas of the central nervous system based on bacterial artificial chromosomes." *Nature* 425.6961 (2003): 917-925)。

【0040】如本文所使用，「表現匣」係指包含生物學上有用的核酸序列(例如，編碼蛋白質、酶或其它有用之基因產物的基因 cDNA；mRNA 等)和與其可操作地連接之調控序列的核酸分子，該調控序列指導或調控該核酸序列及其基因產物的轉錄、轉譯及/或表現。如本文所使用，「可操作地連接的」序列包括與核酸序列鄰接的調控序列和以反式或遠距離作用以控制序列的調控序列二者。此類調控序列一般包括，例如，一種或多種啟動子、增強子、內含子、Kozak 序列、多腺苷酸化序列、及 TATA 訊號。表現匣可包含基因序列之調控序列

上游(5'至)，例如，一種或多種啟動子、增強子、內含子等，及一種或多種增強子、或基因序列之調控序列下游(3'至)，例如，包含多腺苷酸化位點的 3'未轉譯的區域，以及其它元件。在其它具體實施例中，術語「轉基因」係指插入標靶細胞的一種或多種來自外源的 DNA 序列。通常，用於產生病毒載體的此類表現匣含有本文所述基因產物的編碼序列，其兩側是病毒基因體的包裝訊號和其它表現控制序列，例如本文所述的那些。在某些具體實施例中，載體基因體可含有二個或多個表現匣。

【0041】除了用於核酸酶之編碼序列，在某些具體實施例中基因編輯載體尚包括指導該核酸酶在宿主細胞表現的調控序列。在某些具體實施例中，調控元件包括啟動子。在某些具體實施例中，其中該系統設計用於治療以肝細胞中的突變或表型為特徵的代謝性病變，基因編輯載體可設計成使得核酸酶在肝特異性啟動子的控制下表現。本文所述的例示性質體和載體使用肝特異性啟動子甲狀腺素結合球蛋白(TBG)，其特徵在於 SEQ ID NO: 41 的序列。在其它具體實施例中，TBG 的縮短版本，本文稱為 TBG-S1 的變異體，其特徵在於 SEQ ID NO: 11 的序列，是有用的。在另一具體實施例中，使用具有 SEQ ID NO: 12 序列的雜交肝啟動子(HLP)。

【0042】在一些具體實施例中，希望使用具有低轉錄活性的啟動子或弱化啟動子。在一具體實施例中，該啟動子是肝特異性甲狀腺素結合球蛋白(TBG)啟動子的

弱化版本。在一具體實施例中，弱化啟動子在天然啟動子或 TBG-S1 序列的 5'或 3'末端處被截斷。在另一具體實施例中，該啟動子僅保留 TBG-S1 啟動子的 3'末端 113 nt，稱為 F113 (亦稱為 TBG-S1-F113)(SEQ ID NO: 19，nt 206 至 318)。美國臨時專利申請號 63/016,145 (2020 年 4 月 27 日申請)、63/033,738 (2020 年 6 月 2 日申請) 及 63/089,796 (2020 年 10 月 9 日申請)、PCT/US21/29386 及 PCT/US21/29403 (皆於 2021 年 4 月 27 日申請)，發明名稱為「使用具有低轉錄活性的啟動子來降低核酸酶表現和脫靶活性的組成物及方法 (COMPOSITIONS AND METHODS FOR REDUCING NUCLEASE EXPRESSION AND OFF-TARGET ACTIVITY USING A PROMOTER WITH LOW TRANSCRIPTIONAL ACTIVITY) 的各申請案皆藉由引用整體併入本文。

【0043】或者，可使用其它肝特異性啟動子，諸如  $\alpha 1$  抗胰蛋白酶 (A1AT)、人類白蛋白 (Miyatake *et al.*, *J. Virol.*, 71:5124 32 (1997))、及 B 型肝炎病毒核心啟動子 (Sandig *et al.*, *Gene Ther.*, 3:1002 9 (1996), TTR minimal enhancer/ promoter, alpha-antitrypsin promoter, LSP (845 nt)。參見，例如，The Liver Specific Gene Promoter Database, Cold Spring Harbor, <http://rulai.schl.edu/LSPD>。或者，可使用其它組織特異性啟動子，諸如肌肉特異性啟動子，諸如肌肉肌酸激酶 (MCK) 啟動子、或肌肉雜交 (MH) 啟動子。或者，其它啟動子，諸如組成型啟動子 (CMV、CBG、CB7 等等)、可調節的 (可誘導的) 啟動子

[參見，例如，WO 2011/126808 及 WO 2013/049493，藉由引用併入本文]、或反應生理訊號之啟動子可用於本文所述的載體中。可選擇地，若選擇可調節系統，則可能需要第三個載體來提供調節功能。

【0044】除了啟動子，基因編輯匣、表現匣及/或載體可含有一種或多種適當的「調控元件」或「調控序列」，其包含但不限於，增強子；轉錄因子；轉錄終止子；高效 RNA 處理訊號，諸如剪接及多腺苷酸化訊號 (polyA)；穩定細胞質 mRNA 之序列，例如土撥鼠肝炎病毒 (WHP) 轉錄後調控元件 (WPRE)；提高轉譯效率之序列 (即，科扎克 (Kozak) 共通序列)；增強蛋白質穩定性之序列；及當需要時，增強編碼產物分泌之序列。合適的 polyA 序列的實例包括，例如，SV40、牛生長激素 (bGH)、及 TK polyA。合適的增強子的實例包括，例如， $\alpha$  胎兒蛋白增強子、TTR 最小啟動子/增強子、LSP (TH-結合球蛋白啟動子/ $\alpha$ 1-微球蛋白/比庫 (bikunin) 增強子) 等。此等控制序列或調控序列可操作地連接至核酸酶編碼序列或轉基因編碼序列。

【0045】在某些具體實施例中，基因編輯載體包括 TBG 啟動子、一種或多種  $\alpha$  mic/bik 增強子、用於 ARCUS 巨型核酸酶的編碼序列、可選擇地 WPRE、及 polyA。在某些具體實施例中，表現匣包括 SEQ ID NO: 42 之 nt 211 至 nt 2964。

【0046】在一些具體實施例中，基因編輯組分進一步包括使核酸酶導向 PCSK9 標靶基因座中標靶位點的序

列。在某些具體實施例中，諸如對 PCSK9 特異性之如巨型核酸酶，並不需要另外的序列來使核酸酶導向標靶位點。然而，在這種情況下，例如 Cas9，提供稱為「單嚮導 RNA」或「sgRNA」的額外序列，其對於標的序列具有特異性。sgRNA 可以在與 Cas9 相同的載體(順式)或不同的載體(反式)上提供。如本文所使用，sgRNA 具有用於特異性 DNA 結合(即，與標靶 DNA 同源)的至少 20 個鹼基的序列(或約 24-28 個鹼基，有時稱為種子區)，其與 gRNA 支架結合。sgRNA 的轉錄應該精確地從其 5'端開始。當靶向模板 DNA 股時，sgRNA 的鹼基配對區域與轉錄序列具有相同的序列同一性。當靶向非模板 DNA 股時，sgRNA 的鹼基配對區域是轉錄序列的反向互補。可選擇地，基因編輯載體可包含一個以上的 sgRNA。sgRNA 位於被 Cas9(或 Cpf1)酶特異性識別的前間隔序列相鄰基序(PAM)的 5'端。通常，sgRNA 位於 PAM 序列的「立即」5'端，亦即沒有間隔或插入序列。在一具體實施例中，sgRNA「種子」編碼序列為 AAGTTGGTCCCCAAAGTCCC (SEQ ID NO: 8)，其用於藉由 SaCas9 靶向人類外顯子 7 及獼猴 PCSK9。然而，本領域技術之技術人員可設計其它 sgRNA。

【0047】在某些具體實施例中，sgRNA 包括至少 20 個核苷酸並特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處。在一些具體實施例中，種子區與 PCSK9 基因中的標靶位點具有 100%互補性。在其它具體實施

例中，相較於標靶位點，種子區含有 1、2、3、4 或 5 個錯配。

【0048】sgRNA 受 RNA 聚合酶啟動子及/或終止子的控制。在某些具體實施例中，RNA 聚合酶啟動子為 Pol III 啟動子，諸如 U6 啟動子。在另一具體實施例中，啟動子為 H1 啟動子。用於例示的 U6 啟動子的序列可見於 SEQ ID NO: 10。在其它具體實施例中，sgRNA 及 RNA 聚合酶啟動子位於供體載體。

【0049】在其它具體實施例中，例如，其中核酸酶為 Cas9，基因編輯組分進一步包括一種或多種核定位訊號(NLSs)。在一具體實施例中，NLS 位於 Cas9 編碼序列的兩側。在某些具體實施例中，NLS 具有 SEQ ID NO: 5 之 nt 4241 至 4288 的序列。參見，例如，Lu et al. Types of nuclear localization signals and mechanisms of protein import into the nucleus, Cell Commun Signal (May 2021) 19:60，其藉由引用併入本文。

【0050】在某些具體實施例中，提供核酸酶編碼序列作為傳訊 RNA (mRNA)。mRNA 可包括 5'未轉譯區域、3'未轉譯區域、及/或編碼或轉譯序列。在某些具體實施例中，用於 Cas9 之編碼序列以 mRNA 形式提供。

【0051】mRNA 可為天然發生或非天然發生的 mRNA。mRNA 可包括一種或多種修飾的核鹼基、核苷或核苷酸。在一些具體實施例中，本發明組成物中 mRNA 包含至少一種修飾，該修飾賦予核酸增加或增強的穩定性，包括，例如，提高對體內核酸酶消化的抗

性。mRNA 可包括任意數量的鹼基對，包括數十、數百或數千個鹼基對。任何數量(例如，全部、一些或沒有)的核鹼基、核苷或核苷酸可為典型種類的類似物，經取代、修飾的或以其它方式非天然地存在。在某些具體實施例中，可修飾所有特定的核鹼基類型。例如，mRNA 中的所有胞嘧啶可為 5-甲基胞嘧啶。如本文所使用，與本文提供之核酸相關的術語「修飾」及「修飾的」包括至少一種改變，其較佳地增強穩定性並使 mRNA 比野生型或天然存在的 mRNA 更穩定(例如，抗核酸酶消化)。如本文所使用，與本文提供之核酸，且特別是與 mRNA 相關的術語「穩定的」及「穩定性」係指對於例如通常能夠降解此類 mRNA 的核酸酶(即，核酸內切酶或核酸外切酶)降解的抗性增加或增強。增加的穩定性可包括，例如，對內源性酶(例如，核酸內切酶或核酸外切酶)或標靶細胞或組織內條件的水解或其它破壞的敏感性降低，從而增加或增強此類 mRNA 在標靶細胞、組織、受試者及/或細胞質中的駐留。本文提供的穩定化的 mRNA 分子相對於它們天然存在的、未修飾的對應物(例如野生型版本的 mRNA)已證實更長的半衰期。與本發明之 mRNA 相關的術語「修飾」和「修飾的」亦涵蓋改善或增強 mRNA 核酸轉譯的改變，包括例如包含在蛋白質轉譯起始中起作用的序列(例如，Kozak 共通序列)。

【0052】在一些具體實施例中，本文所述之 mRNA 已經接受化學或生物修飾以使其更穩定。對於 mRNA 的

例示性修飾包括鹼基的缺失(例如,藉由刪除或藉由以一個核苷酸替換另一核苷酸)或鹼基的修飾,例如鹼基的化學修飾。如本文所使用的短語「化學修飾」係包括導入不同於天然存在的 mRNA 中所見化學性質的修飾,例如共價修飾,例如導入修飾的核苷酸(例如,核苷酸類似物,或包含在此類 mRNA 分子中非天然存在的側基)。

【0053】在一些具體實施例中, mRNA 序列中 C 及/或 U 殘基的數量減少。在另一具體實施例中,藉由用編碼特定胺基酸的一個密碼子替換編碼相同或相關胺基酸的另一密碼子來減少 C 及/或 U 殘基的數量。對本發明 mRNA 核酸的預期修飾還包括併入假尿苷假尿苷( $\psi$ )或 5-甲基胞嘧啶(m5C)。可藉由本領域技術人員或普通技術人員容易得知的方法對本發明的 mRNA 進行取代及修飾。

【0054】在某些具體實施例中, mRNA 包括 5'端帽結構、鏈終止核苷酸、莖環(stem loop)及/或多腺苷酸化訊號。端帽結構或端帽種類是包括藉由連接子連接的兩個核苷部分的化合物,並可選自天然存在的端帽、非天然存在的端帽或端帽類似物,或抗反向端帽類似物。mRNA 可替代地或額外地包括鏈終止核苷。

【0055】在某些具體實施例中, mRNA 包括莖環,諸如組蛋白莖環。莖環可包括 1、2、3、4、5、6、7、8、或更多個核苷酸鹼基對。莖環可位於 mRNA 的任何區域。例如,莖環可位於未轉譯區域(5'未轉譯區域或

3'未轉譯區域)、編碼區域、或 polyA 序列或尾端之中、之前或之後。

【0056】在某些具體實施例中，mRNA 包括 polyA 序列。polyA 序列可完全或大部分由腺嘌呤核苷酸或其類似物或衍生物組成。在某些具體實施例中，polyA 序列是位於 mRNA 3'未轉譯區域附近的尾部。

【0057】mRNA 可編碼任何感興趣的多肽，例如，核酸酶，包括任何天然或非天然存在的或以其它方式修飾的多肽。藉由 mRNA 編碼的多肽可為任何大小並可具有任何二級結構或活性。在一些具體實施例中，藉由 mRNA 編碼的多肽當在細胞中表現時可具有治療效果。

【0058】

#### [供體載體]

組成物、套組及方法包括供體載體，其提供用於治療性轉基因的編碼序列。在某些具體實施例中，供體載體含有表現匣，該表現匣包含編碼轉基因之核酸序列，及指導該轉基因在標靶細胞中之表現的調控序列。在某些具體實施例中，轉基因編碼在肝臟代謝性病變或其它遺傳疾病中異常表現的蛋白質。轉基因編碼 PCSK9 以外的蛋白質。此類蛋白質包括，但不限於 OTC、低密度脂蛋白受體(LDLr)、因子 IX (諸如 SEQ ID NO: 55 或 56 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列)、及因子 VIII (諸如 SEQ ID NO: 53 或 54 中所示序列，或與其共享至少 70%、75%、

80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列)。

【0059】可經由供體載體遞送的其它例示性基因包括，但不限於，與肝醣儲積症或 1A 型缺乏症(GSD1)相關的葡萄糖-6-磷酸酶、與磷酸烯醇式丙酮酸羧激酶(PEPCK)缺乏症相關的 PEPCK；周期蛋白依賴性激酶樣 5 (CDKL5)，亦已知為與癲癇發作和嚴重神經發育障礙相關的絲胺酸/蘇胺酸激酶 9 (STK9)；與半乳糖血症相關的半乳糖-1 磷酸尿苷轉移酶，諸如 SEQ ID NO: 63 或 64 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列；與苯丙酮尿症(PKU)相關之苯丙胺酸羥化酶(PAH)；與 1 型原發性高草酸鹽尿症相關之基因產物，包括羥基酸氧化酶 1 (GO/HAO1)，諸如 SEQ ID NO: 49 或 50 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列，及 AGXT，諸如 SEQ ID NO: 47 或 48 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列，與楓糖尿症相關之支鏈  $\alpha$ -酮酸脫氫酶，包括 BCKDH、BCKDH-E2、BAKDH-E1a 及 BAKDH-E1b；與 1 型酪胺酸血症相關之富馬醯基乙醯乙酸水解酶；與甲基丙二酸血症相關之甲基丙二醯基-CoA 變位酶；與中鏈乙醯基 CoA 缺乏症相關之中鏈醯基 CoA 脫氫酶；

與鳥胺酸胺甲醯基轉移酶缺乏症相關之鳥胺酸胺甲醯基轉移酶(OTC)；與瓜胺酸血症相關之精胺琥珀酸合成酶(ASS1)，諸如 SEQ ID NO: 69 或 70 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列；卵磷脂-膽固醇醯基轉移酶(LCAT)缺乏症；甲基丙二酸血症(MMA)；與尼曼匹克症(Niemann-Pick disease, C1型)相關之 NPC1；丙酸血症(PA)；與家族性高膽固醇血症(FH)相關之低密度脂蛋白受體(LDLR)蛋白質，諸如 SEQ ID NO: 73 或 74 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列，LDLR 變異體，諸如 WO 2015/164778 中所述者，或具有 SEQ ID NO: 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列；與失智相關之 ApoE 及 ApoC 蛋白質；脂蛋白脂肪酶(LPL)(脂蛋白脂肪酶缺乏症)，諸如 SEQ ID NO: 67 或 68 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列，與克果納傑氏症(Crigler-Najjar disease)相關之 UDP-葡萄糖醛酸轉移酶；與嚴重複合型免疫缺乏症相關之腺苷去胺酶；與痛風及萊施-尼漢二氏症候群(Lesch-Nyan syndrome)相關之鳥嘌呤磷醯基核苷轉移酶；與生物素酶缺乏症相關之生物素酶；與法布瑞氏症

(Fabry disease)相關之  $\alpha$ -半乳糖苷酶 A (a-Gal A)，諸如 SEQ ID NO: 75 或 76 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列；與 GM1 神經節醣苷病相關之  $\beta$ -半乳糖苷酶 (GLB1)；與威爾森氏症 (Wilson's Disease) 相關之 ATP7B；與 2 型及 3 型高歇氏病 (Gaucher disease) 相關之  $\beta$ -葡萄糖腦苷酶，諸如 SEQ ID NO: 51 或 52 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列；與齊威格氏症 (Zellweger syndrome) 相關之過氧化體膜蛋白 70 kDa；與異染性白質失養症 (metachromatic leukodystrophy) 相關之芳基硫酸酯酶 A (ARSA)，與克拉培氏病 (Krabbe disease) 相關之半乳糖腦苷酶 (galactocerebrosidase, GALC) 酵素，與龐貝氏症 (Pompe disease) 相關之  $\alpha$ -葡萄糖苷酶 (GAA)，諸如 SEQ ID NO: 79 或 80 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列；與 A 型尼曼匹克症 (Nieman Pick disease type A) 相關之神經磷脂酶 (sphingomyelinase, SMPD1) 基因；肌肽酶 (carnosinase, CN1)；黃嘌呤-鳥嘌呤磷醣基核甘轉移酶 (hypoxanthine-guanine phosphoribosyltransferase, HGPRT)；紅血球生成素 (EPO)；胺甲醯基磷酸合成酶 (CPS1)、N-乙醯穀胺酸合酶 (NAGS)；精胺基琥珀酸裂

解酶 (ASL)(精胺酸琥珀酸尿症 (Argininosuccinic Aciduria))，諸如 SEQ ID NO: 57 或 58 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列；及精胺酸酶 (AG)；與成人發病 II 型瓜胺酸血症 (CTLN2) 相關之精胺基琥珀酸合成酶 (WO 2018/144709，其藉由引用併入本文)；與尿素循環障礙相關之胺甲醯基磷酸合成酶 1 (CPS1)；與脊髓肌萎縮症相關之運動神經元存活 (survival motor neuron, SMN) 蛋白；與法伯脂肪肉芽腫病 (Farber lipogranulomatosis) 相關之神經醯胺酶；與 GM2 神經節醯苷病與戴薩克斯症 (Tay-Sachs disease) 及山德霍夫症 (Sandhoff disease) 相關的 b-己醯胺酶；與天冬醯胺基胺基葡萄糖尿症 (aspartyl-glucosaminuria) 相關之天冬醯胺基胺基葡萄糖苷酶；與岩藻糖沉積症 (fucosidosis) 相關之  $\alpha$ -岩藻糖苷酶 ( $\alpha$ -fucosidase)；與  $\alpha$ -甘露糖沉積症相關之  $\alpha$ -甘露糖酶；與急性間歇性紫質沈著病 (acute intermittent porphyria, AIP) 相關之膽色素原去胺酶 (porphobilinogen deaminase)；用於治療  $\alpha$ -1 抗胰蛋白酶缺乏症 (肺氣腫) 之  $\alpha$ -1 抗胰蛋白酶，諸如 SEQ ID NO: 77 或 78 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列；用於治療由於地中海貧血或腎功能衰竭引起的貧血之紅血球生成素；用於治療缺血性疾病之血管內皮生長因子、血管生成素-1 和纖維母細胞生長因

子；用於治療例如動脈粥樣硬化、血栓形成或栓塞中所見的血管閉塞之凝血酶調節素(thrombomodulin)及組織因子路徑抑制劑；用於治療帕金森氏病(Parkinson's disease)之芳香族胺基酸去羧基酶(AADC)及酪胺酸羥化酶(TH)；用於治療充血性心衰竭之 $\beta$ 腎上腺素性受體、受磷蛋白(phospholamban)、肌(內)質網三磷酸腺苷酶-2(sarco(endo)plasmic reticulum adenosine triphosphatase-2, SERCA2)和心臟腺苷酸環化酶的反義或突變形式；用於治療各種癌症的腫瘤抑制基因，例如 p53；細胞因子(cytokine)，諸如用於治療炎症和免疫疾病和癌症的各種介白素之一；用於治療肌營養不良之肌肉萎縮蛋白(dystrophin)或小肌肉萎縮蛋白(minidystrophin)及肌營養相關蛋白(utrophin)或小肌營養相關蛋白(miniutrophin)；及用於治療糖尿病之胰島素或 GLP-1。

【0060】用於遞送的合適之轉基因的實例包括，例如，與家族性高膽固醇血症相關者(例如，VLDLr、LDLr、ApoE，參見，例如，WO 2020/132155、WO 2018/152485、WO 2017/100682，其藉由引用而併入本文)、肌營養不良、囊性纖維化及罕見疾病或孤兒疾病(orphan disease)。此類罕見疾病的實例可包括脊髓肌萎縮症(SMA)、亨廷頓氏病(Huntingdon's Disease)、雷特症候群(Rett Syndrome)(例如，甲基-CpG-結合蛋白 2(MeCP2)；UniProtKB – P51608)、肌肉萎縮性脊髓側索硬化症(Amyotrophic Lateral Sclerosis, ALS)、杜氏型肌營養不良症(Duchenne Type Muscular dystrophy)、弗

里德希氏共濟失調 (Friedrichs Ataxia)(例如，共濟蛋白 (frataxin))、顆粒蛋白前體 (progranulin)(PRGN)(與非阿茨海默氏症腦退化相關，包括額顳葉失智症 (FTD)、進行性非流利性失語症 (progressive non-fluent aphasia, PNFA)及語義性失智)等。其它有用的基因產物包括，胺甲醯基合成酶 I、鳥胺酸胺甲醯基轉移酶 (OTC)、精胺基琥珀酸合成酶、用於治療精胺基琥珀酸裂解酶缺乏症之精胺基琥珀酸裂解酶 (ASL)、精胺酸酶、富馬醯基乙醯乙酸水解酶、苯丙胺酸羥化酶、 $\alpha$ -1 抗胰蛋白酶、恆河猴  $\alpha$ -胎蛋白 (AFP)、恆河猴絨毛膜性腺激素 (chorionic gonadotrophin, CG)、葡萄糖-6-磷酸酶，諸如 SEQ ID NO: 59 或 60 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列，與遺傳性血管性水腫相關之血漿蛋白酶 C1 抑制劑 (SERPING1)，諸如 SEQ ID NO: 61 或 62 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列，膽色素原去胺酶、與高胱胺酸尿症相關之胱硫醚  $\beta$ -合酶；諸如 SEQ ID NO: 65 或 66 中所示序列，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 同一性之序列，支鏈酮酸去羧酶，白蛋白，異戊醯基-coA 去氫酶，丙醯基 CoA 羧化酶，甲基丙二醯基 CoA 變位酶，戊二醯基 CoA 去氫酶，胰島素， $\beta$ -葡萄糖苷酶，丙酮酸羧化酶，肝磷

酸化酶，磷酸化酶激酶，甘胺酸去羧酶，H-蛋白質，T-蛋白質，囊性纖維化跨膜調節劑 (cystic fibrosis transmembrane regulator, CFTR) 序列，及肌肉萎縮蛋白基因產物 [例如，小-或微小-肌肉萎縮蛋白]。其它有用的基因產物包括酶，諸如可用於酶替代療法的酶，其可用於由酶活性不足引起的多種病況。例如，含有甘露糖-6-磷酸酯的酶可用於治療溶酶體貯積病 (例如，合適的基因包括編碼  $\beta$ -葡萄糖醛酸 (GUSB) 的基因)。用於遞送之合適的轉基因的實例可包括在 AAV 載體中遞送的人類共濟蛋白，如敘述於，例如，PCT/US20/66167 (2020 年 12 月 18 日)、美國臨時專利申請號 62/950,834 (2019 年 12 月 19 日申請) 及美國臨時申請號 63/136,059 (2021 年 1 月 11 日申請)，其藉由引用而併入本文。用於遞送之合適的轉基因的另一實例可包括在 AAV 載體中遞送的人類酸 - $\alpha$ - 葡萄糖苷酶 (GAA)，如敘述於，例如，PCT/US20/30493 (2020 年 4 月 30 日，現以 WO2020/223362A1 公開)、PCT/US20/30484 (2020 年 4 月 20 日)(現以 WO2020/223356 A1 公開)、美國臨時專利申請號 62/840,911 (2019 年 4 月 30 日)、美國臨時申請號 62.913,401 (2019 年 10 月 10 日)、美國臨時專利申請號 63/024,941 (2020 年 5 月 14 日)、及美國臨時專利申請號 63/109,677 (2020 年 11 月 4 日)，其藉由引用而併入本文。再者，用於遞送之合適的轉基因的另一實例包括在 AAV 載體中遞送的人類  $\alpha$ -L-艾杜糖醛酸酶 (IDUA)，如敘述於，例如，PCT/US2014/025509 (2014

年 3 月 13 日，現以 WO 2014/151341 公開)，及美國臨時專利申請號 61/788,724 (2013 年 3 月 15 日申請)，其藉由引用而併入本文。

【0061】其它有用的治療產物包括那些在肌肉中表現的產物，包括心肌。由轉基因編碼的其它有用的治療產物包括激素和生長及分化因子，包括，但不限於胰島素、升糖素、升糖素樣肽 1 (GLP-1)、生長激素(GH)、副甲狀腺激素(PTH)、生長激素釋放因子(GRF)、濾泡刺激激素(FSH)、黃體激素(LH)、人類絨毛膜促性腺素(hCG)、血管內皮生長因子(VEGF)、血管生成素、血管抑制素、顆粒性白血球聚落刺激因子(GCSF)、紅血球生成素(EPO)、結締組織生長因子(CTGF)、鹼性纖維母細胞生長因子(bFGF)、酸性纖維母細胞生長因子(aFGF)、表皮生長因子(EGF)、轉化生長因子  $\alpha$ (TGF $\alpha$ )、血小板衍生生長因子(PDGF)、胰島素生長因子 I 及 II(IGF-I 及 IGF-II)，任何一種轉化生長因子  $\beta$  超家族(包括 TGF $\beta$ )、活化素、抑制素或任何骨形態發生蛋白(BMP) BMP 1-15、生長因子之調節蛋白(heregulin)/神經調節蛋白(neuregulin)/ARIA/neu 分化因子(NDF)家族中的任何一種、神經生長因子(NGF)、腦源性神經營養因子(BDNF)、神經促素 NT-3 及 NT-4/5、睫狀神經營養因子(CNTF)、膠質細胞株源性神經營養因子(GDNF)、神經秩蛋白(neurturin)、聚集蛋白(agrins)、信號素蛋白(semaphorins)/摺疊蛋白(collapsins)家族中的任何一種、軸突導向因子-1 (netrin-1)和軸突導向因子-2、肝細

胞生長因子 (HGF)、ephrin、noggin、音蝟因子 (sonic hedgehog) 及酪胺酸羥化酶。可用於本文的其它轉基因包括用於治療 I-VII 型黏多糖病的那些 (IDUA、IDS、GNA、HGSNAT、NAGLU、SGSH、GALNS、GLB1、ARSB、GUSB)。可用於治療 MPSI 的例示性序列可見於 WO 2019/010335，其藉由引用併入本文。可用於治療 MPSII 的例示性序列可見於 WO 2019/060662，其藉由引用併入本文。可用於治療 MPSIIIa 的例示性序列可見於 WO 2019/108857，其藉由引用併入本文。可用於治療 MPSIIIb 的例示性序列可見於 WO 2019/108856，其藉由引用併入本文。

【0062】在一些具體實施例中，轉基因匣包括啟動子、轉基因編碼序列及 poly A 序列。在一些具體實施例中，啟動子為肝特異性啟動子，諸如 TBG 啟動子、TBG-S1 啟動子、HLP 啟動子、或本文其它所述的啟動子。在其它具體實施例中，提供不含啟動子之轉基因，且被插入至天然 PSCK9 啟動子之基因體下游中。

【0063】轉基因匣、表現匣及/或載體(編輯或供體)可含有一種或多種適當的「調控元件」或「調控序列」，其包含但不限於增強子；轉錄因子；轉錄終止子；高效 RNA 處理訊號，諸如剪接及多腺苷酸化訊號 (polyA)；穩定細胞質 mRNA 之序列，例如土撥鼠肝炎病毒 (WHP) 轉錄後調控元件 (WPRE)；提高轉譯效率之序列 (即，Kozak 共通序列)；增強蛋白質穩定性之序列；及當需要時，增強編碼產物分泌之序列。合適的 polyA 序

列的實例包括，例如，SV40、牛生長激素(bGH)、及 TK polyA。合適的增強子的實例包括，例如， $\alpha$ 胎兒蛋白增強子、TTR 最小啟動子/增強子、LSP (TH-結合球蛋白啟動子/ $\alpha$ 1-微球蛋白/比庫增強子)等。此等控制序列或調控序列可操作地連接至核酸酶編碼序列或轉基因編碼序列。

【0064】除了轉基因匣，在某些具體實施例中，供體載體亦包括轉基因匣的同源定向重組(HDR)臂 5'及 3'，以促進轉基因之同源定向重組至內源基因體中。同源臂指向標靶 PCSK9 基因座並可具有不同的長度。在一些具體實施例中，HDR 臂長度各為約 100bp 至約 1000bp。在其它具體實施例中，HDR 臂各為約 130bp 至約 500bp。在其它具體實施例中，HDR 臂各為約 100bp 至約 300bp。在其它具體實施例中，HDR 臂各為約 100bp 至約 400bp。在其它具體實施例中，HDR 臂各為約 250bp 至約 500bp。在其它具體實施例中，HDR 臂各為約 300bp 至約 500bp。在某些具體實施例中，HDR 臂各為約 100bp、125bp、150bp、175bp、200bp、225bp、250bp、275bp、300bp、325bp、350bp、375bp、400bp、425bp、450bp、450bp、475bp、或 500bp。在一具體實施例中，HDR 臂為 130bp。在另一具體實施例中，HDR 臂為 137bp。在其它具體實施例中，HDR 臂為約 130bp 至 140bp。在另一具體實施例中，HDR 臂為約 500bp。在另一具體實施例中，不存在 HDR 臂。HDR 臂理想地與標靶 PCSK9 基因座具有高程度的互補性，儘管

它不需要是 100%的互補性。在一些具體實施例中，各 HDR 臂中允許有 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 或更多個錯配。用於靶向 PCSK9 外顯子 7 的合適 HDR 臂序列顯示於 SEQ ID Nos:24-29 中。在一具體實施例中，HDR 臂序列選自 SEQ ID Nos：24-29。

【0065】本文亦提供組成物、套組及方法，用於核酸酶媒介、位點特異性整合 OTC 轉基因匣至基因體中的 PCSK9 安全港中，提供患有 OTC 缺乏症之病患長期治療性益處。提供用於 OTC 的工程化編碼序列，本文稱為 hOTCco2，並顯示在 SEQ ID NO：17 中。提供具有 SEQ ID NO: 17 之序列或與其共享至少 90%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或至少 99.9% 同一性之序列的核酸。在一具體實施例中，該核酸與 SEQ ID NO: 30 中所示之天然 OTC 編碼序列共享少於 80%、少於 79%、少於 78%、少於 77%、少於 76%、少於 75%、少於 74%、少於 73%、少於 72%、少於 71%、或少於 70%同一性。

【0066】可用於治療 OTC 的其它序列描述於 WO 2015/138348 及 WO 2015/138357，其藉由引用併入本文。可用於治療 PKU 的說明性序列描述於 WO 2018/126112，其藉由引用併入本文。其它序列顯示於 SEQ ID NO: 71 或 72，或與其共享至少 70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一性之序列。

**【 0067 】**

## 病毒及非病毒載體

本文所述之(基因編輯及供體)表現匣或編碼序列可工程化至用於遞送至標靶細胞(例如肝細胞)的任何合適的遺傳元件中,例如載體。如本文所使用之「載體」為一種包含核酸序列的生物或化學部分,其可被導入適當的宿主細胞中以複製或表現該核酸序列。一般的載體包括非病毒載體及病毒載體。如本文所使用,非病毒系統可選自奈米顆粒、電穿孔系統及新型生物材料、裸DNA、噬菌體、轉位子、質體、黏接質體(Phillip McClean, [www.ndsu.edu/pubweb/~mcclean/~plsc731/cloning/cloning4.htm](http://www.ndsu.edu/pubweb/~mcclean/~plsc731/cloning/cloning4.htm))及人工染色體(Gong, Shiao-ching, et al. “A gene expression atlas of the central nervous system based on bacterial artificial chromosomes.” *Nature* 425.6961 (2003): 917-925)。在一具體實施例中,如本文所述或本技術領域已知的,經由非病毒載體或脂質奈米顆粒遞送核酸。

**【 0068 】**在某些具體實施例中,基因編輯組分被封裝在脂質奈米顆粒(LNP)中。參見,例如,Conway et al, Non-viral Delivery of Zinc Finger Nuclease mRNA Enables Highly Efficient In Vivo Genome Editing of Multiple Therapeutic Gene Targets, *Molecular Therapy*, 27(4):866-877 (April 2019),其藉由引用併入本文。如本文所使用,短語「脂質奈米顆粒」係指包含一種或多種脂質(例如,陽離子脂質、非陽離子脂質及PEG修飾的脂質)的轉移媒介物。較佳地,脂質奈米顆粒被配製

為將一種或多種 mRNA 遞送至一種或多種標靶細胞(例如，肝臟和/或肌肉)。合適之脂質的實例包括，例如，磷脂醯基化合物(例如，磷脂醯甘油、磷脂醯膽鹼、磷脂醯絲胺酸、磷脂醯乙醇胺、神經鞘質、腦苷脂及神經節苷脂)。亦考慮使用聚合物作為轉移媒介物，無論是單獨使用還是與其它轉移媒介物結合使用。合適的聚合物可包括，例如，聚丙烯酸酯、聚烷基氰基丙烯酸酯、聚丙交酯、聚丙交酯-聚乙交酯共聚物、聚己內酯、葡聚醣、白蛋白、明膠、藻酸鹽、膠原蛋白、幾丁聚醣、環糊精、樹枝狀聚合物及聚乙烯亞胺。在一具體實施例中，轉移媒介物的選擇係基於其促進將 mRNA 轉染至標靶細胞的能力。用於 mRNA 之有用的脂質奈米顆粒包含陽離子脂質以包裹及/或增強將 mRNA 遞送至標靶細胞中，該標靶細胞將充當蛋白質生產的貯庫。如本文所使用，短語「陽離子脂質」係指在選定的 pH (諸如生理 pH)下攜帶淨正電荷的多種脂質種類中的任何一種。可藉由包括採用一種或多種陽離子脂質、非陽離子脂質及 PEG 修飾的脂質的不同比例的多組分脂質混合物來製備所考量的脂質奈米顆粒。文獻中已描述數種陽離子脂質，其中許多是可商購的。參見，例如，WO2014/089486、US 2018/0353616A1 及 US 8,853,377B2，其藉由引用而併入。在某些具體實施例中，使用常規程序進行 LNP 配製，包含膽固醇、可離子化脂質、輔助性脂質、PEG-脂質和在包封的核酸周圍形成脂質雙層的聚合物(Kowalski et al., 2019, Mol. Ther. 27(4):710-

728)。在一些具體實施例中，LNP 包含具有輔助性脂質 DOPE 之陽離子脂質(即，N-[1-(2,3-二油醯基氧基)丙基]-N,N,N-氯化三甲銨(DOTMA)、或 1,2-二油醯基-3-三甲基銨-丙烷(DOTAP))。在一些具體實施例中，LNP 包含離子化脂質 Dlin-MC3-DMA 離子化脂質、或基於二酮哌啉的離子化脂質(cKK-E12)。在一些具體實施例中，聚合物包含聚乙烯亞胺(PEI)、或聚( $\beta$ -胺基)酯(PBAE)。參見，例如，WO2014/089486、US 2018/0353616A1、US2013/0037977A1、WO2015/074085A1、US9670152B2 及 US 8,853,377B2，其藉由引用而併入。在某些具體實施例中，其中基因編輯組分包括 Cas9 mRNA，LNP 亦包括 gRNA。

【 0069 】 本文有用的某些 LNP 包括在 WO 2021/077066 及 WO 2021/055892 中描述的那些，其各藉由引用以其整體併入本文。有用的 LNP 包括那些顯示增強遞送至肝臟的 LNP。可改變 LNP 調配物以增強肝臟遞送。例如，可改變類型和離子化脂質：mRNA 比例、mRNA：sgRNA 比例、離子化脂質、磷脂、膽固醇及 PEG-脂質的莫耳比等。在一具體實施例中，LNP 為 Kauffman, K. J.; Dorkin, J. R.; Yang, J. H.; Heartlein, M. W.; DeRosa, F.; Mir, F. F.; Fenton, O. S.; Anderson, D. G., Optimization of lipid nanoparticle formulations for mRNA delivery in vivo with fractional factorial and definitive screening designs. Nano letters 2015, 15 (11), 7300-7306 所敘述的 LNP，其藉由引用併入本文。在某

些具體實施例中，以離子化脂質：mRNA 之重量比在 5:1 至 25:1 之間變化來設計 LNP。在某些具體實施例中，離子化脂質：mRNA 之重量比為 1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、12.5:1、15:1、20:1、或 25:1。在某些具體實施例中，mRNA：sgRNA 之重量比為 1:1、1:2、2:1、1:4、1:5、5:1、4:1、3:1、或 2:1。

【0070】其它 LNP 已被描述並且在本文中是有用的。參見，例如，WO 2016/118724、US 10,413,618B2、US 10,723,692B2、及 US8754062B2，其各藉由引用併入本文。

【0071】本文中某些實施例說明了在 AAV 載體基因體中使用含有基因編輯組分(核酸酶)編碼序列和轉基因編碼序列的 AAV 載體。然而，本文所述之構建體的使用不限於 AAV 構建體並可用於其它載體。在某些具體實施例中，可將載體基因體包裝至不同的載體中(例如，重組波卡病毒(bocavirus))。在某些具體實施例中，表現匣可被包裝至不同的病毒載體、非病毒載體中、及/或不同的遞送系統中。在某些具體實施例中，基因編輯組分在 LNP 中提供。

【0072】「質體」或「質體載體」在本文中通常是由在載體名稱之前及/或之後的小寫 p 指定。可根據本發明使用的質體、其它選殖和表現載體、其特性以及其構建/操作方法對於本領域技術人員而言是顯而易見的。在一具體實施例中，將如本文所述的核酸序列或如本文

所述的表現匣工程化至合適的遺傳元件(載體)中，該遺傳元件可用於產生病毒載體及/或用於遞送至宿主細胞，例如裸 DNA、噬菌體、轉位子、黏接質體、游離基因體(episome)等，其轉移其上所攜帶的核酸酶序列。選擇的載體可藉由任何合適的方法遞送，包括轉染、電穿孔、微脂體遞送、膜融合技術、高速 DNA 包覆的小球、病毒感染和原生質體融合。用於製造此類構建體的方法是核酸操作技術人員已知的，包括遺傳工程、重組工程和合成技術。參見，例如，Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY。

【0073】在某些具體實施例中，表現匣位於用於包裝到病毒衣殼中的載體基因體中。例如，對於 AAV 載體基因體，表現匣的組分在最末端 5'端和最末端 3'端側接 AAV 反向末端重複序列。例如，5' AAV ITR、表現匣、3' AAV ITR。在其它具體實施例中，可選擇自我互補 AAV。在其它具體實施例中，可使用反轉錄病毒系統、慢病毒載體系統或腺病毒系統。

#### 【0074】

#### AAV 載體

在某些具體實施例中，基因編輯載體及/或供體載體被提供作為重組 AAV。「重組 AAV」或「rAAV」是一種 DNA 酶抗性病毒顆粒，含有兩個元件，一個 AAV 衣殼和一個至少含有包裝在 AAV 衣殼內的非 AAV 編碼序列的載體基因體。除非另有指出，此術語可與短語

「rAAV 載體」或「AAV 載體」互換使用。rAAV 是一種「複製缺陷病毒」或「病毒載體」，因為它缺乏任何功能性 AAV rep 基因或功能性 AAV cap 基因，且不能產生子代。在某些具體實施例中，唯一的 AAV 序列是 AAV 反向末端重複序列 (ITR)，通常位於載體基因體的 5' 和 3' 末端，以便允許將位於 ITR 之間的基因和調控序列包裝在 AAV 衣殼內。

【0075】 AAV 衣殼的來源可為數十種天然存在和可用的腺相關病毒以及工程化的 AAV 中的任何一種。在一具體實施例中，用於基因編輯載體及/或供體載體之 AAV 衣殼的來源是相同的。在另一具體實施例中，用於基因編輯載體及/或供體載體之 AAV 衣殼的來源是不同的。腺相關病毒 (AAV) 病毒載體為一種具有 AAV 蛋白衣殼的 AAV DNase 抗性顆粒，其中包裝了用於遞送至標靶細胞的核酸序列。AAV 衣殼由 60 個衣殼 (cap) 蛋白次單元 VP1、VP2 和 VP3 組成，它們以大約 1:1:10 至 1:1:20 的比例以二十面體對稱排列，具體取決於所選擇的 AAV。可選擇各種 AAV 作為上述 AAV 病毒載體衣殼的來源。參見，例如，美國公開專利申請號 2007-0036760-A1；美國公開專利申請號 2009-0197338-A1；EP 1310571。亦參見，WO 2003/042397 (AAV7 及其它猴 AAV)、美國專利 7790449 及美國專利 7282199 (AAV8)、WO 2005/033321 及 US 7,906,111 (AAV9)、及 WO 2006/110689、WO 2003/042397 (rh.10) 及 WO 2018/160582 (AAVhu68)。這些文件亦描述了可選擇用於生成 AAV 的其它 AAV，且這些文件藉由引用併入。

【0076】除非另有指出，本文所述 AAV 衣殼、ITR 及其它選定的 AAV 組分可以容易地選自任何 AAV，包括，不限於，通常被識別為 AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV8bp、AAV7M8、AAVAnc80、AAVrh10、AAVrh79 及 AAVPHP.B 的 AAV，及任何已知或所述之 AAV 的變異體或尚待發現的 AAV 或其變異體或混合物。參見，例如，WO 2005/033321，其藉由引用併入本文。在一具體實施例中，AAV 衣殼為 AAV1 衣殼或其變異體、AAV8 衣殼或其變異體、AAV9 衣殼或其變異體、AAVhu.68 衣殼或其變異體、AAVrh.10 衣殼或其變異體、AAVrh64R1 衣殼或其變異體、AAVhu.37 衣殼或其變異體、或 AAV3B 或其變異體。於一態樣，衣殼為 AAVhu.37 衣殼。亦參見，WO 2019/168961 及 WO 2019/169004，其藉由引用而整體併入本文中。在其它具體實施例中，AAV 衣殼為 AAVrh79 衣殼或其變異體。在其它具體實施例中，AAV 衣殼為 AAVrh.90 或其變異體。

【0077】在某些具體實施例中，rAAV 包含 AAVhu37 衣殼。AAVhu37 衣殼包含：vp1 蛋白的異質群體，其為編碼 SEQ ID NO: 38 之胺基酸序列的核酸序列的產物、vp2 蛋白的異質群體，其為編碼 SEQ ID NO: 38 的至少約胺基酸 138 至 738 之胺基酸序列的核酸序列的產物、及 vp3 蛋白的異質群體，其為編碼 SEQ ID NO: 38 的至少胺基酸 204 至 738 的核酸序列的產物，其中：vp1、vp2 及 vp3 蛋白含有具有胺基酸修飾的亞群，該胺基酸

修飾包含 SEQ ID NO: 38 中天冬醯胺-甘胺酸對中的至少兩個高度脫醯胺化的天冬醯胺(N)，且可選擇地進一步包含含有其它脫醯胺化胺基酸的亞群，其中脫醯胺導致胺基酸變化。AAVhu37 的特徵在於具有高度脫醯胺化殘基，例如，在基於 AAVhu37 VP1 (SEQ ID NO: 38) 編號之位置 N57、N263、N385、及/或 N514 處。

【0078】如下表所示，以及在例如 2019 年 9 月 6 日公開的 WO 2019/168961 中，已在其它殘基中觀察到脫醯胺作用，該文獻藉由引用併入本文。在某些具體實施例中，AAVhu37 衣殼在下述提供之範圍內的一個或多個以下位置被修飾，如使用具有胰蛋白酶的質譜法測定的。在某些具體實施例中，一個或多個以下位置，或 N 之後的甘胺酸如本文所述被修飾。例如，在某些具體實施例中，G 可被修飾為 S 或 A，例如，在位置 58、264、386、或 515 處。在一具體實施例中，AAVhu37 衣殼在位置 N57/G58 至 N57Q 或 G58A 處被修飾，以提供這此位置處具有減少的脫醯胺的衣殼。在另一具體實施例中，N57/G58 被修改成 NS57/58 或 NA57/58。然而，在某些具體實施例中，當 NG 變為 NS 或 NA 時，觀察到脫醯胺作用增加。在某些具體實施例中，NG 對中的 N 被修飾成 Q，同時保留 G。在某些具體實施例中，NG 對的二個胺基酸皆被修飾。在某些具體實施例中，N385Q 導致在該位置的脫醯胺作用顯著減少。在某些具體實施例中，N499Q 導致在該位置的脫醯胺作用顯著增加。

【0079】在某些具體實施例中，AAVhu37 可使這些或其它殘基脫醯胺化，例如，通常低於 10% 及/或可以具有其它修飾，包括甲基化(例如，~R487)(通常小於 5%，在給定的殘基處更通常小於 1%)、異構化(例如，在 D97 處)(通常小於 5%，在給定的殘基處更通常小於 1%)、磷酸化(例如，當存在時，在約 10 至約 60%、或約 10 至約 30%、或約 20 至約 60%的範圍內)(例如，在 S149、~S153、~S474、~T570、~S665 之一處或多處)、或氧化(例如，在 W248、W307、W307、M405、M437、M473、W480、W480、W505、M526、M544、M561、W621、M637、及/或 W697 之一處或多處)。可選擇地，W 可氧化成犬尿胺酸(kynurenine)。

表 A	
基於 vp1 編號的 AAVhu37 脫醯胺作用	脫醯胺作用%
N57+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
N94+脫醯胺作用	5-15、約 10
~N254+脫醯胺作用	10-20
~N263+脫醯胺作用	75-100
~N305+脫醯胺作用	1-5
~N385+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N410+脫醯胺作用	1-25、
N479+脫醯胺作用	1-5、1-3
~N514+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~Q601+脫醯胺作用	0-1
N653+脫醯胺作用	0-2

【0080】還有其它位置可具有這些或其它修飾(例如，乙醯化或進一步脫醯胺)。在某些具體實施例中，編碼 AAVhu37 vp1 衣殼蛋白的核酸序列提供於 SEQ ID NO: 37 中。在其它具體實施例中，可選擇與 SEQ ID NO: 37 具有 70%至 99.9%同一性的核酸序列來表現 AAVhu37 衣殼蛋白。在某些其它具體實施例中，核酸序列是與 SEQ ID NO: 37 至少約 75%相同、至少 80%相同、至少 85%、至少 90%、至少 95%、至少 97%相同、或至少 99%相同。然而，可選擇編碼 SEQ ID NO: 38 胺基酸序列的其它核酸序列用於製造 rAAVhu37 衣殼。在某些具體實施例中，核酸序列具有 SEQ ID NO: 37 核酸序列，或與編碼 SEQ ID NO: 38 之 SEQ ID NO: 37 有至少 70%至至少 99%相同、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。在某些具體實施例中，核酸序列具有 SEQ ID NO: 37 核酸序列、或與編碼 SEQ ID NO: 38 之 vp2 衣殼蛋白(約 aa 138 至 738)的 SEQ ID NO: 37 之約 nt 412 至約 nt 2214 有至少 70%至 99%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。在某些具體實施例中，核酸序列具有 SEQ ID NO: 37 之約 nt 610 至約 nt 2214 的核酸序列、或與編碼 SEQ ID NO: 38 之 vp3 衣殼蛋白(約 aa 204 至 738)的 SEQ ID NO: 37nt 至少 70%至 99%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。參見，EP 2 345 731 B1 及其中之 SEQ ID NO: 88，其藉由引用而併入。

【0081】在某些具體實施例中，rAAV 包含 AAV8 衣殼。AAV8 衣殼包含：如下表中定義之脫醯胺的 VP 同功型的異質群體，基於使用質譜法測定的衣殼中 VP 蛋白的總量。合適的修飾包括在上文標記為脫醯胺調節的段落中描述的那些，其併入本文。在某些具體實施例中，如使用質譜法所測定，AAV 衣殼在下述提供之範圍內的一個或多個以下位置處被修飾。在某些具體實施例中，一個或多個以下位置，或 N 之後的甘胺酸被如本文所述地修飾。在某些具體實施例中，人工 NG 被導入至與下述確定之位置之一不同的位置中。在某些具體實施例中，一個或多個以下位置，或 N 之後的甘胺酸被如本文所述地修飾。例如，在某些具體實施例中，G 可修飾成 S 或 A，例如，在位置 58、67、95、216、264、386、411、460、500、515、或 541 處。當 NG57/58 變為 NS57/58 或 NA57/58 時，觀察到脫醯胺顯著減少。然而，在某些具體實施例中，當 NG 變為 NS 或 NA 時，觀察到脫醯胺作用增加。在某些具體實施例中，在保留 G 的同時，將 NG 對中的 N 修飾為 Q。在某些具體實施例中，NG 對的二個胺基酸皆被修飾。在某些具體實施例中，N385Q 導致在該位置的脫醯胺作用顯著減少。在某些具體實施例中，N499Q 導致在該位置的脫醯胺作用顯著增加。在某些具體實施例中，NG 突變發生在位於 N263 的對上(例如，至 N263A)。在某些具體實施例中，NG 突變發生在位於 N514 的對上(例如，至 N514A)。在某些具體實施例中，NG 突變發生在位於 N540 的對上(例如，

N540A)。在某些具體實施例中，將包含多個突變和這些位置處的至少一個突變的 AAV 突變體工程化。在某些具體實施例中，在位置 N57 處沒有突變。在某些具體實施例中，在位置 N94 處沒有突變。在某些具體實施例中，在位置 N305 處沒有突變。在某些具體實施例中，在位置 G386 處沒有突變。在某些具體實施例中，在位置 Q467 處沒有突變。在某些具體實施例中，在位置 N479 處沒有突變。在某些具體實施例中，在位置 N653 處沒有突變。在某些具體實施例中，修飾衣殼使其在「NG」對之外的位置處減少「N」或「Q」。殘基編號是基於公佈的 AAV8 序列，在 SEQ ID NO: 36 中再現。

表 B	
基於 VP1 編號的 AAV8 修飾	%
N35+脫醯胺作用	1
N57+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
N66+脫醯胺作用	0-10
N94+脫醯胺作用	1-15
N113+脫醯胺作用	0-10
~Q166+脫醯胺作用	0-10
~N173+脫醯胺作用	0-10
N254/N255+脫醯胺作用	5-45
N263+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N304+脫醯胺作用	0-10
~N305+脫醯胺作用	10-40
N320+脫醯胺作用	0-10
~Q322+脫醯胺作用	0-10
N385+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100

表 B	
基於 VP1 編號的 AAV8 修飾	%
N410+脫醯胺作用	15-70
~Q431+脫醯胺作用	0-10
N438+脫醯胺作用	0-10
~N459+脫醯胺作用	0-10
~Q467+脫醯胺作用	0-10
~N479+脫醯胺作用	0-10
N498/N499+脫醯胺作用	0-10
N502+脫醯胺作用	0-10
N514+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
N517+脫醯胺作用	15-40
N540+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N554+脫醯胺作用	0-10
~Q589+脫醯胺作用	0-10
~N590+脫醯胺作用	0-10
~N599+脫醯胺作用	35-75
~Q601+脫醯胺作用	45-75
~Q610+脫醯胺作用	0-10
Q617+脫醯胺作用	0-10
N630+脫醯胺作用	5-30
Q648+脫醯胺作用	0-10
N653+脫醯胺作用	0-10
N665+脫醯胺作用	5-30
N670+脫醯胺作用	0-10
N693+脫醯胺作用	0-10
~N706+脫醯胺作用	0-10
N718+脫醯胺作用	0-10
N737+脫醯胺作用	0-10

【0082】在某些具體實施例中，rAAV 包含 AAVrh79 衣殼，如敘述於 WO 2019/169004 (2019 年 9 月 6 日公開)，其藉由引用併入本文。在一具體實施例中，AAVrh79 衣殼包含 AAVrh79 VP1 蛋白、AAVrh79 VP2 蛋白、及 AAVrh79 vp3 蛋白之異質群體。在一具體實施例中，藉由從編碼 SEQ ID NO: 34 之 1 至 738 的預測胺基酸序列的核酸序列表現產生 AAVrh79 衣殼。可選擇地，序列共表現來自不包括 vp1-獨特區域(約 aa 1 至 137)或 vp2-獨特區域(約 aa 1 至 203)之核酸序列的 vp3 蛋白、由 SEQ ID NO: 33 產生的 vp1 蛋白、或由與編碼 SEQ ID NO: 34 的 1 至 738 的預測胺基酸序列之 SEQ ID NO: 33 至少 70%相同的核酸序列產生的 vp1 蛋白。在其它具體實施例中，藉由從編碼 SEQ ID NO: 34 的至少約胺基酸 138 至 738 的預測胺基酸序列的核酸序列表現產生的 AAVrh79 vp2 蛋白、由包含 SEQ ID NO: 33 之至少核苷酸 412 至 2214 的序列產生的 vp2 蛋白、或從與編碼 SEQ ID NO: 34 之至少約胺基酸 138 至 738 的預測胺基酸序列之 SEQ ID NO: 33 之至少核苷酸 412 至 2214 有至少 70%相同的核酸序列產生的 vp2 蛋白、藉由從編碼 SEQ ID NO: 34 的至少約胺基酸 204 至 738 的預測胺基酸序列的核酸序列表現產生的 AAVrh79 vp3 蛋白、自包含 SEQ ID NO: 33 之至少核苷酸 610 至 2214 的序列產生的 vp3 蛋白、或從與編碼 SEQ ID NO: 34 之至少約胺基酸 204 至 738 的預測胺基酸序列之 SEQ ID NO: 33 之至少核苷酸 610 至 2214 有至少 70%相同的核酸序列產生的 vp3 蛋白。

【 0083 】 在某些具體實施例中，AAVrh79 衣殼包含：vp1 蛋白之異質群體，該 vp1 蛋白為編碼 SEQ ID NO: 34 之胺基酸序列的核酸序列的產物；vp2 蛋白之異質群體，該 vp2 蛋白為編碼 SEQ ID NO: 34 之至少約胺基酸 138 至 738 的胺基酸序列的核酸序列的產物；及 vp3 蛋白之異質群體，該 vp3 蛋白為編碼 SEQ ID NO: 34 之至少胺基酸 204 至 738 的核酸序列的產物。

【 0084 】 AAVrh79 vp1、vp2 及 vp3 蛋白含有具有胺基酸修飾的亞群，該胺基酸修飾包含在 SEQ ID NO: 34 中天冬醯胺酸-甘胺酸對中的至少二個高度脫醯胺化天冬醯胺酸(N)，且可選擇地進一步包含含有其它脫醯胺化胺基酸的亞群，其中該脫醯胺導致胺基酸變化。相對於 SEQ ID NO: 34 的編號，在 N-G 對 N57、N263、N385 及/或 N514 處觀察到高程度的脫醯胺作用。在其它殘基中已觀察到脫醯胺作用，如下表和實施例所示。在某些具體實施例中，AAVrh79 可具有其它脫醯胺化的殘基，例如，通常低於 10% 及/或可具有其它修飾，包括甲基化(例如，~R487)(通常小於 5%，在給定的殘基處更通常小於 1%)、異構化(例如，在 D97 處)(通常小於 5%，在給定的殘基處更通常小於 1%)、磷酸化(例如，當存在時，在約 10 至約 60%、或約 10 至約 30%、或約 20 至約 60% 的範圍內)(例如，在 S149、~S153、~S474、~T570、~S665 之一處或多處)、或氧化(例如，在 W248、W307、W307、M405、M437、M473、W480、W480、W505、M526、M544、M561、W621、M637 及/

或 W697 之一處或多處)。可選擇地，W 可氧化成犬尿胺酸。

表 C	
基於 VP1 編號的 AAVrh79 脫醯胺作用	脫醯胺作用%
N57+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
N94+脫醯胺作用	5-15、約 10
~N254+脫醯胺作用	10-20
~N263+脫醯胺作用	75-100
~N305+脫醯胺作用	1-5
~N385+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N410+脫醯胺作用	1-25、
N479+脫醯胺作用	1-5、1-3
~N514+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~Q601+脫醯胺作用	0-1
N653+脫醯胺作用	0-2

【0085】在某些具體實施例中，AAVrh79 衣殼在上表中確定的一個或多個位置被修飾，在下述提供的範圍內，如使用胰蛋白酶的質譜法測定。在某些具體實施例中，一個或多個以下位置、或 N 之後的甘胺酸如本文所述被修飾。殘基編號是基於本文提供的 AAVrh79 序列。參見，SEQ ID NO：34。

【0086】在某些具體實施例中，編碼 AAVrh79 vp1 衣殼蛋白的核酸序列提供於 SEQ ID NO: 33 中。在其它具體實施例中，可選擇與 SEQ ID NO: 33 具有 70%至 99.9%同一性的核酸序列來表現 AAVrh79 衣殼蛋白。在某些其它具體實施例中，核酸序列與 SEQ ID NO: 33 至少約 75%相同、至少 80%相同、至少 85%、至少 90%、

至少 95%、至少 97%相同、至少 99%或至少 99.9%相同。然而，可選擇編碼 SEQ ID NO: 34 之胺基酸序列的其它核酸序列來用於生產 rAAV 衣殼。在某些具體實施例中，核酸序列具有 SEQ ID NO: 33 之核酸序列、或與編碼 SEQ ID NO: 34 之 SEQ ID NO: 33 至少 70%至 99%相同、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。在某些具體實施例中，核酸序列具有 SEQ ID NO: 33 之核酸序列、或與編碼 SEQ ID NO: 34 之 vp2 衣殼蛋白(約 aa 138 至 738)的 SEQ ID NO: 33 之約 nt 412 至約 nt 2214 至少 70%至 99%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。在某些具體實施例中，核酸序列具有 SEQ ID NO: 33 之約 nt 610 至約 nt 2214 的核酸序列、或與編碼 SEQ ID NO: 34 之 vp3 衣殼蛋白(約 aa 204 至 738)的 SEQ ID NO: 33nt 至少 70%至 99%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 97%、至少 99%相同之序列。

**【0087】** 本發明亦涵蓋編碼突變體 AAVrh79 的核酸序列，其中一個或多個殘基已被改變以減少脫醯胺作用或本文確認的其它修飾。此類核酸序列可用於製造突變體 rAAVrh79 衣殼。

**【0088】** 在某些具體實施例中，rAAV 包含 AAVrh.90 衣殼，如敘述於 WO 2020/223232 (2020 年 11 月 5 日公開)，其藉由引用併入本文。於另一態樣，提供重組腺相關病毒(rAAV)，其包含：(A) AAVrh.90 衣殼，包含一

種或多種：(1) AAVrh.90 衣殼蛋白，包含：選自下述之 AAVrh.90 vp1 蛋白的異質群體：藉由從編碼 SEQ ID NO: 40 之 1 至 738 的預測胺基酸序列的核酸序列表現所產生的 vp1 蛋白、由 SEQ ID NO: 39 所產生的 vp1 蛋白、或由與編碼 SEQ ID NO: 40 之 1 至 738 的預測胺基酸序列的 SEQ ID NO: 39 至少 70% 相同的核酸序列所產生的 vp1 蛋白；選自下述之 AAVrh.90 vp2 蛋白的異質群體：藉由從編碼 SEQ ID NO: 40 之至少約胺基酸 138 至 738 的預測胺基酸序列的核酸序列表現所產生的 vp2 蛋白、由包含 SEQ ID NO: 39 之至少核苷酸 412 至 2214 的序列所產生的 vp2 蛋白、或由與編碼 SEQ ID NO: 40 之至少約胺基酸 138 至 738 的預測胺基酸序列之 SEQ ID NO: 39 之至少核苷酸 412 至 2214 至少 70% 相同的核酸序列所產生的 vp2 蛋白；選自下述之 AAVrh.90 vp3 蛋白的異質群體：藉由從編碼 SEQ ID NO: 40 之至少約胺基酸 204 至 738 的預測胺基酸序列的核酸序列表現所產生的 vp3 蛋白、由包含 SEQ ID NO: 39 之至少核苷酸 610 至 2214 的序列所產生的 vp3 蛋白、或由與編碼 SEQ ID NO: 40 之至少約胺基酸 204 至 738 的預測胺基酸序列之 SEQ ID NO: 39 之至少核苷酸 610 至 2214 至少 70% 相同的核酸序列所產生的 vp3 蛋白；及/或(2) 為編碼 SEQ ID NO: 40 之胺基酸序列的核酸序列的產物之 vp1 蛋白之異質群體、為編碼 SEQ ID NO: 40 之至少約胺基酸 138 至 738 的胺基酸序列的核酸序列之產物之 vp2 蛋白之異質群體、及為編碼 SEQ ID NO: 40 之至少胺基酸 204 至 738

的核酸序列的產物之 vp3 蛋白異質群體，其中：該 vp1、vp2 及 vp3 蛋白含有具有胺基酸修飾的亞群，該胺基酸修飾包括 SEQ ID NO: 40 中天冬醯胺酸-甘胺酸對中的至少兩個高度脫醯胺化的天冬醯胺酸(N)，且可選擇地進一步包含含有其它脫醯胺化胺基酸的亞群，其中該脫醯胺導致胺基酸變化；及(B) AAVrh.90 衣殼中的載體基因體，該載體基因體包含含有 AAV 反向末端重複序列之核酸分子、及編碼產物的非 AAV 核酸序列，該非 AAV 核酸序列可操作地連接至指導該產物在宿主細胞中表現的序列。

【0089】在某些具體實施例中，AAVrh.90 vp1、vp2 及 vp3 蛋白含有具有胺基酸修飾的亞群，該胺基酸修飾包含在 SEQ ID NO: 40 中的天冬醯胺-甘胺酸對中的至少兩個高度脫醯胺化的天冬醯胺(N)，且可選擇地進一步包含含有其它脫醯胺化胺基酸的亞群，其中脫醯胺導致胺基酸變化。相對於 SEQ ID NO: 40 的編號，在 N-G 對 N57、~N263、~N385 及/或~N514 處觀察到高程度的脫醯胺作用。在其它殘基中觀察到脫醯胺作用，如下表所示。在某些具體實施例中，AAVrh.90 可具有其它脫醯胺化殘基(例如，~N305、~N499、及/或~N599，通常低於 20%)及/或可具有其它修飾，包括磷酸化(例如，當存在時，在約 2 至約 30%、或約 2 至約 20%、或約 2 至約 10%的範圍內)(例如，在 S149 處)、或氧化(例如，在 ~W23、~M204、~M212、W248、W282、M405、M473、W480、W505、M526、~N544、M561 及/或

~M607 的一處或多處)。可選擇地，W 可氧化成犬尿酸酸。

表 D	
基於 VP1 編號的 AAVrh.90 脫醯胺作用	脫醯胺作用%
N57+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
N94+脫醯胺作用	2-15 或 2-5
~N263+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N305+脫醯胺作用	5-30、5-20、或 10-20
~N385+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N499+脫醯胺作用	2-15、2-10、或 5-10
~N514+脫醯胺作用	65-90、70-95、80-95、75-100、80-100、或 90-100
~N599+脫醯胺作用	2-15、2-10、或 5-10

【0090】在某些具體實施例中，AAVrh.90 衣殼在上表中確定的一個或多個位置被修飾，在所提供的範圍內，如使用具有胰蛋白酶的質譜法測定。在某些具體實施例中，一個或多個位置、或 N 之後的甘氨酸如本文所述被修飾。殘基編號是基於本文提供的 AAVrh.90 序列。參見，SEQ ID NO: 40。

【0091】在某些具體實施例中，AAVrh.90 衣殼包含：vp1 蛋白之異質群體，該 vp1 蛋白為編碼 SEQ ID NO: 40 之胺基酸序列的核酸序列的產物；vp2 蛋白之異質群體，該 vp2 蛋白為編碼 SEQ ID NO: 40 之至少約胺基酸 138 to 738 的胺基酸序列之核酸序列的產物；及 vp3 蛋白之異質群體，該 vp3 蛋白為編碼 SEQ ID NO: 40 之至少胺基酸 204 至 738 的核酸序列的產物。

【0092】在某些具體實施例中，選擇微小病毒載體衣殼用於肝向性(liver-tropism)，且接受治療的病患患有肝臟代謝性病變。在某些具體實施例中，選擇微小病毒載體衣殼用於心向性，且接受治療的病患患有心臟病變。在某些具體實施例中，選擇微小病毒載體衣殼用於趨向於骨骼肌中的細胞，且接受治療的病患患有肌肉病變。

【0093】如本文所使用，「載體基因體」係指包裝在形成病毒顆粒的 rAAV 衣殼內部的核酸序列。此核酸序列含有 AAV 反向末端重複序列(ITR)。在本文實施例中，載體基因體最低限度地由 5'至 3'含有 AAV 5' ITR、含有可操作地連接至指導其表現之調控序列的轉基因或編碼序列的表現匣、及 AAV 3' ITR。ITR 是在載體生產過程中負責基因體複製和包裝的遺傳元件，並且是產生 rAAV 所需的唯一病毒順式(cis)元件。在一具體實施例中，ITR 來自不同於提供衣殼的 AAV。在較佳的具體實施例中，可方便使用來自 AAV2 的 ITR 序列，或其刪除版本( $\Delta$ ITR)。然而，可選擇來自其它 AAV 來源的 ITR。在 ITR 的來源來自 AAV2 且 AAV 衣殼來自另一 AAV 來源的情況下，產生的載體可稱為假型的(pseudotyped)。通常，AAV 載體基因體包含 AAV 5' ITR、編碼基因產物及任何調控序列的核酸序列、及 AAV 3' ITR。然而，這些元件的其它配置可為合適的。在一具體實施例中，提供自我互補 AAV。已描述了 5' ITR 的縮短版本，稱為  $\Delta$ ITR，其中刪除了 D-序列和末端解析位點(trs)。在某

些具體實施例中，載體基因體包括 130 個鹼基對之縮短的 AAV2 ITR，其中刪除了外部「a」元件。在使用內部 A 元件作為模板進行載體 DNA 擴增期間，縮短的 ITR 恢復為 145 個鹼基對的野生型長度。在其它具體實施例中，使用全長 AAV 5'和 3' ITR。在其它具體實施例中，可選擇全長或工程化 ITR。可選擇來自 AAV2 的 ITR、與衣殼不同的來源 AAV 或非全長 ITR。ITR 來自與生產期間提供 rep 功能的 AAV 或反向互補 AAV 相同的 AAV 來源。再者，可使用其它 ITR。合適的 ITR 序列之實例顯示於序列表中，例如，SEQ ID NO: 42，nt 1 至 130 及 3052 至 3181。再者，載體基因體含有直接調節基因產物表現的調控序列(例如，直接或間接藉由調節轉錄及/或轉譯)。載體基因體的合適組分在本文中更詳細地討論。

【0094】在某些具體實施例中，基因編輯載體基因體包括 TBG 啟動子、一種或多種  $\alpha$  mic/bik 增強子、用於 ARCUS 巨型核酸酶之編碼序列、可選擇地 WPRE、及 polyA。在某些具體實施例中，表現匣包括 SEQ ID NO: 42 之 nt 211 至 nt 2964，兩側是 5'和 3' ITR。

【0095】為了用於生產 AAV 病毒載體(例如重組(r)AAV)，表現匣可被攜帶於任何合適的載體上，例如質體，其被遞送至包裝宿主細胞。本發明的質體可被工程化，使其適用於在原核細胞、昆蟲細胞、哺乳動物細胞等之中體外地複製及包裝。合適的轉染技術及包裝宿主細胞是已知的及/或可由本領域技術人員容易地設計。

【0096】用於產生及分離適於用作載體的 AAV 的方法是本領域已知的。通常參見，例如，Grieger & Samulski, 2005, “Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications,” *Adv. Biochem. Engin/Biotechnol.* 99: 119-145; Buning *et al.*, 2008, “Recent developments in adeno-associated virus vector technology,” *J. Gene Med.* 10:717-733; 及以下引用的參考文獻，每一篇都藉由引用整體併入本文。為了將轉基因包裝至病毒體中，ITR 是與包含表現匣的核酸分子在同一構建體中順式所需的唯一 AAV 成分。cap 及 rep 基因可以反式提供。

【0097】術語「AAV 中間體」或「AAV 載體中間體」係指組裝的 rAAV 衣殼，其缺少包裝在其中的所需基因體序列。此類亦可稱為「空」衣殼。此類衣殼可不包含表現匣的可檢測基因體序列，或僅包含不足以達到基因產物之表現的部分包裝之基因體序列。這些空衣殼無法將感興趣的基因轉移至宿主細胞中。

【0098】本文所述之重組 AAV 可使用已知技術產生。參見，例如，WO2003/042397; WO2005/033321; WO2006/110689; US7588772 B2。此類方法涉及培養含有編碼 AAV 衣殼之核酸序列的宿主細胞；功能性 rep 基因；至少由 AAV 反向末端重複序列 (ITR) 和轉基因組成的表現匣；和足夠的輔助功能以允許將表現匣包裝至 AAV 衣殼蛋白中。產生衣殼的方法、編碼序列以及產生 rAAV 病毒載體的方法已被描述。參見，例如，Gao, et

al, Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)及 US2013/0045186A1。

【0099】在一具體實施例中，提供可用於生產重組 AAV 的生產細胞培養物。此類細胞培養物含有在宿主細胞中表現 AAV 衣殼蛋白的核酸；適於包裝至 AAV 衣殼中的核酸分子，例如，包含 AAV ITR 和編碼基因產物的非 AAV 核酸序列的載體基因體，該非 AAV 核酸序列可操作地連接至指導產物在宿主細胞中表現的序列；及足以允許將核酸分子包裝至重組 AAV 衣殼中的 AAV rep 功能和腺病毒輔助功能。在一具體實施例中，細胞培養物由哺乳動物細胞(例如人類胚腎 293 細胞等)或昆蟲細胞(例如桿狀病毒)組成。

【0100】可選擇地，rep 功能由除了提供衣殼的 AAV 之外的 AAV 提供。例如，rep 可為，但不限於 AAV1 rep 蛋白、AAV2 rep 蛋白、AAV3 rep 蛋白、AAV4 rep 蛋白、AAV5 rep 蛋白、AAV6 rep 蛋白、AAV7 rep 蛋白、AAV8 rep 蛋白；或 rep 78、rep 68、rep 52、rep 40、rep 68/78 及 rep 40/52；或其片段；或另一種來源。可選擇地，rep 及 cap 序列在細胞培養物中位於相同的遺傳元件上。rep 序列和 cap 基因之間可有間隔子。這些 AAV 或突變 AAV 衣殼序列中的任何一個都可在指導其在宿主細胞中表現的外源調節控制序列的控制下。

【0101】在一具體實施例中，在合適的細胞培養物(例如，HEK 293 細胞)細胞中製造細胞。用於製造本文所述之基因治療載體的方法包括本技術領域熟知的方

法，諸如用於生產基因治療載體的質體 DNA 的產生、載體的產生和載體的純化。在一些具體實施例中，基因治療載體為 AAV 載體，且產生的質體為編碼 AAV 基因體及感興趣的基因的 AAV 順式質體、含 AAV rep 及 cap 基因的 AAV 反式質體、及腺病毒輔助質體。載體產生過程可包括方法步驟，例如細胞培養的開始、細胞繼代、細胞接種、以質體 DNA 轉染細胞、轉染後培養基更換為無血清培養基、及收穫含載體之細胞和培養基。收穫的含有載體的細胞和培養基在本文中稱為粗細胞收穫物。在另一系統中，基因治療載體藉由以基於桿狀病毒的載體感染而被引入昆蟲細胞中。有關這些生產系統的評論，一般參見，例如，Zhang et al., 2009, "Adenovirus-Adeno-associated virus hybrid for large-scale recombinant Adeno-associated virus production," *Human Gene Therapy* 20: 922-929，其藉由引用以其整體併入本文。以下美國專利亦描述製造和使用這些和其它 AAV 生產系統的方法，其各自內容藉由引用以其整體併入本文：5,139,941；5,741,683；6,057,152；6,204,059；6,268,213；6,491,907；6,660,514；6,951,753；7,094,604；7,172,893；7,201,898；7,229,823；及 7,439,065。

【0102】此後，粗細胞收穫物可為主題方法步驟，諸如載體收穫物的濃縮、載體收穫物的滲濾、載體收穫物的微流體化、載體收穫物的核酸酶消化、經過微流化的中間體的過濾、藉由層析的粗純化、藉由超速離心法

的粗純化、藉由切向流過濾進行緩衝液交換及/或調配和過濾以製備大量載體。

【0103】在高鹽濃度下進行兩步驟親和性層析純化，然後使用陰離子交換樹脂層析來純化載體藥物產物並去除空衣殼。此等方法更詳盡的敘述於國際專利公開號 WO 2017/160360，其藉由引用併入本文。關於 AAV8 的純化方法，國際專利公開號 WO 2017/100676，及關於 rh10，國際專利公開號 WO 2017/100704，及關於 AAV1，國際專利公開號 WO 2017/100674，其等藉由引用全部併入本文。

【0104】為了計算空顆粒和完整顆粒的含量，將所選樣品(例如，在本文的實例中經過碘克沙醇(iodixanol)梯度純化的製劑，其中 GC# = 顆粒#)的 vp3 帶體積相對於加載的 GC 顆粒進行作圖。所得線性等式( $y = mx + c$ )用於計算測試品峰值的帶狀體積中的顆粒的數量。然後將加載的每 20  $\mu\text{L}$  顆粒數量(pt)乘以 50，以得到顆粒(pt)/mL。將 Pt/mL 除以 GC/mL 得到顆粒與基因體拷貝的比率(pt/GC)。Pt/mL-GC/mL 得到空 pt/mL。空 pt/mL 除以 pt/mL 並且 $\times 100$ 得到空顆粒的百分比。

【0105】一般而言，用於測定具有包裝的基因體的空衣殼和 AAV 載體顆粒的方法是本技術領域已知的。參見，例如 Grimm et al., Gene Therapy (1999) 6: 1322-1330; Sommer et al., Molec. Ther. (2003) 7: 122-128。為了測試變性的衣殼，該方法包含使經過處理的 AAV 儲料經受 SDS-聚丙烯醯胺凝膠電泳(由能夠分離三種衣殼蛋

白的任何凝膠組成，例如在緩衝液中含有 3-8%三乙酸鹽的梯度凝膠)，然後運行凝膠直到分離出樣品材料，並且將凝膠印漬到尼龍或硝酸纖維素膜(較佳為尼龍)上。然後，將抗 AAV 衣殼抗體用作與變性的衣殼蛋白結合的初級抗體，較佳為抗 AAV 衣殼單株抗體，最佳為 B1 抗 AAV2 單株抗體(Wobus et al., J. Virol.(2000) 74: 9281-9293)。然後使用次級抗體，該次級抗體與初級抗體結合並含有一種用於檢測與初級抗體的結合的裝置，更佳為含有與其共價結合的檢測分子的抗 IgG 抗體，最佳為與辣根過氧化物酶共價連接的綿羊抗小鼠 IgG 抗體。用於檢測結合之方法用於半定量地確定初級抗體與次級抗體之間的結合，較佳為能夠檢測放射性同位素發射、電磁輻射或比色變化的檢測方法，最佳為化學發光檢測套組。例如，對於 SDS-PAGE，可從管柱濾分中提取樣品並在含有還原劑(例如，DTT)的 SDS-PAGE 上樣緩衝液中加熱，並且在預製的梯度聚丙烯醯胺凝膠(例如，Novex)上解析衣殼蛋白。可根據製造商的說明使用 SilverXpress (Invitrogen, CA)或其它合適的染色方法(即 SYPRO 紅寶石色或考馬斯染色)進行銀染色。在一具體實施例中，可藉由定量即時 PCR (Q-PCR)測量管柱濾分中的 AAV 載體基因體(vg)的濃度。將樣品稀釋並用 DNase I (或另一種合適的核酸酶)消化以去除外源性 DNA。在核酸酶去活化後，使用引子和對引子之間的 DNA 序列具有特異性的 TaqMan™ 螢光探針進一步稀釋和擴增樣品。在 Applied Biosystems Prism 7700 序列檢

測系統上測量每種樣品達到定義的螢光水平所需的週期的數量(閾值週期, Ct)。含有與 AAV 載體中所含序列相同的序列的質體 DNA 用於在 Q-PCR 反應中產生標準曲線。從樣品獲得的週期閾值(Ct)的值用於藉由相對於質體標準曲線的 Ct 值對其進行標準化來確定載體基因體力價。亦可使用基於數位 PCR 的端點測定。

【0106】於一態樣，使用優化的 q-PCR 方法，其利用廣譜絲胺酸蛋白酶，例如蛋白酶 K (諸如可從 Qiagen 商購獲得)。更具體而言，優化的 qPCR 基因體力價分析與標準分析相似，除了在 DNase I 消化之後，將樣品以蛋白酶 K 緩衝液稀釋並以蛋白酶 K 處理，然後加熱去活化。適當地，將樣品以等量於樣品大小的蛋白酶 K 緩衝液稀釋。蛋白酶 K 緩衝液可濃縮至 2 倍或更高。通常，蛋白酶 K 處理約為 0.2 mg/mL，但可在 0.1 mg/mL 至約 1 mg/mL 之間變化。處理步驟通常在約 55°C 下進行約 15 分鐘，但可在較低溫度(例如，約 37°C 至約 50°C)下進行較長一段時間(例如，約 20 分鐘至約 30 分鐘)，或在較高溫度(例如，高至約 60°C)下進行較短一段時間(例如，約 5 至 10 分鐘)。類似地，加熱去活化通常在約 95°C 下保持約 15 分鐘，但溫度可降低(例如，約 70°C 至約 90°C)並延長時間(例如，約 20 分鐘至約 30 分鐘)。然後將樣品稀釋(例如，1000 倍)並如標準分析中所述進行 TaqMan 分析。

【 0107 】 另外或可替代地，可使用微滴數位化 PCR(ddPCR)。例如，藉由 ddPCR 確定單股及自我互補 AAV 載體基因體力價的方法已被敘述。參見，例如，M. Lock et al, *Hu Gene Therapy Methods*, *Hum. Gene Ther. Methods*. 2014 Apr;25(2):115-25. Doi: 10.1089/hgtb.2013.131. Epub 2014 Feb 14。ddPCR 方法直接測量包裹載體基因體的濃度。以 DNase I 處理樣品以消化樣品中存在的任何未包殼的 DNA，然後以蛋白酶 K 處理來破壞衣殼。然後稀釋樣品至適合測定範圍。將樣品與 ddPCR Supermix 混合，並使用靶向對 *PCSK9* 基因(M2PCSK9)特異性的巨型核酸酶的序列特異性引子組合對此同一區域雜交的螢光標記探針來完成檢測。在 Bio-Rad 液滴產生器中處理 20 微升 ddPCR 反應混合物，並將 ddPCR 反應混合物分成  $\geq 10,000$  個液滴。液滴生成後，ddPCR 反應混合物進行 PCR 擴增，並使用 Bio-Rad Droplet Reader 讀取擴增的 ddPCR 反應混合物。

【 0108 】 感染單位(IU)測定可用於確定 rAAV 載體在 RC32 細胞(表現 rep2 的 HeLa 細胞)中的生產性攝取和複製。採用與先前公開的類似的 96 孔端點格式。簡而言之，RC32 細胞將被 rAAV BDS 的系列稀釋和 Ad5 的均勻稀釋共同感染，每個 rAAV 稀釋 12 次重複。感染後 72 小時，細胞將被裂解，並進行 qPCR 以檢測 rAAV 載體在輸入上的擴增。將進行終點稀釋 50%組織培養感染劑量(TCID<sub>50</sub>)計算(Spearman-Karber)以確定感染力價，以 IU/mL 表示。由於「感染性」值取決於每個顆粒與細胞

的接觸、受體結合、內化、轉運到細胞核和基因體複製，因此它們受測定幾何形狀以及所用細胞株中存在的適當受體和結合後途徑的影響。受體和結合後途徑通常不保留在永生化細胞株中，因此感染性測定力價不是存在的「感染性」顆粒數量的絕對量度。然而，包裹的 GC 與「感染單位」的比率(描述為 GC/IU 比率)可用作衡量批次間產品一致性的指標。

【0109】簡而言之，將具有包裝基因體序列的 rAAV 顆粒與基因體缺陷型 AAV 中間體分離的方法包括對包含重組 AAV 病毒顆粒和 AAV 衣殼中間體的懸浮液進行快速高效液相層析，其中 AAV 病毒顆粒和 AAV 中間體與在高 pH 平衡的強陰離子交換樹脂結合，AAV 病毒顆粒和 AAV 中間體與在高 pH 平衡的強陰離子交換樹脂結合，並經受鹽梯度，同時監測洗提液在約 260 和約 280 處的紫外線吸光度。可取決於選定的 AAV 調整 pH 值。參見，例如，WO2017/160360 (AAV9)、WO2017/100704 (AAVrh10)、WO 2017/100676 (例如，AAV8)、及 WO 2017/100674 (AAV1)]，其藉由引用而併入本文。在此方法中，當 A260/A280 的比值達到轉折點時，從洗提的濾分中收集 AAV 全衣殼。在一個實例中，對於親和層析步驟，滲濾後的產物可應用於有效捕獲 AAV2 血清型的 Capture Select™ Poros-AVV2/9 親和樹脂 (Life Technologies)。在這些離子條件下，顯著比例的殘留細胞 DNA 和蛋白質流過管柱，而 AAV 顆粒被有效捕獲。

## 【 0110 】

## [雙重載體系統]

在另一態樣中，提供用於治療遺傳疾病的雙重載體系統。該系統包括：(a)基因編輯組分，其包括編碼靶向 PCSK9 之核酸酶的核酸序列及可選擇地指導該核酸酶在包含 PCSK9 基因之標靶細胞中表現的調控序列；及 (b)供體載體，其包含編碼用於從 PCSK9 基因座表現的外源產物的核酸序列，其中該插入的核酸序列並不編碼 PCSK9，且其中該系統進一步包含指導該核酸酶特異性靶向天然 PCSK9 基因座的序列。該系統可選擇地包含允許標靶細胞中的天然 PCSK9 在用雙載體系統給藥後被消融或減少的組分，例如，經由使用具有誘導型啟動子的誘導劑。在一個具體實施例，基因編輯組分包含於含有表現匣之基因編輯載體中，該表現匣包含編碼核酸酶之核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中表現的調控序列。雙重載體之組分如本文所述。

【 0111 】雖然如果基因編輯組分與供體載體的比率為約 1 比約 1，則該系統可能是有效的，但希望供體模板載體的存在量超過基因編輯組分。在一具體實施例中，編輯載體 (a) 與供體載體 (b) 的比例為約 1:3 至約 1:100，或約 1:10。基因編輯酶 (例如，Cas9 或巨型核酸酶) 與供體模板的這種比例可被保持，即使該酶是由 AAV 載體以外的來源額外或替代地提供的。

【 0112 】在一具體實施例中，雙重載體系統包括：包含 AAV 衣殼及第一載體基因體的基因編輯 AAV 載

體，該第一載體基因體包含 5' ITR、編碼在調節序列控制下靶向 PCSK9 的巨型核酸酶的序列(該調控序列指導該巨型核酸酶在包含 PCSK9 基因之標靶細胞中的表現)、及 3' ITR；及(b)包含 AAV 衣殼及第二載體基因體之供體 AAV 載體，其該第二載體基因體包含：5' ITR、5'同源定向重組(HDR)臂、轉基因及指導轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、及 3' ITR。

【0113】在另一具體實施例中，雙重載體系統包括：包含 AAV 衣殼及第一載體基因體的基因編輯 AAV，該第一載體基因體包含 5' ITR、5'核定位訊號(NLS)、編碼 Cas9 之序列及指導 SaCas9 在包含 PCSK9 基因之標靶細胞中表現的調控序列、3' NLS、及 3' ITR；及包含 AAV 衣殼及第二載體基因體供體 AAV 載體，該第二載體基因體包含：5' ITR、5'同源定向重組(HDR)臂、轉基因及指導轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、U6 啟動子、包含至少 20 個核苷酸的 sgRNA (其特異性結合 PCSK9 基因中之標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處)、及 3' ITR。

【0114】在另一具體實施例中，雙重載體系統包括：包含 AAV 衣殼及第一載體基因體的基因編輯 AAV 載體，該第一載體基因體包含 5' ITR、U6 啟動子、包含至少 20 個核苷酸之 sgRNA (其特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處)、5'核定位訊號

(NLS)、編碼 Cas9 之序列及指導 Cas9 在包含 PCSK9 基因之標靶細胞中表現的調控序列、3' NLS、及 3' ITR；及包含 AAV 衣殼及第二載體基因體之供體 AAV 載體，該第二載體基因體包含：5' ITR、5'同源定向重組(HDR)臂、轉基因及指導轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、及 3' ITR。

【0115】在本文所述之系統的某些具體實施例中，基因編輯 AAV 載體及供體 AAV 載體的基因具有相同的 AAV 衣殼。在其它具體實施例中，編輯 AAV 載體及供體 AAV 載體的基因具有不同 AAV 衣殼。在一些具體實施例中，AAV 衣殼選自 AAV8、AAV9、rh10、AAV6.2、AAV3B、hu37、rh79 及 rh64。

【0116】在某些具體實施例中，核酸酶為 Cas9 核酸酶，且該 Cas9 選自金黃色葡萄球菌 Cas9 或釀膿鏈球菌 Cas9。

【0117】在某些具體實施例中，核酸酶及/或轉基因受組織特異性啟動子的控制。在某些具體實施例中，核酸酶及/或轉基因受組成型啟動子的控制。在某些具體實施例中，核酸酶及/或轉基因受誘導型啟動子的控制。在某些具體實施例中，核酸酶及/或轉基因受肝特異性啟動子、可選擇地人類甲狀腺素結合球蛋白(TBG)啟動子、或雜交肝啟動子(HLP)的控制。在某些具體實施例中，系統進一步包含誘導劑。

【0118】在另一具體實施例中，該系統包括：(a)基因編輯組分，其包括編碼靶向 PCSK9 之核酸酶的核酸序

列及指導該核酸酶在標靶細胞中表現的調控序列，該標靶細胞包含包封於 LNP 中的 PCSK9 基因；及 (b) 供體載體，其包含編碼用於從包封於 LNP 中之 PCSK9 基因座表現外源產物的核酸序列，其中該插入的核酸序列並不編碼 PCSK9，且其中該系統進一步包含指導該核酸酶特異性靶向天然 PCSK9 基因座的序列。該系統可選擇地包含允許標靶細胞中的天然 PCSK9 在以雙載體系統給藥後被消融或減少的組分，例如，經由使用具有誘導型啟動子的誘導劑。

【0119】在另一具體實施例中，系統包括：(a) 基因編輯組分，其包括編碼靶向 PCSK9 之核酸酶的核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中表現的調控序列，其中該基因編輯組分經由 AAV 載體提供；及 (b) 供體載體，其包含編碼用於從包封於 LNP 中之 PCSK9 基因座表現外源產物的核酸序列，其中該插入的核酸序列並不編碼 PCSK9，且其中該系統進一步包含指導核酸酶特異性地靶向天然 PCSK9 基因座的序列。該系統可選擇地包含允許標靶細胞中的天然 PCSK9 在用雙載體系統給藥後被消融或減少的組分，例如，經由使用具有誘導型啟動子的誘導劑。

【0120】在另一具體實施例中，系統包括：(a) 基因編輯組分，其包括編碼靶向 PCSK9 之核酸酶的核酸序列及指導該核酸酶在標靶細胞中表現的調控序列，該標靶細胞包含包封於 LNP 中的 PCSK9 基因；及 (b) 供體載體，其包含編碼用於從 PCSK9 基因座表現外源產物的核

酸序列，其中該供體載體為 AAV 載體，其中該插入的核酸序列並不編碼 PCSK9，且其中該系統進一步包含指導該核酸酶特異性靶向天然 PCSK9 基因座的序列。該系統可選擇地包含允許標靶細胞中的天然 PCSK9 在以雙載體系統給藥後被消融或減少的組分，例如，經由使用具有誘導型啟動子的誘導劑。

【0121】在一具體實施例中，雙重載體系統包括：(a) LNP，其包含編碼靶向 PCSK9 之巨型核酸酶的 mRNA，該巨型核酸酶在調控序列控制下指導巨型核酸酶在包含 PCSK9 基因之標靶細胞中的表現；及 (b) 供體 AAV 載體，其包含 AAV 衣殼及第二載體基因體，該第二載體基因體包含：5' ITR、5' 同源定向重組 (HDR) 臂、轉基因及指導轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、及 3' ITR。

【0122】在另一具體實施例中，雙重載體系統包括：(a) LNP，其包含包含編碼 Cas9 之序列的核酸及包含至少 20 個核苷酸的 sgRNA (其特異性結合在 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處)；及 (b) 供體 AAV 載體，其包含 AAV 衣殼及載體基因體，該載體基因體包含：5' ITR、5' 同源定向重組 (HDR) 臂、轉基因及指導轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、及 3' ITR。提供編碼 Cas9 之序列作為 mRNA。

【0123】在另一具體實施例中，雙重載體系統包括：包含 AAV 衣殼及第一載體基因體的基因編輯 AAV

載體，該第一載體基因體包含 5' ITR、U6 啟動子、包含至少 20 個核苷酸之 sgRNA (其特異性結合至 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處)、5'核定位訊號(NLS)、編碼 Cas9 之序列及指導 Cas9 在包含 PCSK9 基因之標靶細胞中表現的調控序列、3' NLS、及 3' ITR；及包含 AAV 衣殼及第二載體基因體之供體 AAV 載體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中表現的調控序列、3' HDR 臂、及 3' ITR。

#### 【 0124 】

#### 醫藥組成物

在另一態樣中，提供一種醫藥組成物，其包含：第一 rAAV 儲料(stock)，其包含 rAAV 基因編輯載體，該 rAAV 基因編輯載體包含表現匣，該表現匣包含編碼靶向 PCSK9 之核酸酶的核酸序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中表現的調控序列；及第二 rAAV 儲料，其包含 rAAV 供體載體，該 rAAV 供體載體包含轉基因匣，該轉基因匣包含編碼轉基因之核酸序列及指導轉基因在標靶細胞中表現的調控序列。醫藥組成物含有可選擇的載劑、賦形劑、及/或防腐劑。在一些具體實施例中，供體載體進一步包括轉基因匣之同源定向重組(HDR)臂 5'及 3'。在一具體實施例中，用於供體載體、基因編輯載體、或二者的 AAV 衣殼為 AAVrh79 衣殼。在另一具體實施例中，用於供體載體、基因編輯

載體、或二者的 AAV 衣殼為 AAVrh.90 衣殼。在另一具體實施例中，用於供體載體、基因編輯載體、或二者的 AAV 衣殼為 AAVhu.37 衣殼。在一具體實施例中，用於供體載體、基因編輯載體、或二者的 AAV 衣殼為 AAV8 衣殼。在一具體實施例中，用於供體載體、基因編輯載體、或二者的 AAV 衣殼為 AAVrh.91 衣殼。在一具體實施例中，用於供體載體、基因編輯載體、或二者的 AAV 衣殼為 AAVhu.68 衣殼。

【0125】如本文所使用，「載劑」包括任何及所有的溶劑、分散介質、媒劑、塗料、稀釋劑、抗細菌及抗真菌劑、等滲及吸收延遲劑、緩衝劑、載劑溶液、懸浮液、膠體等。此類用於醫藥活性物質的介質及試劑的用途為技術領域中所熟知的。補充活性成分亦可摻入組成物中。短語「醫藥上可接受的」係指當投予宿主時不會產生過敏或類似不良反應的分子實體及組成物。遞送媒劑，諸如脂質體、奈米膠囊、微粒、微球、脂質顆粒、囊泡等可用於將本發明之組成物導入適當的宿主細胞中。特別是，可將 rAAV 載體遞送載體基因體調配用於遞送包封在脂質顆粒、脂質體、囊泡、奈米球或奈米顆粒等之中。

【0126】在一具體實施例中，組成物包括適於遞送至受試者的最終調配物，例如，為緩衝至生理學上可相容的 pH 及鹽濃度的水性液體懸浮劑。可選擇地，一或多種界面活性劑可存在於調配物中。在另一具體實施例中，組成物可作為濃縮物輸送，將其稀釋後投予受試

者。在另一具體實施例中，組成物可被凍乾並在投予時重構(reconstituted)。

【0127】本領域熟知的用於製備調配物的方法和試劑已被描述，例如，描述在“Remington's Pharmaceutical Sciences,” Mack Publishing Company, Easton, Pa。調配物可例如含有賦形劑、載劑、穩定劑、或稀釋劑，諸如無菌水、鹽水、聚伸烷基二醇(諸如聚乙二醇)、植物油、或氫化萘、防腐劑(諸如十八烷基二甲基苄基、氯化銨、六甲氫銨(hexamethonium chloride)、氯化烷基二甲基苄基銨(benzalkonium chloride)、氯化苯索寧(benzethonium chloride)、苯酚、丁醇或苄醇、對羥基苯甲酸烷酯(諸如對羥基苯甲酸甲酯或羥基苯甲酸丙酯)、兒茶酚、間苯二酚、環己醇、3-戊醇和間甲酚)、低分子量多肽、蛋白質(諸如血清白蛋白、明膠或免疫球蛋白)、親水聚合物(諸如聚乙烯吡咯啉酮、胺基酸(諸如甘胺酸、麩醯胺酸、天冬醯胺酸、組胺酸、精胺酸及離胺酸)、單醣、二糖、及其它碳水化合物，包括葡萄糖、甘露糖、及糊精、螯合劑(諸如 EDTA)、糖類(諸如蔗糖、甘露醇、海藻糖或山梨糖醇；成鹽反離子，諸如鈉；金屬複合物(例如鋅-蛋白質複合物)；及/或非離子界面活性劑，諸如 TWEEN™、PLURONICS™或聚乙二醇(PEG)。

【0128】活性成分亦可包埋於微膠囊中，例如藉由凝聚技術或藉由界面聚合製備，例如，分別在膠體藥物遞送系統(例如，脂質體、白蛋白微球、微乳液、奈米

顆粒和奈米膠囊)或粗滴乳液中的羥甲基纖維素或明膠-微膠囊和聚-(甲基丙烯酸甲酯)微膠囊。此類技術揭示於 Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)。

【0129】合適的界面活性劑或界面活性劑的組合可選自無毒的非離子界面活性劑。在一具體實施例中，選擇終止於一級羥基的雙官能嵌段共聚物界面活性劑，諸如 Pluronic® F68 [BASF]，亦稱為泊洛沙姆(Poloxamer) 188，其具有中性 pH，平均分子量為 8400。可選擇其它界面活性劑和其它泊洛沙姆，即由兩側是兩個聚氧乙烯(聚(環氧乙烷))親水鏈的聚氧丙烯(聚(環氧丙烷))中央疏水鏈所構成的非離子三嵌段共聚物、SOLUTOL HS 15 (聚(乙烯二醇)-15 (Macrogol-15)羥基硬脂酸酯)、LABRASOL (聚氧基辛基甘油酯 (Polyoxy caprylic glyceride))、聚氧基 10 油基醚、TWEEN(聚氧乙烯山梨聚糖脂肪酸酯)、乙醇和聚乙二醇。在一具體實施例中，調配物包含泊洛沙姆。這些共聚物通常用字母「P」(對泊洛沙姆)跟三個數字命名：前兩個數字 x100 給出聚氧丙烯核心的近似分子量，最後一個數字 x10 給出聚氧乙烯含量百分比。在一個具體實施例中，選擇泊洛沙姆 188。界面活性劑可以以懸浮液的高達約 0.0005%至約 0.001%的量存在。

【0130】以足夠的量將載體投予以轉染細胞並提供足夠程度的基因轉移和表現，以提供治療益處而沒有不適當的副作用，或具有醫學上可接受的生理作用，這可

由醫學領域的技術人員確定。習知和醫藥上可接受的給藥途徑包括，但不限於直接遞送至所需的器官(例如，肝臟(可選擇地經由肝動脈)、肺臟、心臟、眼、腎臟)、經口、吸入、鼻內、鞘內、氣管內、動脈內、眼內、靜脈內、肌肉內、皮下、皮內和其它親代給藥途徑。如果需要，可以組合給藥途徑。

【0131】病毒載體的劑量可取決於例如所欲治療的症狀、患者的年齡、體重和健康等因素，因此可能因患者而異。例如，病毒載體的治療有效人類劑量範圍通常為約 25 至約 1000 微升至約 100 mL 含約  $1 \times 10^9$  至  $1 \times 10^{16}$  基因體病毒載體濃度的溶液。調整劑量以平衡治療益處與任何副作用，並且此類劑量可以根據使用重組載體的治療應用而變化。可監測轉基因產物的表現程度以確定產生病毒載體(較佳為含有袖珍基因之 AAV 載體)的給藥頻率。可選擇地，類似於為治療目的而描述的那些劑量方案可用於使用本發明的組成物進行免疫。

【0132】可將載體組成物配製成劑量單位以包含在約  $1.0 \times 10^9$  GC 至約  $1.0 \times 10^{16}$  GC 範圍內的複製缺陷型病毒的量(以治療平均體重為 70 kg 的受試者)，包括在該範圍內的所有整數或部分量，並且對於人類患者較佳為  $1.0 \times 10^{12}$  GC 至  $1.0 \times 10^{14}$  GC。在一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^9$ 、 $2 \times 10^9$ 、 $3 \times 10^9$ 、 $4 \times 10^9$ 、 $5 \times 10^9$ 、 $6 \times 10^9$ 、 $7 \times 10^9$ 、 $8 \times 10^9$ 、或  $9 \times 10^9$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{10}$ 、

$2 \times 10^{10}$ 、 $3 \times 10^{10}$ 、 $4 \times 10^{10}$ 、 $5 \times 10^{10}$ 、 $6 \times 10^{10}$ 、 $7 \times 10^{10}$ 、 $8 \times 10^{10}$ 、或  $9 \times 10^{10}$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{11}$ 、 $2 \times 10^{11}$ 、 $3 \times 10^{11}$ 、 $4 \times 10^{11}$ 、 $5 \times 10^{11}$ 、 $6 \times 10^{11}$ 、 $7 \times 10^{11}$ 、 $8 \times 10^{11}$ 、或  $9 \times 10^{11}$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{12}$ 、 $2 \times 10^{12}$ 、 $3 \times 10^{12}$ 、 $4 \times 10^{12}$ 、 $5 \times 10^{12}$ 、 $6 \times 10^{12}$ 、 $7 \times 10^{12}$ 、 $8 \times 10^{12}$ 、或  $9 \times 10^{12}$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{13}$ 、 $2 \times 10^{13}$ 、 $3 \times 10^{13}$ 、 $4 \times 10^{13}$ 、 $5 \times 10^{13}$ 、 $6 \times 10^{13}$ 、 $7 \times 10^{13}$ 、 $8 \times 10^{13}$ 、或  $9 \times 10^{13}$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{14}$ 、 $2 \times 10^{14}$ 、 $3 \times 10^{14}$ 、 $4 \times 10^{14}$ 、 $5 \times 10^{14}$ 、 $6 \times 10^{14}$ 、 $7 \times 10^{14}$ 、 $8 \times 10^{14}$ 、或  $9 \times 10^{14}$  GC，包括在該範圍內的所有整數或分數量。在另一具體實施例中，組成物被調配成每劑量含有至少  $1 \times 10^{15}$ 、 $2 \times 10^{15}$ 、 $3 \times 10^{15}$ 、 $4 \times 10^{15}$ 、 $5 \times 10^{15}$ 、 $6 \times 10^{15}$ 、 $7 \times 10^{15}$ 、 $8 \times 10^{15}$ 、或  $9 \times 10^{15}$  GC，包括在該範圍內的所有整數或分數量。在一具體實施例中，對於人類施用，劑量可在範圍為每劑量  $1 \times 10^{10}$  至約  $1 \times 10^{12}$  GC，包括在該範圍內的所有整數或分數量。

【0133】依據待治療區域的大小、使用的病毒力價、給藥途徑和方法的預期效果，這些上述劑量可以各種體積的載劑、賦形劑或緩衝劑調配物給藥，範圍從約

25 至約 1000 微升，或更高的體積，包括在該範圍內的所有數量。

【0134】可選擇任何合適的給藥路徑。因此，醫藥組成物可調配用於任何合適的給藥路徑，例如，液體溶液或懸浮液形式(作為例如用於靜脈投予、經口投予等)。或者，醫藥組成物可為固體形式(例如，錠劑或膠囊形式，例如用於經口投予)。在一些具體實施例中，醫藥組成物可為粉末、滴劑、氣溶膠等形式。

【0135】於一態樣，本文提供一種在調配物緩衝劑中包含微小病毒載體的藥物組成物，該微小病毒載體包含如本文所述的至少一種基因編輯載體及至少一種供體載體。在某些具體實施例中，醫藥組成物包含不同載體群的組合。在一具體實施例中，提供了在調配物緩衝劑中包含本文所述的單一 rAAV 群的醫藥組成物。本文提供的方法提供兩種單獨的含載體之懸浮液的共同給藥。

#### 【0136】

##### 方法

本文所提供之組成物用於治療各種遺傳疾病，包括肝臟代謝性病變。在某些具體實施例中，組成物用於治療鳥胺酸胺甲醯基轉移酶。在其它具體實施例中，組成物用於治療家族性高膽固醇血症。在其它具體實施例中，組成物用於治療苯丙酮尿症。

【0137】可被治療的例示性肝臟疾病或病症包括，但不限於 A 型肝炎、B 型肝炎、C 型肝炎、自體免疫性肝炎、原發性膽汁性膽管炎、原發性硬化性膽管炎、血

色素沉積症、威爾遜氏病(Wilson's disease)、 $\alpha$ -1 抗胰蛋白酶缺乏症、肝癌、膽管癌、肝腺瘤、轉甲狀腺素蛋白(TTR)、前蛋白轉化酶枯草桿菌蛋白酶/kexin 9 型(PCSK9)系疾病或病症，或其任何組合。其它病症包括糖原貯積病或 1A 型缺乏症 (GSD1)、PEPCK 缺乏症、CDKL5 缺乏症、半乳糖血症、苯丙酮尿症(PKU)、1 型原發性高草酸鹽尿症、楓糖尿症，1 型酪胺酸血症、甲基丙二酸血症、中鏈乙醯輔酶 A 缺乏症，鳥胺酸胺甲醯基轉移酶缺乏症、瓜胺酸血症；卵磷脂-膽固醇醯基轉移酶(LCAT)缺乏症、甲基丙二酸酸血症(MMA)、尼曼匹克症、丙酸血症(PA)；家族性高膽固醇血症(FH)、失智、脂蛋白脂肪酶缺乏症、克果納傑氏症、嚴重複合型免疫缺乏症、痛風及萊施-尼漢二氏症候群、生物素酶缺乏症、法布瑞氏症、GM1 神經節醯苷病、威爾森氏症、2 型及 3 型高歇氏病、齊威格氏症、異染性白質失養症、克拉培氏病、龐貝氏症、A 型尼曼匹克症、精胺酸琥珀酸尿症、成人發病 II 型瓜胺酸血症、尿素循環病症；法伯脂肪肉芽腫病、天冬醯胺基胺基葡萄糖尿症、岩藻糖沉積症、 $\alpha$ -甘露糖沉積症、急性間歇性紫質沈著病(AIP)、 $\alpha$ -1 抗胰蛋白酶缺乏症(肺氣腫)、由地中海貧血或腎功能衰竭引起的貧血、缺血性疾病、如動脈粥樣硬化、血栓形成或栓塞中所見的血管閉塞、帕金森氏病、充血性心衰竭、肌肉萎縮症、及糖尿病。

【0138】在本文所述的某些方法中，天然 PCSK9 表現被降低或消融，且轉基因從天然 PCSK9 基因座中的插

入表現。在另一具體實施例中，天然 PCSK9 表現被降低或消融，且轉基因是外源表現的，即沒有整合至受試者的基因體中。

【0139】本文提供一種藉由共同投予本文所述之雙重載體系統來治療人類疾病的方法。

【0140】在一具體實施例中，提供在受試者中治療肝臟代謝性病症的方法。該方法包括對患有肝臟代謝性病症之受試者共同投予：基因編輯 AAV 載體，該基因編輯 AAV 載體包含編碼靶向 PCSK9 之核酸酶的序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中表現的調控序列；及供體 AAV 載體，其包含轉基因及指導轉基因在標靶細胞中之表現的調控序列。在另一具體實施例中，該方法包括對患有肝臟代謝性病症之受試者共同投予 LNP，該 LNP 包含編碼 Cas9 核酸酶及 sgRNA 之序列，其靶向包含 PCSK9 基因之標靶細胞中的 PCSK9；及供體 AAV 載體，其包含轉基因及指導該轉基因在標靶細胞中表現的調控序列。在一具體實施例中，受試者為新生兒。

【0141】在某些具體實施例中，基因編輯 AAV 載體及供體載體經由相同路徑基本上同時遞送。在其它具體實施例中，先遞送基因編輯載體。在其它具體實施例中，先遞送供體載體。

【0142】在一具體實施例中，rAAV 之劑量為每劑量約  $1 \times 10^9$  GC 至約  $1 \times 10^{15}$  基因體拷貝(GC)(治療平均體重 70 kg 的受試者)，且較佳為對於人類病患為  $1.0 \times 10^{12}$

GC 至  $2.0 \times 10^{15}$  GC。在另一具體實施例中，劑量少於約  $1 \times 10^{14}$  GC/kg 受試者體重。在某些具體實施例中，投予病患的劑量為至少約  $1.0 \times 10^9$  GC/kg、約  $1.5 \times 10^9$  GC/kg、約  $2.0 \times 10^9$  GC/g、約  $2.5 \times 10^9$  GC/kg、約  $3.0 \times 10^9$  GC/kg、約  $3.5 \times 10^9$  GC/kg、約  $4.0 \times 10^9$  GC/kg、約  $4.5 \times 10^9$  GC/kg、約  $5.0 \times 10^9$  GC/kg、約  $5.5 \times 10^9$  GC/kg、約  $6.0 \times 10^9$  GC/kg、約  $6.5 \times 10^9$  GC/kg、約  $7.0 \times 10^9$  GC/kg、約  $7.5 \times 10^9$  GC/kg、約  $8.0 \times 10^9$  GC/kg、約  $8.5 \times 10^9$  GC/kg、約  $9.0 \times 10^9$  GC/kg、約  $9.5 \times 10^9$  GC/kg、約  $1.0 \times 10^{10}$  GC/kg、約  $1.5 \times 10^{10}$  GC/kg、約  $2.0 \times 10^{10}$  GC/kg、約  $2.5 \times 10^{10}$  GC/kg、約  $3.0 \times 10^{10}$  GC/kg、約  $3.5 \times 10^{10}$  GC/kg、約  $4.0 \times 10^{10}$  GC/kg、約  $4.5 \times 10^{10}$  GC/kg、約  $5.0 \times 10^{10}$  GC/kg、約  $5.5 \times 10^{10}$  GC/kg、約  $6.0 \times 10^{10}$  GC/kg、約  $6.5 \times 10^{10}$  GC/kg、約  $7.0 \times 10^{10}$  GC/kg、約  $7.5 \times 10^{10}$  GC/kg、約  $8.0 \times 10^{10}$  GC/kg、約  $8.5 \times 10^{10}$  GC/kg、約  $9.0 \times 10^{10}$  GC/kg、約  $9.5 \times 10^{10}$  GC/kg、約  $1.0 \times 10^{11}$  GC/kg、約  $1.5 \times 10^{11}$  GC/kg、約  $2.0 \times 10^{11}$  GC/kg、約  $2.5 \times 10^{11}$  GC/kg、約  $3.0 \times 10^{11}$  GC/kg、約  $3.5 \times 10^{11}$  GC/kg、約  $4.0 \times 10^{11}$  GC/kg、約  $4.5 \times 10^{11}$  GC/kg、約  $5.0 \times 10^{11}$  GC/kg、約  $5.5 \times 10^{11}$  GC/kg、約  $6.0 \times 10^{11}$  GC/kg、約  $6.5 \times 10^{11}$  GC/kg、約  $7.0 \times 10^{11}$  GC/kg、約  $7.5 \times 10^{11}$  GC/kg、約  $8.0 \times 10^{11}$  GC/kg、約  $8.5 \times 10^{11}$  GC/kg、約  $9.0 \times 10^{11}$  GC/kg、約  $9.5 \times 10^{11}$  GC/kg、約  $1.0 \times 10^{12}$  GC/kg、約  $1.5 \times 10^{12}$  GC/kg、約  $2.0 \times 10^{12}$  GC/kg、約  $2.5 \times 10^{12}$  GC/kg、約  $3.0 \times 10^{12}$  GC/kg、約  $3.5 \times 10^{12}$  GC/kg、約

4.0 x 10<sup>12</sup> GC/kg、約 4.5 x 10<sup>12</sup> GC/kg、約 5.0 x 10<sup>12</sup> GC/kg、約 5.5 x 10<sup>12</sup> GC/kg、約 6.0 x 10<sup>12</sup> GC/kg、約 6.5 x 10<sup>12</sup> GC/kg、約 7.0 x 10<sup>12</sup> GC/kg、約 7.5 x 10<sup>12</sup> GC/kg、約 8.0 x 10<sup>12</sup> GC/kg、約 8.5 x 10<sup>12</sup> GC/kg、約 9.0 x 10<sup>12</sup> GC/kg、約 9.5 x 10<sup>12</sup> GC/kg、約 1.0 x 10<sup>13</sup> GC/kg、約 1.5 x 10<sup>13</sup> GC/kg、約 2.0 x 10<sup>13</sup> GC/kg、約 2.5 x 10<sup>13</sup> GC/kg、約 3.0 x 10<sup>13</sup> GC/kg、約 3.5 x 10<sup>13</sup> GC/kg、約 4.0 x 10<sup>13</sup> GC/kg、約 4.5 x 10<sup>13</sup> GC/kg、約 5.0 x 10<sup>13</sup> GC/kg、約 5.5 x 10<sup>13</sup> GC/kg、約 6.0 x 10<sup>13</sup> GC/kg、約 6.5 x 10<sup>13</sup> GC/kg、約 7.0 x 10<sup>13</sup> GC/kg、約 7.5 x 10<sup>13</sup> GC/kg、約 8.0 x 10<sup>13</sup> GC/kg、約 8.5 x 10<sup>13</sup> GC/kg、約 9.0 x 10<sup>13</sup> GC/kg、約 9.5 x 10<sup>13</sup> GC/kg、或約 1.0 x 10<sup>14</sup> GC/kg 受試者體重。

【0143】可使用本文所述之組成物治療的合適疾病的其它實例是家族性高膽固醇血症、肌營養不良、囊性纖維化及罕見疾病或孤兒疾病罕見或孤兒疾病。此類罕見疾病的實例可包括脊髓肌萎縮症、亨廷頓氏病、雷特症候群、肌肉萎縮性脊髓側索硬化症(ALS)、杜氏型肌營養不良症、弗里德希氏共濟失調、2型脊髓小腦性失調症(SCA2)/ALS、顆粒蛋白前體(PRGN)(與非阿茨海默氏症腦退化相關，包括額顳葉失智症(FTD)、進行性非流利性失語症(PNFA)及語義性失智)等。參見，例如，[www.orpha.net/consor/cgi-bin/Disease\\_Search\\_List.php](http://www.orpha.net/consor/cgi-bin/Disease_Search_List.php)；[rarediseases.info.nih.gov/diseases](http://rarediseases.info.nih.gov/diseases)。由本文所述的轉基因指導的其它疾病亦可使用本文所述的方法治療。

【0144】以足夠的投予載體來轉染細胞並提供足夠程度的基因轉移和表現以提供治療益處而沒有過度的副作用，或具有醫學上可接受的生理效應，這可以由醫學領域的技術人員確定。理想的給藥途徑包括直接遞送至所需的器官(例如，肝臟(可選擇地經由肝動脈)、肺臟、心臟、眼、腎臟)、經口、吸入、鼻內、氣管內鞘內、動脈內、眼內、靜脈內、肌肉內、皮下、皮內和其它親代給藥途徑。如果需要，可以組合給藥途徑。

【0145】如果產生足夠量的功能性酶或蛋白質以改善患者的狀況，則本文所述的系統可對治療有用。在某些具體實施例中，低至 5%的健康患者的基因表現水平將為患者提供足夠的治療效果，從而可以用非基因治療方法進行治療。在其它具體實施例中，基因表現水平為在人類(或其它獸類受試者)中觀察到的正常範圍(水平)的至少約 5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、

93%、94%、95%、96%、97%、98%、99%、至高達100%。例如，「功能性酶」意指編碼野生型酶(例如OTCase)的基因，其提供野生型酶或其與疾病無關之天然變體或多晶型物的至少約5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或約相同、或大於100%的生物活性水平。更具體地，由於雜合子病患可能具有低至約50%或更低的酶功能水平，因此有效治療可能不需要將酶活性替換至「正常」或非缺陷病患範圍內的水平。類似地，無可檢測量的酶的病患可藉由將酶功能遞送至低於100%的活性水平來挽救，並可選擇地隨後接受進一步的治療。在某些具體實施例中，在藉由供體模板遞送基因功能的情況下，病患的表達水平可高於「正常」、健康受試者中的水平。在另外其它實施方式中，在需要降低基因表現的情況下，降低多達20%至50%，或多達約100%，

可提供所需的益處。如本文所述，本文所述的治療可與其它治療結合使用，即用於受試者(病患)診斷的護理標準。

【0146】在一具體實施例中，該方法進一步包含向受試者投予免疫抑制聯合療法。此類免疫抑制聯合療法可在遞送 rAAV 或所揭示之組成物之前開始，例如，如果檢測到 AAV 衣殼的中和抗體水平過高。在某些具體實施例中，作為預防措施，也可以在遞送 rAAV 之前開始聯合治療。在某些具體實施例中，免疫抑制聯合療法在 rAAV 遞送後開始，例如，如果在治療後觀察到不希望的免疫反應。

【0147】用於此類聯合療法的免疫抑制劑包括，但不限於糖皮質素、類固醇、抗代謝物、T 細胞抑制劑、巨環內酯(例如，雷帕黴素(rapamycin)或雷帕黴素類似物)、及細胞抑制劑，包括烷化劑、抗代謝物、細胞毒性抗生素、抗體或對免疫親和素有活性之藥劑。免疫抑制劑可包括腎上腺皮質酮(prednisolone)、氮芥(nitrogen mustard)、亞硝脲(nitrosourea)、鉑化合物、胺甲喋呤(methotrexate)、硫唑嘌呤(azathioprine)、巯嘌呤(mercaptopurine)、氟尿嘧啶(flourouracil)、放線菌素(dactinomycin)、蒽環類(anthracycline)、絲裂黴素 C(mitomycin C)、博來黴素(bleomycin)、光輝黴素(mithramycin)、IL-受體-(CD25-)或 CD3-導向的抗體、抗 IL-2 抗體、環孢素(ciclosporin)、他克莫司(tacrolimus)、西羅莫司(sirolimus)、IFN- $\beta$ 、IFN- $\gamma$ 、類

鴉片 (opioid)、或 TNF- $\alpha$ (腫瘤壞死因子- $\alpha$ )結合劑。在某些具體實施例中，可在 rAAV 投予之前的 0、1、2、7 天或更多天或在 rAAV 投予之後的 0、1、2、3、7 或更多天開始免疫抑制治療。此類治療可涉單一藥劑(例如，腎上腺皮質酮)及在同一天共同投予二種或多種藥物(例如，強體松(prednelisone)、嗎替麥考酚酯(micophenolate mofetil, MMF)及/或西羅莫司(即，雷帕黴素))。於基因治療給藥後能以相同劑量或調整劑量繼續使用一種或多種這些藥物。此類療法可根據需要持續約 1 週(7 天)、兩週、三週、約 60 天或更長。在某些具體實施例中，選擇不含他克莫司的方案。

【0148】在另一具體實施例中，該方法包括與標準 OTC 療法共同治療。OTC 缺乏症的治療主要集中在血氨水平的飲食管理上，以避免高氨血症或在高氨血症發作期間從血液中去過量的氨(NORD, 2021)。患有 OTC 缺乏症的個體遵循飲食限制，限制他們的蛋白質攝取量以控制血氨水平。嬰兒的飲食限制必須仔細平衡，他們需要攝入足夠的蛋白質以確保正常生長，同時避免可能引發高氨血症發作的過量蛋白質攝取(Berry and Steiner, 2001)。因此，嬰兒著重高熱量、低蛋白質的飲食，並輔以必需胺基酸。在高氨血症發作中，可在 24 小時內從患者的飲食中去過所有蛋白質(NORD, 2021)。

【0149】有幾種藥物被設計來刺激從血流中去過氮。苯基丁酸鈉(Buphenyl)被美國食品藥物管理局(FDA)批准用於治療 OTC 缺乏症病患的慢性高氨血症。一旦代謝，Buphenyl 轉化為苯乙酸鹽，它與麩醯胺酸結

合形成苯基乙醯基麩醯胺酸，由腎臟排泄，為氮排泄提供了替代途徑。苯丁酸甘油 (glycerol phenylbutyrate)(Ravicti)亦被 FDA 批准用於治療患有尿素循環病症之病患的慢性高氨血症。如同 Buphenyl，Ravicti 被轉化為苯乙酸鹽，並遵循相同的排泄氮的機制 (Lichter-Konecki et al., 1993 ; Gordon, 2003 ; Magellan, 2021)。最後，Ammonul(苯乙酸鈉和苯甲酸鈉)被 FDA 批准作為治療患有尿素循環病症之病患的急性高氨血症的輔助療法。Ammonul 的苯乙酸鈉成分遵循與 Buphenyl 和 Ravicti 產生的苯乙酸鹽代謝物相同的氮排泄機制。Ammonul 的苯甲酸鈉成分與甘胺酸結合形成馬尿酸，馬尿酸由腎臟排出並通過此過程去除氮。苯甲酸鈉亦可以口服製劑提供用於長期維持 OTC 缺乏症，並且因為被認為具有較少的副作用而通常優於 Buphenyl 和 Ravicti (Lichter-Konecki et al., 1993)。

【0150】於一態樣，提供治療患有鳥胺酸胺甲醯基轉移酶(OTC)缺乏症之病患的方法，其使用包含巨型核酸酶編碼序列之核酸酶表現匣，其在如本文所述之啟動子的控制下識別人類 PCSK9 基因內的位點。該方法進一步包括投予攜帶 SEQ ID NO: 17 之 OTC 轉基因、或與其共享至少 90%同一性之序列的表現匣，如本文所述。此類表現匣可經由病毒或非病毒載體遞送。在某些具體實施例中，該表現匣可使用 LNP 遞送。天然人類 OTC 編碼序列顯示於 SEQ ID NO: 30。SEQ ID NO: 17 與 SEQ ID NO: 30 共享約 75.89%同一性。

【0151】在另一態樣中，提供治療患有鳥胺酸胺甲醯基轉移酶 (OTC) 缺乏症之病患的方法，其使用包含 sgRNA 及 Cas9 編碼序列的核酸酶表現匣，其識別人類 PCSK9 基因內的位點。該方法進一步包括投予攜帶 SEQ ID NO: 17 之 OTC 轉基因、或與其共享至少 90% 同一性之序列的表現匣，如本文所述。此類表現匣可經由病毒或非病毒載體遞送。在某些具體實施例中，該表現匣可使用 LNP 遞送。

【0152】存在多種用於測量體外 OTC 表現和活性水平的測定法。參見，例如，X Ye, et al, 1996 Prolonged metabolic correction in adult ornithine transcarbamylase-deficient mice with adenoviral vectors. *J Biol Chem* 271:3639–3646) 或體內。例如，OTC 酶活性可使用液相層析質譜穩定同位素稀釋法檢測標準化為 [1,2,3,4,5-<sup>13</sup>C<sub>5</sub>] 瓜胺酸 (98% <sup>13</sup>C) 的瓜胺酸的形成。該方法從先前開發的用於檢測 N-乙醯麩胺酸鹽合成酶活性的測定法調整 [Morizono H, et al, *Mammalian N-acetylglutamate synthase. Mol Genet Metab.* 2004;81(Suppl 1):S4–11.]。將新鮮冷凍肝臟切片稱重並在含有 10 mM HEPES、0.5% Triton X-100、2.0 mM EDTA 和 0.5 mM DTT 的緩衝劑中簡單地勻漿。調整勻漿緩衝劑的體積以獲得 50 mg/ml 組織。使用含 250 μg 肝組織之 50 mM Tris-乙酸鹽、4 mM 鳥胺酸、5 mM 胺甲醯磷酸鹽，在 pH 8.3 中測量酶活性。加入新鮮製備的溶於 50 mM Tris-乙酸鹽 pH 8.3 中的 50 mM 胺甲醯磷酸鹽開始酶活性，使其在 25°C 下進

行 5 分鐘，並藉由加入等體積的含 5 mM  $^{13}\text{C}_5$ -瓜胺酸之 30%TCA 來淬滅。藉由 5 分鐘的微量離心分離碎片，並將上清液轉移到小瓶中用於質譜分析。在等度條件下將 10  $\mu\text{L}$  樣品注入 Agilent 1100 系列 LC-MS，流動相為 93%溶劑 A (含 1 ml 三氟乙酸之 1 L 水):7%溶劑 B (含 1 ml 三氟乙酸之 1L 的 1:9 水/乙腈)。將對應於瓜胺酸之峰 [176.1 質荷比 (m/z)] 及  $^{13}\text{C}_5$ -瓜胺酸 (181.1 m/z) 量化，並將它們的比率與每次測定運行的瓜胺酸標準曲線獲得的比率進行比較。將樣品標準化為總肝臟組織或使用 Bio-Rad 蛋白質測定套組 (Bio-Rad, Hercules, CA) 測定的蛋白質濃度。亦可使用不需要肝臟生檢的其它測定。一種此類測定是血漿胺基酸測定，其中評估麩醯胺酸與瓜胺酸的比率，如果麩醯胺酸高 (>800 微升/升) 而瓜胺酸低 (例如，個位數)，懷疑是尿素循環缺陷。可測量血漿氨水平，每升約 100 微莫耳的濃度表示 OTCD。如果病患過度換氣，則可以評估血中氣體；呼吸性鹼中毒在 OTCD 中很常見。尿液中的乳清酸 (Orotic acid)，例如每毫莫耳肌酸大於約 20 微莫耳，是 OTCD 的指徵，在異噁呤醇激發試驗後尿乳清酸鹽升高也是如此。OTCD 的診斷標準已被提出於 Tuchman et al, 2008, Urea Cycle Disorders Consortium (UCDC) of the Rare Disease Clinical Research Network (RDCRN)、Tuchman M, et al., Consortium of the Rare Diseases Clinical Research Network. Cross-sectional multicenter study of patients with urea cycle disorders in the United States. Mol Genet

Metab. 2008;94:397-402，其藉由引用併入本文。亦可參見 <http://www.ncbi.nlm.nih.gov/books/NBK154378/>，其討論了目前 OTCD 的護理標準。

【0153】在某些具體實施例中，如本文所述，可投予醫藥組成物中的核酸酶表現匣、非病毒載體、病毒載體(例如，rAAV)或任何相同的載體用於病患的基因編輯。在某些具體實施例中，該方法有用於非胚胎的基因編輯。在某些具體實施例中，病患為嬰兒(例如，從出生至約9個月)。在某些具體實施例中，病患年齡大於嬰兒，例如，12月或更大。

【0154】如本文所使用，「一」、「一種」或「該」可意指一種或大於一種。例如，「一種」細胞可意指單一細胞或多個細胞。

【0155】如本文所使用，術語「特異性」意指核酸酶僅在稱為識別序列之鹼基對的特定序列處或僅在特定的一組識別序列處識別和切割雙股 DNA 分子的能力。該組識別序列將共享某些保留位置或序列基序，但可在一個或多個位置簡併。高度特異性的核酸酶只能切割一個或很少的識別序列。特異性可藉由本領域已知的任何方法確定。

【0156】縮寫「sc」係指自我互補。「自我互補的 AAV」係指其中已經設計由重組 AAV 核酸序列攜帶的編碼區以形成分子內雙股 DNA 模板的構建體。感染後，並非等待細胞媒介的第二股的合成，而是 scAAV 的兩個互補半部將結合形成一個準備立即複製及轉錄的雙股

DNA(dsDNA)單元。參見，例如，D M McCarty et al, “Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis”, Gene Therapy, (August 2001), Vol 8, Number 16, Pages 1248-1254。自我互補 AAV 描述於例如，美國專利號 6,596,535；7,125,717；及 7,456,683，其每一者藉由引用整體併入本文。

【0157】如本文中所使用，術語「可操作地連接」係指相鄰於目的基因的表現控制序列及反向或在遠距離起作用以控制目的基因的表現控制序列二者。

【0158】用於描述核酸序列或蛋白質的術語「外源性」意指核酸或蛋白質並非天然存在於其在染色體或宿主細胞中所存在的位置。外源核酸序列亦指源自並插入相同表現匣或宿主細胞中的序列，但其以非天然狀態存在，例如不同拷貝數，或在不同調控元件的控制下。

【0159】當涉及蛋白質或核酸使用時，術語「異源的」表示蛋白質或核酸包含在自然界中彼此之間沒有相同關係的兩個或更多個序列或子序列。舉例而言，核酸通常是重組產生的，具有二或多個來自無關基因的序列，其排列以產生新的功能性核酸。例如，在一個具體實施例中，該核酸具有來自一個基因的啟動子，其被安排為指導來自不同基因的編碼序列的表現。

【0160】如本文所使用，術語「宿主細胞」可指其中由質體產生的載體(例如，重組 AAV)的包裝細胞株。或者，術語「宿主細胞」可指希望其轉基因的表現之標

的細胞。因此，「宿主細胞」係指含有外源或異源核酸序列的原核或真核細胞，該核酸序列已藉由任何方法導入細胞中，例如，電穿孔、磷酸鈣沉澱、顯微注射、轉化、病毒感染、轉染、微脂體遞送、膜融合技術、高速DNA包覆的小丸、病毒感染和原生質體融合。在本文某些具體實施例中，術語「宿主細胞」係指用於體外評估本文所述組成物的各種哺乳動物物種的細胞培養物。在本文其它具體實施例中，術語「宿主細胞」係指用於產生和包裝病毒載體或重組病毒的細胞。在另一個具體實施例中，術語「宿主細胞」欲指在體內針對本文所述的疾病或症狀進行治療的受試者的標靶細胞。在某些具體實施例中，術語「宿主細胞」為肝臟細胞或肝細胞。

【0161】「受試者」為哺乳動物，例如，人類、小鼠、大鼠、豚鼠、狗、貓、馬、牛、豬、或非人類靈長類動物，諸如猴子、黑猩猩、狒狒或大猩猩。病患係指人類。獸類受試者係指非人類哺乳動物。在某些具體實施例中，受試者的 PCSK9 基因並沒有缺陷。

【0162】「複製缺陷病毒」或「病毒載體」係指合成或人工的病毒顆粒，其中含有感興趣基因的表現匣被包裝於病毒衣殼或套膜中，其中任何病毒基因體序列亦被包裝於病毒衣殼或套膜內為複製缺陷的，即它們不能產生子代病毒體但保留感染標靶細胞的能力。在一個具體實施例中，病毒載體的基因體不包括編碼複製所需的酶的基因(基因體可被工程化為「怯懦的(gutless)」-僅含感興趣的基因，側接人工基因體擴增和包裝所需的訊

號)，但這些基因可在生產過程中提供。因此，其被認為用於基因治療是安全的，因為除非存在複製所需的病毒酶，否則不會發生子代病毒顆粒的複製和感染。

【0163】在核酸序列的情況下，術語「序列同一性」、「序列同一性百分比」或「百分比相同」係指當用於最大對應對齊時，兩個序列中相同的殘基。序列同一性比較的長度可超過基因體的全長、基因編碼序列的全長或至少約 500 至 5000 個核苷酸的片段是受期望的。然而，亦受期望的是較小片段之間的同一性，例如至少約 9 個核苷酸、通常至少約 20 至 24 個核苷酸、至少約 28 至 32 個核苷酸、至少約 36 個或更多個核苷酸。類似地，對於胺基酸序列，可在蛋白質的全長或其片段上容易地確定「序列同一性百分比」。適當地，片段長度為至少約 8 個胺基酸，並可多至約 700 個胺基酸。本文描述了適當片段的實例。

【0164】當提及胺基酸或其片段時，術語「實質上同源」或「實質上相似」表示當利用適當的胺基酸插入或刪除與另一胺基酸(或其互補股)最佳對齊時，經對齊的序列中存在至少約 95%至 99%的胺基酸序列同一性。較佳地，同源是在全長序列、或其蛋白質，例如，cap 蛋白、rep 蛋白，或長度為至少 8 個胺基酸，或更理想地為其長度至少 15 個胺基酸的片段。本文描述了適當片段的實例。

【0165】術語「高度保守的」是指至少 80%同一性，較佳為至少 90%同一性，更佳為大於 97%同一性。

藉由本領域技術人員已知的運算法和電腦程式，本領域技術人員可易於確定同一性。

【0166】一般而言，當於兩不同腺相關病毒之間指「同一性」、「同源性」或「相似性」時，「同一性」、「同源性」或「相似性」係由「排列比對的(aligned)」序列來測定。「排列比對的」序列或「排列比對」係指多個核酸序列或蛋白質(胺基酸)序列，與參考序列相比，通常包含缺失或額外的鹼基或胺基酸的校正。在實例中，使用已公開的 AAV9 序列作為參考點進行 AAV 比對。使用許多公開或市售 Multiple Sequence Alignment Programs 進行排列比對。這類程式之實例包括「Clustal Omega」、「Clustal W」、「CAP Sequence Assembly」、「MAP」及「MEME」，其等可透過網際網路上的 Web 伺服器進行訪問。此類程式的其它來源是本領域技術人員已知的。或者，亦可使用 Vector NTI 公用程式。還有許多技術領域中已知可用於測量核苷酸序列同一性的演算法，包括上述程序中包含的演算法。作為另一實例，可使用 Fasta™ (GCG 版本 6.1 中的程序)比較多核苷酸序列，Fasta™ 提供在查詢和搜尋序列之間最佳重疊區域的排列比對和百分比序列同一性。例如，核酸序列之間的百分比序列同一性可使用 Fasta™ 及其默認參數決定(字組大小為 6 及用於得分矩陣的 NOPAM 因數)，如 GCG 版本 6.1 中所提供，藉由引用併入本文。胺基酸序列也可使用多序列排序比對程式，例如，「Clustal Omega」、「Clustal X」、「MAP」、

「 PIMA 」、「 MSA 」、「 BLOCKMAKER 」、「 MEME 」及「 Match-Box 」程式。一般而言，此等程式之任一者皆於內定值下使用，儘管本項技術領域中具通常知識者可根據需要改變這些設定。或者，熟悉技術者可以利用提供至少由參考的算法及程式提供的同一性的程度或排列比對之另一算法或電腦程式。參見，例如，J. D. Thomson et al, Nucl. Acids. Res., “A comprehensive comparison of multiple sequence alignments”, 27(13):2682-2690 (1999)。

【 0167 】如本文所使用，術語「約」係指與參考整數及其之間的值相差 $\pm 10\%$ 的變體。例如，「約」40 個鹼基對，包括 $\pm 4$ （即，36-44 個，包括整數 36、37、38、39、40、41、42、43、44 個）。對於其它值，尤其是參考百分比時（例如，90%同一性，約 10%變異，或約 36%錯配），術語「約」包括範圍內的所有值，包括整數和分數。

【 0168 】在整個本文中，本發明的各個方面可以範圍的形式呈現。應理解，範圍形式的描述僅為了方便及簡潔，不應該被解釋為對本發明範圍的不可變限制。因此，範圍的描述應被解釋為已特定揭露所有可能的子範圍以及該範圍內的單一數值。諸如，從 1 至 6 的範圍的描述應該被解釋已特定揭露子範圍，諸如 1 至 3、1 至 4、1 至 5、2 至 4、2 至 6、3 至 6 等，以及在該範圍內的個別數字，例如 1、2、2.7、3、4、5、5.3 及 6。無論範圍的廣度如何皆適用。

【0169】如本說明書上下文和申請專利範圍所使用的，術語「包含」、「含有」、「包括」、及其變體包括其它組分、元件、整數、步驟等。相反地，術語「由...組成」及其變體為排除其它組分、元素、整數、步驟等。

【0170】除非在本說明書中另有定義，本文使用的技術和科學術語具有與本領域中具有普通技術人員通常理解的含義相同的含義，並參考已公開內容，這些內容為本領域技術人員提供對本案說明書中使用的許多術語的一般指導。

#### 【0171】

#### 實施例

【0172】鳥胺酸胺甲醯基轉移酶(OTC)缺乏症是一種與高死亡率相關的 X 染色體-連鎖尿素循環病症。儘管對於遲發性 OTC 缺乏症的治療很有希望，但腺相關病毒(AAV)新生兒基因治療僅能提供短期治療效果，因為非整合基因體在肝細胞增殖過程中會丟失。核酸酶媒介、位點特異性整合 OTC 袖珍基因匣至基因體的安全港中將為 OTC 缺乏症患者提供長期治療益處。用於基因靶向的安全港之一是 PCSK9 基因，諸如外顯子 7 區域。核酸酶可以是一種靶向 PCSK9 (ARCUS2)或 CRISPR/Cas9 之工程化巨型核酸酶，具有靶向 PCSK9 的特異性 sgRNA。供體載體含有一袖珍基因，該袖珍基因包括肝特異性啟動子(諸如 TBG 啟動子)、密碼子優化的 hOTC 編碼序列、及 poly A 序列。核酸酶及供體模板皆可藉由 AAV

載體(雙 AAV 載體系統)遞送。在新生非人類靈長類動物(NHP)中單次靜脈注射雙 AAV 載體後 12 週，12%的肝細胞中證實持續的轉基因表現和有效的基因靶向。供體載體中的袖珍基因兩側是同源定向重組(HDR)臂。

【0173】在新生兒或嬰兒單次注射雙 AAV 載體後，首次在 NHP 中證明體內核酸酶媒介的基因靶向 PCSK9 基因座以表現治療性蛋白。用於對人類/NHP PCSK9 基因座的基因靶向之含有 OTC 袖珍基因之供體載體的組成物尚未在臨床上測試用於治療 OTC 缺乏症。我們將在新生 NHP 中測試 hOTC 供體載體的基因靶向效率，並在新生兒轉基因 OTC 缺陷小鼠中測試其功效。

【0174】許多代謝疾病需要早期干預和治療；然而，由於新生兒階段肝臟快速增殖和 AAV 載體的非整合性，AAV 媒介的新生兒基因治療不穩定。在安全港中靶向整合治療性袖珍基因匣將在基因體水平上持續表現治療性基因，並通過細胞分裂維持治療效果。對於許多代謝疾病，諸如 OTC 缺乏症，需要在肝臟中達到足夠的轉導效率才能獲得臨床益處。

【0175】我們描述了一種用於治療鳥胺酸胺甲醯基轉移酶缺乏症(OTCD)的基因體編輯方法，該方法可導致嬰兒期致命的高氨血症發作。基因體編輯的目標是使治療效果在所有 OTCD 病患中持久並達成，而與他們的突變無關。我們建議藉由使用兩種 AAV 載體治療倖存的新生兒來實現此一目標：一個是遞送核酸酶以在安全港位點產生雙股斷裂，而第二個是遞送 OTC 袖珍基因以敲

入此位點。我們的假設是，新生兒肝臟的肝細胞分裂將有助於有效敲入 OTC 基因，並將通過稀釋消除未整合的輸入載體基因體。我們決定使用 PCSK9 基因作為安全港位點和一種稱為 ARCUS 的巨型核酸酶來靶向它，這是基於我們之前在成年獼猴中的研究，該研究顯示在 AAV 遞送 ARCUS 後 PCSK9 安全、有效且穩定地減少。我們對 OTCD 基因體編輯的初步研究是在通過對內源性 PCSK9 基因的外顯子 7 進行生殖細胞系修飾而對 PCSK9 ARCUS 核酸酶敏感的 OTC 缺陷小鼠中進行的。將這兩種載體注射至新生兒小鼠體內導致有效敲入人類 OTC 袖珍基因，並在受到高蛋白飲食挑戰時防護致死性高氨血症。在準備臨床研究時，我們評估了新生兒和嬰兒獼猴的關鍵安全性和有效性參數。共有 24 隻動物接受 AAV 載體的治療，分析包括在 3 個月和 12 個月時檢測肝臟生檢。在這些研究中，我們評估了以下參數對編輯效率和毒性的影響：轉基因(人類 IX 因子及人類 OTC)、驅動 ARCUS 的啟動子、分支群 E 衣殼、轉基因側翼的供體長度及給藥時獼猴的年齡。我們在此報告 16/24 隻動物的初步數據，其中包括至少 3 個月的生檢結果。我們發現 AAV 載體的注射是非常安全的，在任何受 ARCUS 治療的動物中都沒有轉胺酶升高或肝臟組織病理學的證據。靈長類動物模型中療效的關鍵量度是藉由原位雜交和免疫染色測量的轉導效率，以分別檢測表現人類 OTC mRNA 和蛋白質的細胞。在第一個載體中使用帶有 TBG 啟動子的新型進化枝 E 衣殼驅動 ARCUS，並在供體載

體上使用 500 bp 側翼同源臂，使用載體獲得最高且最一致的結果。以此組合，我們達到  $10.0 \pm 6.4\%$  (N=6) 轉導，其高於我們認為可以為病患提供實質益處的閾值，即約 5% OTC 表現細胞。初步數據表明，編輯水平穩定超過一年，且當注射至 3 個月大的獼猴中時，可達到有效的靶向插入。PCSK9 標靶基因座之分子分析表明，絕大多數的載體基因體敲入是通過非同源末端連接 (NHEJ) 而不是同源定向修復 (HDR)。總之，OTCD 的新生兒形式的大量未滿足需求值得考慮實驗性療法，例如本報告中描述的基因體編輯。

#### 【 0176 】

實施例 1-材料與方法

#### 【 0177 】

材料與方法方法

【 0178 】 AAV 載體是根據先前建立的程序和製造商的說明構建的。AAVhu37 衣殼用於如本文所述的實驗，其中指明。

【 0179 】 所有動物程序均按照賓夕法尼亞大學機構動物護理和使用委員會 (the Institutional Animal Care and Use Committee of the University of Pennsylvania) 批准的方案進行。

#### 【 0180 】

實施例 2-先期研究：在新生 NHP 中藉由 ARCUS2 或 SACAS9 在 PCSK9 基因座中敲入 HFIX 袖珍基因

【0181】在此研究中，我們評估在新生非人類靈長類動物(NHP)中的靶向(PSK9) SaCas9 或 ARCUS 媒介的基因編輯和 hFIX 或 OTC 袖珍基因敲入效率。圖 1 顯示 rhPCSK9 基因座的示意圖，其顯示外顯子 7 內的供體剪接位點，以及包含感興趣的供體模板的 HDR 供體載體，例如 hFIX、hOTC。此外，圖 3A 至 3C 顯示用於 SaCas9 或 ARCUS 媒介的基因校正的雙 AAV 載體系統之示意圖。圖 3A 顯示用於 ARCUS2 媒介的基因校正的雙 AAVhu37 載體系統的示意圖，其中該 AAVhu37-供體載體包含 hOTC 供體模板序列。圖 3B 顯示用於 Sa-Cas9 媒介的基因校正(反式；AAVhu37-SaCas9)的雙 AAVhu37 載體系統的示意圖，其中 AAV.hu37.shRNA-供體載體包含 hOTC 供體模板序列。圖 3C 顯示用於 Sa-Cas9 媒介的基因校正的雙 AAVhu37 載體系統(順式；AAVhu37.PCSK9-sgRN.SaCas9)的示意圖，其中 AAV.hu37-供體載體包含 hOTC 供體模板序列。

【0182】上述包含基因編輯核酸酶和供體模板的雙 AAVhu37 載體用於新生 NHP，以檢測由 SaCas9 或 ARCUS2 媒介的 PCSK9 基因座中的 hFIX 袖珍基因敲入。基因編輯 AAVhu37 載體以  $1 \times 10^{13}$  GC/kg 的劑量遞送，且供體模板 AAVhu37 載體以  $3 \times 10^{13}$  GC/kg 的劑量遞送。整體而言，有三個治療 NHP 組：1) AAVhu37.EGFP 及 AAVhu37.供體-HDR-hFIX.U6.sgR；2) AAVhu37.ARCUS2 及 AAVhu37.供體-HDR-hFIX；3) AAVhu37.SaCas9 及 AAVhu37.供體-HDR-hFIX.U6.sgR。

圖 2 顯示一項先期研究的時間線，該研究包含 ARCUS2 或 SaCas9 在新生 NHP 中的 PCSK9 基因座中的 hFIX 袖珍基因敲入。在此研究中，在第 0 天注射 NHP，每 2-4 週收集一次血液樣品(檢查血清化學、血漿中 hFIX 表現、血清中 PCSK9 水平、LDL 水平和中和抗體(NAb)水平)，在第 84 天進行第一次肝臟生檢(檢測載體基因體水平、基因表現水平、標靶及脫靶編輯和組織學)。

【0183】在新生兒和嬰兒 NHP 中進行核酸酶媒介的基因靶向的體內測試。向動物投予  $1 \times 10^{13}$  GC/kg 之 AAVhu37.ARCUS2.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR，或  $1 \times 10^{13}$  GC/kg 之 AAVhu37.SaCas9.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR.U6.sgR，或  $1 \times 10^{13}$  GC/kg 之 AAVhu37.GFP.WPRE 及  $3 \times 10^{13}$  GC/kg 之 AAVhu37.hFIXco-HDR.U6.sgR，如圖 4A、4B 及 5G 中所示。圖 4C 顯示在治療後第 0 天至第 13 個月的指定時間點的 hFIX 水平(繪製為 ng/mL)。圖 4D 顯示在治療後第 0 天至第 12 個月的指定時間點的 PCSK9 水平(繪製為第 0 天的基線百分比)。圖 4E 顯示治療後從第 0 天至第 196 天的指定時間點(繪製為 U/L)的 ALT (丙胺酸轉胺酶)水平。圖 4F 顯示治療後從第 0 天到第 196 天的指定時間點的抗 FIX IgG 水平(繪製為稀釋因子，1/稀釋度)。圖 4G 顯示治療後從第 0 天到第 196 天的指定時間點的 PCSK9 水平(繪製為 ng/mL)。圖 4H 顯示治療後從第 0 天到第 196 天的指定時間點所量測的重量(繪製為 g)。圖

5A 顯示在嬰兒 NHP 中於指定時間點的 hFIX 水平(繪製為 ng/mL)。圖 5B 顯示在嬰兒 NHP 中於指定時間點的 PCSK9 水平(繪製為第 0 天的基線百分比)。圖 5C 顯示在嬰兒 NHP 中於指定時間點的 ALT (丙胺酸轉胺酶)水平(繪製為 U/L)。圖 5D 顯示在嬰兒 NHP 中於指定時間點的抗 FIX IgG 水平(繪製為稀釋因子，1/稀釋度)。圖 5E 顯示在嬰兒 NHP 中於指定時間點的 PCSK9 水平(繪製為 ng/mL)。圖 5F 顯示在嬰兒 NHP 中於指定時間點所量測的重量(繪製為 g)。圖 5G 為總表，顯示圖 4A-5F 中描述的實驗數據。圖 5H 顯示測試的新生兒和嬰兒 NHP 之間的各种數據比較。

【0184】圖 6A 至 6E 顯示在 NHP 中治療後幾天收集的肝臟生檢樣品中的載體轉導(GC)和轉基因表現。圖 6A 顯示肝臟生檢樣品中的載體轉導水平，繪製為每個二倍體細胞的 AAV 基因體拷貝(GC)。圖 6B 顯示肝臟生檢樣品中轉基因 RNA 的相對表現。圖 6C 顯示使用特異性探針檢測肝臟生檢中的 FIX 和 ARCUS 的雙重原位雜交(ISH)。圖 6D 顯示用於轉導量化的數位 ISH 圖像。圖 6E 顯示藉由 ISH 量化的 FIX 轉基因的轉導效率，並繪製為轉導百分比。

【0185】圖 7A 至 7L 顯示使用特異性探針檢測 NHP 治療後 84 天收集的肝臟生檢中的 FIX 和 ARCUS 的雙重原位雜交(ISH)；顯示在各種放大視圖(以 AAVhu37.ARCUS2 及 AAVhu37.供體-HDR-hFIX 治療的

NHP)。圖 7A 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS。圖 7B 顯示放大 4 倍觀察到的肝肝臟生檢中 ISH 檢測的 hFIX。圖 7C 顯示放大 4 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7D 顯示放大 4 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI(細胞核染色)的疊加圖像。圖 7E 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS。圖 7F 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 7G 顯示放大 10 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7H 顯示放大 10 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI(細胞核染色)的疊加圖像。圖 7I 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 ARCUS 表現。圖 7J 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 7K 顯示放大 20 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 的疊加圖像。圖 7L 顯示放大 20 倍觀察到的 ISH 檢測的 ARCUS 和 hFIX 作為具有 DAPI(細胞核染色)的疊加圖像。載體轉導(GC/二倍體基因體)的總結如下表 1 所示。

表 1.

	GC/二倍體基因體
hFIX	0.63
ARCUS	0.13
比率(FIX/ARCUS)	4.8
載體劑量比率(FIX/ARCUS)	3.0

【0186】圖 8A 至 8M 顯示使用特異性探針檢測 NHP 治療後 84 天收集的肝臟生檢中的 FIX 和 ARCUS 之雙重原位雜交 (ISH)；顯示在各種放大視圖 (以 AAVhu37.EGFP 及 AAVhu37.Donor-HDR-hFIX.U6.sgR 治療的 NHP)。圖 8A 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE。圖 8B 顯示放大 4 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8C 顯示放大 4 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8D 顯示放大 4 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。圖 8E 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE。圖 8F 顯示放大 10 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8G 顯示放大 10 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8H 顯示放大 10 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。圖 8I 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 GFP-WRPE 表現。圖 8J 顯示放大 20 倍觀察到的肝臟生檢中 ISH 檢測的 hFIX。圖 8K 顯示放大 20 倍觀察到的 ISH 檢測的 GFP-WRPE 和 hFIX 的疊加圖像。圖 8L 顯示放大 20 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。圖 8M 顯示在未治療的對照中放大 20 倍觀察到的 ISH 檢測的 GFP-WPRE 和 hFIX 作為具有 DAPI (細胞核染色) 的疊加圖像。載體轉導 (GC/二倍體基因體) 的總結如下表 2 所示。

表 2.

	GC/二倍體基因體
hFIX	0.017
ARCUS	0.006
比率(FIX/ARCUS)	2.7
載體劑量比率(FIX/ARCUS)	3.0

【0187】圖 9 顯示以 AAVhu37.ARCUS2 和 AAVhu37.供體-HDR-hFIX 治療的 NHP 中 ARCUS 媒介的靶向編輯。在治療後 84 天，收集肝臟生檢樣品，並計算存在的標靶區域中總插入缺失的百分比。此外，以 AAVhu37.ARCUS2 和 AAVhu37.供體-HDR-hFIX 處理的 NHP 中 ARCUS 媒介的靶向編輯。在治療後 84 天，收集肝臟生檢樣品，並計算存在的標靶區域中總插入缺失的頻率，繪製成相對於標靶的唯一 UMI OT 讀數的頻率。藉由擴增子測序量化的插入缺失總結如下表 3 所示。

表 3.

ID	20-196 (NB, d84)	20-196 (NB, d366)	AH0120 (3m, d98)	RA3567 (ctl)
插入% (Ins %)	1.1	1.3	1.2	0.00
缺失% (Del %)	9.2	10.9	13.2	0.04
總插入或缺失% (Total Indel %)	10.2	12.1	14.4	0.04
HDR% (LMU-seq)	2.8	2.0	1.2	

## 【0188】

實施例 3-新生 NHP 中 ARCUS2 媒介的 H0TC 基因靶向

【0189】新生兒(1-16 日齡)或嬰兒(3-4 月齡)恆河猴用於非 GLP 順應性的 POC 藥理學研究。M2PCSK9 巨型核酸酶靶向存在於人類及恆河猴獼猴 *PCSK9* 基因的 22-

bp 序列。因此，恆河猴獼猴用於評估靶向編輯 (on-target editing)(藥理學)及安全性/毒理學。此外，新生和嬰兒恆河猴具有與人類嬰兒相似的解剖學和生理學特徵，將允許用於預期的臨床 ROA (IV)用途。預計解剖結構和 ROA 的相似性將產生具有代表性的矢量分佈和轉導輪廓，從而能夠更準確地評估測試物的藥理學和毒性，包括靶向和脫靶編輯以及臨床病理學，這在新生小鼠中是不可能的。

【 0190 】 在此項研究中，新生 NHP 被投予了 ARCUS2 核酸酶載體及具有不同長度 HDR 臂(500bp 臂或短 HDR 臂)的供體載體。載體圖示於圖 11I。圖 11A 是顯示來自實驗的數據匯總表。所有 14 隻新生獼猴都良好地耐受載體輸注(即，沒有明顯的臨床後遺症)並且隨著時間的推移體重增加(圖 11E)。肝酶水平在正常範圍內，除了在第 14 天一些動物的 ALT 水平短暫和適度升高(圖 11C)。

【 0191 】 對給藥前從新生兒採集的第 0 天血漿樣品的分析顯示 3 隻動物(21-111、21-113、21-122)對於 AAVrh79 具有高水平( $\geq 400$ )的結合抗體(圖 11A)。此等預先存在的抗 AAVrh79 抗體將阻斷 AAV 基因轉移。

【 0192 】 隨著時間的推移，在所有新生動物中追蹤 PCSK9 水平，包括僅供體的對照動物。第 0 天的 PCSK9 水平在新生兒之間有所不同(圖 11B)。九隻動物在載體投予後表現出 PCSK9 水平降低的趨勢，包括一隻僅供體的對照動物，而其餘五隻動物在給藥後表現出 PCSK9 水平持續或短暫升高(圖 11B)。

【0193】在第 84 天，經由剖腹手術進行肝臟生檢。肝臟中 hOTC 的轉導效率藉由使用 hOTC-及 M2PCSK9-特異性探針的雙重 ISH 檢測轉基因 mRNA，並藉由 OTC 免疫螢光檢測人類 OTC 蛋白，然後在掃描的載玻片上定量(圖 11D)。在給藥時具有預先存在的抗 AAVrh79 結合抗體的三隻動物(21-111、21-113 和 21-122)藉由兩種方法確實顯示出任何 OTC 陽性肝細胞。兩隻僅供體的對照動物表現出低水平( $\leq 1\%$ )的 hOTC 轉導。在接受 AAVrh79.TBG.PI.M2PCSK9.WPRE.bGH 和 AAVrh79.rhHDR.TBG.hOTCco.bGH 供體載體(G6)的兩隻動物中檢測到最高的轉導效率(OTC 免疫螢光法分別為 11.9%及 18.6%)。亦發現陽性的表現 hOTC 的肝細胞以群集存在。此等水平高於使病患受益的閾值，即約 5%的 OTC 表現細胞。

【0194】在第 84 天對來自每隻動物的肝臟生檢樣品進行分子分析以測量每個二倍體基因體的轉基因拷貝數、mRNA 表現水平、靶向編輯及脫靶編輯(圖 11F-11H)。與轉導效率分析一致，第 6 組中的兩隻動物(21-157 及 21-175)具有最高的 hOTC 載體 GC(圖 11F)、hOTC mRNA (圖 11G)及靶向插入或缺失%(圖 11H)。動物中的 M2PCSK9 載體 GC 比 hOTC 載體 GC 低 2 倍至 7 倍，而 M2PCSK9 mRNA 水平比 hOTC mRNA 水平低 23 倍和 765 倍(圖 11F 及 11G)。

【0195】在本研究的第 84 天肝臟生檢樣本中，藉由 ITR-seq 評估的脫靶活性確定了 2 到 40 個潛在的脫靶。在多隻動物中檢測到一些脫靶位點，包括分別在研究 2

和研究 3 中的 hFIX 嬰兒和 hFIX 新生動物。脫靶編輯將進一步以潛在脫靶位點上的擴增子序列為特徵。

【0196】總之，我們確定一種 M2PCSK9 載體及 hOTCco 供體載體組合，當共同投予新生獼猴時，在給藥後 3 個月可在肝臟中達到 12-18.6% 的轉導效率，二者均高於使病患受益的閾值，即約 5% OTC-表現肝細胞。正在對本研究中的動物進行長期效率和安全性評估。我們將在給藥後 1 年進行第二次肝臟生檢，以評估 hOTC 轉導的穩定性、肝臟組織病理學及肝臟中的靶向和脫靶。

#### 【0197】

實施例 4-PCSK9-HE7-KI 小鼠模型

【0198】由於人類和獼猴 *PCSK9* 基因中的 M2PCSK9 靶向序列與鼠類 *Pcsk9* 基因並不保守，我們不能使用 M2PCSK9 在小鼠基因體基因座中進行基因體編輯。因此，我們委託 Jackson Laboratory 生成敲入(knock-in)小鼠模型，該模型將包括鼠類 *Pcsk9* 基因外顯子 7 的區域置換成包含外顯子 7 的人類 *PCSK9* 基因的區域，命名為 *PCSK9-hE7-KI* 小鼠(圖 10A-10B)。此模型可用於評估體內基因體編輯和基因靶向效率。然後，我們將 *PCSK9-hE7-KI* 小鼠與 sparse fur ash (*spf<sup>ash</sup>*)小鼠雜交。*spf<sup>ash</sup>* 小鼠在 *Otc* 基因外顯子 4 末端處的剪接供體位點具有 G 到 A 的點突變，其導致 *Otc* mRNA 的異常剪接和 *OTC* mRNA 和蛋白質表現皆減少 20 倍(Hodges and Rosenberg, 1989)。受影響的動物有 5-10% 的殘留 OTC 活性，並可

通過食物生存，但牠們會出現高氨血症，在高蛋白飲食時可能是致命的 (Yang et al., 2016)。

【0199】 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠模型可用於評估人類 OTC 體內基因靶向的功效，並證實靶向效率和功效的相關性。然而，由於新生小鼠體型較小，血液臨床病理學和基因靶向臨床療效的評價只能在小鼠斷奶後，一旦牠們達到足夠的體重，並作為終末程序進行。

【0200】圖 12 顯示 265 bp 序列的序列比對代表人類 PCSK9 序列的 PCSK9-hE7 敲入等位基因、小鼠 PCSK9 (mPCSK9) 及恆河猴 PCSK9 (rhPCSK9)。在此 265 bp 區域中，人類和恆河猴序列之間有 6 個錯配。由於插入了各種 LINE 和 LTR，嚙齒動物和靈長類動物的序列分歧超出了這個窗口。人類和小鼠之間的外顯子 7 存在 2 個胺基酸差異。藉由測定的 ELISA，hE7-KI 小鼠表現正常水平的 mPCSK9。

#### 【0201】

實施例 5-在 PCSK9-HE7-KI.SPF<sup>ASH</sup> PUPS 中靶向 *PCSK9* 基因座的體內 OTC 基因

【0202】這項正在進行的非 GLP 順應性的藥理學研究旨在評估在新生 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠中，人類 OTC 基因的 M2PCSK9 巨型核酸酶媒介的敲入是否可在經由預期的臨床 ROA (IV) 的單次共同投予 M2PCSK9 核酸酶表現載體與人類 OTC 供體載體後在標靶組織中達到治療性人類 OTC 表現來用於治療 OTC 缺乏症 (肝臟)。實驗設計的示意圖顯示於圖 14A 中，劑量組顯示於圖 14B 中。

【0203】在第0天，將劑量為  $1.0 \times 10^{13}$  GC/kg 的表現 M2PCSK9 巨型核酸酶之 AAVrh79 載體 (AAVrh79.TBG.PI.M2PCSK9.WPRE.bGH) 與劑量為  $3.0 \times 10^{13}$  GC/kg 的三種不同 AAVrh79 hOTCco 供體載體中的一種的組合以 IV 共同投予新生 (PND 1-2) 雄性 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠。在此研究中評估的 M2PCSK9 巨型核酸酶表現載體 (AAVrh79.TBG.PI.M2PCSK9.WPRE.bGH) 與主要臨床候選者相同，而各 hOTCco 供體載體與主要臨床候選者相同，除了 HDR 臂之外。具體而言，雖然臨床候選者包括人類 HDR 序列的長版本 (AAVrh79.hHDR.TBG.hOTCco.bGH)，但此研究中評估的 hOTCco 供體載體包括小鼠-人類雜交 HDR 序列 (AAVrh79.mhHDR.TBG.hOTCco.bGH)、人類 HDR 序列的較短版本 (AAVrh79.shHDR.TBG.hOTCco.bGH) 或沒有 HDR 序列 (AAVrh79.TBG.hOTCco.bGH)。圖 13 顯示 HDR 臂與人類、敲入小鼠和 NHP 序列的同源性比較。作為陰性對照，向另外的年齡匹配的 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠投予不表現巨型核酸酶的 AAVrh79 載體 (AAVrh79.TBG.PI.EGFP.WPRE.bGH) 與 AAVrh79.mhHDR.TBG.hOTCco.bGH 的組合。

【0204】生存期間的評估包括每天進行的活力監測、體重測量、高蛋白飲食挑戰後的血漿 PCSK9 及血漿 NH<sub>3</sub> 和尿乳清酸水平之評估，以及在第 120 天進行部分肝切除術，以評估在三分之二肝臟部分切除術後的人類 *OTC* 轉導的穩定性。在第 49 天及第 170 天，各群組的

一個子群接受為期 10 天的高蛋白飲食挑戰，然後在挑戰結束時進行屍檢。在屍檢時，收集肝臟以評估人類 *OTC* 基因的敲入，包括評估人類 *OTC* mRNA 表現(原位雜交)、*OTC* 蛋白表現(免疫染色)和藉由染色及/或酶活性測定所評估的 *OTC* 酶活性。分離肝臟 DNA 以評估靶向編輯(擴增子序列、牛津奈米孔長讀取定序(Oxford nanopore long-read sequencing))和評估載體基因體拷貝。

【0205】初步結果顯示，以具有 mhHDR 臂的載體給藥的小鼠顯示出與野生型小鼠相當的存活率，在 10 天的高蛋白飲食挑戰後，以 shHDR 治療的小鼠達到 80% 的存活率(圖 14C)。所有經治療的小鼠都比未經治療的 *KI-spf-ash* 小鼠維持更好的體重(圖 14D)。與未經治療的小鼠相比，經 mHDR 治療的小鼠的血漿氨水平顯著地降低(圖 14E)。

【0206】在第 48 天測量 mPCSK9 水平，所有治療的小鼠都顯示出降低(圖 14F)。插入或缺失百分比在 HDR 類型中相當一致(圖 14G)。在以 shHDR 及 mhHDR 治療的小鼠中 h*OTC* 水平增加(圖 14H)。

#### 【0207】

實施例 6-靶向新生恆河猴 *PCSK9* 基因座的體內 *OTC* 基因

【0208】這項正在進行的非 GLP 順應性的藥理學研究旨在評估在新生恆河猴中，人類 *OTC* 基因的 M2PCSK9 巨型核酸酶媒介的敲入是否可在經由預期的

臨床 ROA (IV)的單次共同投予 M2PCSK9 巨型核酸酶表現載體與人類 OTC 供體載體後在標靶組織中達到治療性人類 OTC 表現來用於治療 OTC 缺乏症(肝臟)。

【0209】在第 0 天，將劑量為  $1.0 \times 10^{13}$  GC/kg 的二種不同的表現 M2PCSK9 巨型核酸酶的載體中的一種與劑量為  $3.0 \times 10^{13}$  GC/kg 的二種不同 AAV hOTCco 供體載體中的一種的組合以 IV 共同投予新生(1 至 16 日齡)恆河猴。僅接受劑量為  $3.0 \times 10^{13}$  GC/kg 的 AAV hOTCco 供體載體的非核酸酶組被包括作為僅供體的對照組。

【0210】對於靶向 PCSK9 基因的 AAV 載體，我們比較了兩種在肝臟中表現 M2PCSK9 的 AAV 載體構建體。AAV.TBG.PI.M2PCSK9.WPRE.bGH 含有全長 TBG 啟動子及增強子元件的二個拷貝，且 WPRE 表現的核酸酶水平高於 AAV.TBG-S1-F113.PI.M2PCSK9.bGH (其含有一個短而弱的啟動子)。對於 hOTC 供體載體，我們比較了兩個 AAV.hOTCco 供體載體，它們在 hOTCco 轉基因匣兩側的同源臂長度不同。

【0211】NHP 在第 0 天靜脈投予兩種載體，且每天監測其活力。生存期間評估包括體重測量、血液臨床病理學和血漿基因編輯分析。計劃進行兩次剖腹手術以分離肝組織，用於分析基因體編輯效率、載體基因體拷貝、轉基因表現、組織病理學、免疫染色和 RNA ISH 染色。NHP 將被長期追蹤並進行屍檢(日期待定)，屆時將收集來自肝臟和其它主要器官的組織，用於評估基因體編輯效率、載體基因體拷貝、轉基因表現、組織病理學、免疫染色及 RNA ISH 染色。

## 【 0212 】

實施例 7-評估在 PCSK9-HE7-KI.SPF<sup>ASH</sup> PUPS 中的功效及確定載體之比例

【 0213 】 這項計劃的非 GLP 順應性的藥理學研究旨在評估經由預期的臨床 ROA (IV)的單次共同投予 M2PCSK9 核酸酶表現載體與人類 *OTC* 供體載體後，在新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠中達到用於治療 *OTC* 缺乏症(肝臟)之人類 *OTC* 基因的 M2PCSK9 巨型核酸酶媒介的敲入在最高功效時所需的載體組分比例。

【 0214 】 在第 0 天，將三種劑量中的一種的表現 M2PCSK9 巨型核酸酶的 AAVrh79 載體 (AAVrh79.TBG.PI.M2PCSK9.WPRE.bGH)與三種劑量中的一種的包括小鼠-人類雜交 HDR 序列的 hOTCco 供體載體 (AAVrh79.mhHDR.TBG.hOTCco.bGH)的組合以 IV 共同投予新生 (PND 1-2)雄性 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠。在此研究中所評估的 M2PCSK9 巨型核酸酶表現載體 (AAVrh79.TBG.PI.M2PCSK9.WPRE.bH)與臨床候選者相同，而 hOTCco 供體載體與臨床候選者相同，除了 HDR 臂之外。具體而言，雖然臨床候選者包括人類 HDR 序列的長版本 (AAVrh79.hHDR.TBG.hOTCco.bGH)，但此研究中評估的 hOTCco 供體載體包括小鼠-人類雜交 HDR 序列 (AAVrh79.mhHDR.TBG.hOTCco.bGH)。

【 0215 】 選擇在供體載體中的小鼠-人類雜交 HDR 序列 (AAVrh79.mhHDR.TBG.hOTCco.bGH)用於此研究，以評估這種方法的藥理學，其中供體序列與 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠中的序列直接同源。

【0216】生存期間的評估包括每天進行的活力監測、體重測量、高蛋白飲食挑戰後血漿  $\text{NH}_3$  和尿乳清酸水平的評估。在第 81 天，小鼠將接受為期 10 天的高蛋白飲食挑戰，然後在挑戰結束時進行屍檢。在屍檢時，收集肝臟以評估人類 *OTC* 基因的敲入，包括評估人類 *OTC* mRNA 表現(原位雜交)、*OTC* 蛋白表現(免疫染色)、及藉由染色及/或酶活性測定所評估的 *OTC* 酶活性。亦分離肝臟 DNA 以評估靶向編輯(擴增子序列)及評估載體基因體拷貝。

【0217】

實施例 8-評估在 PCSK9-HE7-KI.SPF<sup>ASH</sup> PUPS 中的功效及確定最小有效劑量

【0218】這項計劃的 GLP 順應性的藥理學研究旨在評估新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠模型中 IV 投予 AAV 的療效並確定 MED。表現 M2PCSK9 巨型核酸酶的 AAVrh79 載體 (AAVrh79.TBG.PI.M2PCSK9.WPRE.bGH) 將是為計劃的 GLP 順應性的毒理學研究所製造的毒理學載體批次。本研究不使用包括人類 HDR 序列長版本 (AAVrh79.hHDR.TBG.hOTCco.bGH) 測試物，而是利用包括小鼠-人類雜交 HDR 序列的 hOTCco 供體載體 (AAVrh79.mhHDR.TBG.hOTCco.bGH)。此載體將以與臨床候選者之毒理學載體批次相當的方法製造。

【0219】我們已選擇於本研究中使用供體載體中具有小鼠-人類雜交 HDR 序列 (AAVrh79.mhHDR.TBG.hOTCco.bGH)，以使我們能夠有效地研究這種方法的藥

理學，其中供體序列與 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠中的序列直接同源。

【0220】此研究將評估 N=60 隻新生兒(PND 1-2)新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠且 N=15 隻年齡匹配的雄性 PCSK9-hE7-KI.WT (野生型)作為對照。研究將包括一個屍檢時間點(90 天)。對於功效評估，小鼠將從第 81 天到第 90 天接受為期 10 天的高蛋白飲食。將評估生存、身體狀況及生物標誌物變化。將使用 IV 投予來評估 AAV 的三種劑量水平。將根據先前非臨床研究中評估的劑量範圍選擇劑量水平。評估的劑量水平將以括號包含預期的臨床劑量。

【0221】生存期間的評估將包括每天的活力檢查、存活監測、體重測量、高蛋白飲食挑戰後血清 PCSK9 水平、血漿 NH<sub>3</sub> 及尿乳清酸水平的評估。屍檢將在第 90 天進行。在屍檢時，將收集血液用於 CBC/差異和血清臨床化學分析。將收集表列的組織用於組織病理學評估。收集肝臟以評估人類 *OTC* 基因的敲入，包括評估人類 *OTC* mRNA 表現(原位雜交)、*OTC* 蛋白表現(免疫染色)、及藉由染色及/或酶活性測定所評估的 *OTC* 酶活性。亦分離肝臟 DNA 以評估靶向編輯(擴增子序列)及評估載體基因體拷貝。

【0222】MED 將基於下述而確定：相較於媒劑治療的新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 對照小鼠，以 AAV 治療的新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠高蛋白飲食後的存活、高蛋白飲食挑戰結束時的血漿 NH<sub>3</sub> 水平、人類 *OTC* mRNA 及蛋白質表現、*OTC* 酶活性、及靶向編輯。

## 【 0223 】

實施例 9-PCSK9-HE7-KI.SPF<sup>ASH</sup> PUPS 中的毒理學研究

【 0224 】 將在新生 (PND 1-2) PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠中進行為期 6 個月的 GLP 順應性安全性研究，以研究 IV 共同給藥後測試物的安全性、耐受性、藥理學和藥物動力學。期間分析，包括靶向編輯、脫靶編輯、轉基因表現和組織病理學分析，將在第 60 天和第 180 天進行，因為這些時間點將使核酸酶依賴性基因插入有足夠的時間在給藥後達到穩定的平台水平 (plateau level)。新生 PCSK9-hE7-KI.*spf<sup>ash</sup>* 小鼠將接受三種劑量水平中的一種的測試物 ( $1.0 \times 10^{12}$  GC/kg 核酸酶載體與  $3.0 \times 10^{12}$  GC/kg 供體載體、 $3.3 \times 10^{12}$  GC/kg 核酸酶載體與  $1.0 \times 10^{13}$  GC/kg 供體載體、或  $1.0 \times 10^{13}$  GC/kg 核酸酶載體與  $3.0 \times 10^{13}$  GC/kg；每劑量 N=20) 或媒劑 (磷酸鹽緩衝食鹽水 [PBS]；N=20)。在測試物或媒劑投予後，生存期間的評估將包括臨床觀察每天監測痛苦和異常行為的跡象、體重測量及血液臨床血清化學 (特別是 ALT、AST 和總膽紅素)。

【 0225 】 在測試物投予後第 60 天，將群組 1、3、5 和 7 安樂死，並將對包括但不限於腦、脊髓、心臟、肝臟、脾臟、腎臟、肺臟、生殖器官、腎上腺和淋巴結的綜合組織列表進行組織病理學分析。器官將酌情稱重。

【 0226 】 將收集並分析肝臟樣品用於藉由擴增子序列和 AMP 序列的靶向編輯、藉由 ITR 序列和擴增子序列的脫靶編輯、載體生物分佈和轉基因表現。在肝臟樣

品中，生物分佈將藉由 PCR 評估，巨型核酸酶 RNA 表現將藉由 RT-PCR 進行分析。將對高度灌注的器官進行巨型核酸酶 RNA 分析，並對具有可檢測到巨型核酸酶 RNA 表現的組織進行評估，以便通過擴增子序列進行靶向編輯。將進一步評估具有可檢測之靶向編輯的組織以進行脫靶編輯。

【0227】對於載體生物分布，將開發特定於雙重載體 M2PCSK9 及 hOTCco 的轉基因的 qPCR 檢測。將使用 AAV 順式質體作為標的序列的替代物來評估測定的效率、線性、精密度、再現性和檢測限度。測定的定量下限 (LLOQ) 將在對測試組織或排泄物進行測定之前確定。將實施驗證計劃，以將轉基因特異性測定與先前進行的驗證研究聯繫起來。測試的基質將包括預期目標，用於生物分佈的肝臟。基質效應將基於在生物分佈研究過程中從測試的所有樣品中加料標靶對照的回收率以及從先前進行的驗證研究中減去的數據進行進一步評估。

#### 【0228】

實施例 10- 在 PCSK9-HE7-KI.LDLR<sup>-</sup>/LDLR<sup>-</sup>.APOBEC<sup>-</sup>/APOBEC<sup>-</sup>PUPS (HOFH 模型) 中藉由 SaCas9 的 HDLR 袖珍基因敲入 PCSK9 基因座

【0229】本研究旨在評估在藉由預期的臨床 ROA (IV) 的單次共同投予 SaCas9 核酸酶表現載體與人類 LDLR 供體載體後，人類 LDLR 基因之 Cas9 媒介的敲入在新生 PCSK9-hE7-KI.lldlr<sup>-</sup>/ldlr<sup>-</sup>.apobec<sup>-</sup>/apobec 小鼠是否可在治療家族性高膽固醇血症的標靶組織(肝臟)中實

現治療性人類 LDLR 表現。使用圖 15 中的實驗設計生成小鼠模型。在小鼠模型中，小鼠 PCSK9 外顯子 7 以人類 PCSK9 外顯子 7 置換，其包含 SaCas9 靶向序列。

【0230】在第 0 天，將劑量為  $1.0 \times 10^{13}$  GC/kg 的表現 Cas9 的 AAVrh79 載體 (AAVrh79.U6.sgR3.PSCK9.APB2.HLP.SaCas9.bGH) 與劑量為  $3.0 \times 10^{13}$  GC/kg 的二種不同 AAVrh79 hLDLR 供體載體中之一種的組合 IV 共同投予新生 *PCSK9-hE7-KI.ldr<sup>-</sup>/ldr<sup>-</sup>.apobec<sup>-</sup>/apobec* 小鼠。圖 16 顯示了使用的載體的示意圖。具體而言，此研究中評估的供體載體之一包括小鼠-人類雜交 HDR 序列 (AAVrh79.mhHDR.hLDLR011)，另一個包括較短版本的人類 HDR 序列 (AAVrh79.shHDR.hLDLR011)。作為陰性對照，其它年齡匹配的 *PCSK9-hE7-KI.spf<sup>ash</sup>* 小鼠投予不表現 saCas9 之 AAVrh79 載體與 AAVrh79.shHDR.hLDLR011 的組合。

【0231】生存期間的評價包括每天進行活力監測，並在第 42、63、90、120 和 150 天評估血清 LDL-c 水平。在第 63 天進行部分肝切除術以評估人類 LDLR 轉導的穩定性，並在第 150 天進行屍檢。在屍檢時，收集肝臟以評估人類 LDLR 基因的敲入，包括評估人類人類 LDLR mRNA 表現(原位雜交)、LDLR 蛋白表現(免疫染色)。分離肝臟 DNA 以評估靶向編輯(擴增子序列，牛津奈米孔長讀取定序)和評估載體基因體拷貝。實驗設計顯示於圖 17。

【0232】初步結果顯示，以 saCas9 與供體載體給藥的小鼠具有顯著降低的血清 LDL 水平。在 2/3 肝部分切除術後 LDL 並沒有變化，表明穩定的整合(圖 18A)。使用 mhHDR 和 shHDR 供體載體時，插入或缺失是一致的(圖 18B)。在第 63 天，shHDR 治療的小鼠顯示出稍高的 hLDLR 水平(圖 18C)，而 mhHDR 和 shHDR(具有 saCas9)載體的血清 LDL 水平相似(圖 18D)。

【0233】圖 19 顯示部分肝切除術後第 63 天肝臟中 hLDLR 表現的免疫組織化學評價。

【0234】本說明書中引用的所有文件藉由引用併入本文，與本文一起提出的序列表的序列和正文藉由引用併入本文。美國臨時專利申請號 63/180,603 (2021 年 4 月 27 日申請)、63/242,474 (2021 年 9 月 9 日申請)、63/244,205 (2021 年 9 月 14 日申請)、63/301,933 (2022 年 1 月 21 日申請)、63/331,385 (2022 年 4 月 15 日申請)各藉由引用以其整體併入本文。儘管已參照特定具體實施例描述本發明，但應理解，可在不背離本發明之精神的情況下進行修改。此類修改旨在落入所附申請專利範圍的範圍內。

#### 【符號說明】

無。

## 序列表

- <110> 賓州大學委員會(The Trustees of the University of Pennsylvania)
- <120> 用於治療遺傳疾病的體內核酸酶媒介的基因靶向之組成物及方法
- <130> UPN-21-9650. PCT
- <140> TW 111116003  
<141> 2022-04-27
- <150> 63/180, 603  
<151> 2021-04-27
- <150> 63/242, 474  
<151> 2021-09-09
- <150> 63/244, 205  
<151> 2021-09-14
- <150> 63/301, 933  
<151> 2022-01-21
- <150> 63/331, 385  
<151> 2022-04-15
- <160> 80
- <170> PatentIn版本3.5
- <210> 1  
<211> 7429  
<212> DNA  
<213> 人工序列
- <220>  
<223> 生產質體  
U6. sgR. PCSK9. ABP2. TBG-S1. hSaCas9 (pX601+3NLS) . bGH
- <220>  
<221> LTR  
<222> (1)..(168)  
<223> AAV ITR
- <220>  
<221> terminator  
<222> (193)..(199)  
<223> U6終止子
- <220>  
<221> misc\_feature  
<222> (199)..(275)  
<223> gRNA支架
- <220>  
<221> misc\_feature  
<222> (276)..(295)  
<223> PCSK9 sgRNA

<220>  
 <221> promoter  
 <222> (297).. (545)  
 <223> U6啟動子

<220>  
 <221> enhancer  
 <222> (557).. (656)  
 <223>  $\alpha$  mic/bik增強子

<220>  
 <221> enhancer  
 <222> (663).. (762)  
 <223>  $\alpha$  mic/bik增強子

<220>  
 <221> promoter  
 <222> (777).. (952)  
 <223> TBG-S1

<220>  
 <221> misc\_feature  
 <222> (956).. (958)  
 <223> 終止密碼子突變

<220>  
 <221> misc\_feature  
 <222> (973).. (1017)  
 <223> SV40 NLS

<220>  
 <221> primer\_bind  
 <222> (1160).. (1179)  
 <223> SA2-Seq1

<220>  
 <221> primer\_bind  
 <222> (1647).. (1666)  
 <223> SA2-Seq2

<220>  
 <221> primer\_bind  
 <222> (2170).. (2189)  
 <223> SA2-Seq3

<220>  
 <221> primer\_bind  
 <222> (2680).. (2700)  
 <223> SA2-Seq4

<220>  
 <221> primer\_bind  
 <222> (3130).. (3149)  
 <223> SA2-Seq5

<220>

<221> misc\_feature  
 <222> (4174).. (4221)  
 <223> NLS

<220>  
 <221> polyA\_signal  
 <222> (4233).. (4435)  
 <223> bGH poly A

<220>  
 <221> LTR  
 <222> (4456).. (4619)  
 <223> AAV ITR

<220>  
 <221> misc\_feature  
 <222> (5382).. (6239)  
 <223> Amp-R

<220>  
 <221> misc\_feature  
 <222> (6413).. (7001)  
 <223> 起始序列

<400> 1  
 ctgocgcctc gctogctcac tgaggccgcc cgggcaaagc cggggcgtcg ggcgaccttt 60  
 ggtogcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120  
 aggggttctt ttagttaa gattaaccgg ccatgctact tatctactta aggctatttc 180  
 tagctctaaa acaaaaaaat ctgccaaca agttgacgag ataaacacgg cttttgcct 240  
 tgttttagta gattctgttt ccagagtact aaaacgggac tttggggacc aacttcgggtg 300  
 tttogtcctt tccacaagat atataagcc aagaaatoga aatactttca agttacggta 360  
 agcatatgat agtccatttt aaaacataat tttaaaactg caaactacc cagaattat 420  
 tactttctac gtcacgtatt ttgtactaat atctttgtgt ttacagtcaa attaattcca 480  
 attatctctc taacagcctt gtatcgtata tgcaaatatg aaggaatcat gggaaatagg 540  
 ccctccttaa gctagcaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg 600  
 cagcatttac tctctctgtt tgctctggtt aataatctca ggagcacaaa cattccagat 660  
 ccaggttaat ttttaaaaag cagtcaaaaag tccaagtggc ctttggcagc atttactctc 720  
 totgtttgct ctggttaata atctcaggag cacaaacatt ccagatccgg cgcgccactc 780  
 aaagttcaaa ccttatcatt ttttgctttg ttctcttgg ccttggtttt gtacatcagc 840  
 tttgaaaata ccatcccagg gttaatgctg gggttaattt ataactaaga gtgctctagt 900  
 tttgcaatac aggacatgct ataaaaatgg aaagatgttg ctttctgaga gacagtgacc 960

gccaccatgg	ccccaaagaa	gaagcggag	gtcggtatcc	acggagtccc	agcagccaag	1020
cggaactaca	tccatgggct	ggacatcggc	atcaccagcg	tgggctacgg	catcatcgac	1080
tacgagacac	gggacgtgat	cgatgccggc	gtgcggctgt	tcaaagaggc	caacgtggaa	1140
aacaacgagg	gcagcgggag	caagagaggc	gccagaaggc	tgaagcggcg	gaggcggcat	1200
agaatccaga	gagtgaagaa	gctgctgttc	gactacaacc	tgctgaccga	ccacagcgag	1260
ctgagcggca	tcaaccocct	cgaggccaga	gtgaagggcc	tgagccagaa	gctgagcgag	1320
gaagagttct	ctgccgcctt	gctgcacctg	gccaagagaa	gaggcgtgca	caacgtgaac	1380
gaggtggaag	aggacaccgg	caacgagctg	tccaccaaaag	agcagatcag	cgggaacagc	1440
aaggccctgg	aagagaaata	cgtggccgaa	ctgcagctgg	aacggctgaa	gaaagacggc	1500
gaagtgcggg	gcagcatcaa	cagattcaag	accagcagct	acgtgaaaga	agccaaacag	1560
ctgctgaagg	tgcagaaggc	ctaccaccag	ctggaccaga	gcttcatcga	cacctacatc	1620
gacctgctgg	aaaccocggc	gacctactat	gagggacctg	gcgagggcag	ccccttcggc	1680
tggaaggaca	tcaaagaatg	gtacgagatg	ctgatgggcc	actgcacctt	cttccccgag	1740
gaactgcgga	gcgtgaagta	cgctacaac	gccgacctgt	acaacgccct	gaacgacctg	1800
aacaatctcg	tgatcaccag	ggacgagaac	gagaagctgg	aatattacga	gaagttccag	1860
atcatcgaga	acgtgttcaa	gcagaagaag	aagccccacc	tgaagcagat	cgccaaagaa	1920
atcctcgtga	acgaagagga	tattaagggc	tacagagtga	ccagcaccgg	caagcccagag	1980
ttaccaaac	tgaagggtga	ccacgacatc	aaggacatta	cgcgccggaa	agagattatt	2040
gagaacgccg	agctgctgga	tcagattgcc	aagatcctga	ccatctacca	gagcagcgag	2100
gacatccagg	aagaactgac	caatctgaac	tccgagctga	cccaggaaga	gatcagcgag	2160
atctctaate	tgaagggcta	taccggcacc	cacaacctga	gcctgaaggc	catcaacctg	2220
atcctggacg	agctgtggca	caccaacgac	aaccagatcg	ctatcttcaa	cgggctgaag	2280
ctggtgcccc	agaagggtga	cctgtcccag	cagaaagaga	tccccaccac	cctggtggac	2340
gacttcatcc	tgagccccgt	cgtgaagaga	agcttcatcc	agagcatcaa	agtgatcaac	2400
gccatcatca	agaagtacgg	cctgccccaac	gacatcatta	tcgagctggc	cgcgagaag	2460
aactccaagg	acgccagaa	aatgatcaac	gagatgcaga	agcggaaaccg	gcagaccaac	2520
gagcggatcg	aggaaatcat	ccggaccacc	ggcaaagaga	acgccaagta	cctgatcgag	2580
aagatcaagc	tgcaacgacat	gcaggaaggc	aagtgcctgt	acagcctgga	agccatccct	2640
ctggaagatc	tgctgaacaa	ccccttcaac	tatgaggtgg	accacatcat	ccccagaagc	2700

gtgtccttcg	acaacagctt	caacaacaag	gtgtctcgtga	agcaggaaga	aaacagcaag	2760
aagggcaacc	ggacccatt	ccagtacctg	agcagcagcg	acagcaagat	cagctacgaa	2820
accttcaaga	agcacatcct	gaatctggcc	aagggcaagg	gcagaatcag	caagaccaag	2880
aaagagtatc	tgctggaaga	acgggacatc	aacaggttct	ccgtgcagaa	agacttcatc	2940
aaccggaacc	tggtggatac	cagatacgcc	accagaggcc	tgatgaacct	gctgcggagc	3000
tacttcagag	tgaacaacct	ggacgtgaaa	gtgaagtcca	tcaatggcgg	cttcaccagc	3060
tttctgcggc	ggaagtggaa	gtttaagaaa	gagcggaaaca	aggggtacaa	gcaccacgcc	3120
gaggacgccc	tgatcattgc	caacgccgat	ttcatcttca	aagagtggaa	gaaactggac	3180
aaggccaaaa	aagtgatgga	aaaccagatg	ttcgaggaaa	agcaggccga	gagcatgccc	3240
gagatcgaaa	ccgagcagga	gtacaaagag	atcttcatca	ccccccacca	gatcaagcac	3300
attaaggact	tcaaggacta	caagtacagc	caccgggtgg	acaagaagcc	taatagagag	3360
ctgattaacg	acaccctgta	ctccaccogg	aaggacgaca	agggcaacac	cctgatcgtg	3420
aaaatctga	acggcctgta	cgacaaggac	aatgacaagc	tgaaaaagct	gatcaacaag	3480
agccccgaaa	agctgctgat	gtaccaccac	gacccccaga	cctaccagaa	actgaagctg	3540
attatggaac	agtacggcga	cgagaagaat	cccctgtaca	agtactacga	ggaaaccggg	3600
aactacctga	ccaagtactc	caaaaaggac	aacggccccg	tgatcaagaa	gattaagtat	3660
tacggcaaca	aactgaacgc	ccatctggac	atcaccgacg	actaccccaa	cagcagaaac	3720
aaggtcgtga	agctgtccct	gaagccctac	agattcgacg	tgtacctgga	caatggcgtg	3780
tacaagttcg	tgaccgtgaa	gaatctggat	gtgatcaaaa	aagaaaacta	ctacgaagtg	3840
aatagcaagt	gctatgagga	agctaagaag	ctgaagaaga	tcagcaacca	ggccgagttt	3900
atgcctctct	tctacaacaa	cgatctgata	aagatcaacg	gcgagctgta	tagagtgatc	3960
ggcgtgaaca	acgacctgct	gaaccggatc	gaagtgaaca	tgatcgacat	cacctaccgc	4020
gagtacctgg	aaaacatgaa	cgacaagagg	ccccccagga	tcattaagac	aatgcctcc	4080
aagaccgaga	gcattaagaa	gtacagcaca	gacattctgg	gcaacctgta	tgaagtgaaa	4140
totaagaagc	accctcagat	catcaaaaag	ggcaaaaaggc	ctgctgccac	caaaaaggcc	4200
ggccaggcaa	aaaagaaaaa	gtaatgacta	gtgcctcgac	tgtgccttct	agttgccagc	4260
catctgttgt	ttgccctcc	cccgtgcctt	ccttgacct	ggaaggtgcc	actcccactg	4320
tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	4380

tgggggtgg ggtgggag gacagcaagg gggaggattg ggaagacaat agcaggcatg	4440
ctggggactc gagtagataa gtagcatggc gggttaatca ttaactacaa ggaaccoccta	4500
gtgatggagt tggccactcc ctctctgogc gctcgcctgc tcaactgaggc cgggagacca	4560
aaggtcgccc gacgcccggg ctttgcccgg gcgccctcag tgagcgagcg agcgcgcagc	4620
cttaattaac ctaattcact ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc	4680
gttaccacac ttaatgcct tgcagcacat cccctttcg ccagctggcg taatagcgaa	4740
gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc tgaatggcga atgggagcg	4800
ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca	4860
cttgccagcg ccctagcgcc cgtctctttc gctttcttcc cttcctttct cggcacgttc	4920
gocggctttc ccogtcaagc tctaaatcgg gggctccctt tagggttcog atttagtgct	4980
ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg	5040
ccctgataga cggtttttcg ccttttgacg ttggagtcca cgttctttaa tagtggactc	5100
ttgttccaaa ctggaacaac actcaacct atctcggctt attcttttga tttataaggg	5160
attttgccga tttoggccta ttggttaaaa aatgagctga tttacacaaa atttaacgag	5220
aattttaaca aaatattaac gcttacaatt taggtggcac ttttcgggga aatgtgcgag	5280
gaaccocctat ttgtttatit ttctaaatac attcaaatat gtatccgctc atgagacaat	5340
aacctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc	5400
gtgtcgccct tattcccttt tttgogcat tttgccttcc tgtttttgct caccagaaa	5460
cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgac	5520
tggatctcaa cagcgtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga	5580
tgagcacttt taaagttctg ctatgtggcg cggattatc ccgtattgac gccgggcaag	5640
agcaactcgg tcgccgata cactattctc agaatgactt ggttgagtac tcaccagtca	5700
cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca	5760
tgagtataa cactgoggcc aactacttc tgacaacgat cggaggaccg aaggagctaa	5820
ccgctttttt gcacaacatg ggggatcatg taactgcct tgatcgttgg gaaccggagc	5880
tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa	5940
cgttgcgcaa actattaact ggogaactac ttactctagc ttcccggcaa caattaatag	6000
actggatgga ggcggataaa gttgcaggac cactctgog ctcggccctt ccggctggct	6060
ggtttattgc tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac	6120

tggggccaga tggaagccc tcccgatatcg tagttatcta cacgacgggg agtcaggcaa 6180  
 ctatggatga acgaaataga cagatcgctg agatagggtc ctcaactgatt aagcattggt 6240  
 aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt cttttttaat 6300  
 ttaaaaggat ctaggatgaag atcctttttg ataatctcat gaccaaaatc cottaacgtg 6360  
 agttttcgtt ccaactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 6420  
 ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg 6480  
 tttgtttgcc ggatcaagag ctaccaactc tttttcogaa ggtaactggc ttcagcagag 6540  
 cgcagatacc aaatactggt cttctagtgt agccgtagtt aggccaccac ttcaagaact 6600  
 ctgtagcacc gcctacatac ctgcctctgc taatcctggt accagtggct gctgccagtg 6660  
 gcgataagtc gtgtottacc gggttggact caagacgata gttaccggat aaggcgcagc 6720  
 ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 6780  
 aactgagata cctacagcgt gagctatgag aaagcggcac gcttcccga gggagaaagg 6840  
 cggacaggta tcoggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 6900  
 ggggaaacgc ctggtatctt tatagtctcg tcgggtttcg ccacctctga cttgagcgtc 6960  
 gatttttgtg atgctogtca ggggggogga gcctatggaa aaacgccagc aacgcggcct 7020  
 ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 7080  
 ctgattctgt ggataaccgt attaccgctt ttgagtgagc tgataccgct cgcgcagcc 7140  
 gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac 7200  
 cgctctctcc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact 7260  
 ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcaactcat taggcacccc 7320  
 aggctttaca ctttatgctt ccggctogta tgttgtgtgg aattgtgagc ggataacaat 7380  
 ttcacacagg aaacagctat gaccatgatt acgccagatt taattaagg 7429

<210> 2  
 <211> 4619  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 載體基因體 U6. sgR. PCSK9. ABP2. TBG-S1. hSaCas9 (pX601+3NLS). bGH

<220>  
 <221> LTR

<222> (1)..(168)  
 <223> AAV ITR  
  
 <220>  
 <221> terminator  
 <222> (193)..(199)  
 <223> U6終止子  
  
 <220>  
 <221> misc\_feature  
 <222> (199)..(275)  
 <223> shRNA 支架  
  
 <220>  
 <221> misc\_feature  
 <222> (276)..(295)  
 <223> PCSK9 sgRNA  
  
 <220>  
 <221> promoter  
 <222> (297)..(545)  
 <223> U6 啟動子  
  
 <220>  
 <221> enhancer  
 <222> (557)..(656)  
 <223>  $\alpha$  mic/bik 增強子  
  
 <220>  
 <221> enhancer  
 <222> (663)..(762)  
 <223>  $\alpha$  mic/bik 增強子  
  
 <220>  
 <221> promoter  
 <222> (777)..(952)  
 <223> TBG-S1  
  
 <220>  
 <221> misc\_feature  
 <222> (956)..(958)  
 <223> 終止密碼子突變  
  
 <220>  
 <221> CDS  
 <222> (967)..(4221)  
 <223> SaCas9 (px601+3LNS)  
  
 <220>  
 <221> misc\_feature  
 <222> (973)..(1017)  
 <223> SV40 NLS  
  
 <220>  
 <221> primer\_bind  
 <222> (1160)..(1179)  
 <223> Sa2-Seq1

<220>  
 <221> primer\_bind  
 <222> (1647)..(1666)  
 <223> Sa2-Seq2

<220>  
 <221> primer\_bind  
 <222> (2170)..(2189)  
 <223> Sa2-Seq3

<220>  
 <221> primer\_bind  
 <222> (2680)..(2700)  
 <223> Sa2-Seq4

<220>  
 <221> primer\_bind  
 <222> (3130)..(3149)  
 <223> Sa2-Seq5

<220>  
 <221> misc\_feature  
 <222> (4174)..(4221)  
 <223> NLS

<220>  
 <221> polyA\_signal  
 <222> (4233)..(4435)  
 <223> bGH polyA

<220>  
 <221> LTR  
 <222> (4456)..(4619)  
 <223> AAV ITR

<400> 2  
 ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggctcog ggcgaccttt 60  
 ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120  
 aggggttccct ttagttaa gattaaccgg ccatgctact tatctactta aggctatttc 180  
 tagctctaaa acaaaaaaat ctgccaaca agttgacgag ataaacacgg cattttgcct 240  
 tgttttagta gattctgttt ccagagtact aaaacgggac tttggggacc aacttcggtg 300  
 tttcgtcctt tccacaagat atataagcc aagaaatoga aatactttca agttacggtg 360  
 agcatatgat agtccatttt aaaacataat tttaaaactg caaactacc aagaaattat 420  
 tactttctac gtcacgtatt ttgtactaat atctttgtgt ttacagtcaa attaattcca 480  
 attatctctc taacagcctt gtatcgtata tgcaaatatg aaggaatcat gggaaatagg 540  
 ccctccttaa gctagcaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg 600

cagcatttac tctctctggtt tgctctggtt aataatctca ggagcacaaa cattccagat 660  
ccaggttaat ttttaaaaag cagtcaaaaag tccaagtggc ccttggcagc atttactctc 720  
tctgtttgct ctggttaata atctcaggag cacaaacatt ccagatccgg cgcgccactc 780  
aaagttcaaa ccttatcatt ttttgctttg ttctctttgg ccttggtttt gtacatcagc 840  
tttgaaaata ccatcccagg gttaatgctg gggttaattt ataactaaga gtgctctagt 900  
tttgcaatac aggacatgct ataaaaatgg aaagatgttg ctttctgaga gacagtgacc 960  
gccacc atg gcc cca aag aag aag cgg aag gtc ggt atc cac gga gtc 1008  
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val  
1 5 10  
cca gca gcc aag cgg aac tac atc ctg ggc ctg gac atc ggc atc acc 1056  
Pro Ala Ala Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr  
15 20 25 30  
agc gtg ggc tac ggc atc atc gac tac gag aca cgg gac gtg atc gat 1104  
Ser Val Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp  
35 40 45  
gcc ggc gtg cgg ctg ttc aaa gag gcc aac gtg gaa aac aac gag ggc 1152  
Ala Gly Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly  
50 55 60  
agg cgg agc aag aga ggc gcc aga agg ctg aag cgg cgg agg cgg cat 1200  
Arg Arg Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His  
65 70 75  
aga atc cag aga gtg aag aag ctg ctg ttc gac tac aac ctg ctg acc 1248  
Arg Ile Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr  
80 85 90  
gac cac agc gag ctg agc ggc atc aac ccc tac gag gcc aga gtg aag 1296  
Asp His Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys  
95 100 105 110  
ggc ctg agc cag aag ctg agc gag gaa gag ttc tct gcc gcc ctg ctg 1344  
Gly Leu Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu  
115 120 125  
cac ctg gcc aag aga aga ggc gtg cac aac gtg aac gag gtg gaa gag 1392  
His Leu Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu  
130 135 140  
gac acc ggc aac gag ctg tcc acc aaa gag cag atc agc cgg aac agc 1440  
Asp Thr Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser  
145 150 155  
aag gcc ctg gaa gag aaa tac gtg gcc gaa ctg cag ctg gaa cgg ctg 1488  
Lys Ala Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu  
160 165 170  
aag aaa gac ggc gaa gtg cgg ggc agc atc aac aga ttc aag acc agc 1536

Lys 175	Lys	Asp	Gly	Glu	Val 180	Arg	Gly	Ser	Ile	Asn 185	Arg	Phe	Lys	Thr	Ser 190	
gac Asp	tac Tyr	gtg Val	aaa Lys	gaa Glu	gcc Ala	aaa Lys	cag Gln	ctg Leu	ctg Leu	aag Lys	gtg Val	cag Gln	aag Lys	gcc Ala	tac Tyr	1584
195			200										205			
cac His	cag Gln	ctg Leu	gac Asp	cag Gln	agc Ser	ttc Phe	atc Ile	gac Asp	acc Thr	tac Tyr	atc Ile	gac Asp	ctg Leu	ctg Leu	gaa Glu	1632
210			215										220			
acc Thr	cgg Arg	cgg Arg	acc Thr	tac Tyr	tat Tyr	gag Glu	gga Gly	cct Pro	ggc Gly	gag Glu	ggc Gly	agc Ser	ccc Pro	ttc Phe	ggc Gly	1680
225						230						235				
tgg Trp	aag Lys	gac Asp	atc Ile	aaa Lys	gaa Glu	tgg Trp	tac Tyr	gag Glu	atg Met	ctg Leu	atg Met	ggc Gly	cac His	tgc Cys	acc Thr	1728
240						245					250					
tac Tyr	ttc Phe	ccc Pro	gag Glu	gaa Glu	ctg Leu	cgg Arg	agc Ser	gtg Val	aag Lys	tac Tyr	gcc Ala	tac Tyr	aac Asn	gcc Ala	gac Asp	1776
255					260					265					270	
ctg Leu	tac Tyr	aac Asn	gcc Ala	ctg Leu	aac Asn	gac Asp	ctg Leu	aac Asn	aat Asn	ctc Leu	gtg Val	atc Ile	acc Thr	agg Arg	gac Asp	1824
			275						280					285		
gag Glu	aac Asn	gag Glu	aag Lys	ctg Leu	gaa Glu	tat Tyr	tac Tyr	gag Glu	aag Lys	ttc Phe	cag Gln	atc Ile	atc Ile	gag Glu	aac Asn	1872
			290					295					300			
gtg Val	ttc Phe	aag Lys	cag Gln	aag Lys	aag Lys	aag Lys	ccc Pro	acc Thr	ctg Leu	aag Lys	cag Gln	atc Ile	gcc Ala	aaa Lys	gaa Glu	1920
		305					310					315				
atc Ile	ctc Leu	gtg Val	aac Asn	gaa Glu	gag Glu	gat Asp	att Ile	aag Lys	ggc Gly	tac Tyr	aga Arg	gtg Val	acc Thr	agc Ser	acc Thr	1968
		320				325					330					
ggc Gly	aag Lys	ccc Pro	gag Glu	ttc Phe	acc Thr	aac Asn	ctg Leu	aag Lys	gtg Val	tac Tyr	cac His	gac Asp	atc Ile	aag Lys	gac Asp	2016
335					340					345					350	
att Ile	acc Thr	gcc Ala	cgg Arg	aaa Lys	gag Glu	att Ile	att Ile	gag Glu	aac Asn	gcc Ala	gag Glu	ctg Leu	ctg Leu	gat Asp	cag Gln	2064
			355						360					365		
att Ile	gcc Ala	aag Lys	atc Ile	ctg Leu	acc Thr	atc Ile	tac Tyr	cag Gln	agc Ser	agc Ser	gag Glu	gac Asp	atc Ile	cag Gln	gaa Glu	2112
			370					375					380			
gaa Glu	ctg Leu	acc Thr	aat Asn	ctg Leu	aac Asn	tcc Ser	gag Glu	ctg Leu	acc Thr	cag Gln	gaa Glu	gag Glu	atc Ile	gag Glu	cag Gln	2160
		385					390					395				
atc Ile	tct Ser	aat Asn	ctg Leu	aag Lys	ggc Gly	tat Tyr	acc Thr	ggc Gly	acc Thr	cac His	aac Asn	ctg Leu	agc Ser	ctg Leu	aag Lys	2208

400	405	410	
gcc atc aac ctg atc ctg gac gag ctg tgg cac acc aac gac aac cag Ala Ile Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln 415 420 425 430			2256
atc gct atc ttc aac cgg ctg aag ctg gtg ccc aag aag gtg gac ctg Ile Ala Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu 435 440 445			2304
tcc cag cag aaa gag atc ccc acc acc ctg gtg gac gac ttc atc ctg Ser Gln Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu 450 455 460			2352
agc ccc gtc gtg aag aga agc ttc atc cag agc atc aaa gtg atc aac Ser Pro Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn 465 470 475			2400
gcc atc atc aag aag tac ggc ctg ccc aac gac atc att atc gag ctg Ala Ile Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu 480 485 490			2448
gcc cgc gag aag aac tcc aag gac gcc cag aaa atg atc aac gag atg Ala Arg Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met 495 500 505 510			2496
cag aag cgg aac cgg cag acc aac gag cgg atc gag gaa atc atc cgg Gln Lys Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg 515 520 525			2544
acc acc ggc aaa gag aac gcc aag tac ctg atc gag aag atc aag ctg Thr Thr Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu 530 535 540			2592
cac gac atg cag gaa ggc aag tgc ctg tac agc ctg gaa gcc atc cct His Asp Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro 545 550 555			2640
ctg gaa gat ctg ctg aac aac ccc ttc aac tat gag gtg gac cac atc Leu Glu Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile 560 565 570			2688
atc ccc aga agc gtg tcc ttc gac aac agc ttc aac aac aag gtg ctc Ile Pro Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu 575 580 585 590			2736
gtg aag cag gaa gaa aac agc aag aag ggc aac cgg acc cca ttc cag Val Lys Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln 595 600 605			2784
tac ctg agc agc agc gac agc aag atc agc tac gaa acc ttc aag aag Tyr Leu Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys 610 615 620			2832
cac atc ctg aat ctg gcc aag ggc aag ggc aga atc agc aag acc aag His Ile Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys 625 630 635			2880

aaa gag tat ctg ctg gaa gaa cgg gac atc aac agg ttc tcc gtg cag Lys Glu Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln 640 645 650	2928
aaa gac ttc atc aac cgg aac ctg gtg gat acc aga tac gcc acc aga Lys Asp Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg 655 660 665 670	2976
ggc ctg atg aac ctg ctg cgg agc tac ttc aga gtg aac aac ctg gac Gly Leu Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp 675 680 685	3024
gtg aaa gtg aag tcc atc aat ggc ggc ttc acc agc ttt ctg cgg cgg Val Lys Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg 690 695 700	3072
aag tgg aag ttt aag aaa gag cgg aac aag ggg tac aag cac cac gcc Lys Trp Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala 705 710 715	3120
gag gac gcc ctg atc att gcc aac gcc gat ttc atc ttc aaa gag tgg Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp 720 725 730	3168
aag aaa ctg gac aag gcc aaa aaa gtg atg gaa aac cag atg ttc gag Lys Lys Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu 735 740 745 750	3216
gaa aag cag gcc gag agc atg ccc gag atc gaa acc gag cag gag tac Glu Lys Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr 755 760 765	3264
aaa gag atc ttc atc acc ccc cac cag atc aag cac att aag gac ttc Lys Glu Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe 770 775 780	3312
aag gac tac aag tac agc cac cgg gtg gac aag aag cct aat aga gag Lys Asp Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu 785 790 795	3360
ctg att aac gac acc ctg tac tcc acc cgg aag gac gac aag ggc aac Leu Ile Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn 800 805 810	3408
acc ctg atc gtg aac aat ctg aac ggc ctg tac gac aag gac aat gac Thr Leu Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp 815 820 825 830	3456
aag ctg aaa aag ctg atc aac aag agc ccc gaa aag ctg ctg atg tac Lys Leu Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr 835 840 845	3504
cac cac gac ccc cag acc tac cag aaa ctg aag ctg att atg gaa cag His His Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln 850 855 860	3552

tac ggc gac gag aag aat ccc ctg tac aag tac tac gag gaa acc ggg Tyr Gly Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly 865 870 875	3600
aac tac ctg acc aag tac tcc aaa aag gac aac ggc ccc gtg atc aag Asn Tyr Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys 880 885 890	3648
aag att aag tat tac ggc aac aaa ctg aac gcc cat ctg gac atc acc Lys Ile Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr 895 900 905 910	3696
gac gac tac ccc aac agc aga aac aag gtc gtg aag ctg tcc ctg aag Asp Asp Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys 915 920 925	3744
ccc tac aga ttc gac gtg tac ctg gac aat ggc gtg tac aag ttc gtg Pro Tyr Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val 930 935 940	3792
acc gtg aag aat ctg gat gtg atc aaa aaa gaa aac tac tac gaa gtg Thr Val Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val 945 950 955	3840
aat agc aag tgc tat gag gaa gct aag aag ctg aag aag atc agc aac Asn Ser Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn 960 965 970	3888
cag gcc gag ttt atc gcc tcc ttc tac aac aac gat ctg atc aag atc Gln Ala Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile 975 980 985 990	3936
aac ggc gag ctg tat aga gtg atc ggc gtg aac aac gac ctg ctg aac Asn Gly Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn 995 1000 1005	3984
ogg atc gaa gtg aac atg atc gac atc acc tac cgc gag tac ctg Arg Ile Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu 1010 1015 1020	4029
gaa aac atg aac gac aag agg ccc ccc agg atc att aag aca atc Glu Asn Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile 1025 1030 1035	4074
gcc tcc aag acc cag agc att aag aag tac agc aca gac att ctg Ala Ser Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu 1040 1045 1050	4119
ggc aac ctg tat gaa gtg aaa tct aag aag cac cct cag atc atc Gly Asn Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile 1055 1060 1065	4164
aaa aag ggc aaa agg cct gct gcc acc aaa aag gcc ggc cag gca Lys Lys Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala 1070 1075 1080	4209
aaa aag aaa aag taatgactag tgcctcgact gtgccttcta gttgccagcc	4261

Lys Lys Lys Lys  
1085

atctgttgtt tgcccctccc ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt 4321  
cctttcctaa taaaatgagg aaattgcacg gcattgtctg agtaggtgtc attctattct 4381  
ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata gcaggcatgc 4441  
tggggactcg agtagataag tagcatggcg ggtaatcat taactacaag gaaccctag 4501  
tgatggagtt ggccactccc tctctgcgcg ctgctcgct cactgaggcc gggcgaccaa 4561  
aggtgcgccg acgcccgggc tttgccggg cggcctcagt gagcgagcga gcgcgcag 4619

<210> 3  
<211> 1085  
<212> PRT  
<213> 人工序列

<220>  
<223> 合成構建體

<400> 3

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala  
1 5 10 15

Ala Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val  
20 25 30

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly  
35 40 45

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg  
50 55 60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile  
65 70 75 80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His  
85 90 95

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu  
100 105 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu  
115 120 125

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

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala  
 145 150 155 160

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys  
 165 170 175

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr  
 180 185 190

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln  
 195 200 205

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg  
 210 215 220

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys  
 225 230 235 240

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe  
 245 250 255

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr  
 260 265 270

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn  
 275 280 285

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe  
 290 295 300

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu  
 305 310 315 320

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys  
 325 330 335

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr  
 340 345 350

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala  
 355 360 365

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu  
 370 375 380

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser  
 385 390 395 400

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile  
 405 410 415

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala  
 420 425 430

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln  
 435 440 445

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro  
 450 455 460

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile  
 465 470 475 480

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg  
 485 490 495

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys  
 500 505 510

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr  
 515 520 525

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp  
 530 535 540

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu  
 545 550 555 560

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro  
 565 570 575

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys

580                      585                      590  
 Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu  
               595                      600                      605  
 Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile  
       610                      615                      620  
 Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu  
  625                      630                      635  
 Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp  
                   645                      650                      655  
 Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu  
                   660                      665                      670  
 Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys  
                   675                      680                      685  
 Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp  
       690                      695                      700  
 Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp  
  705                      710                      715                      720  
 Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys  
                   725                      730                      735  
 Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys  
                   740                      745                      750  
 Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu  
                   755                      760                      765  
 Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp  
       770                      775                      780  
 Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile  
  785                      790                      795                      800  
 Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu  
                   805                      810                      815

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu  
820 825 830

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His  
835 840 845

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly  
850 855 860

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr  
865 870 875 880

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile  
885 890 895

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp  
900 905 910

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr  
915 920 925

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val  
930 935 940

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser  
945 950 955 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala  
965 970 975

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly  
980 985 990

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile  
995 1000 1005

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn  
1010 1015 1020

Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser  
1025 1030 1035

Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn  
 1040 1045 1050

Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys  
 1055 1060 1065

Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys  
 1070 1075 1080

Lys Lys  
 1085

<210> 4  
 <211> 7507  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 生產質體 U6. sgR3. PCSK9. ABP2. HLP. hSaCas9 (pX601+3NLS) bGH

<220>  
 <221> LTR  
 <222> (1)..(168)  
 <223> AAV ITR

<220>  
 <221> terminator  
 <222> (193)..(199)  
 <223> U6終止子

<220>  
 <221> misc\_feature  
 <222> (199)..(275)  
 <223> gRNA 支架

<220>  
 <221> misc\_feature  
 <222> (276)..(295)  
 <223> PCSK9 sgRNA

<220>  
 <221> promoter  
 <222> (297)..(545)  
 <223> U6啟動子

<220>  
 <221> enhancer  
 <222> (557)..(656)  
 <223>  $\alpha$  mic/bik增強子

<220>

<221> enhancer  
 <222> (663)..(762)  
 <223>  $\alpha$  mic/bik  
  
 <220>  
 <221> promoter  
 <222> (777)..(1027)  
 <223> HLP  
  
 <220>  
 <221> misc\_feature  
 <222> (1034)..(4288)  
 <223> SaCas9 (px601+3NLS)  
  
 <220>  
 <221> misc\_feature  
 <222> (1040)..(1084)  
 <223> SV40 NLS  
  
 <220>  
 <221> primer\_bind  
 <222> (1227)..(1246)  
 <223> Sa2-Seq1  
  
 <220>  
 <221> primer\_bind  
 <222> (1714)..(1733)  
 <223> Sa2-Seq2  
  
 <220>  
 <221> primer\_bind  
 <222> (2237)..(2256)  
 <223> Sa2-Seq3  
  
 <220>  
 <221> primer\_bind  
 <222> (2747)..(2767)  
 <223> Sa2-Seq4  
  
 <220>  
 <221> primer\_bind  
 <222> (3197)..(3216)  
 <223> Sa2-Seq5  
  
 <220>  
 <221> misc\_feature  
 <222> (4241)..(4288)  
 <223> NLS  
  
 <220>  
 <221> LTR  
 <222> (4253)..(4686)  
 <223> AAV ITR  
  
 <220>  
 <221> polyA\_signal  
 <222> (4300)..(4502)

<223> bGH

<220>

<221> misc\_feature

<222> (5449).. (6306)

<223> Amp-R

<220>

<221> misc\_feature

<222> (6480).. (7068)

<223> 起始序列

<400> 4

```

ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc cggggcgtcg ggcgacctt      60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact    120
aggggttctc ttagttaat gattaaccgc ccatgctact tatctactta aggctatttc    180
tagctctaaa acaaaaaaat ctgcgaaca agttgacgag ataaacacgg cattttgcct    240
tgttttagta gattctgttt ccagagtact aaaacgggac tttggggacc aacttcgggtg    300
tttctctcct tccacaagat atataaagcc aagaaatoga aatactttca agttacggta    360
agcatatgat agtccatttt aaaacataat tttaaaactg caaactacc aagaaattat    420
tactttctac gtcacgtatt ttgtactaat atctttgtgt ttacagtcaa attaattcca    480
attatctctc taacagcctt gtatcgtata tgcaaatatg aaggaatcat gggaaatagg    540
ccctccttaa gctagcaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg    600
cagcatttac tctctctgtt tgctctggtt aataatctca ggagcacaaa cattccagat    660
ccaggttaat ttttaaaaag cagtcaaaaag tccaagtggc ctttggcagc atttactctc    720
totgtttgct ctggttaata atctcaggag cacaaacatt ccagatccgg cgcgcctgtt    780
tgctgcttgc aatgtttgcc cattttaggg tggacacagg acgctgtggt ttctgagcca    840
gggggcgact cagatcccag ccagtggact tagccccgtt ttgctcctcc gataactggg    900
gtgaccttgg ttaatattca ccagcagcct cccccgttgc ccctctggat ccactgctta    960
aatacggacg aggacagggc cctgtctcct cagcttcagg caccaccact gacctgggac   1020
agtgaatgcc accatggccc caaagaagaa gcggaaggtc ggtatccaag gagtcccagc   1080
agccaagcgg aactacatcc tgggcctgga catcggcatc accagcgtgg gctacggcat   1140
catcgactac gagacacggg acgtgatcga tgccggcgtg cggctgttca aagaggccaa   1200
cgtggaaaac aacgagggca ggcggagcaa gagaggcgcc agaaggctga agcggcggag   1260
gcggcataga atccagagag tgaagaagct gctgttcgac tacaacctgc tgaccgacca   1320

```

cagcgagctg agcggcatca acccctaoga gccagagtg aagggcctga gccagaagct	1380
gagcgaggaa gaggttctctg ccgccctgct gcacctggcc aagagaagag gcgtgcacaa	1440
cgtgaacgag gtggaagagg acaccggcaa cgagctgtcc accaaagagc agatcagccg	1500
gaacagcaag gccctggaag agaaatacgt ggccgaactg cagctggaac ggctgaagaa	1560
agacggcgaa gtgcggggca gcatcaacag attcaagacc agcgactacg tgaagaagc	1620
caaacagctg ctgaaggtgc agaaggccta ccaccagctg gaccagagct tcatcgacac	1680
ctacatcgac ctgctggaaa ccggggggac ctactatgag ggacctggcg agggcagccc	1740
cttcggctgg aaggacatca aagaatggta cgagatgctg atgggccaact gcacctactt	1800
ccccgaggaa ctgcgagcgc tgaagtaogc ctacaacgcc gacctgtaca acgccctgaa	1860
cgacctgaac aatctctgta tcaccagga cgagaacgag aagctggaat attacgagaa	1920
gttcagatc atogagaacg tgttcaagca gaagaagaag cccaccctga agcagatcgc	1980
caaagaaatc ctogtgaacg aagaggatat taagggtac agagtgacca gcaccggcaa	2040
gcccagagttc accaacctga aggtgtacca cgacatcaag gacattaccg cccggaaaga	2100
gattattgag aacgocgagc tgctggatca gattgccaag atcttgacca tctaccagag	2160
cagcgaggac atccaggaag aactgaccaa tctgaaactc gagctgaccg aggaagagat	2220
cgagcagatc tctaacttga agggctatac cggcaccac aacctgagcc tgaaggccat	2280
caacctgac ctggacgagc tgtggcacac caacgacaac cagatcgcta tcttcaaccg	2340
gctgaagctg gtgccaaga aggtggacct gtcccagcag aaagagatcc ccaccacct	2400
ggtggacgac ttcatcctga gccccgtctg gaagagaagc ttcatccaga gcatcaaagt	2460
gatcaacgcc atcatcaaga agtaaggcct gcccaacgac atcattatcg agctggcccg	2520
cgagaagaac tccaaggacg cccagaaaat gatcaacgag atgcagaagc ggaaccggca	2580
gaccaacgag cggatcgagg aatcatccg gaccaccggc aaagagaacg ccaagtacct	2640
gatcgagaag atcaagctgc acgacatgca ggaaggcaag tgccctgtaca gcctggaagc	2700
catccctctg gaagatctgc tgaacaaccc cttcaactat gaggtggacc acatcatccc	2760
cagaagcgtg tccttcgaca acagcttcaa caacaaggtg ctctgtaagc aggaagaaaa	2820
cagcaagaag ggcaaccgga cccattcca gtacctgagc agcagcgaca gcaagatcag	2880
ctacgaaacc ttcaagaagc acatcctgaa tctggccaag ggcaagggca gaatcagcaa	2940
gaccaagaaa gagtatctgc tggagaacg ggacatcaac aggttctccg tgcagaaaga	3000
cttcatcaac cggaacctgg tggataccag atacgccacc agaggcctga tgaacctgct	3060

gggagctac ttcagagtga acaacctgga cgtgaaagtg aagtccatca atggcggctt 3120  
 caccagcttt ctgcggcgga agtggaagtt taagaaagag cggaacaagg ggtacaagca 3180  
 ccacgccgag gacgccctga tcattgcca cgccgatttc atcttcaaag agtggaagaa 3240  
 actggacaag gccaaaaaag tgatggaaaa ccagatgttc gaggaaaagc aggccgagag 3300  
 catgcccgag atcgaaaccg agcaggagta caaagagatc ttcatcacc cccaccagat 3360  
 caagcacatt aaggacttca aggactacaa gtacagccac cgggtggaca agaagcctaa 3420  
 tagagagctg attaacgaca cctgtactc ccccgggaag gacgacaagg gcaacaccct 3480  
 gatogtgaac aatctgaacg gcctgtacga caaggacaat gacaagctga aaaagctgat 3540  
 caacaagagc ccgaaaaagc tgctgatgta ccaccaogac cccagacct accagaaact 3600  
 gaagctgatt atggaacagt acggcgacga gaagaatccc ctgtacaagt actacgagga 3660  
 aaccgggaac tacctgacca agtactccaa aaaggacaac ggccccgtga tcaagaagat 3720  
 taagtattac ggcaacaaac tgaacgcca tctggacatc accgacgact accccaacag 3780  
 cagaaacaag gtogtgaagc tgtccctgaa gccctacaga ttcgacgtgt acctggacaa 3840  
 tggogtgtac aagttogtga cgtgaagaa tctggatgtg atcaaaaaag aaaactacta 3900  
 cgaagtgaat agcaagtgtc atgaggaagc taagaagctg aagaagatca gcaaccaggc 3960  
 cgagtttata gcctccttct acaacaacga tctgatcaag atcaacggcg agctgtatag 4020  
 agtgatcggc gtgaacaacg acctgctgaa ccggatogaa gtgaacatga tcgacatcac 4080  
 ctaccgcgag tacctggaaa acatgaacga caagaggccc cccaggatca ttaagacaat 4140  
 cgctccaag acccagagca ttaagaagta cagcacagac attctgggca acctgtatga 4200  
 agtgaatct aagaagcacc ctcagatcat caaaaagggc aaaaggcctg ctgccaccaa 4260  
 aaaggccggc caggcaaaaa agaaaaagta atgactagtg cctcgactgt gccttctagt 4320  
 tgccagccat ctgttgtttg ccctcccc gtgccttctc tgaccctgga aggtgccact 4380  
 cccactgtcc tttcctaata aatgaggaa attgcatgc attgtctgag taggtgtcat 4440  
 totattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc 4500  
 aggcattgct gggactcgag tagataagta gcatggcggg ttaatcatta actacaagga 4560  
 acccctagt atggagtgg ccactccctc tctgcgcgt cgctcgtca ctgaggccgg 4620  
 gcgaccaaag gtcgcccgac gccgggctt tgcccgggog gcctcagtga gcgagcgagc 4680  
 gcgcagcctt aattaacctt attcaactggc cgtcgtttta caacgtcgtg actgggaaaa 4740

ccctggcggtt acccaactta atcgcccttgc agcacatccc cctttcgcca gctggcgtaa 4800  
 tagcgaagag gcccgaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg 4860  
 ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 4920  
 cgctacactt gccagcgcgc tagcgcgcgc tcctttcgct ttcttccctt cctttctcgc 4980  
 cacgttcgcc ggctttcccc gtcaagctct aatcggggg ctcccttag ggttccgatt 5040  
 tagtgcttta cggcacctcg acccaaaaa acttgattag ggtgatggtt cacgtagtgg 5100  
 gccatgccc tgatagacgg tttttcgccc tttagcgttg gagtccacgt tctttaatag 5160  
 tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 5220  
 ataagggtt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt 5280  
 taacgcgaat ttttaaaaa tattaacgtt tacaatttag gtggcacttt tcggggaat 5340  
 gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg 5400  
 agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa 5460  
 catttccgtg tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac 5520  
 ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcaag agtgggttac 5580  
 atogaactgg atctcaacag cggtaagatc cttgagagtt ttgccccga agaacgtttt 5640  
 ccaatgatga gcacttttaa agttctgcta tgtggcggg tattatcccg tattgacgcc 5700  
 gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggg tgagtactca 5760  
 ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc 5820  
 ataaccatga gtgataacac tgcggccaac ttacttotga caacgatcgg aggaccgaag 5880  
 gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa 5940  
 ccggagctga atgaagccat accaaaacgac gagcgtgaca ccacgatgcc tgtagcaatg 6000  
 gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa 6060  
 ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg 6120  
 gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt 6180  
 gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 6240  
 caggcaacta tggatgaacg aaatagacag atcgtctgaga taggtgcctc actgattaag 6300  
 cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 6360  
 ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 6420  
 taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 6480

```

tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca      6540
gcggtgggtt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc      6600
agcagagcgc agataccaaa tactgttctt ctagttagc cgtagttagg ccaccacttc      6660
aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct      6720
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag      6780
gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg gogaacgacc      6840
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg      6900
agaaaggcgg acaggtatcc ggtaagcggc agggctcggaa caggagagcg cacgagggag      6960
cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt      7020
gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac      7080
gcggcctttt taaggttcct ggcccttttg tggccttttg ctacatggt ctttctgctg      7140
ttatcccctg attctgtgga taacogtatt accgcctttg agtgagctga taccgctcgc      7200
cgcagccgaa cgaccgagcg cagcagtgca gtgagcaggg aagcgggaaga gcgcccaata      7260
cgcaaaccgc ctctccccgc gcgttgcccg attcattaat gcagctggca cgacaggttt      7320
cccactgga aagcgggcag tgagcgaac gcaattaatg tgagttagct cactcattag      7380
gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcggg      7440
taacaatttc acacaggaag cagctatgac catgattagc ccagatttaa ttaaggcctt      7500
aattagg                                           7507

```

<210> 5  
<211> 4686  
<212> DNA  
<213> 人工序列

<220>  
<223> 載體基因體 U6. sgR3. PCSK9. ABP2. HLP. hSaCas9 (pX601+3NLS) bGH

<220>  
<221> LTR  
<222> (1)..(168)  
<223> AAV ITR

<220>  
<221> terminator  
<222> (193)..(199)  
<223> U6終止子

<220>  
 <221> misc\_feature  
 <222> (199).. (275)  
 <223> gRNA 支架

<220>  
 <221> misc\_feature  
 <222> (276).. (295)  
 <223> PCSK9 sgRNA

<220>  
 <221> promoter  
 <222> (297).. (545)  
 <223> U6 啟動子

<220>  
 <221> enhancer  
 <222> (557).. (656)  
 <223>  $\alpha$  mic/bik 增強子

<220>  
 <221> enhancer  
 <222> (663).. (762)  
 <223>  $\alpha$  mic/bik 增強子

<220>  
 <221> promoter  
 <222> (777).. (1027)  
 <223> HLP

<220>  
 <221> CDS  
 <222> (1034).. (4288)  
 <223> SaCas9 (px601+3NLS)

<220>  
 <221> misc\_feature  
 <222> (1040).. (1084)  
 <223> SV40 NLS

<220>  
 <221> primer\_bind  
 <222> (1227).. (1246)  
 <223> Sa2-Seq1

<220>  
 <221> primer\_bind  
 <222> (1714).. (1733)  
 <223> Sa2-Seq2

<220>  
 <221> primer\_bind  
 <222> (2237).. (2256)  
 <223> Sa2-Seq3

<220>  
 <221> primer\_bind

<222> (2747).. (2767)  
 <223> Sa2-Seq4

<220>  
 <221> primer\_bind  
 <222> (3197).. (3216)  
 <223> Sa2-Seq5

<220>  
 <221> misc\_feature  
 <222> (4241).. (4288)  
 <223> NLS

<220>  
 <221> LTR  
 <222> (4253).. (4686)  
 <223> AAV ITR

<220>  
 <221> polyA\_signal  
 <222> (4300).. (4502)  
 <223> bGH polyA

<400> 5  
 ctgocgctc gctogctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60  
 ggtogcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120  
 aggggttctt thtagttaat gattaaccog ccatgctact tatctactta aggctatttc 180  
 tagctctaaa acaaaaaaat ctgcacaaca agttgaocgag ataaacacgg cattttgcct 240  
 tgttttagta gattctgttt ccagagtact aaaacgggac tttggggacc aacttcgggtg 300  
 tttogtcttt tccacaagat atataagcc aagaaatoga aatactttca agttacggta 360  
 agcatatgat agtccatttt aaaacataat tttaaaactg caaactacc aagaaattat 420  
 tactttctac gtcacgtatt ttgtactaat atctttgtgt ttacagtcaa attaattcca 480  
 attatctctc taacagcctt gtatcgtata tgcaaatatg aaggaatcat gggaaatagg 540  
 ccctccttaa gctagcaggt taatttttaa aaagcagtoa aaagtccaag tggcccttgg 600  
 cagcatttac tctctctggt tgctctgggt aataatctca ggagcacaaa cattccagat 660  
 ccaggttaat ttttaaaaag cagtcaaaaag tccaagtggc ctttggcagc atttactctc 720  
 totgtttgct ctggttaata atctcaggag cacaaacatt ccagatccgg cgcgcctggt 780  
 tgctgcttgc aatgtttgcc cattttagggt tggacacagg acgctgtggt ttctgagcca 840  
 gggggcgact cagatcccag ccagtggact tagccccgtt ttgctcctcc gataactggg 900  
 gtgaccttgg ttaatattca ccagcagcct cccccgttgc ccctctggat ccactgctta 960  
 aatacggacg aggacagggc cctgtctctc cagcttcagg caccaccact gacctgggac 1020

agtgaatgcc acc atg gcc cca aag aag aag cgg aag gtc ggt atc cac 1069  
                   Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His  
                   1                                  5                                  10

gga gtc cca gca gcc aag cgg aac tac atc ctg ggc ctg gac atc ggc 1117  
 Gly Val Pro Ala Ala Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly  
                   15                                  20                                  25

atc acc agc gtg ggc tac ggc atc atc gac tac gag aca cgg gac gtg 1165  
 Ile Thr Ser Val Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val  
                   30                                  35                                  40

atc gat gcc ggc gtg cgg ctg ttc aaa gag gcc aac gtg gaa aac aac 1213  
 Ile Asp Ala Gly Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn  
                   45                                  50                                  55                                  60

gag ggc agg cgg agc aag aga ggc gcc aga agg ctg aag cgg cgg agg 1261  
 Glu Gly Arg Arg Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg  
                   65                                  70                                  75

cgg cat aga atc cag aga gtg aag aag ctg ctg ttc gac tac aac ctg 1309  
 Arg His Arg Ile Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu  
                   80                                  85                                  90

ctg acc gac cac agc gag ctg agc ggc atc aac ccc tac gag gcc aga 1357  
 Leu Thr Asp His Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg  
                   95                                  100                                  105

gtg aag ggc ctg agc cag aag ctg agc gag gaa gag ttc tct gcc gcc 1405  
 Val Lys Gly Leu Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala  
                   110                                  115                                  120

ctg ctg cac ctg gcc aag aga aga ggc gtg cac aac gtg aac gag gtg 1453  
 Leu Leu His Leu Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val  
                   125                                  130                                  135                                  140

gaa gag gac acc ggc aac gag ctg tcc acc aaa gag cag atc agc cgg 1501  
 Glu Glu Asp Thr Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg  
                   145                                  150                                  155

aac agc aag gcc ctg gaa gag aaa tac gtg gcc gaa ctg cag ctg gaa 1549  
 Asn Ser Lys Ala Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu  
                   160                                  165                                  170

cgg ctg aag aaa gac ggc gaa gtg cgg ggc agc atc aac aga ttc aag 1597  
 Arg Leu Lys Lys Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys  
                   175                                  180                                  185

acc agc gac tac gtg aaa gaa gcc aaa cag ctg ctg aag gtg cag aag 1645  
 Thr Ser Asp Tyr Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys  
                   190                                  195                                  200

gcc tac cac cag ctg gac cag agc ttc atc gac acc tac atc gac ctg 1693  
 Ala Tyr His Gln Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu  
                   205                                  210                                  215                                  220

ctg gaa acc cgg cgg acc tac tat gag gga cct ggc gag ggc agc ccc Leu Glu Thr Arg Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro 225 230 235	1741
ttc ggc tgg aag gac atc aaa gaa tgg tac gag atg ctg atg ggc cac Phe Gly Trp Lys Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His 240 245 250	1789
tgc acc tac ttc ccc gag gaa ctg cgg agc gtg aag tac gcc tac aac Cys Thr Tyr Phe Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn 255 260 265	1837
gcc gac ctg tac aac gcc ctg aac gac ctg aac aat ctc gtg atc acc Ala Asp Leu Tyr Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr 270 275 280	1885
agg gac gag aac gag aag ctg gaa tat tac gag aag ttc cag atc atc Arg Asp Glu Asn Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile 285 290 295 300	1933
gag aac gtg ttc aag cag aag aag aag ccc acc ctg aag cag atc gcc Glu Asn Val Phe Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala 305 310 315	1981
aaa gaa atc ctc gtg aac gaa gag gat att aag ggc tac aga gtg acc Lys Glu Ile Leu Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr 320 325 330	2029
agc acc ggc aag ccc gag ttc acc aac ctg aag gtg tac cac gac atc Ser Thr Gly Lys Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile 335 340 345	2077
aag gac att acc gcc cgg aaa gag att att gag aac gcc gag ctg ctg Lys Asp Ile Thr Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu 350 355 360	2125
gat cag att gcc aag atc ctg acc atc tac cag agc agc gag gac atc Asp Gln Ile Ala Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile 365 370 375 380	2173
cag gaa gaa ctg acc aat ctg aac tcc gag ctg acc cag gaa gag atc Gln Glu Glu Leu Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile 385 390 395	2221
gag cag atc tct aat ctg aag ggc tat acc ggc acc cac aac ctg agc Glu Gln Ile Ser Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser 400 405 410	2269
ctg aag gcc atc aac ctg atc ctg gac gag ctg tgg cac acc aac gac Leu Lys Ala Ile Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp 415 420 425	2317
aac cag atc gct atc ttc aac cgg ctg aag ctg gtg ccc aag aag gtg Asn Gln Ile Ala Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val 430 435 440	2365
gac ctg tcc cag cag aaa gag atc ccc acc acc ctg gtg gac gac ttc	2413

Asp 445	Leu	Ser	Gln	Gln	Lys 450	Glu	Ile	Pro	Thr	Thr 455	Leu	Val	Asp	Asp	Phe 460	
atc Ile	ctg Leu	agc Ser	ccc Pro	gtc Val	gtg Val	aag Lys	aga Arg	agc Ser	ttc Phe	atc Ile	cag Gln	agc Ser	atc Ile	aaa Lys	gtg Val	2461
atc Ile	aac Asn	gcc Ala	atc Ile	atc Ile	aag Lys	aag Lys	tac Tyr	ggc Gly	ctg Leu	ccc Pro	aac Asn	gac Asp	atc Ile	att Ile	atc Ile	2509
gag Glu	ctg Leu	gcc Ala	cgc Arg	gag Glu	aag Lys	aac Asn	tcc Ser	aag Lys	gac Asp	gcc Ala	cag Gln	aaa Lys	atg Met	atc Ile	aac Asn	2557
gag Glu	atg Met	cag Gln	aag Lys	cgg Arg	aac Asn	cgg Arg	cag Gln	acc Thr	aac Asn	gag Glu	cgg Arg	atc Ile	gag Glu	gaa Glu	atc Ile	2605
atc Ile	cgg Arg	acc Thr	acc Thr	ggc Gly	aaa Lys	gag Glu	aac Asn	gcc Ala	aag Lys	tac Tyr	ctg Leu	atc Ile	gag Glu	aag Lys	atc Ile	2653
aag Lys	ctg Leu	cac His	gac Asp	atg Met	cag Gln	gaa Glu	ggc Gly	aag Lys	tgc Cys	ctg Leu	tac Tyr	agc Ser	ctg Leu	gaa Glu	gcc Ala	2701
atc Ile	cct Pro	ctg Leu	gaa Glu	gat Asp	ctg Leu	ctg Leu	aac Asn	aac Asn	ccc Pro	ttc Phe	aac Asn	tat Tyr	gag Glu	gtg Val	gac Asp	2749
cac His	atc Ile	atc Ile	ccc Pro	aga Arg	agc Ser	gtg Val	tcc Ser	ttc Phe	gac Asp	aac Asn	agc Ser	ttc Phe	aac Asn	aac Asn	aag Lys	2797
gtg Val	ctc Leu	gtg Val	aag Lys	cag Gln	gaa Glu	gaa Glu	aac Asn	agc Ser	aag Lys	aag Lys	ggc Gly	aac Asn	cgg Arg	acc Thr	cca Pro	2845
ttc Phe	cag Gln	tac Tyr	ctg Leu	agc Ser	agc Ser	agc Ser	gac Asp	agc Ser	aag Lys	atc Ile	agc Ser	tac Tyr	gaa Glu	acc Thr	ttc Phe	2893
aag Lys	aag Lys	cac His	atc Ile	ctg Leu	aat Asn	ctg Leu	gcc Ala	aag Lys	ggc Gly	aag Lys	ggc Gly	aga Arg	atc Ile	agc Ser	aag Lys	2941
acc Thr	aag Lys	aaa Lys	gag Glu	tat Tyr	ctg Leu	ctg Leu	gaa Glu	gaa Glu	cgg Arg	gac Asp	atc Ile	aac Asn	agg Arg	ttc Phe	tcc Ser	2989
gtg Val	cag Gln	aaa Lys	gac Asp	ttc Phe	atc Ile	aac Asn	cgg Arg	aac Asn	ctg Leu	gtg Val	gat Asp	acc Thr	aga Arg	tac Tyr	gcc Ala	3037
acc Thr	aga Arg	ggc Gly	ctg Leu	atg Met	aac Asn	ctg Leu	ctg Leu	cgg Arg	agc Ser	tac Tyr	ttc Phe	aga Arg	gtg Val	aac Asn	aac Asn	3085

670	675	680	
ctg gac gtg aaa gtg aag tcc atc aat ggc ggc ttc acc agc ttt ctg			3133
Leu Asp Val Lys Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu			
685	690	695	700
cgg cgg aag tgg aag ttt aag aaa gag cgg aac aag ggg tac aag cac			3181
Arg Arg Lys Trp Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His			
	705	710	715
cac gcc gag gac gcc ctg atc att gcc aac gcc gat ttc atc ttc aaa			3229
His Ala Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys			
	720	725	730
gag tgg aag aaa ctg gac aag gcc aaa aaa gtg atg gaa aac cag atg			3277
Glu Trp Lys Lys Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met			
	735	740	745
ttc gag gaa aag cag gcc gag agc atg ccc gag atc gaa acc gag cag			3325
Phe Glu Glu Lys Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln			
	750	755	760
gag tac aaa gag atc ttc atc acc ccc cac cag atc aag cac att aag			3373
Glu Tyr Lys Glu Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys			
	765	770	780
gac ttc aag gac tac aag tac agc cac cgg gtg gac aag aag cct aat			3421
Asp Phe Lys Asp Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn			
	785	790	795
aga gag ctg att aac gac acc ctg tac tcc acc cgg aag gac gac aag			3469
Arg Glu Leu Ile Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys			
	800	805	810
ggc aac acc ctg atc gtg aac aat ctg aac ggc ctg tac gac aag gac			3517
Gly Asn Thr Leu Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp			
	815	820	825
aat gac aag ctg aaa aag ctg atc aac aag agc ccc gaa aag ctg ctg			3565
Asn Asp Lys Leu Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu			
	830	835	840
atg tac cac cac gac ccc cag acc tac cag aaa ctg aag ctg att atg			3613
Met Tyr His His Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met			
	845	850	860
gaa cag tac ggc gac gag aag aat ccc ctg tac aag tac tac gag gaa			3661
Glu Gln Tyr Gly Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu			
	865	870	875
acc ggg aac tac ctg acc aag tac tcc aaa aag gac aac ggc ccc gtg			3709
Thr Gly Asn Tyr Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val			
	880	885	890
atc aag aag att aag tat tac ggc aac aaa ctg aac gcc cat ctg gac			3757
Ile Lys Lys Ile Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp			
	895	900	905

atc acc gac gac tac ccc aac agc aga aac aag gtc gtg aag ctg tcc 3805  
 Ile Thr Asp Asp Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser  
 910 915 920

ctg aag ccc tac aga ttc gac gtg tac ctg gac aat ggc gtg tac aag 3853  
 Leu Lys Pro Tyr Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys  
 925 930 935 940

ttc gtg acc gtg aag aat ctg gat gtg atc aaa aaa gaa aac tac tac 3901  
 Phe Val Thr Val Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr  
 945 950 955

gaa gtg aat agc aag tgc tat gag gaa gct aag aag ctg aag aag atc 3949  
 Glu Val Asn Ser Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile  
 960 965 970

agc aac cag gcc gag ttt atc gcc tcc ttc tac aac aac gat ctg atc 3997  
 Ser Asn Gln Ala Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile  
 975 980 985

aag atc aac ggc gag ctg tat aga gtg atc ggc gtg aac aac gac ctg 4045  
 Lys Ile Asn Gly Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu  
 990 995 1000

ctg aac cgg atc gaa gtg aac atg atc gac atc acc tac cgc gag 4090  
 Leu Asn Arg Ile Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu  
 1005 1010 1015

tac ctg gaa aac atg aac gac aag agg ccc ccc agg atc att aag 4135  
 Tyr Leu Glu Asn Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys  
 1020 1025 1030

aca atc gcc tcc aag acc cag agc att aag aag tac agc aca gac 4180  
 Thr Ile Ala Ser Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp  
 1035 1040 1045

att ctg ggc aac ctg tat gaa gtg aaa tot aag aag cac cct cag 4225  
 Ile Leu Gly Asn Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln  
 1050 1055 1060

atc atc aaa aag ggc aaa agg cct gct gcc acc aaa aag gcc gcc 4270  
 Ile Ile Lys Lys Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly  
 1065 1070 1075

cag gca aaa aag aaa aag taatgactag tgccctgact gtgccttcta 4318  
 Gln Ala Lys Lys Lys Lys  
 1080 1085

gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg gaaggtgcca 4378

ctcccactgt cctttcctaa taaaatgagg aaattgcac gcattgtctg agtaggtgtc 4438

attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata 4498

gcaggcatgc tggggactcg agtagataag tagcatggcg ggttaatcat taactacaag 4558

gaaccctag tgatggagtt ggccactccc tctctgcgcg ctgctcgt cactgaggcc 4618  
 gggcgaccaa aggtcgcccg acgcccgggc ttgcccggg cggcctcagt gagcgagcga 4678  
 gcgcgag 4686

<210> 6  
 <211> 1085  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 合成構建體

<400> 6

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala  
 1 5 10 15

Ala Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val  
 20 25 30

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly  
 35 40 45

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg  
 50 55 60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile  
 65 70 75 80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His  
 85 90 95

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu  
 100 105 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu  
 115 120 125

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

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala  
 145 150 155 160

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys  
165 170 175

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr  
180 185 190

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln  
195 200 205

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg  
210 215 220

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys  
225 230 235 240

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe  
245 250 255

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr  
260 265 270

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn  
275 280 285

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe  
290 295 300

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu  
305 310 315 320

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys  
325 330 335

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr  
340 345 350

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala  
355 360 365

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu  
370 375 380

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser



Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu  
625 630 635 640

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

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu  
660 665 670

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys  
675 680 685

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp  
690 695 700

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp  
705 710 715 720

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys  
725 730 735

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys  
740 745 750

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu  
755 760 765

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp  
770 775 780

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile  
785 790 795 800

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu  
805 810 815

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu  
820 825 830

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His  
835 840 845

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly  
850 855 860

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr  
865 870 875 880

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile  
885 890 895

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp  
900 905 910

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr  
915 920 925

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val  
930 935 940

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser  
945 950 955 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala  
965 970 975

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly  
980 985 990

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile  
995 1000 1005

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn  
1010 1015 1020

Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser  
1025 1030 1035

Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn  
1040 1045 1050

Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys  
1055 1060 1065

Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys  
 1070 1075 1080

Lys Lys  
 1085

<210> 7  
 <211> 3255  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 核酸序列 SaCas9 (pX601+3NLS)

<400> 7  
 atggcccaa agaagaagcg gaaggtcgg atccacggag tcccagcagc caagcggaac 60  
 tacatcctgg gcctggacat cggcatcacc agcgtgggct acggcatcat cgactacgag 120  
 acaogggagc tgatogatgc cggcgtgogg ctgttcaaag aggccaacgt ggaaaacaac 180  
 gagggcaggg ggagcaagag aggcgccaga aggctgaagc ggcggaggcg gcatagaatc 240  
 cagagagtga agaagctgct gttcgactac aacctgctga ccgaccacag cgagctgagc 300  
 ggcatacaacc cctacgagggc cagagtgaag ggccctgagcc agaagctgag cgaggaagag 360  
 ttctctgccg ccctgctgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg 420  
 gaagaggaca ccggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc 480  
 ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagtg 540  
 cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg 600  
 aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg 660  
 ctggaaaccc ggcggacctc ctatgagggg cctggcggag gcagcccctt cggctggaag 720  
 gacatcaaag aatggtacga gatgctgatg ggccaactgca cctacttccc cgaggaactg 780  
 cggagcgtga agtacgccta caacgccgac ctgtacaacg ccctgaacga cctgaacaat 840  
 ctctgatca ccagggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc 900  
 gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatcctc 960  
 gtgaacgaag aggatattaa gggctacaga gtgaccagca ccggcaagcc cgagttcacc 1020  
 aacctgaagg tgtaccacga catcaaggac attaccgcc ggaaagagat tattgagaac 1080  
 gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc 1140  
 caggaagaac tgaccaatct gaactccgag ctgaccagg aagagatcga gcagatctct 1200

aatctgaagg gctataccgg cacccacaac ctgagcctga aggccatcaa cctgatcctg	1260
gacgagctgt ggcacaccaa cgacaaccag atcgctatct tcaaccggct gaagctgggtg	1320
cccaagaagg tggacctgtc ccagcagaaa gagatcccca ccaccctggt ggacgacttc	1380
atcctgagcc ccgtcgtgaa gagaagcttc atccagagca tcaaagtgat caacgccatc	1440
atcaagaagt acggcctgcc caacgacatc attatcgagc tggcccgcga gaagaactcc	1500
aaggacgccc agaaaatgat caacgagatg cagaagcggga accggcagac caacgagcgg	1560
atcgaggaaa tcatccggac caccggcaaa gagaacgcca agtacctgat cgagaagatc	1620
aagctgcacg acatgcagga aggcaagtgc ctgtacagcc tggaaagccat ccctctggaa	1680
gatctgctga acaaccctt caactatgag gtggaccaca tcatccccag aagcgtgtcc	1740
ttcgacaaca gcttcaaca caaggtgctc gtgaagcagg aagaaaacag caagaagggc	1800
aaccggaccc cattccagta cctgagcagc agcgacagca agatcagcta cgaaaccttc	1860
aagaagcaca tcctgaatct ggccaagggc aagggcagaa tcagcaagac caagaaagag	1920
tatctgctgg aagaacggga catcaacagg ttctccgtgc agaaagactt catcaaccgg	1980
aacctggtgg ataccagata cgccaccaga ggccctgatga acctgctgcg gagctacttc	2040
agagtgaaca acctggacgt gaaagtgaag tccatcaatg gcggcttcac cagctttctg	2100
cggcgggaagt ggaagtttaa gaaagagcgg aacaaggggt acaagcacca cgccgaggac	2160
gccctgatca ttgccaacgc cgatttcatc ttcaaagagt ggaagaaact ggacaaggcc	2220
aaaaaagtga tggaaaacca gatgttcgag gaaaagcagg ccgagagcat gcccgagatc	2280
gaaaccgagc aggagtacaa agagatcttc atccccccc accagatcaa gcacattaag	2340
gacttcaagg actacaagta cagccaccgg gtggacaaga agcctaatag agagctgatt	2400
aacgacaccc tgtactccac ccggaaggac gacaagggca acaccctgat cgtgaacaat	2460
ctgaacggcc tgtacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc	2520
gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg	2580
gaacagtacg gcgacgagaa gaatcccctg tacaagtact acgaggaaac cgggaactac	2640
ctgaccaagt actccaaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc	2700
aacaaactga acgcccctct ggacatcacc gacgactacc ccaacagcag aaacaaggtc	2760
gtgaagctgt ccctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
ttcgtgaccg tgaagaatct ggatgtgatc aaaaaagaaa actactacga agtgaatagc	2880
aagtgctatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940

tcctttctaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg 3000  
 aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac 3060  
 ctggaaaaca tgaacgacaa gaggccccc aggatcatta agacaatcgc ctccaagacc 3120  
 cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag 3180  
 aagcacctc agatcatcaa aaagggcaaa aggcctgctg ccaccaaaaa ggccggccag 3240  
 gcaaaaaaga aaaag 3255

<210> 8  
 <211> 20  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> PCSK9 sgRNA

<400> 8  
 aagtttgtcc ccaaagtccc 20

<210> 9  
 <211> 77  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> sgRNA 支架

<400> 9  
 atctcgccaa caagttgacg agataaacac ggcattttgc cttgttttag tagattctgt 60  
 ttccagagta ctaaaac 77

<210> 10  
 <211> 249  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> U6 啟動子

<400> 10  
 ggtgtttcgt cctttccaca agatatataa agccaagaaa tcgaaatact ttcaagttac 60  
 ggtaagcata tgatagtcca ttttaaaaca taattttaaa actgcaaact acccaagaaa 120  
 ttattacttt ctacgtcacg tttttgtac taatatcttt gtgtttacag tcaaattaat 180  
 tccaattatc tctotaacag ccttgatcgc tatatgcaaa tatgaaggaa tcatgggaaa 240

taggcctc 249

<210> 11  
 <211> 176  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> TBG-S1 啟動子

<400> 11  
 actcaaagtt caaaccttat cttttttgctc ttgttctc ttggccttg tttgtacat 60  
 cagctttgaa aataccatcc cagggttaat gctggggtta attataact aagagtgctc 120  
 tagttttgca atacaggaca tgctataaaa atggaaagat gttgctttct gagaga 176

<210> 12  
 <211> 251  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> HLP 啟動子

<400> 12  
 tgtttgctgc ttgcaatggt tgcccatttt aggggtggaca caggacgctg tggtttctga 60  
 gccagggggc gactcagatc ccagccagtg gacttagccc ctgtttgctc ctccgataac 120  
 tggggtgacc ttggttaata ttcaccagca gcctcccccg ttgccctct ggatccactg 180  
 ottaaatacg gacgaggaca gggccctgtc tctcagctt caggcaccac cactgacctg 240  
 ggacagtgaa t 251

<210> 13  
 <211> 417  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> rhPCSK9 基因體序列-外顯子 7 核酸序列

<220>  
 <221> misc\_feature  
 <222> (214)..(397)  
 <223> 外顯子 7

<220>  
 <221> misc\_feature  
 <222> (256)..(261)  
 <223> PAM (SaCas9)

<220>  
 <221> misc\_RNA  
 <222> (262)..(281)  
 <223> sgRNA-SaCas9

<220>  
 <221> misc\_feature  
 <222> (293)..(314)  
 <223> PCS 7-8 靶向序列

<220>  
 <221> misc\_feature  
 <222> (305)..(306)  
 <223> 巨型核酸酶切割位點

<400> 13  
 cctggctatg gtagggacag aggggagcac cagggcgggg caggggtgcc agagttctgc 60  
 ctggggagtc agattttctt taggagggga catttgaatg ggacccaaac aggtgtatag 120  
 cagttgttca gccagctgg caaggcctca gtttgcttct gcaaccctc ttttgggctc 180  
 ctttctgtgc cacctacctc ctacacttc caggtcatca cagttggggc caccaatgcc 240  
 caggaccagc cggtgaccct ggggactttg gggaccaact ttggccgctg tgtggacctc 300  
 tttgccccag gggaggacat catttgggtgcc tccagcgact gcagcacctg ctttgtgtca 360  
 cggagtggga catcgcaggc tgctgccac gtggctggta agtcaccacc ccactgc 417

<210> 14  
 <211> 8924  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 生產質體 TBG. h0TCco2. G4SVPA. LB

<220>  
 <221> repeat\_region  
 <222> (1)..(105)  
 <223> ITR

<220>  
 <221> promoter  
 <222> (164)..(880)  
 <223> TBG 啟動子

<220>  
 <221> enhancer  
 <222> (183)..(282)  
 <223> 增強子

<220>

<221> TATA\_signal  
 <222> (847)..(850)  
 <223> TATA 訊號  
  
 <220>  
 <221> misc\_feature  
 <222> (972)..(2039)  
 <223> h0TCco2  
  
 <220>  
 <221> polyA\_signal  
 <222> (2042)..(2178)  
 <223> G4SVPA  
  
 <220>  
 <221> repeat\_region  
 <222> (2207)..(2374)  
 <223> ITR  
  
 <220>  
 <221> misc\_feature  
 <222> (3898)..(5700)  
 <223> tmp1  
  
 <220>  
 <221> misc\_feature  
 <222> (6866)..(7726)  
 <223> Amp-R  
  
 <220>  
 <221> misc\_feature  
 <222> (7897)..(8485)  
 <223> 起始序列

<400> 14  
 ctgctgctc gctctctcac tgaggccgcc cgggcaaagc cggggcgtcg ggcgacctt 60  
 ggtgccccg cctcagtgag cgagcgagcg cgagagagg gaggtagcc atgctctagg 120  
 aagatcaatt caattcacgc gtggtacct gaactatagc tagaattcgc ccttaagcta 180  
 gcaggttaat ttttaaaaag cagtcaaaaag tccaagtggc cttggcagc atttactctc 240  
 totgtttgct ctggttaata atctcaggag cacaaacatt ccagatccag gttaatTTTT 300  
 aaaaagcagt caaaagtcca agtggccctt ggcagcattt actctctctg tttgctctgg 360  
 ttaataatct caggagcaca aacattccag atccggcgcg ccagggtctgg aagctacct 420  
 tgacatcatt tcctctgcga atgcatgtat aattctaca gaacctatta gaaaggatca 480  
 cccagcctct gctttgtac aactttccct taaaaaactg ccaattccac tgctgtttgg 540  
 cccaatagtg agaactTTTT cctgctgcct cttggtgott ttgcctatgg cccctattct 600  
 gcctgctgaa gacactcttg ccagcatgga cttaaaccoc tccagctctg acaatcctct 660

ttctcttttg ttttacatga agggctctggc agccaaagca atcactcaaa gttcaaacct	720
tatcattttt tgctttgttc ctcttggcct tggttttgta catcagcttt gaaaatacca	780
tcccagggtt aatgctgggg ttaatttata actaagagtg ctctagtttt gcaatacagg	840
acatgctata aaaatggaaa gatgttgctt tctgagagac agctttattg cggtagttta	900
tcacagttaa attgctaacg cagtcagtgc ttctgacaca acagtctcga acttaagctg	960
cattagccac catgctgttc aacctgcgga tcttgcgtgaa caacgccgcc ttccggaacg	1020
gacacaactt catggtgcgg aacttccggt gcggacagcc actgcagaac aaggtgcagc	1080
tgaagggacg ggacctgctg accctgaaga acttcacogg agaggagatc aagtacatgc	1140
tgtggctgag gcgcgatctg aagttccgga tcaagcagaa gggagagtac ctgccactgc	1200
tgcagggaaa gagcctggga atgatcttcg agaagcggag caccgccgacc cggctgagca	1260
ccgagaccgg attogccctg ctgggaggac acccatgctt cctgaccacc caggacatcc	1320
acctgggagt gaatgagagc ctgaccgaca ccgcccggt gctgagcagc atggccgacg	1380
ccgtgctggc ccgggtgtac aagcagagcg acctggatac cctggccaag gaggccagca	1440
toccaatcat caacggactg agcgaacctg accacccaat ccagatcctg gccgattacc	1500
tgaccctgca ggagcactac agcagcctga agggactgac cctgagctgg atcggagacg	1560
gaaacaacat cctgcacagc atcatgatga gcgccgcaa gttcggaatg cacctgcagg	1620
ccgccacccc aaagggatac gagccagatg ccagcgtgac caagctggcc gagcagtagc	1680
ccaaggagaa cggaaccaag ctgctgctga ccaacgacc actggaggcc gccacggag	1740
gaaacgtgct gatcacgat acctggatca gcatgggaca ggaggaggag aagaagaagc	1800
ggctgcaggc cttccagga taccaggtga ccatgaagac cgccaaggtg gccgccagcg	1860
attggacctt cctgcactgc ctgccacgga agccagagga ggtggacgac gaggtgttct	1920
acagcccacg gagcctggtg ttcccagagg ccgagaaccg gaagtggacc atcatggccg	1980
tgatggtgag cctgctgacc gattacagcc cacagctgca gaagccaaag ttctgataag	2040
cgccgcctat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	2100
ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggaggtgt	2160
gggaggtttt ttaggcatcg ataaggatct tcctagagca tggctacgta gataagtagc	2220
atggcgggtt aatcattaac tacaaggaac ccctagtgat ggagttggcc actccctctc	2280
tgcgcgctcg ctgcctcact gaggcggggc gaccaaaggt cgcccagcgc ccgggctttg	2340
ccggggcggc ctcaagtgagc gagcagcgc gcagccttaa ttaacctaat tcaactggccg	2400

tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag	2460
cacatcccc ttctgccagc tggcgtaata gcgaagaggc ccgcaccgat cgccttcccc	2520
aacagttgcg cagcctgaat ggcgaaatggg acgcgccctg tagcggcgca ttaagcgcgg	2580
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgccgctc	2640
ctttcgcttt cttcccttcc tttctgccca cgttcgccgg ctttccccgt caagctctaa	2700
atcgggggct ccccttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac	2760
ttgattaggg tgatggttca cgtgagttac gtgcgaggtc ttccctgcag tgtgggattt	2820
acgctgattc aggaatgggt tgttccctgg gatatggttc taacgcggga ggagcttgta	2880
atcctgagga agtgtatgca cgtgtgcctg tgttgtgcca acattgatat catgacgagc	2940
atgatgatcc atggttacga gtccctgggt ctccaactgtc attgttccag tcccgttcc	3000
ctgcagtgta tagccggcgg gcaggttttg gccagctggt ttaggatggt ggtggatggc	3060
gccatgttta atcagaggtt tatatggtac cgggaggtgg tgaattacaa catgccaaaa	3120
gaggtaatgt ttatgtccag cgtgtttatg aggggtogcc acttaatcta cctgcgcttg	3180
tggtatgatg gccacgtggg ttctgtggtc cccgccatga gctttggata cagcgcttg	3240
cactgtggga ttttgaacaa tattgtggtg ctgtgctgca gttactgtgc tgatttaagt	3300
gagatcaggg tgcgctgctg tgcccggagg acaaggcgc ttatgctgcg ggcggtgca	3360
atcatcgctg aggagaccac tgccatgttg tattcctgca ggacggagcg gcggcggcag	3420
cagtttattc gcgcgctgct gcagcaccac cgcctatcc tgatgcacga ttatgactct	3480
accccatgt aggcgtggac ttctccttgc cgcgccgta agcaaccgca agttggacag	3540
cagcctgtgg ctcagcagct ggacagcagc atgaacttaa gtgagctgcc cggggagttt	3600
attaatatca ctgatgagcg tttggctoga caggaaaccg tgtggaatat aacacctaag	3660
aatatgtctg ttaccatga tatgatgctt ttttaaggcca gccggggaga aaggactgtg	3720
tactctgtgt gttgggaggg aggtggcagg ttgaatacta gggttctgtg agtttgatta	3780
aggtacggtg atctgtataa gctatgtggt ggtggggcta tactactgaa tgaaaaatga	3840
cttgaaattt tctgcaattg aaaaataaac acgttgaaac ataacacaaa cgattcttta	3900
ttcttgggca atgtatgaaa aagtgaaga ggatgtggca aatatttcat taatgtagtt	3960
gtggccagac cagtccatg aaaatgacat agagtatgca cttggagttg tgtctcctgt	4020
ttcctgtgta ccgtttagtg taatggttag tgttacaggt ttagttttgt ctccgtttaa	4080

gtaaacttga ctgacaatgt tacttttggc agttttaccg tgagattttg gataagctga 4140  
 taggttaggc ataaatccaa cagcgtttgt ataggctgtg ccttcagtaa gatctccatt 4200  
 tctaaagttc caatattctg ggtccaggaa ggaattgttt agtagcactc cattttcgtc 4260  
 aaatcttata ataagatgag cactttgaac tgttccagat attggagcca aactgccttt 4320  
 aacagccaaa actgaaactg tagcaagtat ttgactgcca cattttgtta agaccaaagt 4380  
 gagtttagca tctttctctg catttagtct acagtttaga gatggagctg gtgtggtcca 4440  
 caaagttagc ttatcattat ttttgtttcc tactgtaatg gcacctgtgc tgtcaaaact 4500  
 aaggccagtt cctagtttag gaaccatagc cttgtttgaa tcaaattcta ggccatggcc 4560  
 aatttttgtt ttgaggggat ttgtgtttgg tgcattaggt gaaccaaatt caagcccatc 4620  
 tcttgcatta atggctatgg ctgtagcgtc aaacatcaac cccttggcag tgcttaggtt 4680  
 aacctcaagc tttttggaat tgtttgaagc tgtaaacaag taaaggcctt tgttttagtt 4740  
 aatatccaag ttgtgggctg agtttataaa aagagggccc tgtcctagtc ttagatttag 4800  
 ttggttttga gcatcaaacg gataactaac atcaagtata aggcgtctgt tttgagaatc 4860  
 aatccttagt cctcctgcta cattaagttg catattgcoct tgtgaatcaa aaccaaggc 4920  
 tccagtaact ttagtttgca aggaagtatt attaatagtc acacctggac cagttgctac 4980  
 ggtcaaagtg ttaggtctgt ctgttacatg caaaggagcc cgtacttta gtccctagttt 5040  
 tccattttgt gtataaatgg gctctttcaa gtcaatgcc aagctaccag tggcagtagt 5100  
 tagagggggt gaggcagtga tagtaagggt actgctatcg gtggtggtga gggggcctga 5160  
 tgtttgcagg gctagctttc cttctgacac tgtgaggggt ccttgggtgg caatgctaag 5220  
 tttggagtgc tgcacggtta gcggggcctg tgattgcatg gtgagtgtgt tgcccgcgac 5280  
 cattagaggt gcggcggcag ccacagttag ggcttctgag gtaactgtga ggggtgcaga 5340  
 tatttccagg tttatgtttg acttggtttt tttgagaggt gggctcacag tggttacatt 5400  
 ttgggaggta aggttgcggg cctcgtccag agagaggccg ttgccattt tgagcgcaag 5460  
 catgccattg gaggtaaact gaggttcgga taggcgcaaa gagagtacc cagggggact 5520  
 ctcttgaaac ccattggggg atacaaaggg aggagtaaga aaaggcacag ttggaggacc 5580  
 ggtttccgtg tcatatggat acacggggtt gaaggtatct tcagacggtc ttgcgcgctt 5640  
 catctgcaac aacatgaaga tagtgggtgc ggatggacag gaacaggagg aaactgacat 5700  
 tccattttaga ttgtggagaa agtttgcagc caggaggaag ctgcaatacc agagctggga 5760  
 ggagggcaag gaggtgctgc tgaataaact ggacagaaat ttgctaactg attttaagta 5820

agtgatgctt tattatTTTT tttattagt taaaggaat aagatccccg ggtactctag 5880  
 ttaattaact agaggatctt gatgtaatcc aaggtagga cagttgcaaa tcacagttag 5940  
 aacacagggt cccctgtccc gctcaactag cagggggcgc tgggtaaact cccgaatcag 6000  
 gctacgggca agctctccct gggcggtaag ccggacgcgc tgcgccgggc cctcgatatg 6060  
 atcctcgggc aattcaaagt agcaaaactc accggagtgc cgggcaaagc acttgtggcg 6120  
 gcgacagtgg accaggtggt tcaggcgcag ttgctctgcc tctccactta acattcagtc 6180  
 gtagccgtcc gccagatcct ttaccgcgtc aaagtagga ataaattgat ccggatagtg 6240  
 gccgggagggt cccgagaagg ggttaaagta gaccgatggc acaaactcct caataaattg 6300  
 cagagttcca atgcctccag agcgcggctc agaggacgag gtctgcagag ttaggattgc 6360  
 ctgacgaggc gtgaatgaag gacggccggc gccgccgato tgaatgtcc cgtccggacg 6420  
 gagaccaagc gaggagctca ccgactcgtc gttgagctga atacctgcc ctctgattgt 6480  
 caggtgagtt ataccctgcc cgggcgaccg cacgtgggccc atcgccctga tagacggttt 6540  
 ttgcccttt gacgttggag tccaagtctt ttaatagtg actcttgttc caaactggaa 6600  
 caaactcaa cccatctcgc gtctattctt ttgatttata agggattttg ccgatttcgg 6660  
 cctattggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat 6720  
 taacgcttac aatttaggtg gcacttttcg gggaaatgtg cgcggaacct ctatttgttt 6780  
 atttttctaa atacattcaa atagtatcc gctcatgaga caataacct gataaatgct 6840  
 tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc 6900  
 cttttttgcg gcattttgcc ttctgtttt tgctcaccga gaaacgctgg tgaaagtaaa 6960  
 agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg 7020  
 taagatcctt gagagttttc gccccgaaga acgttttcca atgatgagca cttttaagt 7080  
 totgctatgt ggcgcgggtat tatcccgat tgacgccggg caagagcaac tcggtcgcgc 7140  
 catacactat tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac 7200  
 ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc 7260  
 ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaaa 7320  
 catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc 7380  
 aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt 7440  
 aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcggg 7500

taaagttgca ggaccacttc tgcgctoggc cttccggct ggctggttta ttgctgataa	7560
atctggagcc ggtgagcgtg ggtctcggg tatcattgca gcaactggggc cagatggtaa	7620
gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa	7680
tagacagatc gctgagatag gtgcctcact gattaagcat tggtactgt cagaccaagt	7740
ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt	7800
gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg	7860
agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgct	7920
aatctgctgc ttgcaaaaa aaaaaccacc gctaccagcg gtggtttgtt tgccgatca	7980
agagctacca actctttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac	8040
tgttcttcta gtgtagccgt agttaggcca ccactcaag aactctgtag caccgcctac	8100
atacctcgct ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct	8160
tacogggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg	8220
gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga gatacctaca	8280
gcgtgagcta tgagaaagcg ccacgcttc cgaagggaga aaggcggaca ggtatccggt	8340
aagcggcagg gtcggaacag gagagcgac gagggagctt ccagggggaa acgcctggt	8400
tctttatagt cctgtogggt ttgccacct ctgacttgag cgtcgatttt tgtgatgctc	8460
gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttcttggc	8520
cttttcttg ccttttctc acatgttctt tctgcgtta tcccctgatt ctgtggataa	8580
ccgtattacc gcctttgagt gagctgatac cgctcgcgc agccgaacga ccgagcgcag	8640
cgagtcagtg agcaggaag cggaaagcg cccaatacgc aaaccgcctc tccccgcgcg	8700
ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaag cgggcagtga	8760
gcgcaacgca attaattgta gttagctcac tcattaggca ccccaggctt tacactttat	8820
gottccggct cgtatgttgt gtggaattgt gagcggataa caatttcaca caggaaacag	8880
ctatgaccat gattacgcca gatttaatta aggccttaat tagg	8924

<210> 15  
 <211> 2374  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 載體基因體 TBG. hOTCco2. G4SVPA. LB

<220>  
 <221> repeat\_region  
 <222> (1)..(105)  
 <223> ITR

<220>  
 <221> promoter  
 <222> (164)..(880)  
 <223> TBG 啟動子

<220>  
 <221> enhancer  
 <222> (183)..(282)  
 <223> 增強子

<220>  
 <221> TATA\_signal  
 <222> (847)..(850)  
 <223> TATA

<220>  
 <221> CDS  
 <222> (972)..(2039)  
 <223> h0TCco2

<220>  
 <221> polyA\_signal  
 <222> (2042)..(2178)  
 <223> G4SVPA

<220>  
 <221> repeat\_region  
 <222> (2207)..(2374)  
 <223> ITR

<400> 15  
 ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggctcg ggcgacctt 60  
 ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtgtagcc atgctctagg 120  
 aagatcaatt caattcacgc gtggtacctt gaactatagc tagaattcgc ccttaagcta 180  
 gcaggttaat ttttaaaaag cagtcaaaaag tccaagtggc ccttggcagc atttactctc 240  
 totgtttgct ctggttaata atctcaggag cacaaacatt ccagatccag gttaatTTTT 300  
 aaaaagcagt caaaagtcca agtggccctt ggcagcattt actctctctg tttgctctgg 360  
 ttaataatct caggagcaca aacattccag atccggcgcg ccagggtctgg aagctacctt 420  
 tgacatcatt tcctctgcga atgcatgtat aatttctaca gaacctatta gaaaggatca 480  
 cccagcctct gcttttgtac aactttccct taaaaaactg ccaattccac tgctgtttgg 540  
 cccaatagtg agaactTTTT cctgctgcct cttggtgctt ttgcctatgg cccctattct 600

gcctgctgaa gacactcttg ccagcatgga cttaaaccoc tccagctctg acaatcctct 660  
 ttctcttttg ttttacatga agggctctggc agccaaagca atcactcaaa gttcaaacct 720  
 tatcattttt tgctttgttc ctcttgccct tggttttgta catcagcttt gaaaatacca 780  
 tcccagggtt aatgctgggg ttaatttata actaagagtg ctctagtitt gcaatacagg 840  
 acatgctata aaaatggaaa gatgttgctt tctgagagac agctttattg cggtagttta 900  
 tcacagttaa attgctaacg cagtcagtgc ttctgacaca acagtctoga acttaagctg 960  
 cattagccac c atg ctg ttc aac ctg cgg atc ctg ctg aac aac gcc gcc 1010  
           Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala  
           1                  5                                  10  
 ttc cgg aac gga cac aac ttc atg gtg cgg aac ttc cgg tgc gga cag 1058  
 Phe Arg Asn Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln  
           15                  20                                  25  
 cca ctg cag aac aag gtg cag ctg aag gga cgg gac ctg ctg acc ctg 1106  
 Pro Leu Gln Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu  
           30                  35                                  40                  45  
 aag aac ttc acc gga gag gag atc aag tac atg ctg tgg ctg agc gcc 1154  
 Lys Asn Phe Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala  
                           50                                  55                  60  
 gat ctg aag ttc cgg atc aag cag aag gga gag tac ctg cca ctg ctg 1202  
 Asp Leu Lys Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu Leu  
                           65                                  70                  75  
 cag gga aag agc ctg gga atg atc ttc gag aag cgg agc acc cgg acc 1250  
 Gln Gly Lys Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg Thr  
                           80                                  85                  90  
 cgg ctg agc acc gag acc gga ttc gcc ctg ctg gga gga cac cca tgc 1298  
 Arg Leu Ser Thr Glu Thr Gly Phe Ala Leu Leu Gly Gly His Pro Cys  
           95                  100                                  105  
 ttc ctg acc acc cag gac atc cac ctg gga gtg aat gag agc ctg acc 1346  
 Phe Leu Thr Thr Gln Asp Ile His Leu Gly Val Asn Glu Ser Leu Thr  
           110                  115                                  120                  125  
 gac acc gcc cgg gtg ctg agc agc atg gcc gac gcc gtg ctg gcc cgg 1394  
 Asp Thr Ala Arg Val Leu Ser Ser Met Ala Asp Ala Val Leu Ala Arg  
                           130                                  135                  140  
 gtg tac aag cag agc gac ctg gat acc ctg gcc aag gag gcc agc atc 1442  
 Val Tyr Lys Gln Ser Asp Leu Asp Thr Leu Ala Lys Glu Ala Ser Ile  
                           145                                  150                  155  
 cca atc atc aac gga ctg agc gac ctg tac cac cca atc cag atc ctg 1490  
 Pro Ile Ile Asn Gly Leu Ser Asp Leu Tyr His Pro Ile Gln Ile Leu  
           160                  165                                  170  
 gcc gat tac ctg acc ctg cag gag cac tac agc agc ctg aag gga ctg 1538

Ala Asp Tyr Leu Thr Leu Gln Glu His Tyr Ser Ser Leu Lys Gly Leu 175 180 185	
acc ctg agc tgg atc gga gac gga aac aac atc ctg cac agc atc atg Thr Leu Ser Trp Ile Gly Asp Gly Asn Asn Ile Leu His Ser Ile Met 190 195 200 205	1586
atg agc gcc gcc aag ttc gga atg cac ctg cag gcc gcc acc cca aag Met Ser Ala Ala Lys Phe Gly Met His Leu Gln Ala Ala Thr Pro Lys 210 215 220	1634
gga tac gag cca gat gcc agc gtg acc aag ctg gcc gag cag tac gcc Gly Tyr Glu Pro Asp Ala Ser Val Thr Lys Leu Ala Glu Gln Tyr Ala 225 230 235	1682
aag gag aac gga acc aag ctg ctg ctg acc aac gac cca ctg gag gcc Lys Glu Asn Gly Thr Lys Leu Leu Leu Thr Asn Asp Pro Leu Glu Ala 240 245 250	1730
gcc cac gga gga aac gtg ctg atc acc gat acc tgg atc agc atg gga Ala His Gly Gly Asn Val Leu Ile Thr Asp Thr Trp Ile Ser Met Gly 255 260 265	1778
cag gag gag gag aag aag aag cgg ctg cag gcc ttc cag gga tac cag Gln Glu Glu Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr Gln 270 275 280 285	1826
gtg acc atg aag acc gcc aag gtg gcc gcc agc gat tgg acc ttc ctg Val Thr Met Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe Leu 290 295 300	1874
cac tgc ctg cca cgg aag cca gag gag gtg gac gac gag gtg ttc tac His Cys Leu Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe Tyr 305 310 315	1922
agc cca cgg agc ctg gtg ttc cca gag gcc gag aac cgg aag tgg acc Ser Pro Arg Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp Thr 320 325 330	1970
atc atg gcc gtg atg gtg agc ctg ctg acc gat tac agc cca cag ctg Ile Met Ala Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln Leu 335 340 345	2018
cag aag cca aag ttc tga taa gggccgcta tttgtgaaat ttgtgatgct Gln Lys Pro Lys Phe 350	2069
attgctttat ttgtaacctat tataagctgc aataaacaag ttaacaacaa caattgcatt	2129
cattttatgt ttcaggttca gggggaggtg tgggaggttt tttaggcatc gataaggatc	2189
ttcctagagc atggctacgt agataagtag catggcgggt taatcattaa ctacaaggaa	2249
ccoctagtga tggagttggc cactccctct ctgcgcgctc gctcgtcac tgaggccggg	2309
cgaccaaagg tcgcccgcg cccgggcttt gcccgggcgg cctcagtgag cgagcgagcg	2369

cgcag

2374

<210> 16  
 <211> 354  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 合成構建體

<400> 16

Met Leu Phe Asn Leu Arg Ile Leu Leu Asn Asn Ala Ala Phe Arg Asn  
 1 5 10 15

Gly His Asn Phe Met Val Arg Asn Phe Arg Cys Gly Gln Pro Leu Gln  
 20 25 30

Asn Lys Val Gln Leu Lys Gly Arg Asp Leu Leu Thr Leu Lys Asn Phe  
 35 40 45

Thr Gly Glu Glu Ile Lys Tyr Met Leu Trp Leu Ser Ala Asp Leu Lys  
 50 55 60

Phe Arg Ile Lys Gln Lys Gly Glu Tyr Leu Pro Leu Leu Gln Gly Lys  
 65 70 75 80

Ser Leu Gly Met Ile Phe Glu Lys Arg Ser Thr Arg Thr Arg Leu Ser  
 85 90 95

Thr Glu Thr Gly Phe Ala Leu Leu Gly Gly His Pro Cys Phe Leu Thr  
 100 105 110

Thr Gln Asp Ile His Leu Gly Val Asn Glu Ser Leu Thr Asp Thr Ala  
 115 120 125

Arg Val Leu Ser Ser Met Ala Asp Ala Val Leu Ala Arg Val Tyr Lys  
 130 135 140

Gln Ser Asp Leu Asp Thr Leu Ala Lys Glu Ala Ser Ile Pro Ile Ile  
 145 150 155 160

Asn Gly Leu Ser Asp Leu Tyr His Pro Ile Gln Ile Leu Ala Asp Tyr  
 165 170 175

Leu Thr Leu Gln Glu His Tyr Ser Ser Leu Lys Gly Leu Thr Leu Ser  
180 185 190

Trp Ile Gly Asp Gly Asn Asn Ile Leu His Ser Ile Met Met Ser Ala  
195 200 205

Ala Lys Phe Gly Met His Leu Gln Ala Ala Thr Pro Lys Gly Tyr Glu  
210 215 220

Pro Asp Ala Ser Val Thr Lys Leu Ala Glu Gln Tyr Ala Lys Glu Asn  
225 230 235 240

Gly Thr Lys Leu Leu Leu Thr Asn Asp Pro Leu Glu Ala Ala His Gly  
245 250 255

Gly Asn Val Leu Ile Thr Asp Thr Trp Ile Ser Met Gly Gln Glu Glu  
260 265 270

Glu Lys Lys Lys Arg Leu Gln Ala Phe Gln Gly Tyr Gln Val Thr Met  
275 280 285

Lys Thr Ala Lys Val Ala Ala Ser Asp Trp Thr Phe Leu His Cys Leu  
290 295 300

Pro Arg Lys Pro Glu Glu Val Asp Asp Glu Val Phe Tyr Ser Pro Arg  
305 310 315 320

Ser Leu Val Phe Pro Glu Ala Glu Asn Arg Lys Trp Thr Ile Met Ala  
325 330 335

Val Met Val Ser Leu Leu Thr Asp Tyr Ser Pro Gln Leu Gln Lys Pro  
340 345 350

Lys Phe

<210> 17  
<211> 1068  
<212> DNA  
<213> 人工序列

<220>  
<223> 工程化核酸序列 h0TCco2

<400> 17

```

atgctgttca acctgoggat cctgctgaac aacgccgcct tccggaacgg acacaacttc      60
atggtgcgga acttccgggtg cggacagcca ctgcagaaca aggtgcagct gaagggacgg      120
gacctgctga cctgaagaa cttcaccgga gaggagatca agtacatgct gtggctgagc      180
gccgatctga agttccggat caagcagaag ggagagtacc tgccactgct gcagggaaag      240
agcctgggaa tgatcttcga gaagcggagc acccggaccg ggctgagcac cgagaccgga      300
ttgccctgc tgggaggaca cccatgcttc ctgaccaccg aggacatcca cctgggagtg      360
aatgagagcc tgaccgacac cgcccgggtg ctgagcagca tggccgacgc cgtgctggcc      420
cgggtgtaca agcagagcga cctggatacc ctggccaagg aggccagcat cccaatcatc      480
aacggactga gcgacctgta ccaccaatc cagatcctgg ccgattacct gaccctgcag      540
gagcactaca gcagcctgaa gggactgacc ctgagctgga tcggagacgg aaacaacatc      600
ctgcacagca tcatgatgag cgccgccaag ttcggaatgc acctgcaggc cgccacccca      660
aagggatacg agccagatgc cagcgtgacc aagctggccg agcagtacgc caaggagaac      720
ggaaccaagc tgctgctgac caacgaccca ctggaggccg cccacggagg aaacgtgctg      780
atcaccgata cctggatcag catgggacag gaggaggaga agaagaagcg gctgcaggcc      840
ttccagggat accaggtgac catgaagacc gccaaagtg gcccagcga ttggaccttc      900
ctgcaactgcc tgccaocggaa gccagaggag gtggacgacg aggtgttcta cagcccacgg      960
agcctggtgt tcccagaggc cgagaaccgg aagtggacca tcatggccgt gatggtgagc     1020
ctgctgaccg attacagccc acagctgcag aagccaaagt tctgataa                       1068

```

<210> 18  
<211> 4687  
<212> DNA  
<213> 人工序列

<220>  
<223> 生產質體 TBG-S1-F113.PCS7-8L.197.bGH

<220>  
<221> repeat\_region  
<222> (1)..(168)  
<223> ITR

<220>  
<221> promoter  
<222> (206)..(318)  
<223> TBG S1 啟動子

<220>

<221> misc\_feature  
 <222> (330)..(1424)  
 <223> PCS7-8L  
  
 <220>  
 <221> polyA\_signal  
 <222> (1435)..(1649)  
 <223> bGH polyA 訊號  
  
 <220>  
 <221> repeat\_region  
 <222> (1699)..(1866)  
 <223> ITR  
  
 <220>  
 <221> misc\_feature  
 <222> (2629)..(3486)  
 <223> Amp-R  
  
 <220>  
 <221> misc\_feature  
 <222> (3660)..(4248)  
 <223> 起始序列

<400> 18  
 ctgocgcctc gctogctcac tgaggccgcc cgggcaaagc cggggcgtcg ggcgaccttt 60  
 ggtogcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120  
 aggggttctt ttagttaaatt gattaaccgg ccatgctact tatctacgta gccatgctct 180  
 aggaagatcg gaattogccc ttaagctttg aaaataccat cccagggtta atgctgggggt 240  
 taatttataa ctaagagtgc tctagttttg caatacagga catgctataa aatggaaag 300  
 atgttgcttt ctgagagaca gggccgccca tggcacogaa gaagaagcgc aaggtgcata 360  
 tgaatacaaa atataataaa gaggttttac tctacttagc agggtttgta gacgggtgacg 420  
 gttccatctt tgccaggatc aagcctagtc aacgtagtaa gttcaagcac aagctgcatc 480  
 togttttcgc tgtctatcag aagacacagc gccgttggtt cctcgacaag ctggtggacg 540  
 agatcgggtg gggttacgtg ctggactctg gcagcgtctc cttttactcg ctgtccgaga 600  
 tcaagccttt gcataatfff ttaacacaac tacaaccttt tctaaaacta aaacaaaaac 660  
 aagcaaattt agttttaaaa attattgaac aacttccgtc agcaaaaagaa tccccggaca 720  
 aattcttaga agttttgtaca tgggtggatc aaattgcagc tctgaatgat tcgaagacgc 780  
 gtaaaacaac ttctgaaacc gttcgtgctg tgctagacag tttaccagga tccgtgggag 840  
 gtctatcgcc atctcaggca tccagcgccg catcctcggc ttctcaagc ccgggttcag 900  
 ggatctccga agcaactcaga gctggagcag gttccggcac tggatacaac aaggaattcc 960

tgctctacct	ggcgggcttc	gtcgacgggg	acggctccat	ctatgcccgt	atcaagccgg	1020
ttcagcgggc	taagttcaag	cacgagctgg	ttctcgggtt	cgatgtcaact	cagaagacac	1080
agcgcggttg	gttctctgac	aagctggtgg	acgagatcgg	tgtgggttac	gtgtatgaca	1140
agggcagcgt	ctccgcgtac	cgtctgtccc	agatcaagcc	tctgcacaac	ttcctgaccc	1200
agctccagcc	cttctggaag	ctcaagcaga	agcaggccaa	cctcgtgctg	aagatcatcg	1260
agcagctgcc	ctccgccaaag	gaatccccgg	acaagttcct	ggaggtgtgc	acctgggtgg	1320
accagatcgc	cgctctgaac	gactccaaga	cccgcaagac	cacttccgaa	accgtccgcg	1380
cogttctaga	cagtctctcc	gagaagaaga	agtcgtcccc	ctaaggtacg	atctgcctcg	1440
actgtgcctt	ctagttgccca	gccatctggt	gtttgcccct	ccccctgcc	ttccttgacc	1500
ctggaaggtg	ccactcccac	tgtcctttcc	taataaaatg	aggaaattgc	atgcatttgt	1560
ctgagtaggt	gtcattctat	tctgggggggt	gggggtggggc	aggacagcaa	gggggaggat	1620
tgggaagaca	atagcaggca	tgctggggac	tcgagttaag	ggcgaattcc	cgataaggat	1680
cttctagag	catggctacg	tagataagta	gcatggcggg	ttaatcatta	actacaagga	1740
accctagtg	atggagttgg	ccactccctc	tctgcgcgct	cgctcgtca	ctgaggccgg	1800
gcgaccaaag	gtcgcccgac	gcccgggctt	tgcccggggc	gcctcagtga	gcgagcgagc	1860
gcgcagcctt	aattaaccta	attcaactggc	cgtcgtttta	caacgtcgtg	actgggaaaa	1920
ccctggcggtt	acccaactta	atgccttgc	agcacatccc	cctttcgcca	gctggcgtaa	1980
tagcgaagag	gcccgcaccg	atgcacctc	ccaacagttg	cgcagcctga	atggcgaatg	2040
ggacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	gcagcgtgac	2100
cgctacactt	gccagcggcc	tagcgcggcg	tcctttogct	ttcttccctt	cctttctcgc	2160
cacgttcgcc	ggctttcccc	gtcaagctct	aaatcggggg	ctccctttag	ggttccgatt	2220
tagtgcttta	cggcacctcg	acccccaaaa	acttgattag	ggtgatggtt	cacgtagtgg	2280
gccatcgccc	tgatagacgg	ttttcgccc	tttgacgttg	gagtccacgt	tctttaatag	2340
tggactcttg	ttccaaactg	gaacaacact	caaccctatc	tcggtctatt	cctttgattt	2400
ataagggatt	ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt	2460
taacgcgaat	tttaacaaaa	tattaacgct	tacaatttag	gtggcacttt	tcggggaat	2520
gtgcgcggaa	cccctatttg	tttatttttc	taaatacatt	caaatatgta	tcgcctcatg	2580
agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	2640
catttccgtg	tcgcccttat	tccttttttt	gcggcatttt	gccttccctg	ttttgctcac	2700

ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac 2760  
atogaactgg atotcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt 2820  
ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgacgcc 2880  
gggcaagagc aactcggctc cgcatacac tattctcaga atgacttggg tgagtactca 2940  
ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc 3000  
ataaccatga gtgataacac tgcggccaac ttactctga caacgatcgg aggaccgaag 3060  
gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa 3120  
ccggagctga atgaagccat accaaaacgac gagcgtgaca ccacgatgcc tntagcaatg 3180  
gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa 3240  
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg 3300  
gctggctggt ttattgctga taaatctgga gccgggtgagc gtgggtctcg cggtatcatt 3360  
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 3420  
caggcaacta tggatgaacg aaatagacag atcgtctgaga taggtgcctc actgattaag 3480  
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 3540  
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 3600  
taacgtgagt tttogttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 3660  
tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaac accgctacca 3720  
gcgggtggtt gtttgccgga tcaagagcta ccaactottt ttccgaaggt aactggcttc 3780  
agcagagcgc agataccaaa tactgttctt ctagttagc cgtagttagg ccaccacttc 3840  
aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct 3900  
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag 3960  
gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg gcgaacgacc 4020  
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg 4080  
agaaaggcgg acaggtatcc ggtaagcggc agggctcggaa caggagagcg cacgagggag 4140  
cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt 4200  
gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaa cgcagcaac 4260  
gcggcctttt tacggttcct ggccctttgc tggccttttg ctcatatgtt ctttctgcg 4320  
ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc 4380

cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga ggcaccaata 4440  
 cgcaaaccgc ctctccccgc gcgttgccg attcattaat gcagctggca cgacaggttt 4500  
 cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag 4560  
 gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga 4620  
 taacaatttc acacaggaaa cagctatgac catgattacg ccagatttaa ttaaggcctt 4680  
 aattagg 4687

<210> 19  
 <211> 1866  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 載體基因體 TBG-S1-F113. PCS7-8L. 197. bGH

<220>  
 <221> repeat\_region  
 <222> (1)..(168)  
 <223> ITR

<220>  
 <221> promoter  
 <222> (206)..(318)  
 <223> TBG S1

<220>  
 <221> CDS  
 <222> (330)..(1424)  
 <223> PCS7-8L

<220>  
 <221> polyA\_signal  
 <222> (1435)..(1649)  
 <223> bGH polyA

<220>  
 <221> repeat\_region  
 <222> (1691)..(1736)  
 <223> ITR

<400> 19  
 ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc cggggcgtcg ggcgaccttt 60  
 ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120  
 aggggttctt ttagttaat gattaaccgg ccatgctact tatctacgta gccatgctct 180  
 aggaagatcg gaattcgccc ttaagctttg aaaataccat cccagggtta atgctggggt 240  
 taatttataa ctaagagtgc tctagttttg caatacagga catgctataa aatggaaag 300

atgttgcttt ctgagagaca gcggccgcc atg gca ccg aag aag aag cgc aag 353  
Met Ala Pro Lys Lys Lys Arg Lys  
1 5

gtg cat atg aat aca aaa tat aat aaa gag ttc tta ctc tac tta gca 401  
Val His Met Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala  
10 15 20

ggg ttt gta gac ggt gac ggt tcc atc ttt gcc agg atc aag cct agt 449  
Gly Phe Val Asp Gly Asp Gly Ser Ile Phe Ala Arg Ile Lys Pro Ser  
25 30 35 40

caa cgt agt aag ttc aag cac aag ctg cat ctc gtt ttc gct gtc tat 497  
Gln Arg Ser Lys Phe Lys His Lys Leu His Leu Val Phe Ala Val Tyr  
45 50 55

cag aag aca cag cgc cgt tgg ttc ctc gac aag ctg gtg gac gag atc 545  
Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile  
60 65 70

ggt gtg ggt tac gtg ctg gac tct ggc agc gtc tcc ttt tac tcg ctg 593  
Gly Val Gly Tyr Val Leu Asp Ser Gly Ser Val Ser Phe Tyr Ser Leu  
75 80 85

tcc gag atc aag cct ttg cat aat ttt tta aca caa cta caa cct ttt 641  
Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe  
90 95 100

cta aaa cta aaa caa aaa caa gca aat tta gtt tta aaa att att gaa 689  
Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu  
105 110 115 120

caa ctt ccg tca gca aaa gaa tcc ccg gac aaa ttc tta gaa gtt tgt 737  
Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys  
125 130 135

aca tgg gtg gat caa att gca gct ctg aat gat tcg aag acg cgt aaa 785  
Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys  
140 145 150

aca act tct gaa acc gtt cgt gct gtg cta gac agt tta cca gga tcc 833  
Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Pro Gly Ser  
155 160 165

gtg gga ggt cta tcg cca tct cag gca tcc agc gcc gca tcc tcg gct 881  
Val Gly Gly Leu Ser Pro Ser Gln Ala Ser Ser Ala Ala Ser Ser Ala  
170 175 180

tcc tca agc ccg ggt tca ggg atc tcc gaa gca ctc aga gct gga gca 929  
Ser Ser Ser Pro Gly Ser Gly Ile Ser Glu Ala Leu Arg Ala Gly Ala  
185 190 195 200

ggt tcc gcc act gga tac aac aag gaa ttc ctg ctc tac ctg gcg gcc 977  
Gly Ser Gly Thr Gly Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly  
205 210 215

ttc gtc gac ggg gac ggc tcc atc tat gcc cgt atc aag ccg gtt cag 1025  
 Phe Val Asp Gly Asp Gly Ser Ile Tyr Ala Arg Ile Lys Pro Val Gln  
 220 225 230

cgg gct aag ttc aag cac gag ctg gtt ctc ggg ttc gat gtc act cag 1073  
 Arg Ala Lys Phe Lys His Glu Leu Val Leu Gly Phe Asp Val Thr Gln  
 235 240 245

aag aca cag cgc cgt tgg ttc ctc gac aag ctg gtg gac gag atc ggt 1121  
 Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly  
 250 255 260

gtg ggt tac gtg tat gac aag ggc agc gtc tcc gcg tac cgt ctg tcc 1169  
 Val Gly Tyr Val Tyr Asp Lys Gly Ser Val Ser Ala Tyr Arg Leu Ser  
 265 270 275 280

cag atc aag cct ctg cac aac ttc ctg acc cag ctc cag ccc ttc ctg 1217  
 Gln Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu  
 285 290 295

aag ctc aag cag aag cag gcc aac ctc gtg ctg aag atc atc gag cag 1265  
 Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln  
 300 305 310

ctg ccc tcc gcc aag gaa tcc ccg gac aag ttc ctg gag gtg tgc acc 1313  
 Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr  
 315 320 325

tgg gtg gac cag atc gcc gct ctg aac gac tcc aag acc cgc aag acc 1361  
 Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr  
 330 335 340

act tcc gaa acc gtc cgc gcc gtt cta gac agt ctc tcc gag aag aag 1409  
 Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys  
 345 350 355 360

aag tcg tcc ccc taa ggtacgatct gctctgactg tgccttctag ttgccagcca 1464  
 Lys Ser Ser Pro

totgtttgttt gccctcccc cgtgccttcc ttgacctgg aaggtgccac tcccactgtc 1524

otttctaat aaaatgagga aattgcatcg cattgtotga gtaggtgtca ttctattctg 1584

ggggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct 1644

ggggactcga gtaagggcg aattccgat aaggatcttc ctagagcatg gctacgtaga 1704

taagtagcat ggcgggttaa tcattaacta caaggaacct ctagtgatgg agttggccac 1764

tccctctctg cgcgctcgct cgtcactga ggccggggcga ccaaaggtcg cccgacgcc 1824

gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc ag 1866

<210> 20

<211> 364

<212> PRT  
 <213> 人工序列

<220>  
 <223> 合成構建體

<400> 20

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn  
 1 5 10 15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser  
 20 25 30

Ile Phe Ala Arg Ile Lys Pro Ser Gln Arg Ser Lys Phe Lys His Lys  
 35 40 45

Leu His Leu Val Phe Ala Val Tyr Gln Lys Thr Gln Arg Arg Trp Phe  
 50 55 60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Leu Asp Ser  
 65 70 75 80

Gly Ser Val Ser Phe Tyr Ser Leu Ser Glu Ile Lys Pro Leu His Asn  
 85 90 95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala  
 100 105 110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser  
 115 120 125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala  
 130 135 140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala  
 145 150 155 160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln  
 165 170 175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile  
 180 185 190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys

195

200

205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile  
210 215 220

Tyr Ala Arg Ile Lys Pro Val Gln Arg Ala Lys Phe Lys His Glu Leu  
225 230 235 240

Val Leu Gly Phe Asp Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu  
245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp Lys Gly  
260 265 270

Ser Val Ser Ala Tyr Arg Leu Ser Gln Ile Lys Pro Leu His Asn Phe  
275 280 285

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn  
290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro  
305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu  
325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val  
340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro  
355 360

<210> 21  
<211> 1095  
<212> DNA  
<213> 人工序列

<220>  
<223> 核酸序列 PCS7-8L

<400> 21  
atggcaccga agaagaagcg caaggtgcat atgaatacaa aatataataa agagttctta 60  
ctctacttag cagggtttgt agacggtgac ggttccatct ttgccaggat caagcctagt 120  
caacgtagta agttcaagca caagctgcat ctcgttttog ctgtctatca gaagacacag 180

cgccgttggg tcctcgacaa gctggtggac gagatcggg tgggttacgt gctggactct 240  
 ggcagcgtct ccttttactc gctgtccgag atcaagcctt tgcataattt tttaacacaa 300  
 ctacaacctt ttctaaaact aaaacaaaaa caagcaaatt tagttttaaa aattattgaa 360  
 caacttccgt cagcaaaaaga atccccggac aaattcttag aagtttgtac atgggtggat 420  
 caaattgcag ctctgaatga ttogaagacg cgtaaaacaa cttctgaaac cgttcgtgct 480  
 gtgctagaca gtttaccagg atccgtggga ggtctatcgc catctcaggc atccagcggc 540  
 gcatcctcgg ctctctcaag cccgggttca gggatctccg aagcactcag agctggagca 600  
 gtttccggca ctggatacaa caaggaattc ctgctctacc tggcgggctt cgtcgacggg 660  
 gacggctcca tctatgcccg tatcaagccg gttcagcggg ctaagttcaa gcacgagctg 720  
 gttctcgggt tcgatgtcac tcagaagaca cagcgcggtt gtttctctga caagctgggtg 780  
 gacgagatcg gtgtgggtta cgtgtatgac aagggcagcg tctccgcgta cgtctgtcc 840  
 cagatcaagc ctctgcacaa cttcctgacc cagctccagc ccttctgaa gctcaagcag 900  
 aagcaggcca acctcgtgct gaagatcctc gagcagctgc cctccgcca ggaatccccg 960  
 gacaagttcc tggaggtgtg cacctgggtg gaccagatcg ccgctctgaa cgactccaag 1020  
 accgcaaga ccacttccga aaccgtccgc gccgttctag acagtctctc cgagaagaag 1080  
 aagtctccc cctaa 1095

<210> 22  
 <211> 2079  
 <212> DNA  
 <213> 智人

<220>  
 <221> CDS  
 <222> (1)..(2079)  
 <223> PCSK9 CDS

<400> 22  
 atg ggc acc gtc agc tcc agg cgg tcc tgg tgg ccg ctg cca ctg ctg 48  
 Met Gly Thr Val Ser Ser Arg Arg Ser Trp Trp Pro Leu Pro Leu Leu  
 1 5 10 15  
 ctg ctg ctg ctg ctg ctc ctg ggt ccc gcg ggc gcc cgt gcg cag gag 96  
 Leu Leu Leu Leu Leu Leu Leu Gly Pro Ala Gly Ala Arg Ala Gln Glu  
 20 25 30  
 gac gag gac ggc gac tac gag gag ctg gtg cta gcc ttg cgt tcc gag 144  
 Asp Glu Asp Gly Asp Tyr Glu Glu Leu Val Leu Ala Leu Arg Ser Glu  
 35 40 45

gag gac ggc ctg gcc gaa gca ccc gag cac gga acc aca gcc acc ttc Glu Asp Gly Leu Ala Glu Ala Pro Glu His Gly Thr Thr Ala Thr Phe 50 55 60	192
cac cgc tgc gcc aag gat ccg tgg agg ttg cct ggc acc tac gtg gtg His Arg Cys Ala Lys Asp Pro Trp Arg Leu Pro Gly Thr Tyr Val Val 65 70 75 80	240
gtg ctg aag gag gag acc cac ctc tcg cag tca gag cgc act gcc cgc Val Leu Lys Glu Glu Thr His Leu Ser Gln Ser Glu Arg Thr Ala Arg 85 90 95	288
cgc ctg cag gcc cag gct gcc cgc cgg gga tac ctc acc aag atc ctg Arg Leu Gln Ala Gln Ala Ala Arg Arg Gly Tyr Leu Thr Lys Ile Leu 100 105 110	336
cat gtc ttc cat ggc ctt ctt cct ggc ttc ctg gtg aag atg agt ggc His Val Phe His Gly Leu Leu Pro Gly Phe Leu Val Lys Met Ser Gly 115 120 125	384
gac ctg ctg gag ctg gcc ttg aag ttg ccc cat gtc gac tac atc gag Asp Leu Leu Glu Leu Ala Leu Lys Leu Pro His Val Asp Tyr Ile Glu 130 135 140	432
gag gac tcc tct gtc ttt gcc cag agc atc ccg tgg aac ctg gag cgg Glu Asp Ser Ser Val Phe Ala Gln Ser Ile Pro Trp Asn Leu Glu Arg 145 150 155 160	480
att acc cct cca cgg tac cgg gcg gat gaa tac cag ccc ccc gac gga Ile Thr Pro Pro Arg Tyr Arg Ala Asp Glu Tyr Gln Pro Pro Asp Gly 165 170 175	528
ggc agc ctg gtg gag gtg tat ctc cta gac acc agc ata cag agt gac Gly Ser Leu Val Glu Val Tyr Leu Leu Asp Thr Ser Ile Gln Ser Asp 180 185 190	576
cac cgg gaa atc gag ggc agg gtc atg gtc acc gac ttc gag aat gtg His Arg Glu Ile Glu Gly Arg Val Met Val Thr Asp Phe Glu Asn Val 195 200 205	624
ccc gag gag gac ggg acc cgc ttc cac aga cag gcc agc aag tgt gac Pro Glu Glu Asp Gly Thr Arg Phe His Arg Gln Ala Ser Lys Cys Asp 210 215 220	672
agt cat ggc acc cac ctg gca ggg gtg gtc agc ggc cgg gat gcc ggc Ser His Gly Thr His Leu Ala Gly Val Val Ser Gly Arg Asp Ala Gly 225 230 235 240	720
gtg gcc aag ggt gcc agc atg cgc agc ctg cgc gtg ctc aac tgc caa Val Ala Lys Gly Ala Ser Met Arg Ser Leu Arg Val Leu Asn Cys Gln 245 250 255	768
ggg aag ggc acg gtt agc ggc acc ctc ata ggc ctg gag ttt att cgg Gly Lys Gly Thr Val Ser Gly Thr Leu Ile Gly Leu Glu Phe Ile Arg 260 265 270	816

aaa agc cag ctg gtc cag cct gtg ggg cca ctg gtg gtg ctg ctg ccc Lys Ser Gln Leu Val Gln Pro Val Gly Pro Leu Val Val Leu Leu Pro	864
275 280 285	
ctg gcg ggt ggg tac agc cgc gtc ctc aac gcc gcc tgc cag cgc ctg Leu Ala Gly Gly Tyr Ser Arg Val Leu Asn Ala Ala Cys Gln Arg Leu	912
290 295 300	
gcg agg gct ggg gtc gtg ctg gtc acc gct gcc gcc aac ttc cgg gac Ala Arg Ala Gly Val Val Leu Val Thr Ala Ala Gly Asn Phe Arg Asp	960
305 310 315 320	
gat gcc tgc ctc tac tcc cca gcc tca gct ccc gag gtc atc aca gtt Asp Ala Cys Leu Tyr Ser Pro Ala Ser Ala Pro Glu Val Ile Thr Val	1008
325 330 335	
ggg gcc acc aat gcc caa gac cag ccg gtg acc ctg ggg act ttg ggg Gly Ala Thr Asn Ala Gln Asp Gln Pro Val Thr Leu Gly Thr Leu Gly	1056
340 345 350	
acc aac ttt ggc cgc tgt gtg gac ctc ttt gcc cca ggg gag gac atc Thr Asn Phe Gly Arg Cys Val Asp Leu Phe Ala Pro Gly Glu Asp Ile	1104
355 360 365	
att ggt gcc tcc agc gac tgc agc acc tgc ttt gtg tca cag agt ggg Ile Gly Ala Ser Ser Asp Cys Ser Thr Cys Phe Val Ser Gln Ser Gly	1152
370 375 380	
aca tca cag gct gct gcc cac gtg gct ggc att gca gcc atg atg ctg Thr Ser Gln Ala Ala Ala His Val Ala Gly Ile Ala Ala Met Met Leu	1200
385 390 395 400	
tct gcc gag ccg gag ctc acc ctg gcc gag ttg agg cag aga ctg atc Ser Ala Glu Pro Glu Leu Thr Leu Ala Glu Leu Arg Gln Arg Leu Ile	1248
405 410 415	
cac ttc tct gcc aaa gat gtc atc aat gag gcc tgg ttc cct gag gac His Phe Ser Ala Lys Asp Val Ile Asn Glu Ala Trp Phe Pro Glu Asp	1296
420 425 430	
cag cgg gta ctg acc ccc aac ctg gtg gcc gcc ctg ccc ccc agc acc Gln Arg Val Leu Thr Pro Asn Leu Val Ala Ala Leu Pro Pro Ser Thr	1344
435 440 445	
cat ggg gca ggt tgg cag ctg ttt tgc agg act gta tgg tca gca cac His Gly Ala Gly Trp Gln Leu Phe Cys Arg Thr Val Trp Ser Ala His	1392
450 455 460	
tcg ggg cct aca cgg atg gcc aca gcc gtc gcc cgc tgc gcc cca gat Ser Gly Pro Thr Arg Met Ala Thr Ala Val Ala Arg Cys Ala Pro Asp	1440
465 470 475 480	
gag gag ctg ctg agc tgc tcc agt ttc tcc agg agt ggg aag cgg cgg Glu Glu Leu Leu Ser Cys Ser Ser Phe Ser Arg Ser Gly Lys Arg Arg	1488
485 490 495	
ggc gag cgc atg gag gcc caa ggg ggc aag ctg gtc tgc cgg gcc cac	1536

Gly	Glu	Arg	Met	Glu	Ala	Gln	Gly	Gly	Lys	Leu	Val	Cys	Arg	Ala	His		
			500					505					510				
aac	gct	ttt	ggg	ggt	gag	ggt	gtc	tac	gcc	att	gcc	agg	tgc	tgc	ctg		1584
Asn	Ala	Phe	Gly	Gly	Glu	Gly	Val	Tyr	Ala	Ile	Ala	Arg	Cys	Cys	Leu		
		515					520					525					
cta	ccc	cag	gcc	aac	tgc	agc	gtc	cac	aca	gct	cca	cca	gct	gag	gcc		1632
Leu	Pro	Gln	Ala	Asn	Cys	Ser	Val	His	Thr	Ala	Pro	Pro	Ala	Glu	Ala		
	530					535					540						
agc	atg	ggg	acc	cgt	gtc	cac	tgc	cac	caa	cag	ggc	cac	gtc	ctc	aca		1680
Ser	Met	Gly	Thr	Arg	Val	His	Cys	His	Gln	Gln	Gly	His	Val	Leu	Thr		
545					550				555						560		
ggc	tgc	agc	tcc	cac	tgg	gag	gtg	gag	gac	ctt	ggc	acc	cac	aag	ccg		1728
Gly	Cys	Ser	Ser	His	Trp	Glu	Val	Glu	Asp	Leu	Gly	Thr	His	Lys	Pro		
				565					570					575			
cct	gtg	ctg	agg	cca	cga	ggt	cag	ccc	aac	cag	tgc	gtg	ggc	cac	agg		1776
Pro	Val	Leu	Arg	Pro	Arg	Gly	Gln	Pro	Asn	Gln	Cys	Val	Gly	His	Arg		
			580					585					590				
gag	gcc	agc	atc	cac	gct	tcc	tgc	tgc	cat	gcc	cca	ggt	ctg	gaa	tgc		1824
Glu	Ala	Ser	Ile	His	Ala	Ser	Cys	Cys	His	Ala	Pro	Gly	Leu	Glu	Cys		
		595					600					605					
aaa	gtc	aag	gag	cat	gga	atc	cgg	gcc	cct	cag	gag	cag	gtg	acc	gtg		1872
Lys	Val	Lys	Glu	His	Gly	Ile	Pro	Ala	Pro	Gln	Glu	Gln	Val	Thr	Val		
	610					615					620						
gcc	tgc	gag	gag	ggc	tgg	acc	ctg	act	ggc	tgc	agt	gcc	ctc	cct	ggg		1920
Ala	Cys	Glu	Glu	Gly	Trp	Thr	Leu	Thr	Gly	Cys	Ser	Ala	Leu	Pro	Gly		
625					630					635					640		
acc	tcc	cac	gtc	ctg	ggg	gcc	tac	gcc	gta	gac	aac	acg	tgt	gta	gtc		1968
Thr	Ser	His	Val	Leu	Gly	Ala	Tyr	Ala	Val	Asp	Asn	Thr	Cys	Val	Val		
				645					650					655			
agg	agc	cgg	gac	gtc	agc	act	aca	ggc	agc	acc	agc	gaa	ggg	gcc	gtg		2016
Arg	Ser	Arg	Asp	Val	Ser	Thr	Thr	Gly	Ser	Thr	Ser	Glu	Gly	Ala	Val		
			660					665					670				
aca	gcc	gtt	gcc	atc	tgc	tgc	cgg	agc	cgg	cac	ctg	gcg	cag	gcc	tcc		2064
Thr	Ala	Val	Ala	Ile	Cys	Cys	Arg	Ser	Arg	His	Leu	Ala	Gln	Ala	Ser		
		675					680						685				
cag	gag	ctc	cag	tga													2079
Gln	Glu	Leu	Gln														
		690															

<210> 23  
 <211> 692  
 <212> PRT  
 <213> 智人

&lt;400&gt; 23

Met Gly Thr Val Ser Ser Arg Arg Ser Trp Trp Pro Leu Pro Leu Leu  
1 5 10 15

Leu Leu Leu Leu Leu Leu Leu Gly Pro Ala Gly Ala Arg Ala Gln Glu  
20 25 30

Asp Glu Asp Gly Asp Tyr Glu Glu Leu Val Leu Ala Leu Arg Ser Glu  
35 40 45

Glu Asp Gly Leu Ala Glu Ala Pro Glu His Gly Thr Thr Ala Thr Phe  
50 55 60

His Arg Cys Ala Lys Asp Pro Trp Arg Leu Pro Gly Thr Tyr Val Val  
65 70 75 80

Val Leu Lys Glu Glu Thr His Leu Ser Gln Ser Glu Arg Thr Ala Arg  
85 90 95

Arg Leu Gln Ala Gln Ala Ala Arg Arg Gly Tyr Leu Thr Lys Ile Leu  
100 105 110

His Val Phe His Gly Leu Leu Pro Gly Phe Leu Val Lys Met Ser Gly  
115 120 125

Asp Leu Leu Glu Leu Ala Leu Lys Leu Pro His Val Asp Tyr Ile Glu  
130 135 140

Glu Asp Ser Ser Val Phe Ala Gln Ser Ile Pro Trp Asn Leu Glu Arg  
145 150 155 160

Ile Thr Pro Pro Arg Tyr Arg Ala Asp Glu Tyr Gln Pro Pro Asp Gly  
165 170 175

Gly Ser Leu Val Glu Val Tyr Leu Leu Asp Thr Ser Ile Gln Ser Asp  
180 185 190

His Arg Glu Ile Glu Gly Arg Val Met Val Thr Asp Phe Glu Asn Val  
195 200 205

Pro Glu Glu Asp Gly Thr Arg Phe His Arg Gln Ala Ser Lys Cys Asp  
210 215 220

Ser His Gly Thr His Leu Ala Gly Val Val Ser Gly Arg Asp Ala Gly  
225 230 235 240

Val Ala Lys Gly Ala Ser Met Arg Ser Leu Arg Val Leu Asn Cys Gln  
245 250 255

Gly Lys Gly Thr Val Ser Gly Thr Leu Ile Gly Leu Glu Phe Ile Arg  
260 265 270

Lys Ser Gln Leu Val Gln Pro Val Gly Pro Leu Val Val Leu Leu Pro  
275 280 285

Leu Ala Gly Gly Tyr Ser Arg Val Leu Asn Ala Ala Cys Gln Arg Leu  
290 295 300

Ala Arg Ala Gly Val Val Leu Val Thr Ala Ala Gly Asn Phe Arg Asp  
305 310 315 320

Asp Ala Cys Leu Tyr Ser Pro Ala Ser Ala Pro Glu Val Ile Thr Val  
325 330 335

Gly Ala Thr Asn Ala Gln Asp Gln Pro Val Thr Leu Gly Thr Leu Gly  
340 345 350

Thr Asn Phe Gly Arg Cys Val Asp Leu Phe Ala Pro Gly Glu Asp Ile  
355 360 365

Ile Gly Ala Ser Ser Asp Cys Ser Thr Cys Phe Val Ser Gln Ser Gly  
370 375 380

Thr Ser Gln Ala Ala Ala His Val Ala Gly Ile Ala Ala Met Met Leu  
385 390 395 400

Ser Ala Glu Pro Glu Leu Thr Leu Ala Glu Leu Arg Gln Arg Leu Ile  
405 410 415

His Phe Ser Ala Lys Asp Val Ile Asn Glu Ala Trp Phe Pro Glu Asp  
420 425 430

Gln Arg Val Leu Thr Pro Asn Leu Val Ala Ala Leu Pro Pro Ser Thr  
435 440 445

His Gly Ala Gly Trp Gln Leu Phe Cys Arg Thr Val Trp Ser Ala His  
 450 455 460

Ser Gly Pro Thr Arg Met Ala Thr Ala Val Ala Arg Cys Ala Pro Asp  
 465 470 475 480

Glu Glu Leu Leu Ser Cys Ser Ser Phe Ser Arg Ser Gly Lys Arg Arg  
 485 490 495

Gly Glu Arg Met Glu Ala Gln Gly Gly Lys Leu Val Cys Arg Ala His  
 500 505 510

Asn Ala Phe Gly Gly Glu Gly Val Tyr Ala Ile Ala Arg Cys Cys Leu  
 515 520 525

Leu Pro Gln Ala Asn Cys Ser Val His Thr Ala Pro Pro Ala Glu Ala  
 530 535 540

Ser Met Gly Thr Arg Val His Cys His Gln Gln Gly His Val Leu Thr  
 545 550 555 560

Gly Cys Ser Ser His Trp Glu Val Glu Asp Leu Gly Thr His Lys Pro  
 565 570 575

Pro Val Leu Arg Pro Arg Gly Gln Pro Asn Gln Cys Val Gly His Arg  
 580 585 590

Glu Ala Ser Ile His Ala Ser Cys Cys His Ala Pro Gly Leu Glu Cys  
 595 600 605

Lys Val Lys Glu His Gly Ile Pro Ala Pro Gln Glu Gln Val Thr Val  
 610 615 620

Ala Cys Glu Glu Gly Trp Thr Leu Thr Gly Cys Ser Ala Leu Pro Gly  
 625 630 635 640

Thr Ser His Val Leu Gly Ala Tyr Ala Val Asp Asn Thr Cys Val Val  
 645 650 655

Arg Ser Arg Asp Val Ser Thr Thr Gly Ser Thr Ser Glu Gly Ala Val  
 660 665 670

Thr Ala Val Ala Ile Cys Cys Arg Ser Arg His Leu Ala Gln Ala Ser

675

680

685

Gln Glu Leu Gln  
690

<210> 24  
<211> 132  
<212> DNA  
<213> 智人

<400> 24  
tgggctcctt tctctgccac ccacctcctc acctttccag gtcatcacag ttggggccac 60  
caatgcccaa gaccagccgg tgaccctggg gactttgggg accaactttg gccgctgtgt 120  
ggacctcttt gc 132

<210> 25  
<211> 137  
<212> DNA  
<213> 智人

<400> 25  
ttgccccagg ggaggacatc attggtgcoct ccagcgactg cagcacctgc tttgtgtcac 60  
agagtgggac atcacaggct gctgcccacg tggctggtaa gtcaccaccc cactgcoctcg 120  
gccaccgtga tgctaac 137

<210> 26  
<211> 500  
<212> DNA  
<213> 智人

<400> 26  
acctaggtcc ccccgcccta tcaaggcttc cctggcggcc gaatttaaag gcatcaagca 60  
aaciaagccc aacacatctc tgccttgtcc tctcagtttc cccccgtggc acttagaacc 120  
acttgataca ccgaatagtt tccggtctat ctccccact aggatgtaaa ctccacaggg 180  
gcattgggaa tgctgcoctgg ctatggtagg gacagagggg agcaccaggg cggggcaggg 240  
gtgocagagt tctgcoctggg cagtcagatt ttcttagga ggggacattt gagtgggacc 300  
caaacaggtg tatagcagtt gtccagocca gctggcaagg cctgagtctg cctctgcaac 360  
ccctctcttg ggctcctttc tctgccaccc acctcctcac ctttccaggt catcacagtt 420  
ggggccacca atgccagga ccagccggtg accctgggga ctttggggac caactttggc 480  
cgctgtgtgg acctctttgc 500

<210> 27  
 <211> 500  
 <212> DNA  
 <213> 智人

<400> 27  
 ttgcccagg ggaggacatc attggtgcoct ccagcgactg cagcacctgc tttgtgtcac 60  
 agagtgggac atcacaggct gctgcccacg tggctggtaa gtcaccaccc cactgcctcg 120  
 gccaccgtga tgctaacagc ccctttggca gtcagggtct gtgccgggac ctccagtgcc 180  
 aggctctgtg cagggggacc agagatgaag taggcctgat ggtgccttca aggacactca 240  
 gtctgatgag ggaggcgagt gcacagaggg aacacagaggt cagggctgta ttagagggag 300  
 ccagaggag gcacctgccc agcccagagg tcagagaagg catcttggag gagggacatt 360  
 tgatcgggag cttgatggat gaataggagt ttacctggcc gataagacag caactaccaa 420  
 ggcttagagg tgtgagagga ggctgtotta cctcactgag taaggactgc aggcggctta 480  
 cttcgagaa gagagcttag 500

<210> 28  
 <211> 516  
 <212> DNA  
 <213> 恆河獼猴

<400> 28  
 aagcatacct gacagcggta acccagatcc tccttggcoct cggaggcttc cctggcagcc 60  
 caatttaaag gcatcaagca aacaaagccc aacacatctc tgccctgccc tctcagtctc 120  
 ccccgctggc acttagagcc acctgataca ccgagtagtt tcctatctcc ctactagaa 180  
 tgtaaactcc acaggggcat tgggaatgct gcctggctat ggtagggaca gaggggagca 240  
 ccagggcggg gcaggggtgc cagagttctg cctggggagt cagattttcc ttaggagggg 300  
 acatttgaat gggacccaaa caggtgtata gcagttgttc agcccagctg gcaaggcctc 360  
 agtttgcttc tgcaaccocct ctcttgggct cttttotgtg ccacctacct cctcaccttt 420  
 ccaggtcatc acagttgggg ccaccaatgc ccaggaccag ccggtgacct tggggacttt 480  
 ggggaccaac tttggccgct gtgtggacct ctttgc 516

<210> 29  
 <211> 510  
 <212> DNA  
 <213> 恆河獼猴

<400> 29

```

ttgcccagg ggaggacatc attggtgcoct ccagcgactg cagcacctgc tttgtgtcac      60
ggagtgggac atcgcaggct gctgcccacg tggctggtaa gtcaccaccc cactgccttg      120
gccaccgtga tgctaacagc cttttggcgg tcagggtctg tgccaggacc tccagtgcc      180
ggctctgtgc agggggacca gagatgagac aggcctgatg gtgccttcat ggacactcag      240
tctgatgagg gagacgagtg cacagtggga acacgaggtt agggctgtat tagagggagc      300
ccagaggagg cacctgcccc gcccgagggt cagagaaggc ttcttggaca agggacattt      360
gatctggagc ttgatggatg aataggagtc cacctggccg ataagacagc aactaccaag      420
gcttagaggt gtgacggaag gctgtcttac ctactgtgt gagggactgc aggcggctta      480
ccttctagaa gagagcttgg tgtctctggt      510

```

```

<210> 30
<211> 1062
<212> DNA
<213> 智人

```

```

<220>
<221> misc_feature
<222> (1)..(1062)
<223> 用於 hOTC 的 CDS

```

```

<400> 30
atgctgttta atctgaggat cctgttaaac aatgcagctt ttagaaatgg tcacaacttc      60
atggttcgaa attttcggtg tggacaacca ctacaaaata aagtgcagct gaagggccgt      120
gaccttctca ctotaaaaaa ctttaccgga gaagaaatta aatatatgct atggctatca      180
gcagatctga aathtaggat aaaacagaaa ggagagtatt tgcccttatt gcaagggaa      240
tccttaggca tgatTTTTga gaaaagaagt actcgaacaa gattgtctac agaaacaggc      300
tttgcacttc tgggaggaca tccttgtttt cttaccacac aagatattca tttgggtgtg      360
aatgaaagtc tcacggacac ggcccgtgta ttgtctagca tggcagatgc agtattggct      420
cgagtgtata aacaatcaga tttggacacc ctggctaaag aagcatccat cccaattatc      480
aatgggctgt cagatttcta ccatcctatc cagatcctgg ctgattacct cacgctccag      540
gaacactata gctctctgaa aggtcttacc ctgagctgga tcggggatgg gaacaatatc      600
ctgcactcca tcatgatgag cgcagcgaaa ttcggaatgc accttcaggc agctactcca      660
aagggttatg agccggatgc tagtctaacc aagttggcag agcagtatgc caaagagaat      720
ggtaccaagc tgttctgac aatgatcca ttggaagcag cgcctggagg caatgtatta      780

```

attacagaca cttggataag catgggacaa gaagaggaga agaaaaagcg gctccaggct 840  
 ttccaaggtt accaggttac aatgaagact gctaaagttg ctgcctctga ctggacattt 900  
 ttacactgct tgcccagaaa gccagaagaa gtggatgatg aagtctttta ttctcctcga 960  
 tcactagtgt tcccagaggc agaaaacaga aagtgacaaa tcatggctgt catggtgtcc 1020  
 ctgctgacag attactcacc tcagctccag aagcctaaat tt 1062

<210> 31  
 <211> 9  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> CRE基序

<400> 31

Leu Ala Gly Leu Ile Asp Ala Asp Gly  
 1 5

<210> 32  
 <211> 22  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> CRE 識別序列

<400> 32  
 caaacgtcg tgagacagtt tg 22

<210> 33  
 <211> 2214  
 <212> DNA  
 <213> 腺相關病毒 rh79

<220>  
 <221> CDS  
 <222> (1).. (2214)

<400> 33  
 atg gct gct gac ggt tat ctt oca gat tgg ctc gag gac aac ctc tct 48  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15

gag ggc att cgc gag tgg tgg gac ctg aaa cct gga gcc ccc aag ccc 96  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30

aag gcc aac cag cag aag cag gac gac ggc cgg ggt ctg gtg ctt cct 144

Lys	Ala	Asn	Gln	Gln	Lys	Gln	Asp	Asp	Gly	Arg	Gly	Leu	Val	Leu	Pro		
		35					40					45					
ggc	tac	aag	tac	ctc	gga	ccc	ttc	aac	gga	ctc	gac	aag	ggg	gag	ccc		192
Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Phe	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro		
	50					55					60						
gtc	aac	gag	gcg	gac	gcc	gcg	gcc	ctc	gag	cac	gac	aag	gcc	tac	gac		240
Val	Asn	Glu	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp		
	65				70					75					80		
cag	cag	ctc	aaa	gcg	ggt	gac	aat	ccg	tac	ctg	cgg	tat	aac	cac	gcc		288
Gln	Gln	Leu	Lys	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Arg	Tyr	Asn	His	Ala		
			85						90					95			
gac	gcc	gag	ttt	cag	gag	cgt	ctg	caa	gaa	gat	acg	tct	ttt	ggg	ggc		336
Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly		
			100					105					110				
aac	ctc	ggg	cga	gca	gtc	ttc	cag	gcc	aag	aag	cgg	gtt	ctc	gaa	cct		384
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro		
		115					120					125					
ctc	ggt	ctg	gtt	gag	gaa	gct	gct	aag	acg	gct	cct	gga	aag	aag	aga		432
Leu	Gly	Leu	Val	Glu	Glu	Ala	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg		
	130					135					140						
cgg	gta	gaa	cgg	tca	cct	cag	cga	tcc	ccc	gac	tcc	tcc	acg	ggc	atc		480
Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile		
	145				150					155					160		
ggc	aaa	aaa	ggc	cag	cag	ccc	gcg	aga	aag	aga	ctg	aac	ttt	ggg	cag		528
Gly	Lys	Lys	Gly	Gln	Gln	Pro	Ala	Arg	Lys	Arg	Leu	Asn	Phe	Gly	Gln		
			165						170				175				
act	ggc	gac	tca	gag	tca	gtc	ccc	gac	cct	caa	cca	atc	gga	gaa	cca		576
Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro		
			180					185					190				
cca	gca	ggc	ccc	tct	ggt	ctg	gga	tct	ggt	aca	atg	gct	gca	ggc	ggt		624
Pro	Ala	Gly	Pro	Ser	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly		
		195					200					205					
ggc	gct	cca	atg	gca	gac	aat	aac	gaa	ggc	gcc	gac	gga	gtg	ggt	agt		672
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser		
	210					215					220						
tcc	tca	gga	aat	tgg	cat	tgc	gat	tcc	aca	tgg	ctg	ggc	gac	aga	gtc		720
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val		
	225				230					235					240		
atc	acc	acc	agc	acc	cga	acc	tgg	gcc	ctg	ccc	acc	tac	aac	aac	cac		768
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His		
			245						250					255			
ctc	tac	aag	caa	atc	tcc	aat	ggg	aca	tgc	gga	gga	agc	acc	aac	gac		816
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp		

260			265			270										
aac	acc	tac	ttc	ggc	tac	agc	acc	ccc	tgg	ggg	tat	ttt	gac	ttc	aac	864
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	
		275					280					285				
aga	ttc	cac	tgt	cac	ttc	tca	cca	cgt	gac	tgg	cag	aga	ctc	atc	aac	912
Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	
	290					295					300					
aac	aac	tgg	gga	ttc	cgg	ccc	aag	aga	ctc	agc	ttc	aag	ctc	ttc	aac	960
Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn	
305					310					315					320	
atc	cag	ggt	aag	gag	gtc	acg	cag	aat	gaa	ggc	acc	aag	acc	atc	gcc	1008
Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	
				325					330					335		
aat	aac	ctt	acc	agc	acg	att	cag	gta	ttt	acg	gac	tcg	gaa	tac	cag	1056
Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	
			340					345					350			
ctg	cgc	tac	gtc	ctc	ggc	tcc	gcg	cac	cag	ggc	tgc	ctg	cct	cgc	ttc	1104
Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	
		355					360					365				
cgc	gcg	gat	gtc	ttc	atg	att	ccc	cag	tac	ggc	tac	ctg	aca	ctg	aac	1152
Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	
	370					375					380					
aac	gga	agt	caa	gcc	gta	ggc	cgt	tcc	tca	ttc	tac	tgc	ctg	gaa	tat	1200
Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	
385					390					395					400	
ttt	cca	tct	caa	atg	ctg	cgg	act	gga	aac	aac	ttt	gaa	ttt	agc	tac	1248
Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	
				405					410					415		
acc	ttt	gag	gac	gtg	ccc	ttc	cac	agc	agc	tac	gca	cac	agc	cag	agc	1296
Thr	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	
			420					425					430			
ctg	gac	cgg	ctg	atg	aac	cct	ctc	atc	gac	cag	tac	ctg	tat	tac	cta	1344
Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	
		435					440					445				
tcc	aga	act	cag	tcc	aca	gga	gga	act	caa	ggt	aca	cag	caa	ttg	tta	1392
Ser	Arg	Thr	Gln	Ser	Thr	Gly	Gly	Thr	Gln	Gly	Thr	Gln	Gln	Leu	Leu	
	450					455					460					
ttt	tct	caa	gcc	ggg	cct	gca	aat	atg	tcg	gct	cag	gcc	aag	aac	tgg	1440
Phe	Ser	Gln	Ala	Gly	Pro	Ala	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	
465					470					475					480	
cta	cct	gga	cct	tgc	tac	cgg	cag	cag	cga	gtc	tcc	acg	aca	ctg	tcg	1488
Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	
				485					490					495		

caa aac aac aac agc aac ttt gct tgg act ggt gcc acg aaa tat cat	1536
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His	
500 505 510	
ctg aac gga aga gac tct ttg gtg aat ccc ggt gtt gct atg gca acg	1584
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr	
515 520 525	
cat aag gac gac gag gaa cgt ttc ttt cca tcg agc gga gtc ctg atg	1632
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met	
530 535 540	
ttt gga aaa cag ggt gct gga aga gac aat gtg gac tat agc agc gtt	1680
Phe Gly Lys Gln Gly Ala Gly Arg Asp Asn Val Asp Tyr Ser Ser Val	
545 550 555 560	
atg cta acc agc gag gaa gaa atc aag acc act aac cct gta gcc act	1728
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr	
565 570 575	
gaa caa tac ggc gtg gtg gct gat aac ttg cag caa acc aat aca gga	1776
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Thr Gly	
580 585 590	
cct atc gtg gga aat gtc aac agc caa gga gcc tta cct ggc atg gtc	1824
Pro Ile Val Gly Asn Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val	
595 600 605	
tgg cag aac cga gac gtg tac ctg cag ggt ccc att tgg gcc aag att	1872
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile	
610 615 620	
cct cac acg gac ggc aac ttt cac ccg tct cct ctg atg ggc ggc ttt	1920
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe	
625 630 635 640	
gga ctg aaa cac ccg cct cct caa atc ctg atc aag aac act ccc gtt	1968
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val	
645 650 655	
cct gcg gat cct cca acg acg ttc agc cag gcg aaa ttg gct tcc ttc	2016
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe	
660 665 670	
atc acg cag tat agt acc ggc cag gtc agc gtg gag atc gag tgg gag	2064
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu	
675 680 685	
ctg cag aag gag aac agc aag cgc tgg aac cca gaa att cag tat act	2112
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr	
690 695 700	
tcc aac tac tac aaa tct aca aat gtg gac ttt gct gtc aat acc gag	2160
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu	
705 710 715 720	

ggt aca tat tca gag cct cgc ccc att gga act cgt tac ctc acc cgt 2208  
 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735

aat ctg 2214  
 Asn Leu

<210> 34  
 <211> 738  
 <212> PRT  
 <213> 腺相關病毒 rh79

<400> 34

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln



Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
405 410 415

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
435 440 445

Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu  
450 455 460

Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp  
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
500 505 510

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
530 535 540

Phe Gly Lys Gln Gly Ala Gly Arg Asp Asn Val Asp Tyr Ser Ser Val  
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Thr Gly  
580 585 590

Pro Ile Val Gly Asn Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
725 730 735

Asn Leu

<210> 35  
<211> 2214  
<212> DNA  
<213> 腺相關病毒 8

<220>  
<221> CDS  
<222> (1).. (2214)

<400> 35  
atg gct gcc gat ggt tat ctt cca gat tgg ctc gag gac aac ctc tct 48  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15  
gag ggc att cgc gag tgg tgg gcg ctg aaa cct gga gcc ccg aag ccc 96  
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30  
aaa gcc aac cag caa aag cag gac gac ggc cgg ggt ctg gtg ctt cct 144  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45  
ggc tac aag tac ctc gga ccc ttc aac gga ctc gac aag ggg gag ccc 192

Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Phe	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro	
50						55					60					
gtc	aac	gcg	gcg	gac	gca	gcg	gcc	ctc	gag	cac	gac	aag	gcc	tac	gac	240
Val	Asn	Ala	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp	
65					70					75					80	
cag	cag	ctg	cag	gcg	ggt	gac	aat	ccg	tac	ctg	cgg	tat	aac	cac	gcc	288
Gln	Gln	Leu	Gln	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Arg	Tyr	Asn	His	Ala	
				85					90					95		
gac	gcc	gag	ttt	cag	gag	cgt	ctg	caa	gaa	gat	acg	tct	ttt	ggg	ggc	336
Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	
			100					105					110			
aac	ctc	ggg	cga	gca	gtc	ttc	cag	gcc	aag	aag	cgg	gtt	ctc	gaa	cct	384
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	
		115					120					125				
ctc	ggt	ctg	gtt	gag	gaa	ggc	gct	aag	acg	gct	cct	gga	aag	aag	aga	432
Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	
	130					135					140					
cgg	gta	gag	cca	tca	ccc	cag	cgt	tct	cca	gac	tcc	tct	acg	ggc	atc	480
Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	
					150					155					160	
ggc	aag	aaa	ggc	caa	cag	ccc	gcc	aga	aaa	aga	ctc	aat	ttt	ggt	cag	528
Gly	Lys	Lys	Gly	Gln	Gln	Pro	Ala	Arg	Lys	Arg	Leu	Asn	Phe	Gly	Gln	
				165					170				175			
act	ggc	gac	tca	gag	tca	gtt	cca	gac	cct	caa	cct	ctc	gga	gaa	cct	576
Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro	
			180					185					190			
cca	gca	gcg	ccc	tct	ggt	gtg	gga	cct	aat	aca	atg	gct	gca	ggc	ggt	624
Pro	Ala	Ala	Pro	Ser	Gly	Val	Gly	Pro	Asn	Thr	Met	Ala	Ala	Gly	Gly	
		195				200						205				
ggc	gca	cca	atg	gca	gac	aat	aac	gaa	ggc	gcc	gac	gga	gtg	ggt	agt	672
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	
	210					215					220					
tcc	tcg	gga	aat	tgg	cat	tgc	gat	tcc	aca	tgg	ctg	ggc	gac	aga	gtc	720
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	
					230					235					240	
atc	acc	acc	agc	acc	cga	acc	tgg	gcc	ctg	ccc	acc	tac	aac	aac	cac	768
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	
				245					250					255		
ctc	tac	aag	caa	atc	tcc	aac	ggg	aca	tcg	gga	gga	gcc	acc	aac	gac	816
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ala	Thr	Asn	Asp	
			260					265					270			
aac	acc	tac	ttc	ggc	tac	agc	acc	ccc	tgg	ggg	tat	ttt	gac	ttt	aac	864
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	

275	280	285	
aga ttc cac tgc cac ttt tca cca cgt gac tgg cag cga ctc atc aac			912
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn			
290	295	300	
aac aac tgg gga ttc cgg ccc aag aga ctc agc ttc aag ctc ttc aac			960
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn			
305	310	315	320
atc cag gtc aag gag gtc acg cag aat gaa ggc acc aag acc atc gcc			1008
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala			
325	330	335	
aat aac ctc acc agc acc atc cag gtg ttt acg gac tcg gag tac cag			1056
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln			
340	345	350	
ctg ccg tac gtt ctc ggc tct gcc cac cag ggc tgc ctg cct ccg ttc			1104
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe			
355	360	365	
ccg gcg gac gtg ttc atg att ccc cag tac ggc tac cta aca ctc aac			1152
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn			
370	375	380	
aac ggt agt cag gcc gtg gga cgc tcc tcc ttc tac tgc ctg gaa tac			1200
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr			
385	390	395	400
ttt cct tcg cag atg ctg aga acc ggc aac aac ttc cag ttt act tac			1248
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Thr Tyr			
405	410	415	
acc ttc gag gac gtg cct ttc cac agc agc tac gcc cac agc cag agc			1296
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser			
420	425	430	
ttg gac cgg ctg atg aat cct ctg att gac cag tac ctg tac tac ttg			1344
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu			
435	440	445	
tct cgg act caa aca aca gga ggc acg gca aat acg cag act ctg gcc			1392
Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly			
450	455	460	
ttc agc caa ggt ggg cct aat aca atg gcc aat cag gca aag aac tgg			1440
Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp			
465	470	475	480
ctg cca gga ccc tgt tac cgc caa caa cgc gtc tca acg aca acc ggg			1488
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly			
485	490	495	
caa aac aac aat agc aac ttt gcc tgg act gct ggg acc aaa tac cat			1536
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His			
500	505	510	

ctg aat gga aga aat tca ttg gct aat cct ggc atc gct atg gca aca Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr 515 520 525	1584
cac aaa gac gac gag gag cgt ttt ttt ccc agt aac ggg atc ctg att His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile 530 535 540	1632
ttt ggc aaa caa aat gct gcc aga gac aat gcg gat tac agc gat gtc Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val 545 550 555 560	1680
atg ctc acc agc gag gaa gaa atc aaa acc act aac cct gtg gct aca Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr 565 570 575	1728
gag gaa tac ggt atc gtg gca gat aac ttg cag cag caa aac acg gct Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala 580 585 590	1776
cct caa att gga act gtc aac agc cag ggg gcc tta ccc ggt atg gtc Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val 595 600 605	1824
tgg cag aac cgg gac gtg tac ctg cag ggt ccc atc tgg gcc aag att Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile 610 615 620	1872
cct cac acg gac ggc aac ttc cac ccg tct ccg ctg atg ggc ggc ttt Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe 625 630 635 640	1920
ggc ctg aaa cat cct ccg cct cag atc ctg atc aag aac acg cct gta Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val 645 650 655	1968
cct gcg gat cct ccg acc acc ttc aac cag tca aag ctg aac tct ttc Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe 660 665 670	2016
atc acg caa tac agc acc gga cag gtc agc gtg gaa att gaa tgg gag Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 675 680 685	2064
ctg cag aag gaa aac agc aag cgc tgg aac ccc gag atc cag tac acc Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr 690 695 700	2112
tcc aac tac tac aaa tct aca agt gtg gac ttt gct gtt aat aca gaa Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu 705 710 715 720	2160
ggc gtg tac tct gaa ccc cgc ccc att ggc acc cgt tac ctc acc cgt Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg 725 730 735	2208

aat ctg  
Asn Leu

2214

<210> 36  
<211> 738  
<212> PRT  
<213> 腺相關病毒 8  
  
<400> 36

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Gln Gln Leu Gln Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro

			180						185							190			
Pro	Ala	Ala	Pro	Ser	Gly	Val	Gly	Pro	Asn	Thr	Met	Ala	Ala	Gly	Gly				
		195					200					205							
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser				
	210					215					220								
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val				
225					230					235					240				
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His				
				245					250					255					
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ala	Thr	Asn	Asp				
			260					265					270						
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn				
		275					280					285							
Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn				
	290					295					300								
Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn				
305					310					315					320				
Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala				
				325					330					335					
Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln				
			340					345					350						
Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe				
		355					360					365							
Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn				
	370					375					380								
Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr				
385					390					395					400				
Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Thr	Tyr				
				405					410					415					

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
435 440 445

Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly  
450 455 460

Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp  
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly  
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His  
500 505 510

Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr  
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile  
530 535 540

Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val  
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
565 570 575

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

Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe  
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu  
705 710 715 720

Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
725 730 735

Asn Leu

<210> 37  
<211> 2214  
<212> DNA  
<213> 腺相關病毒 hu37

<220>  
<221> CDS  
<222> (1)..(2214)

<400> 37  
atg gct gct gac ggt tat ctt cca gat tgg ctc gag gac aac ctc tct 48  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

gag ggc att cgc gag tgg tgg gac ctg aaa cct gga gcc ccc aag ccc 96  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

aag gcc aac cag cag aag cag gac gac ggc cgg ggt ctg gtg ctt cct 144  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

ggc tac aag tac ctc gga ccc ttc aac gga ctc gac aag ggg gag ccc 192  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

gtc aac gcg gcg gac gca gcg gcc ctc gag cac gac aag gcc tac gac 240

Val 65	Asn	Ala	Ala	Asp	Ala 70	Ala	Ala	Leu	Glu	His 75	Asp	Lys	Ala	Tyr	Asp 80	
cag Gln	cag Gln	ctc Leu	aaa Lys	gcg Ala	ggt Gly	gac Asp	aat Asn	ccg Pro	tac Tyr	ctg Leu	cgg Arg	tat Tyr	aac Asn	cac His	gcc Ala	288
				85					90					95		
gac Asp	gcc Ala	gag Glu	ttt Phe	cag Gln	gag Glu	cgt Arg	ctg Leu	caa Gln	gaa Glu	gat Asp	acg Thr	tct Ser	ttt Phe	ggg Gly	ggc Gly	336
			100					105					110			
aac Asn	ctc Leu	ggg Gly	cga Arg	gca Ala	gtc Val	ttc Phe	cag Gln	gcc Ala	aag Lys	aag Lys	cgg Arg	gtt Val	ctc Leu	gaa Glu	cct Pro	384
		115					120					125				
ctc Leu	ggt Gly	ctg Leu	gtt Val	gag Glu	gaa Glu	gct Ala	gct Ala	aag Lys	acg Thr	gct Ala	cct Pro	gga Gly	aag Lys	aag Lys	aga Arg	432
	130					135					140					
cgg Pro	gta Val	gaa Glu	cgg Pro	tca Ser	cct Pro	cag Gln	cgt Arg	tcc Ser	ccc Pro	gac Asp	tcc Ser	tcc Ser	acg Thr	ggc Gly	atc Ile	480
	145				150					155					160	
ggc Gly	aag Lys	aaa Lys	ggc Gly	cag Gln	cag Gln	ccc Pro	gct Ala	aaa Lys	aag Lys	aga Arg	ctg Leu	aac Asn	ttt Phe	ggt Gly	cag Gln	528
				165					170				175			
act Thr	ggc Gly	gac Asp	tca Ser	gag Glu	tca Ser	gtc Val	ccc Pro	gac Asp	cct Pro	caa Gln	cca Pro	atc Ile	gga Gly	gaa Glu	cca Pro	576
			180					185					190			
cca Pro	gca Ala	ggc Gly	ccc Pro	tct Ser	ggt Gly	ctg Leu	gga Gly	tct Ser	ggt Gly	aca Thr	atg Met	gct Ala	gca Ala	ggc Gly	ggt Gly	624
		195				200					205					
ggc Gly	gct Ala	cca Pro	atg Met	gca Ala	gac Asp	aat Asn	aac Asn	gaa Glu	ggc Gly	gcc Ala	gac Asp	gga Gly	gtg Val	ggt Gly	agt Ser	672
	210					215					220					
tcc Ser	tca Ser	gga Gly	aat Asn	tgg Trp	cat His	tgc Cys	gat Asp	tcc Ser	aca Thr	tgg Trp	ctg Leu	ggc Gly	gac Asp	aga Arg	gtc Val	720
	225				230					235					240	
atc Ile	acc Thr	acc Thr	agc Ser	acc Thr	cga Arg	acc Thr	tgg Trp	gcc Ala	ctg Leu	ccc Pro	acc Thr	tac Tyr	aac Asn	aac Asn	cac His	768
				245					250				255			
ctc Leu	tac Tyr	aag Lys	caa Gln	ata Ile	tcc Ser	aat Asn	ggg Gly	aca Thr	tcg Ser	gga Gly	gga Gly	agc Ser	acc Thr	aac Asn	gac Asp	816
			260					265					270			
aac Asn	acc Thr	tac Tyr	ttc Phe	ggc Gly	tac Tyr	agc Ser	acc Thr	ccc Pro	tgg Trp	ggg Gly	tat Tyr	ttt Phe	gac Asp	ttc Phe	aac Asn	864
		275					280					285				
aga Arg	ttc Phe	cac His	tgc Cys	cac His	ttc Phe	tca Ser	cca Pro	cgt Arg	gac Asp	tgg Trp	cag Gln	cga Arg	ctc Leu	atc Ile	aac Asn	912

290	295	300	
aac aac tgg gga ttc cgg cca aaa aga ctc agc ttc aag ctc ttc aac Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn 305 310 315 320			960
atc cag gtc aag gag gtc acg cag aat gaa ggc acc aag acc atc gcc Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala 325 330 335			1008
aat aac ctt acc agc acg att cag gta ttt acg gac tcg gaa tac cag Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln 340 345 350			1056
ctg ccg tac gtc ctc ggc tcc gcg cac cag ggc tgc ctg cct ccg ttc Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe 355 360 365			1104
ccg gcg gac gtc ttc atg att ccc cag tac ggc tac ctt aca ctg aac Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn 370 375 380			1152
aat gga agt caa gcc gta ggc cgt tcc tcc ttc tac tgc ctg gaa tat Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr 385 390 395 400			1200
ttt cca tct caa atg ctg cga act gga aac aat ttt gaa ttc agc tac Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr 405 410 415			1248
acc ttc gag gac gtg cct ttc cac agc agc tac gca cac agc cag agc Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser 420 425 430			1296
ttg gac cga ctg atg aat cct ctc atc gac cag tac ctg tac tac tta Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu 435 440 445			1344
tcc aga act cag tcc aca gga gga act caa ggt acc cag caa ttg tta Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu 450 455 460			1392
ttt tct caa gct ggg cct gca aac atg tcg gct cag gct aag aac tgg Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp 465 470 475 480			1440
cta cct gga cct tgc tac cgg cag cag cga gtc tct acg aca ctg tcg Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser 485 490 495			1488
caa aac aac aac agc aac ttt gct tgg act ggt gcc acc aaa tat cac Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His 500 505 510			1536
ctg aac gga aga gac tct ttg gta aat ccc ggt gtc gcc atg gca acc Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr 515 520 525			1584

cac aag gac gac gag gaa cgc ttc ttc ccg tcg agt gga gtc ctg atg His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met 530 535 540	1632
ttc gga aaa cag ggt gct gga aga gac aat gtg gac tac agc agc gtt Phe Gly Lys Gln Gly Ala Gly Arg Asp Asn Val Asp Tyr Ser Ser Val 545 550 555 560	1680
atg cta acc agc gaa gaa gaa att aaa acc act aac ccc gta gcc aca Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr 565 570 575	1728
gaa caa tac ggt gtg gtg gct gac aac ttg cag caa acc aat aca ggg Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Thr Gly 580 585 590	1776
cct att gtg gga aat gtc aac agc caa gga gcc tta cct ggc atg gtc Pro Ile Val Gly Asn Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val 595 600 605	1824
tgg cag aac cga gac gtg tac ctg cag ggt ccc atc tgg gcc aag att Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile 610 615 620	1872
cct cac acg gac ggc aac ttc cac cct tca ccg cta atg gga gga ttt Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe 625 630 635 640	1920
gga ctg aag cac cca cct cct cag atc ctg atc aag aac acg ccg gta Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val 645 650 655	1968
cct gcg gat cct cca aca acg ttc agc cag gcg aaa ttg gct tcc ttc Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe 660 665 670	2016
att acg cag tac agc acc gga cag gtc agc gtg gaa atc gag tgg gag Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 675 680 685	2064
ctg cag aag gag aac agc aaa cgc tgg aac cca gag att cag tac act Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr 690 695 700	2112
tca aac tac tac aaa tct aca aat gtg gac ttt gct gtc aat aca gag Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu 705 710 715 720	2160
gga act tat tct gag cct cgc ccc att ggt act cgt tac ctc acc cgt Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg 725 730 735	2208
aat ctg Asn Leu	2214

<210> 38  
 <211> 738  
 <212> PRT  
 <213> 腺相關病毒 hu37

<400> 38

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly

195                                      200                                      205  
 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
     210                                      215                                      220  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
     225                                      230                                      235                                      240  
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
                                     245                                      250                                      255  
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
                                     260                                      265                                      270  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
                                     275                                      280                                      285  
 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
                                     290                                      295                                      300  
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn  
     305                                      310                                      315  
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
                                     325                                      330                                      335  
 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
                                     340                                      345                                      350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
                                     355                                      360                                      365  
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
                                     370                                      375                                      380  
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
     385                                      390                                      395                                      400  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
                                     405                                      410                                      415  
 Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
                                     420                                      425                                      430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445

Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu  
 450 455 460

Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp  
 465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
 485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
 500 505 510

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540

Phe Gly Lys Gln Gly Ala Gly Arg Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Thr Gly  
 580 585 590

Pro Ile Val Gly Asn Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
 645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735

Asn Leu

<210> 39  
 <211> 2217  
 <212> DNA  
 <213> 腺相關病毒 rh. 90

<220>  
 <221> CDS  
 <222> (1).. (2217)

<400> 39	
atg gct gcc gat ggt tat ctt cca gat tgg ctc gag gac aac ctc tct	48
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser	
1 5 10 15	
gag ggc att cgc gag tgg tgg gac ctg aaa cct gga gcc cca aaa ccc	96
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro	
20 25 30	
aaa gcc aac cag caa aag cag gac gac ggc cgg ggt ctg gtg ctt cct	144
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro	
35 40 45	
ggc tac aag tac ctc gga ccc ttc aac gga ctc gac aag ggg gag ccc	192
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro	
50 55 60	
gtc aac gcg gcg gac gca gcg gcc ctc gag cac gac aag gcc tac gac	240
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp	
65 70 75 80	
cag cag ctg cag gcg ggt gat aat ccg tac ctg cgg tat aac cac gcc	288

Gln	Gln	Leu	Gln	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Arg	Tyr	Asn	His	Ala		
				85					90					95			
gac	gcc	gag	ttt	cag	gag	cgt	ctg	caa	gaa	gat	acg	tca	ttt	ggg	ggc		336
Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly		
			100					105					110				
aac	ctc	ggg	cga	gca	gtc	ttc	cag	gcc	aag	aag	cgg	gtt	ctc	gaa	cct		384
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro		
		115					120					125					
ctc	ggt	ctg	gtt	gag	gaa	ggc	gct	aag	acg	gct	cct	gga	aag	aag	aga		432
Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg		
	130					135					140						
cag	gta	gag	cca	tca	cca	cag	cgt	tcc	ccc	gac	tcc	tcc	acg	ggc	atc		480
Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile		
	145				150					155					160		
ggc	aag	aaa	ggc	cag	cag	ccc	gcc	aga	aag	aga	ctc	aat	ttc	ggt	cag		528
Gly	Lys	Lys	Gly	Gln	Gln	Pro	Ala	Arg	Lys	Arg	Leu	Asn	Phe	Gly	Gln		
				165					170					175			
act	ggc	gac	tca	gag	tca	gtc	ccc	gac	cct	caa	cct	ctc	gga	gaa	cct		576
Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro		
			180					185					190				
cca	gca	gcg	ccc	tct	agt	gtg	gga	tct	ggt	aca	atg	gct	gca	ggc	ggt		624
Pro	Ala	Ala	Pro	Ser	Ser	Val	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly		
		195					200					205					
ggc	gca	cca	atg	gca	gac	aat	aac	gaa	ggt	gcc	gac	gga	gtg	ggt	agt		672
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser		
	210					215					220						
tcc	tcg	gga	aat	tgg	cat	tgc	gat	tcc	aca	tgg	ctg	ggc	gac	aga	gtc		720
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val		
					230					235					240		
atc	acc	acc	agc	acc	cga	acc	tgg	gcc	ctg	ccc	acc	tac	aac	aac	cac		768
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His		
				245					250					255			
ctc	tac	aag	caa	atc	tcc	aac	ggg	acc	tcg	gga	ggc	agc	acc	aac	gac		816
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp		
			260					265					270				
aac	acc	tac	ttc	ggc	tac	agc	acc	ccc	tgg	ggg	tat	ttt	gac	ttt	aac		864
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn		
		275					280					285					
aga	ttc	cac	tgc	cac	ttc	tca	cca	cgt	gac	tgg	cag	cga	ctt	atc	aac		912
Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn		
	290					295					300						
aac	aac	tgg	gga	ttc	cgg	ccc	aag	aga	ctc	agc	ttc	aag	ctc	ttc	aac		960
Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn		



ttt ggc aag cag ggg gct gga aaa gac aac gtg gac tat agc aac gtg Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Asn Val 545 550 555 560	1680
atg cta acc agc gag gaa gaa atc aag acc act aac cct gtg gcc aca Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr 565 570 575	1728
gaa cag tat ggt gtg gtg gcg gat aac ctg cag cag caa aac aca gct Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala 580 585 590	1776
cct att gtg ggg gcc gtc aac agc caa gga gcc tta cct ggc atg gtt Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val 595 600 605	1824
tgg cag aac cgg gac gtg tac ctg cag ggt ccc atc tgg gcc aag att Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile 610 615 620	1872
cct cac acg gat ggt aac ttt cac ccg tct cct ctc atg ggc ggc ttt Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe 625 630 635 640	1920
gga ctt aaa cat ccg cct cct cag atc ctg atc aag aac act ccc gtt Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val 645 650 655	1968
cct gcg gat cct cca acg gcg ttc aac cag gcc aag ctg aac tct ttc Pro Ala Asp Pro Pro Thr Ala Phe Asn Gln Ala Lys Leu Asn Ser Phe 660 665 670	2016
atc acg cag tac agc acc gga caa gtc agc gtg gag atc gag tgg gag Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 675 680 685	2064
ctg cag aag gag aac agc aag cgc tgg aac cca gag att cag tat acc Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr 690 695 700	2112
tcc aac tac tac aaa tct aca aat gtg gac ttt gct gtt aat act gag Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu 705 710 715 720	2160
ggt gtt tac tct gag cct cgc ccc att ggc acc cgt tac ctc acc cgt Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg 725 730 735	2208
aat ctg taa Asn Leu	2217

<210> 40  
<211> 738  
<212> PRT

<213> 腺相關病毒 rh. 90

<400> 40

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Gln Gln Leu Gln Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro  
180 185 190

Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly  
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser



Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
450 455 460

Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp  
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
500 505 510

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
515 520 525

Asn Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Ile Leu Met  
530 535 540

Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Asn Val  
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala  
580 585 590

Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
645 650 655

Pro Ala Asp Pro Pro Thr Ala Phe Asn Gln Ala Lys Leu Asn Ser Phe  
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720

Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735

Asn Leu

<210> 41  
 <211> 434  
 <212> DNA  
 <213> 智人

<220>  
 <221> misc\_feature  
 <222> (1)..(434)  
 <223> TBG 啟動子

<400> 41  
 tgtataattt ctacagaacc tattagaaag gatcaccag cctctgcttt tgtacaactt 60  
 tcccttaaaa aactgccaat tccaactgctg tttggcccaa tagtgagaac tttttcctgc 120  
 tgccctttgg tgcttttgcc tatggccctt attctgcctg ctgaagacac tcttgccagc 180  
 atggacttaa acccctccag ctctgacaat cctctttctc ttttgtttta catgaagggt 240  
 ctggcagcca aagcaatcac tcaaagttca aacctatca ttttttgctt tgttcctctt 300  
 ggcccttggtt ttgtacatca gctttgaaaa taccatocca gggttaatgc tggggttaat 360  
 ttataactaa gagtgcctta gttttgcaat acaggacatg ctataaaaaat ggaaagatgt 420  
 tgctttctga gaga 434

<210> 42  
 <211> 3181  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 構建序列

<220>  
 <221> repeat\_region  
 <222> (1)..(130)  
 <223> 5' ITR

<220>  
 <221> promoter  
 <222> (211)..(907)  
 <223> 具有增強子之TBG

<220>  
 <221> promoter  
 <222> (211)..(310)  
 <223>  $\alpha$  mic/bik 增強子

<220>  
 <221> promoter  
 <222> (317)..(419)  
 <223>  $\alpha$  mic/bik 增強子

<220>  
 <221> promoter  
 <222> (474)..(907)  
 <223> TBG 啟動子

<220>  
 <221> Intron  
 <222> (939)..(1071)

<220>  
 <221> CDS  
 <222> (1089)..(2183)  
 <223> ARCUS

<220>  
 <221> misc\_feature  
 <222> (2022)..(2743)  
 <223> WPRE

<220>  
 <221> polyA\_signal  
 <222> (2750)..(2964)  
 <223> BGH polyA

<220>  
 <221> repeat\_region  
 <222> (3052)..(3181)  
 <223> 3' ITR

<400> 42  
 ctgctgctc gctgctcac tgaggccgcc cgggcaaagc ccgggctcg ggcgacctt 60  
 ggtgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120  
 aggggttct ttagttaat gattaaccg ccatgctact tatctacgta gccatgctct 180

```

aggaagatcg gaattogccc ttaagctagc aggttaattt ttaaaaagca gtcaaaagtc      240
caagtggccc ttggcagcat ttaactctctc tgtttgctct ggtaataat ctccaggagca      300
caaacattcc agatccaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg      360
cagcatttac tctctctggt tgctctgggt aataatctca ggagcacaaa cattccagat      420
ccggcgcgcc agggctggaa gctaccttg acatcatttc ctctgcgaat gcatgtataa      480
tttctacaga acctattaga aaggatcacc cagcctctgc ttttgtacaa ctttccctta      540
aaaaactgcc aattccactg ctgtttggcc caatagttag aactttttcc tgctgcctct      600
tggtgctttt gcctatggcc cctattctgc ctgctgaaga cactcttggc agcatggact      660
taaaccctc cagctctgac aatcctcttt ctcttttgtt ttacatgaag ggtctggcag      720
ccaaagcaat cactcaaagt tcaaacctta tcattttttg ctttgttcct cttggccttg      780
gttttgtaca tcagctttga aaataccatc ccagggttaa tgctgggggt aatttataac      840
taagagtgtc ctagttttgc aatacaggac atgctataaa aatggaaaga tgttgctttc      900
tgagagactg cagaagttgg tcgtgaggca ctgggcaggt aagtatcaag gttacaagac      960
aggtttaagg agaccaatag aaactgggct tgtcgagaca gagaagactc ttgcgtttct    1020
gataggcacc tattggtctt actgacatcc actttgcctt tctctccaca ggtgtccagg    1080
cggcgcgcc atg gca ccg aag aag aag cgc aag gtg cat atg aat aca aaa      1130
      Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys
      1           5           10

tat aat aaa gag ttc tta ctc tac tta gca ggg ttt gta gac ggt gac      1178
Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp
15           20           25           30

ggt tcc atc ttt gcc agg atc aag cct agt caa cgt agt aag ttc aag      1226
Gly Ser Ile Phe Ala Arg Ile Lys Pro Ser Gln Arg Ser Lys Phe Lys
35           40           45

cac aag ctg cat ctc gtt ttc gct gtc tat cag aag aca cag cgc cgt      1274
His Lys Leu His Leu Val Phe Ala Val Tyr Gln Lys Thr Gln Arg Arg
50           55           60

tgg ttc ctc gac aag ctg gtg gac gag atc ggt gtg ggt tac gtg ctg      1322
Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Leu
65           70           75

gac tct ggc agc gtc tcc ttt tac tcg ctg tcc gag atc aag cct ttg      1370
Asp Ser Gly Ser Val Ser Phe Tyr Ser Leu Ser Glu Ile Lys Pro Leu
80           85           90

cat aat ttt tta aca caa cta caa cct ttt cta aaa cta aaa caa aaa      1418
His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys

```

95		100		105		110	
caa gca aat tta gtt tta aaa att att gaa caa ctt ccg tca gca aaa							1466
Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys							
	115			120		125	
gaa tcc ccg gac aaa ttc tta gaa gtt tgt aca tgg gtg gat caa att							1514
Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile							
	130			135		140	
gca gct ctg aat gat tcg aag acg cgt aaa aca act tct gaa acc gtt							1562
Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val							
	145			150		155	
cgt gct gtg cta gac agt tta cca gga tcc gtg gga ggt cta tcg cca							1610
Arg Ala Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro							
	160			165		170	
tct cag gca tcc agc gcc gca tcc tcg gct tcc tca agc ccg ggt tca							1658
Ser Gln Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser							
	175			180		185	
ggg atc tcc gaa gca ctc aga gct gga gca ggt tcc ggc act gga tac							1706
Gly Ile Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr							
	195			200		205	
aac aag gaa ttc ctg ctc tac ctg gcg ggc ttc gtc gac ggg gac ggc							1754
Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly							
	210			215		220	
tcc atc tat gcc cgt atc aag ccg gtt cag cgg gct aag ttc aag cac							1802
Ser Ile Tyr Ala Arg Ile Lys Pro Val Gln Arg Ala Lys Phe Lys His							
	225			230		235	
gag ctg gtt ctc ggg ttc gat gtc act cag aag aca cag cgc cgt tgg							1850
Glu Leu Val Leu Gly Phe Asp Val Thr Gln Lys Thr Gln Arg Arg Trp							
	240			245		250	
ttc ctc gac aag ctg gtg gac gag atc ggt gtg ggt tac gtg tat gac							1898
Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp							
	255			260		265	
aag ggc agc gtc tcc gcg tac cgt ctg tcc cag atc aag cct ctg cac							1946
Lys Gly Ser Val Ser Ala Tyr Arg Leu Ser Gln Ile Lys Pro Leu His							
	275			280		285	
aac ttc ctg acc cag ctc cag ccc ttc ctg aag ctc aag cag aag cag							1994
Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln							
	290			295		300	
gcc aac ctc gtg ctg aag atc atc gag cag ctg ccc tcc gcc aag gaa							2042
Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu							
	305			310		315	
tcc ccg gac aag ttc ctg gag gtg tgc acc tgg gtg gac cag atc gcc							2090
Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala							
	320			325		330	

gct ctg aac gac tcc aag acc cgc aag acc act tcc gaa acc gtc cgc 2138  
 Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg  
 335 340 345 350

gcc gtt cta gac agt ctc tcc gag aag aag aag tcc tcc ccc taa 2183  
 Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro  
 355 360

ggtaccagcg gccgatccaa tcaacctctg gattacaaaa tttgtgaaag attgactggt 2243

attottaact atgttgctcc ttttagcta tgtggatacg ctgctttaat gcctttgtat 2303

catgctattg cttcccgat ggctttcatt ttctcctcct tgtataaatc ctggttgctg 2363

tctotttatg aggagtgtg gcccgttgtc aggcaacgtg gcgtgggtgtg cactgtgttt 2423

gctgacgcaa cccccactgg ttggggcatt gccaccaact gtcagctcct ttccgggact 2483

ttogctttcc ccctccctat tgccaaggcg gaactcatcg ccgcctgcct tgcccgtgc 2543

tggacagggg ctoggctgtt gggcactgac aattccgtgg tgttgcggg gaaatcatcg 2603

tcctttcctt ggctgctcgc ctgtgttggc acctggatto tgcgcgggac gtccttctgc 2663

taagtccctt cggccctcaa tccagggac cttccttccc gcggcctgct gccggctctg 2723

cggcctcttc cgcgtcttcg agatctgctt cgactgtgcc ttctagttgc cagccatctg 2783

ttgtttgcc ccccccgctg ccttccttga ccttgggaagg tgccactccc actgtccttt 2843

cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg 2903

gtgggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg 2963

actcgagtta agggcgaatt cccgataagg atcttcctag agcatggcta cgtagataag 3023

tagcatggcg ggttaatcat taactacaag gaaccctag tgatggagtt ggccactccc 3083

tctctgcgcg ctgcctcgtc cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc 3143

tttggccggg cggcctcagt gagcgagcga gcgcgcag 3181

- <210> 43
- <211> 364
- <212> PRT
- <213> 人工序列

- <220>
- <223> 合成構建體

<400> 43

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn  
 1 5 10 15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser  
20 25 30

Ile Phe Ala Arg Ile Lys Pro Ser Gln Arg Ser Lys Phe Lys His Lys  
35 40 45

Leu His Leu Val Phe Ala Val Tyr Gln Lys Thr Gln Arg Arg Trp Phe  
50 55 60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Leu Asp Ser  
65 70 75 80

Gly Ser Val Ser Phe Tyr Ser Leu Ser Glu Ile Lys Pro Leu His Asn  
85 90 95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala  
100 105 110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser  
115 120 125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala  
130 135 140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala  
145 150 155 160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln  
165 170 175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile  
180 185 190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys  
195 200 205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile  
210 215 220

Tyr Ala Arg Ile Lys Pro Val Gln Arg Ala Lys Phe Lys His Glu Leu  
225 230 235 240

Val Leu Gly Phe Asp Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu  
 245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp Lys Gly  
 260 265 270

Ser Val Ser Ala Tyr Arg Leu Ser Gln Ile Lys Pro Leu His Asn Phe  
 275 280 285

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn  
 290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro  
 305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu  
 325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val  
 340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro  
 355 360

<210> 44  
 <211> 184  
 <212> DNA  
 <213> 智人

<220>  
 <221> misc\_feature  
 <222> (80)..(101)  
 <223> ARCUS 靶向序列

<400> 44  
 gtcacacag ttggggccac caatgccaa gaccagcgg tgaccctggg gactttgggg 60  
 accaactttg gccgctgtgt ggacctttt gcccagggg aggacatcat tggcctcc 120  
 agcactgca gcacctgctt tgtgtcacag agtgggacat cacaggctgc tgcccacgtg 180  
 gctg 184

<210> 45  
 <211> 184  
 <212> DNA  
 <213> 恆河獼猴

<400> 45  
gtcatcacag ttggggccac caatgccag gaccagccgg tgaccctggg gactttgggg 60  
accaactttg gccgctgtgt ggacctcttt gccccagggg aggacatcat tggcctcc 120  
agcgactgca gcacctgctt tgtgtcacgg agtgggacat cgcaggctgc tgcccacgtg 180  
gctg 184

<210> 46  
<211> 184  
<212> DNA  
<213> 小鼠

<400> 46  
gtcatcacag tcggggccac gaatgccag gaccagccag ttaccttggg gactttgggg 60  
actaattttg gacgctgtgt ggatctcttt gccccggga aggacatcat cggagcgtcc 120  
agtgactgca gcacatgctt catgtcacag agtgggacct cacaggctgc tgcccacgtg 180  
gccg 184

<210> 47  
<211> 1179  
<212> DNA  
<213> 智人

<400> 47  
atggcctctc acaagctgct ggtgacccc cccaaggccc tgctcaagcc cctctccatc 60  
ccaaccagc tcctgctggg gcctggtcct tccaacctgc ctctcgcacat catggcagcc 120  
ggggggctgc agatgatcgg gtccatgagc aaggatatgt accagatcat ggacgagatc 180  
aaggaaggca tccagtacgt gttccagacc aggaaccac tcacactggt catctctggc 240  
tcgggacact gtgcctgga gccgcacctg gtcaatgtgc tggagcctgg ggactccttc 300  
ctggttgggg ccaatggcat ttgggggcag cgagccgtgg acatcgggga gcgcatagga 360  
gcccgagtgc accgatgac caaggacct ggaggccact acacactgca ggaggtggag 420  
gagggcctgg ccagcacia gccagtgtg ctgttttaa cccacgggga gtcgtccacc 480  
ggcgtgctgc agccccttga tgcttcggg gaactctgcc acaggtacaa gtgcctgctc 540  
ctggtggatt cgggtggcat cctgggcggg accccccttt acatggaccg gcaaggcatc 600  
gacatcctgt actcgggctc ccagaaggcc ctgaacgcc ctccaggac ctgcctcatc 660  
tccttcagtg acaaggccaa aaagaagatg tactcccga agacgaagcc cttctccttc 720  
taoctggaca tcaagtggct ggccaacttc tggggctgtg acgaccagcc caggatgtac 780

catcacacaa tccccgtcat cagcctgtac agcctgagag agagcctggc cctcattgog 840  
 gaacagggcc tggagaacag ctggcgccag caccgcgagg ccgcgggcgt tctgcatggg 900  
 cgctgcagg cactggggct gcagctcttc gtgaaggacc cggcgctccg gcttcccaca 960  
 gtcaccactg tggctgtacc cgctggctat gactggagag acatcgtcag ctacgtcata 1020  
 gaccacttcg acattgagat catgggtggc cttgggcoct ccacggggaa ggtgctgcgg 1080  
 atcggcctgc tgggctgcaa tgccaccgc gagaatgtgg accgcgtgac ggaggcoctg 1140  
 agggcggccc tgcagcactg cccaagaag aagctgtga 1179

<210> 48  
 <211> 392  
 <212> PRT  
 <213> 智人

<400> 48

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

Pro Leu Ser Ile Pro Asn Gln Leu Leu Leu Gly Pro Gly Pro Ser Asn  
20 25 30

Leu Pro Pro Arg Ile Met Ala Ala Gly Gly Leu Gln Met Ile Gly Ser  
35 40 45

Met Ser Lys Asp Met Tyr Gln Ile Met Asp Glu Ile Lys Glu Gly Ile  
50 55 60

Gln Tyr Val Phe Gln Thr Arg Asn Pro Leu Thr Leu Val Ile Ser Gly  
65 70 75 80

Ser Gly His Cys Ala Leu Glu Ala Ala Leu Val Asn Val Leu Glu Pro  
85 90 95

Gly Asp Ser Phe Leu Val Gly Ala Asn Gly Ile Trp Gly Gln Arg Ala  
100 105 110

Val Asp Ile Gly Glu Arg Ile Gly Ala Arg Val His Pro Met Thr Lys  
115 120 125

Asp Pro Gly Gly His Tyr Thr Leu Gln Glu Val Glu Glu Gly Leu Ala  
130 135 140

Gln His Lys Pro Val Leu Leu Phe Leu Thr His Gly Glu Ser Ser Thr  
145 150 155 160

Gly Val Leu Gln Pro Leu Asp Gly Phe Gly Glu Leu Cys His Arg Tyr  
165 170 175

Lys Cys Leu Leu Leu Val Asp Ser Val Ala Ser Leu Gly Gly Thr Pro  
180 185 190

Leu Tyr Met Asp Arg Gln Gly Ile Asp Ile Leu Tyr Ser Gly Ser Gln  
195 200 205

Lys Ala Leu Asn Ala Pro Pro Gly Thr Ser Leu Ile Ser Phe Ser Asp  
210 215 220

Lys Ala Lys Lys Lys Met Tyr Ser Arg Lys Thr Lys Pro Phe Ser Phe  
225 230 235 240

Tyr Leu Asp Ile Lys Trp Leu Ala Asn Phe Trp Gly Cys Asp Asp Gln  
245 250 255

Pro Arg Met Tyr His His Thr Ile Pro Val Ile Ser Leu Tyr Ser Leu  
260 265 270

Arg Glu Ser Leu Ala Leu Ile Ala Glu Gln Gly Leu Glu Asn Ser Trp  
275 280 285

Arg Gln His Arg Glu Ala Ala Ala Tyr Leu His Gly Arg Leu Gln Ala  
290 295 300

Leu Gly Leu Gln Leu Phe Val Lys Asp Pro Ala Leu Arg Leu Pro Thr  
305 310 315 320

Val Thr Thr Val Ala Val Pro Ala Gly Tyr Asp Trp Arg Asp Ile Val  
325 330 335

Ser Tyr Val Ile Asp His Phe Asp Ile Glu Ile Met Gly Gly Leu Gly  
340 345 350

Pro Ser Thr Gly Lys Val Leu Arg Ile Gly Leu Leu Gly Cys Asn Ala  
355 360 365

Thr Arg Glu Asn Val Asp Arg Val Thr Glu Ala Leu Arg Ala Ala Leu  
 370 375 380

Gln His Cys Pro Lys Lys Lys Leu  
 385 390

<210> 49  
 <211> 1113  
 <212> DNA  
 <213> 智人

<400> 49  
 atgctcccc ggctaatttg tatcaatgat tatgaacaac atgctaaatc agtacttcca 60  
 aagtctatat atgactatta caggctctggg gcaaatgatg aagaaacttt ggctgataat 120  
 attgcagcat tttccagatg gaagctgtat ccaaggatgc tccggaatgt tgctgaaaca 180  
 gatctgtcga cttctgtttt aggacagagg gtcagcatgc caatatgtgt gggggctacg 240  
 gccatgcagc gcatggctca tgtggaocggc gagcttgcca ctgtgagagc ctgtcagtc 300  
 ctgggaacgg gcatgatggt gagttcctgg gccacctoct caattgaaga agtggcggaa 360  
 gctggtcctg aggcaacttg ttggctgcaa ctgtatatct acaaggaccg agaagtcacc 420  
 aagaagctag tgcggcagggc agagaagatg ggctacaagg ccatatttgt gacagtggac 480  
 acaccttacc tgggcaaccg tctggatgat gtgcgtaaca gattcaaact gccgccacaa 540  
 ctcaggatga aaaattttga aaccagtact ttatcatttt ctctgagga aaattttggg 600  
 gacgacagtg gaattgctgc atatgtggct aaagcaatag accatctat cagctgggaa 660  
 gatatcaaat ggctgagaag actgacatca ttgccaattg ttgcaaaggg cattttgaga 720  
 ggtgatgatg ccaggggaggc tgtaaacat ggcttgaatg ggatcttggg gtogaatcat 780  
 ggggctcgac aactogatgg ggtgccagcc actattgatg ttctgccaga aattgtggag 840  
 gctgtggaag ggaaggtgga agtcttctctg gacgggggtg tgcggaaagg cactgatgtt 900  
 ctgaaagctc tggctcttgg cggcaaggct gtgtttgtgg ggagaccaat cgtttggggc 960  
 ttagctttcc agggggagaa aggtgttcaa gatgtcctcg agatactaaa ggaagaattc 1020  
 cggttggcca tggctctgag tgggtgccag aatgtgaaag tcatcgacaa gacattggtg 1080  
 aggaaaaatc ctttggccgt ttccaagatc tga 1113

<210> 50  
 <211> 370  
 <212> PRT

&lt;213&gt; 智人

&lt;400&gt; 50

Met Leu Pro Arg Leu Ile Cys Ile Asn Asp Tyr Glu Gln His Ala Lys  
 1 5 10 15

Ser Val Leu Pro Lys Ser Ile Tyr Asp Tyr Tyr Arg Ser Gly Ala Asn  
 20 25 30

Asp Glu Glu Thr Leu Ala Asp Asn Ile Ala Ala Phe Ser Arg Trp Lys  
 35 40 45

Leu Tyr Pro Arg Met Leu Arg Asn Val Ala Glu Thr Asp Leu Ser Thr  
 50 55 60

Ser Val Leu Gly Gln Arg Val Ser Met Pro Ile Cys Val Gly Ala Thr  
 65 70 75 80

Ala Met Gln Arg Met Ala His Val Asp Gly Glu Leu Ala Thr Val Arg  
 85 90 95

Ala Cys Gln Ser Leu Gly Thr Gly Met Met Leu Ser Ser Trp Ala Thr  
 100 105 110

Ser Ser Ile Glu Glu Val Ala Glu Ala Gly Pro Glu Ala Leu Arg Trp  
 115 120 125

Leu Gln Leu Tyr Ile Tyr Lys Asp Arg Glu Val Thr Lys Lys Leu Val  
 130 135 140

Arg Gln Ala Glu Lys Met Gly Tyr Lys Ala Ile Phe Val Thr Val Asp  
 145 150 155 160

Thr Pro Tyr Leu Gly Asn Arg Leu Asp Asp Val Arg Asn Arg Phe Lys  
 165 170 175

Leu Pro Pro Gln Leu Arg Met Lys Asn Phe Glu Thr Ser Thr Leu Ser  
 180 185 190

Phe Ser Pro Glu Glu Asn Phe Gly Asp Asp Ser Gly Leu Ala Ala Tyr  
 195 200 205

Val Ala Lys Ala Ile Asp Pro Ser Ile Ser Trp Glu Asp Ile Lys Trp

210	215	220
Leu Arg Arg Leu Thr Ser Leu Pro Ile Val Ala Lys Gly Ile Leu Arg 225 230 235 240		
Gly Asp Asp Ala Arg Glu Ala Val Lys His Gly Leu Asn Gly Ile Leu 245 250 255		
Val Ser Asn His Gly Ala Arg Gln Leu Asp Gly Val Pro Ala Thr Ile 260 265 270		
Asp Val Leu Pro Glu Ile Val Glu Ala Val Glu Gly Lys Val Glu Val 275 280 285		
Phe Leu Asp Gly Gly Val Arg Lys Gly Thr Asp Val Leu Lys Ala Leu 290 295 300		
Ala Leu Gly Ala Lys Ala Val Phe Val Gly Arg Pro Ile Val Trp Gly 305 310 315 320		
Leu Ala Phe Gln Gly Glu Lys Gly Val Gln Asp Val Leu Glu Ile Leu 325 330 335		
Lys Glu Glu Phe Arg Leu Ala Met Ala Leu Ser Gly Cys Gln Asn Val 340 345 350		
Lys Val Ile Asp Lys Thr Leu Val Arg Lys Asn Pro Leu Ala Val Ser 355 360 365		
Lys Ile 370		

<210> 51  
 <211> 1611  
 <212> DNA  
 <213> 智人

<400> 51  
 atggagtttt caagtccttc cagagaggaa tgtccaagc ctttgagtag ggtaagcatc 60  
 atggctggca gcctacacagg attgcttcta cttcaggcag tgtcgtgggc atcaggtgcc 120  
 cgcacctgca tcctaaaag cttcggctac agctcgggtg tgtgtgtctg caatgccaca 180  
 tactgtgact cctttgacct cccgacctt cctgcccttg gtaccttcag ccgctatgag 240

agtacacgca gtggcgacg gatggagctg agtatggggc ccatccaggc taatcacacg 300  
 ggcacaggcc tgctactgac cctgcagcca gaacagaagt tccagaaagt gaagggattt 360  
 ggagggggcca tgacagatgc tgctgtcttc aacatccttg ccctgtcacc ccctgccccaa 420  
 aatttgctac ttaaattcgtta cttctctgaa gaaggaatcg gatataacat catccgggta 480  
 cccatggcca gctgtgactt ctccatccgc acctacacct atgcagacac ccctgatgat 540  
 ttccagttgc acaacttcag cctcccagag gaagatacca agctcaagat acccctgatt 600  
 caccgagccc tgcagttggc ccagcgtccc gtttcactcc ttgccagccc ctggacatca 660  
 cccacttggc tcaagaccaa tggagcgggtg aatgggaagg ggtcactcaa gggacagccc 720  
 ggagacatct accaccagac ctgggcccaga tactttgtga agttcctgga tgcctatgct 780  
 gagcacaagt tacagttctg ggcagtgaca gctgaaaatg agccttctgc tgggctgttg 840  
 agtggatacc ccttccagtg cctgggcttc acccctgaac atcagcgaga cttcattgcc 900  
 cgtgacctag gtccatccct cgccaacagt actcaccaca atgtccgcct actcatgctg 960  
 gatgaccaac gcttgcctgt gccccactgg gcaaagggtg tactgacaga cccagaagca 1020  
 gctaaatatg ttcattggcat tgctgtacat tggtaacctg actttctggc tccagccaaa 1080  
 gccaccctag gggagacaca ccgcctgttc cccaacacca tgctctttgc ctcagaggcc 1140  
 tgtgtgggct ccaagttctg ggagcagagt gtgcggctag gctcctggga tcgagggatg 1200  
 cagtacagcc acagcatcat cacgaacctc ctgtaccatg tggtcggctg gaccgactgg 1260  
 aaccttggcc tgaaccccgga aggaggacc c aattgggtgc gtaactttgt cgacagtccc 1320  
 atcattgtag acatcaccaa ggacacgttt tacaacagc ccatgttcta ccaccttggc 1380  
 cacttcagca agttcattcc tgagggtccc cagagagtgg ggctggttgc cagtcagaag 1440  
 aacgacctgg acgcagtggc actgatgcat cccgatggct ctgctgttgt ggtcgtgcta 1500  
 aaccgctcct ctaaggatgt gcctottacc atcaaggatc ctgctgtggg cttcctggag 1560  
 acaatctcac ctggctactc cattcacacc tacctgtggc gtcgccagtg a 1611

<210> 52  
 <211> 536  
 <212> PRT  
 <213> 智人

<400> 52

Met Glu Phe Ser Ser Pro Ser Arg Glu Glu Cys Pro Lys Pro Leu Ser  
 1 5 10 15

Arg Val Ser Ile Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln  
20 25 30

Ala Val Ser Trp Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe  
35 40 45

Gly Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser  
50 55 60

Phe Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu  
65 70 75 80

Ser Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln  
85 90 95

Ala Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln  
100 105 110

Lys Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala  
115 120 125

Ala Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu  
130 135 140

Lys Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val  
145 150 155 160

Pro Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp  
165 170 175

Thr Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp  
180 185 190

Thr Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln  
195 200 205

Arg Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu  
210 215 220

Lys Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro  
225 230 235 240

Gly Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu  
 245 250 255

Asp Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu  
 260 265 270

Asn Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu  
 275 280 285

Gly Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly  
 290 295 300

Pro Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu  
 305 310 315 320

Asp Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr  
 325 330 335

Asp Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr  
 340 345 350

Leu Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg  
 355 360 365

Leu Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser  
 370 375 380

Lys Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met  
 385 390 395 400

Gln Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly  
 405 410 415

Trp Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp  
 420 425 430

Val Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp  
 435 440 445

Thr Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys  
 450 455 460

Phe Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys



gacottggac agttttact gttttgtcat atctcttccc accaacatga tggcatggaa	1020
gcttatgtca aagtagacag ctgtccagag gaaccccaac tacgaatgaa aaataatgaa	1080
gaagcgggaag actatgatga tgatcttact gattctgaaa tggatgtggt caggtttgat	1140
gatgacaact ctcccttccct tatccaaatt cgctcagttg ccaagaagca tcctaaaact	1200
tgggtacatt acattgtctg tgaagaggag gactgggact atgctccctt agtcctcgcc	1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg	1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa cctttaagac tcgtgaagct	1380
atcagcatg aatcaggaat cttgggacct ttactttatg gggaagttgg agacacactg	1440
ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact	1500
gatgtccgtc ctttgtattc aaggagatta ccaaaagggtg taaaacattt gaaggatttt	1560
ccaattctgc caggagaaat attcaaatat aatgggacag tgactgtaga agatgggcca	1620
actaaatcag atccctoggtg cctgaccogc tattactota gtttcgttaa tatggagaga	1680
gatctagcct caggactcat tggccctctc ctcatctgct acaaagaatc tgtagatcaa	1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag	1800
aaccgaagct ggtacctcac agagaatata caacgcttcc tcccacatcc agctggagtg	1860
cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatggt	1920
tttgatagtt tgcagttgtc agtttgtttg catgaggtgg catactggta cattctaagc	1980
attggagcac agactgactt cctttctgtc ttcttctctg gatatacctt caaacacaaa	2040
atggtctatg aagacacact caccctatcc ccattctcag gagaaactgt cttcatgtcg	2100
atggaaaacc caggctctatg gattctgggg tgccacaact cagactttcg gaacagaggc	2160
atgaccgctt tactgaaggt ttctagttgt gacaagaaca ctggtgatta ttacgaggac	2220
agttatgaag atatttcagc ataacttgctg agtaaaaaa atgccattga accaagaagc	2280
ttctcccaga attcaagaca ccctagcact aggcaaaagc aatttaatgc caccacaatt	2340
ccagaaaatg acatagagaa gactgacctt tggtttgac acagaacacc tatgcctaaa	2400
atacaaaatg tctcctctag tgatttggtg atgctcttgc gacagagtcc tactccacat	2460
gggctatcct tatctgatct ccaagaagcc aatatgaga ctttttctga tgatccatca	2520
cctggagcaa tagacagtaa taacagcctg tctgaaatga cacacttcag gccacagctc	2580
catcacagtg gggacatggt atttaccctt gagtcaggcc tccaattaag attaaatgag	2640
aaactgggga caactgcagc aacagagttg aagaaacttg atttcaaagt ttctagtaca	2700

tcaaataatc	tgatttcaac	aattccatca	gacaatttgg	cagcaggta	tgataatata	2760
agttccttag	gaccccccaag	tatgccagtt	cattatgata	gtcaattaga	taccactcta	2820
tttggcaaaa	agtcattctcc	ccttactgag	tctggtaggac	ctctgagctt	gagtgaagaa	2880
aataatgatt	caaagttggt	agaatcaggt	ttaatgaata	gccaagaaag	ttcatgggga	2940
aaaaatgtat	cgtaacaga	gagtgtagg	ttatttaaag	ggaaaagagc	tcatggacct	3000
gctttgttga	ctaaagataa	tgccattatc	aaagttagca	tctctttgtt	aaagacaaa	3060
aaaacttcca	ataattcagc	aactaataga	aagactcaca	ttgatggccc	atcattatta	3120
attgagaata	gtccatcagt	ctggcaaaa	atattagaaa	gtgacactga	gtttaaaaaa	3180
gtgacacctt	tgattcatga	cagaatgctt	atggacaaaa	atgctacagc	tttgaggcta	3240
aatcatatgt	caaataaaa	tacttcatca	aaaaacatgg	aatgggtcca	acagaaaaaa	3300
gagggcccca	ttccaccaga	tgacaaaaat	ccagatatgt	cgttctttaa	gatgctattc	3360
ttgccagaat	cagcaaggtg	gatacaaagg	actcatggaa	agaactctct	gaactctggg	3420
caaggcccca	gtccaaagca	attagtatcc	ttaggaccag	aaaaatctgt	ggaaggtcag	3480
aatttcttgt	ctgagaaaaa	caaagtggta	gtaggaaagg	gtgaatttac	aaaggacgta	3540
ggactcaaag	agatggtttt	tccaagcagc	agaaacctat	ttcttactaa	cttggataat	3600
ttacatgaaa	ataatacaca	caatcaagaa	aaaaaaatc	aggaagaaat	agaaaagaag	3660
gaaacattaa	tccaagagaa	tgtagttttg	cctcagatac	atacagtgac	tggcactaag	3720
aatttcatga	agaacctttt	cttactgagc	actaggcaaa	atgtagaagg	ttcatatgac	3780
ggggcatatg	ctccagtact	tcaagatfff	aggtcattaa	atgattcaac	aaatagaaca	3840
aagaaacaca	cagctcattt	ctcaaaaaaa	ggggaggaag	aaaacttggg	aggcttggga	3900
aatcaaacca	agcaaattgt	agagaaatat	gcatgcacca	caaggatata	tcctaataca	3960
agccagcaga	atfffgtcac	gcaacgtagt	aagagagctt	tgaaacaatt	cagactccca	4020
ctagaagaaa	cagaacttga	aaaaaggata	attgtggatg	acacctcaac	ccagtgggtc	4080
aaaaacatga	aacatttgac	cccagcacc	ctcacacaga	tagactacaa	tgagaaggag	4140
aaaggggcca	ttactcagtc	tcccttatca	gattgcotta	cgaggagtca	tagcatccct	4200
caagcaaata	gatctccatt	acctattgca	aaggtatcat	catttccatc	tattagacct	4260
atatacttga	ccagggtcct	attccaagac	aactcttctc	atcttccagc	agcatcttat	4320
agaaagaaag	attctgggggt	ccaagaaagc	agtcatttct	tacaaggagc	caaaaaaat	4380

aacctttctt tagccattct aaccttgag atgactggg atcaaagaga gtttggctcc	4440
ctggggacaa gtgccacaaa ttcagtcaca tacaagaaag ttgagaacac tgttctcccg	4500
aaaccagact tgcccaaaac atctggcaaa gttgaattgc ttccaaaagt tcacatttat	4560
cagaaggacc tattccctac ggaaactagc aatgggtctc ctggccatct ggatctcgtg	4620
gaaggaggcc ttcttcaggg aacagaggga gcgattaagt ggaatgaagc aaacagacct	4680
ggaaaagtcc cttttctgag agtagcaaca gaaagctctg caaagactcc ctccaagcta	4740
ttggatcctc ttgcttggga taaccactat ggtactcaga taccaaaaga agagtggaaa	4800
tccaagaga agtcaccaga aaaaacagct ttaagaaaa aggataccat tttgtccctg	4860
aacgcttgtg aaagcaatca tgcaatagca gcaataaatg agggacaaaa taagcccga	4920
atagaagtca cctgggcaaa gcaaggtagg actgaaaggc tgtgctctca aaaccacca	4980
gtcttgaaac gccatcaacg ggaaataact cgtactactc ttcagtcaga tcaagaggaa	5040
attgactatg atgataccat atcagttgaa atgaagaagg aagattttga cttttatgat	5100
gaggatgaaa atcagagccc cgcagcttt caaaagaaaa cacgacacta ttttattgct	5160
gcagtggaga ggctctggga ttatgggatg agtagctccc cacatgttct aagaaacagg	5220
gctcagagtg gcagtgctcc tcagttcaag aaagttgttt tccaggaatt tactgatggc	5280
tcctttactc agcccttata cgtggagaa ctaaataaac atttgggact cctggggcca	5340
tatataagag cagaagttga agataatata atggtaactt tcagaaatca ggctctcgt	5400
ccctattcct tctattctag ccttatttct tatgaggaag atcagaggca aggagcagaa	5460
cctagaaaaa actttgtcaa gcctaataaa accaaaactt acttttggaa agtgcaacat	5520
catatggcac ccaactaaaga tgagtttgac tgcaaaagcct gggcttattt ctctgatgtt	5580
gacctggaag aagatgtgca ctcaggcctg attggacccc ttctggtctg ccacactaac	5640
acactgaacc ctgctcatgg gagacaagtg acagtacagg aatttgctct gtttttcacc	5700
atctttgatg agacaaaaag ctggtacttc actgaaaata tggaaagaaa ctgcagggct	5760
ccctgcaata tccagatgga agatcccact ttaaaagaga attatcgctt ccatgcaatc	5820
aatggctaca taatggatac actacctggc ttagtaatgg ctccagatca aaggattcga	5880
tggtatctgc tcagcatggg cagcaatgaa aacatccatt ctattcattt cagtggacat	5940
gtgttccactg tacgaaaaaa agaggagtat aaaatggcac tgtacaatct ctatccaggt	6000
gtttttgaga cagtggaaat gttaccatcc aaagctggaa tttggcgggt ggaatgcctt	6060
attggcgagc atctacatgc tgggatgagc acacttttct tggtgtacag caataagtgt	6120

cagactcccc tgggaatggc ttctggacac attagagatt ttcagattac agcttcagga 6180  
 caatatggac agtgggcccc aaagctggcc agacttcatt attccggatc aatcaatgcc 6240  
 tggagcacca aggagccctt ttcttggatc aaggtggatc tgttggcacc aatgattatt 6300  
 cacggcatca agaccaggg tgcccgtcag aagttctcca gcctctacat ctctcagttt 6360  
 atcatcatgt atagtottga tgggaagaag tggcagactt atcgaggaaa ttccactgga 6420  
 acctaatgg tcttctttgg caatgtggat tcctctggga taaaacacaa tatttttaac 6480  
 cctccaatta ttgctogata catccgtttg cacccaactc attatagcat tcgcagcact 6540  
 ottogcatgg agttgatggg ctgtgattta aatagttgca gcatgccatt gggaatggag 6600  
 agtaaagcaa tateagatgc acagattact gcttcatcct actttaccaa tatgtttgcc 6660  
 acctggtctc cttcaaaagc tcgacttcac ctccaaggga ggagtaatgc ctggagacct 6720  
 caggtgaata atccaaaaga gtggctgcaa gtggacttcc agaagacaat gaaagtcaca 6780  
 ggagtaacta ctcagggagt aaaatctctg cttaccagca tgtatgtgaa ggagttcctc 6840  
 atctccagca gtcaagatgg ccatcagtgg actctctttt ttcagaatgg caaagtaaag 6900  
 gtttttcagg gaaatcaaga ctccctcaca cctgtgggta actctctaga cccaccgtta 6960  
 ctgactcgtc accttogaat tcacccccag agttgggtgc accagattgc cctgaggatg 7020  
 gaggttctgg gctgcgaggc acaggacctc tactga 7056

<210> 54  
 <211> 2351  
 <212> PRT  
 <213> 智人

<400> 54

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe  
 1 5 10 15

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser  
 20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg  
 35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val  
 50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile  
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln  
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser  
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser  
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp  
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu  
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser  
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile  
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr  
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly  
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp  
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr  
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val  
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile  
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser



Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp  
530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg  
545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu  
565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val  
580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu  
595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp  
610 615 620

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val  
625 630 635 640

Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp  
645 650 655

Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe  
660 665 670

Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr  
675 680 685

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro  
690 695 700

Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly  
705 710 715 720

Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp  
725 730 735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys  
740 745 750

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro  
755 760 765

Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp  
770 775 780

Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys  
785 790 795 800

Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser  
805 810 815

Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr  
820 825 830

Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn  
835 840 845

Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly  
850 855 860

Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu  
865 870 875 880

Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys  
885 890 895

Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn  
900 905 910

Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met  
915 920 925

Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys  
930 935 940

Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu  
945 950 955 960

Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu  
965 970 975

Ser Ser Trp Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe  
 980 985 990

Lys Gly Lys Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala  
 995 1000 1005

Leu Phe Lys Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser  
 1010 1015 1020

Asn Asn Ser Ala Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser  
 1025 1030 1035

Leu Leu Ile Glu Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu  
 1040 1045 1050

Ser Asp Thr Glu Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg  
 1055 1060 1065

Met Leu Met Asp Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met  
 1070 1075 1080

Ser Asn Lys Thr Thr Ser Ser Lys Asn Met Glu Met Val Gln Gln  
 1085 1090 1095

Lys Lys Glu Gly Pro Ile Pro Pro Asp Ala Gln Asn Pro Asp Met  
 1100 1105 1110

Ser Phe Phe Lys Met Leu Phe Leu Pro Glu Ser Ala Arg Trp Ile  
 1115 1120 1125

Gln Arg Thr His Gly Lys Asn Ser Leu Asn Ser Gly Gln Gly Pro  
 1130 1135 1140

Ser Pro Lys Gln Leu Val Ser Leu Gly Pro Glu Lys Ser Val Glu  
 1145 1150 1155

Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys Val Val Val Gly Lys  
 1160 1165 1170

Gly Glu Phe Thr Lys Asp Val Gly Leu Lys Glu Met Val Phe Pro  
 1175 1180 1185

Ser Ser Arg Asn Leu Phe Leu Thr Asn Leu Asp Asn Leu His Glu

1190		1195		1200										
Asn 1205	Asn	Thr	His	Asn	Gln	Glu 1210	Lys	Lys	Ile	Gln	Glu 1215	Glu	Ile	Glu
Lys 1220	Lys	Glu	Thr	Leu	Ile	Gln 1225	Glu	Asn	Val	Val	Leu 1230	Pro	Gln	Ile
His 1235	Thr	Val	Thr	Gly	Thr	Lys 1240	Asn	Phe	Met	Lys	Asn 1245	Leu	Phe	Leu
Leu 1250	Ser	Thr	Arg	Gln	Asn	Val 1255	Glu	Gly	Ser	Tyr	Asp 1260	Gly	Ala	Tyr
Ala 1265	Pro	Val	Leu	Gln	Asp	Phe 1270	Arg	Ser	Leu	Asn	Asp 1275	Ser	Thr	Asn
Arg 1280	Thr	Lys	Lys	His	Thr	Ala 1285	His	Phe	Ser	Lys	Lys 1290	Gly	Glu	Glu
Glu 1295	Asn	Leu	Glu	Gly	Leu	Gly 1300	Asn	Gln	Thr	Lys	Gln 1305	Ile	Val	Glu
Lys 1310	Tyr	Ala	Cys	Thr	Thr	Arg 1315	Ile	Ser	Pro	Asn	Thr 1320	Ser	Gln	Gln
Asn 1325	Phe	Val	Thr	Gln	Arg	Ser 1330	Lys	Arg	Ala	Leu	Lys 1335	Gln	Phe	Arg
Leu 1340	Pro	Leu	Glu	Glu	Thr	Glu 1345	Leu	Glu	Lys	Arg	Ile 1350	Ile	Val	Asp
Asp 1355	Thr	Ser	Thr	Gln	Trp	Ser 1360	Lys	Asn	Met	Lys	His 1365	Leu	Thr	Pro
Ser 1370	Thr	Leu	Thr	Gln	Ile	Asp 1375	Tyr	Asn	Glu	Lys	Glu 1380	Lys	Gly	Ala
Ile 1385	Thr	Gln	Ser	Pro	Leu	Ser 1390	Asp	Cys	Leu	Thr	Arg 1395	Ser	His	Ser
Ile 1400	Pro	Gln	Ala	Asn	Arg	Ser 1405	Pro	Leu	Pro	Ile	Ala 1410	Lys	Val	Ser

Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu Phe  
1415 1420 1425

Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys  
1430 1435 1440

Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys  
1445 1450 1455

Lys Asn Asn Leu Ser Leu Ala Ile Leu Thr Leu Glu Met Thr Gly  
1460 1465 1470

Asp Gln Arg Glu Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser  
1475 1480 1485

Val Thr Tyr Lys Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp  
1490 1495 1500

Leu Pro Lys Thr Ser Gly Lys Val Glu Leu Leu Pro Lys Val His  
1505 1510 1515

Ile Tyr Gln Lys Asp Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser  
1520 1525 1530

Pro Gly His Leu Asp Leu Val Glu Gly Ser Leu Leu Gln Gly Thr  
1535 1540 1545

Glu Gly Ala Ile Lys Trp Asn Glu Ala Asn Arg Pro Gly Lys Val  
1550 1555 1560

Pro Phe Leu Arg Val Ala Thr Glu Ser Ser Ala Lys Thr Pro Ser  
1565 1570 1575

Lys Leu Leu Asp Pro Leu Ala Trp Asp Asn His Tyr Gly Thr Gln  
1580 1585 1590

Ile Pro Lys Glu Glu Trp Lys Ser Gln Glu Lys Ser Pro Glu Lys  
1595 1600 1605

Thr Ala Phe Lys Lys Lys Asp Thr Ile Leu Ser Leu Asn Ala Cys  
1610 1615 1620

Glu Ser Asn His Ala Ile Ala Ala Ile Asn Glu Gly Gln Asn Lys  
1625 1630 1635

Pro Glu Ile Glu Val Thr Trp Ala Lys Gln Gly Arg Thr Glu Arg  
1640 1645 1650

Leu Cys Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu  
1655 1660 1665

Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr  
1670 1675 1680

Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile  
1685 1690 1695

Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys  
1700 1705 1710

Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr  
1715 1720 1725

Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser  
1730 1735 1740

Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr  
1745 1750 1755

Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu  
1760 1765 1770

His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp  
1775 1780 1785

Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser  
1790 1795 1800

Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly  
1805 1810 1815

Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr  
1820 1825 1830

Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu  
1835 1840 1845

Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu  
1850 1855 1860

Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His  
1865 1870 1875

Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln  
1880 1885 1890

Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp  
1895 1900 1905

Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn  
1910 1915 1920

Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His  
1925 1930 1935

Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met  
1940 1945 1950

Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser  
1955 1960 1965

Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr  
1970 1975 1980

Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr  
1985 1990 1995

Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly  
2000 2005 2010

Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly  
2015 2020 2025

Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro  
2030 2035 2040

Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala

2045		2050		2055
Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His	2060	2065	2070	
Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser	2075	2080	2085	
Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile	2090	2095	2100	
Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser	2105	2110	2115	
Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr	2120	2125	2130	
Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn	2135	2140	2145	
Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile	2150	2155	2160	
Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg	2165	2170	2175	
Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys	2180	2185	2190	
Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln	2195	2200	2205	
Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser	2210	2215	2220	
Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp	2225	2230	2235	
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe	2240	2245	2250	
Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys	2255	2260	2265	

Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser  
2270 2275 2280

Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys  
2285 2290 2295

Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val  
2300 2305 2310

Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His  
2315 2320 2325

Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu  
2330 2335 2340

Gly Cys Glu Ala Gln Asp Leu Tyr  
2345 2350

<210> 55  
<211> 1386  
<212> DNA  
<213> 智人

<400> 55  
atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta 60  
ggatatctac tcagtctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt 120  
ctgaatcggc caaagaggta taattcaggt aattggaag agtttgttca agggaacctt 180  
gagagagaat gtatggaaga aaagttagt tttgaagaag cacgagaagt ttttgaaaac 240  
actgaaagaa caactgaatt ttggaagcag tatgttgatg gagatcagtg tgagtccaat 300  
ccatgtttta atggcggcag ttgcaaggat gacattaatt cctatgaatg ttggtgtccc 360  
tttgatttg aaggaaagaa ctgtgaatta gatgtaacat gtaacattaa gaatggcaga 420  
tgcgagcagt tttgtaaaaa tagtctgat aacaaggagg tttgctcctg tactgaggga 480  
tatcgacttg cagaaaacca gaagtcctgt gaaccagcag tgccatttcc atgtggaaga 540  
gtttctgttt cacaaacttc taagtcacc cgtgctgaga ctgtttttcc tgatgtggac 600  
tatgtaaatt ctaactgaagc tgaaccatt ttggataaca tcaactcaaag cacccaatca 660  
tttaatgact tcaactcgggt tgttggtgga gaagatgcca aaccaggcca attcccttgg 720  
caggttgttt tgaatggtaa agttgatgca ttctgtggag gctctatcgt taatgaaaaa 780

tggattgtaa ctgctgcccc ctgtgttgaa actgggtgta aaattacagt tgtcgcaggt 840  
 gaacataata ttgaggagac agaacataca gagcaaaaagc gaaatgtgat tcgaattatt 900  
 cctcaccaca actacaatgc agctattaat aagtacaacc atgacattgc ccttctggaa 960  
 ctggacgaac ccttagtgct aacagctac gttacaccta tttgcattgc tgacaaggaa 1020  
 tacacgaaca tcttctcaa atttggatct ggctatgtaa gtggctgggg aagagtcttc 1080  
 cacaaaggga gatcagcttt agttcttcag taccttagag ttccacttgt tgaccgagcc 1140  
 acatgtcttc gatctacaaa gttcaccatc tataacaaca tgttctgtgc tggcttccat 1200  
 gaaggagga gagattcatg tcaaggagat agtgggggac cccatgttac tgaagtggaa 1260  
 gggaccagtt tcttaactgg aattattagc tggggtgaag agtgtgcaat gaaaggcaaa 1320  
 tatggaatat ataccaaggt atcccggat gtcaactgga ttaaggaaaa aacaaagctc 1380  
 acttaa 1386

<210> 56  
 <211> 461  
 <212> PRT  
 <213> 智人

<400> 56

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr  
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu  
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn  
35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys  
50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn  
65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly Asp Gln  
85 90 95

Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp Asp Ile  
100 105 110

Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys Asn Cys  
 115 120 125

Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe  
 130 135 140

Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly  
 145 150 155 160

Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val Pro Phe  
 165 170 175

Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala  
 180 185 190

Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala Glu  
 195 200 205

Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe  
 210 215 220

Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp  
 225 230 235 240

Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile  
 245 250 255

Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly  
 260 265 270

Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu  
 275 280 285

His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn  
 290 295 300

Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu  
 305 310 315 320

Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile  
 325 330 335

Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr  
340 345 350

Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val  
355 360 365

Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg  
370 375 380

Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His  
385 390 395 400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val  
405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly  
420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser  
435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr  
450 455 460

<210> 57  
<211> 1395  
<212> DNA  
<213> 智人

<400> 57  
atggcctcgg agagtgggaa gctttggggt ggccggtttg tgggtgcagt ggaccccatc 60  
atggagaagt tcaacgcgtc cattgcctac gaccggcacc tttgggaggt ggatgttcaa 120  
ggcagcaaag cctacagcag gggcctggag aaggcagggc tcctcaccaa ggccgagatg 180  
gaccagatac tccatggcct agacaaggtg gctgaggagt gggcccaggg caccttcaaa 240  
ctgaactcca atgatgagga catccacaca gccaatgagc gccgcctgaa ggagctcatt 300  
ggtgcaacgg caggggaagct gcacacggga cggagccgga atgaccaggt ggtcacagac 360  
ctcaggctgt ggatgoggca gacctgctcc acgctctogg gcctcctctg ggagctcatt 420  
aggaccatgg tggatcgggc agaggcggaa cgtgatgttc tcttcccggg gtacacccat 480  
ttgcagaggg cccagcccat ccgctggagc cactggattc tgagccacgc cgtggcactg 540

acccgagact ctgagcggct gctggagggt cggaagcga tcaatgtcct gccctgggg 600  
 agtggggcca ttgcaggcaa tcccctgggt gtggaccgag agctgtccg agcagaactc 660  
 aactttgggg ccatcactct caacagcatg gatgccacta gtgagcggga ctttgtggcc 720  
 gaggttcctgt tctgggcttc gctgtgatg acccatctca gcaggatggc cgaggacctc 780  
 atcctctact gcaccaagga attcagcttc gtgcagctct cagatgccta cagcacggga 840  
 agcagcctga tgcccagaa gaaaaacccc gacagtttg agctgatccg gagcaaggct 900  
 gggcgtgtgt ttggcggtg tgccgggctc ctgatgacc tcaagggact tcccagcacc 960  
 tacaacaaag acttacagga ggacaaggaa gctgtgtttg aagtgtcaga cactatgagt 1020  
 gccgtgctcc aggtggccac tggcgtcadc tctacgtgc agattacca agagaacatg 1080  
 ggacaggctc tcagccccga catgctggcc actgacctg cctattacct ggtccgcaa 1140  
 gggatgccat tccgccaggc ccacgaggcc tccgggaaag ctgtgttcat ggccgagacc 1200  
 aagggggtcg ccctcaacca gctgtcactg caggagctgc agaccatcag ccccctgttc 1260  
 togggcgacg tgatctcgt gtgggactac gggcacagtg tggagcagta tggtgccctg 1320  
 ggccgactg cgcgctccag cgtcactgg cagatccgc aggtgcgggc gctactgcag 1380  
 gcacagcagg cctag 1395

<210> 58  
 <211> 464  
 <212> PRT  
 <213> 智人

<400> 58

Met Ala Ser Glu Ser Gly Lys Leu Trp Gly Gly Arg Phe Val Gly Ala  
 1 5 10 15

Val Asp Pro Ile Met Glu Lys Phe Asn Ala Ser Ile Ala Tyr Asp Arg  
 20 25 30

His Leu Trp Glu Val Asp Val Gln Gly Ser Lys Ala Tyr Ser Arg Gly  
 35 40 45

Leu Glu Lys Ala Gly Leu Leu Thr Lys Ala Glu Met Asp Gln Ile Leu  
 50 55 60

His Gly Leu Asp Lys Val Ala Glu Glu Trp Ala Gln Gly Thr Phe Lys  
 65 70 75 80

Leu Asn Ser Asn Asp Glu Asp Ile His Thr Ala Asn Glu Arg Arg Leu  
85 90 95

Lys Glu Leu Ile Gly Ala Thr Ala Gly Lys Leu His Thr Gly Arg Ser  
100 105 110

Arg Asn Asp Gln Val Val Thr Asp Leu Arg Leu Trp Met Arg Gln Thr  
115 120 125

Cys Ser Thr Leu Ser Gly Leu Leu Trp Glu Leu Ile Arg Thr Met Val  
130 135 140

Asp Arg Ala Glu Ala Glu Arg Asp Val Leu Phe Pro Gly Tyr Thr His  
145 150 155 160

Leu Gln Arg Ala Gln Pro Ile Arg Trp Ser His Trp Ile Leu Ser His  
165 170 175

Ala Val Ala Leu Thr Arg Asp Ser Glu Arg Leu Leu Glu Val Arg Lys  
180 185 190

Arg Ile Asn Val Leu Pro Leu Gly Ser Gly Ala Ile Ala Gly Asn Pro  
195 200 205

Leu Gly Val Asp Arg Glu Leu Leu Arg Ala Glu Leu Asn Phe Gly Ala  
210 215 220

Ile Thr Leu Asn Ser Met Asp Ala Thr Ser Glu Arg Asp Phe Val Ala  
225 230 235 240

Glu Phe Leu Phe Trp Ala Ser Leu Cys Met Thr His Leu Ser Arg Met  
245 250 255

Ala Glu Asp Leu Ile Leu Tyr Cys Thr Lys Glu Phe Ser Phe Val Gln  
260 265 270

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

Asn Pro Asp Ser Leu Glu Leu Ile Arg Ser Lys Ala Gly Arg Val Phe  
290 295 300

Gly Arg Cys Ala Gly Leu Leu Met Thr Leu Lys Gly Leu Pro Ser Thr  
305 310 315 320

Tyr Asn Lys Asp Leu Gln Glu Asp Lys Glu Ala Val Phe Glu Val Ser  
325 330 335

Asp Thr Met Ser Ala Val Leu Gln Val Ala Thr Gly Val Ile Ser Thr  
340 345 350

Leu Gln Ile His Gln Glu Asn Met Gly Gln Ala Leu Ser Pro Asp Met  
355 360 365

Leu Ala Thr Asp Leu Ala Tyr Tyr Leu Val Arg Lys Gly Met Pro Phe  
370 375 380

Arg Gln Ala His Glu Ala Ser Gly Lys Ala Val Phe Met Ala Glu Thr  
385 390 395 400

Lys Gly Val Ala Leu Asn Gln Leu Ser Leu Gln Glu Leu Gln Thr Ile  
405 410 415

Ser Pro Leu Phe Ser Gly Asp Val Ile Cys Val Trp Asp Tyr Gly His  
420 425 430

Ser Val Glu Gln Tyr Gly Ala Leu Gly Gly Thr Ala Arg Ser Ser Val  
435 440 445

Asp Trp Gln Ile Arg Gln Val Arg Ala Leu Leu Gln Ala Gln Gln Ala  
450 455 460

<210> 59  
<211> 1074  
<212> DNA  
<213> 智人

<400> 59  
atggaggaag gaatgaatgt tctccatgac tttgggatcc agtcaacaca ttacctcag 60  
gtgaattacc aagactocca ggactggttc atcttgggtgt ccgtgatcgc agacctcagg 120  
aatgccttct acgtoctctt ccccatctgg ttccatcttc aggaagctgt gggcattaaa 180  
ctcctttggg tagctgtgat tggagactgg ctcaacctog tctttaagtg gattctcttt 240  
ggacagcgtc catactggtg ggttttggat actgactact acagcaacac ttccgtgcc 300  
ctgataaagc agttocctgt aacctgtgag actggaccag ggagcccctc tggccatgcc 360

atgggcacag cagggtgata ctacgtgatg gtcacatcta ctctttccat ctttcagggg 420  
 aagataaagc cgacctacag atttcgggtgc ttgaatgtca ttttgtggtt gggattctgg 480  
 gctgtgcagc tgaatgtctg tctgtcaoga atctacottg ctgctcattt tcctcatcaa 540  
 gttgttgctg gagtcctgtc aggcattgct gttgcagaaa ctttcagcca catccacagc 600  
 atctataatg ccagcctcaa gaaatatttt ctattacct tcttctgtt cagcttcgcc 660  
 atcggatttt atctgctgct caagggactg ggtgtagacc tcctgtggac tctggagaaa 720  
 gccagaggt ggtgcgagca gccagaatgg gtccacattg acaccacacc ctttgccagc 780  
 ctctcaaga acctggggcac gctctttggc ctggggctgg ctctcaactc cagcatgtac 840  
 agggagagct gcaaggggaa actcagcaag tggctcccat tccgcctcag ctctattgta 900  
 gcctccctcg tcctcctgca cgtctttgac tccttgaaac ccccatccca agtcgagctg 960  
 gtcttctacg tcttgtcctt ctgcaagagt gcggtagtgc ccctggcatc cgtcagtgtc 1020  
 atcccctact gcctcgccca ggtcctgggc cagccgcaca agaagtcggt gtaa 1074

<210> 60  
 <211> 359  
 <212> PRT  
 <213> 智人

<400> 60

Met Arg Met Glu Glu Gly Met Asn Val Leu His Asp Phe Gly Ile Gln  
1 5 10 15

Ser Thr His Tyr Leu Gln Val Asn Tyr Gln Asp Ser Gln Asp Trp Phe  
20 25 30

Ile Leu Val Ser Val Ile Ala Asp Leu Arg Asn Ala Phe Tyr Val Leu  
35 40 45

Phe Pro Ile Trp Phe His Leu Gln Glu Ala Val Gly Ile Lys Leu Leu  
50 55 60

Trp Val Ala Val Ile Gly Asp Trp Leu Asn Leu Val Phe Lys Trp Ile  
65 70 75 80

Leu Phe Gly Gln Arg Pro Tyr Trp Trp Val Leu Asp Thr Asp Tyr Tyr  
85 90 95

Ser Asn Thr Ser Val Pro Leu Ile Lys Gln Phe Pro Val Thr Cys Glu  
 100 105 110

Thr Gly Pro Gly Ser Pro Ser Gly His Ala Met Gly Thr Ala Gly Val  
 115 120 125

Tyr Tyr Val Met Val Thr Ser Thr Leu Ser Ile Phe Gln Gly Lys Ile  
 130 135 140

Lys Pro Thr Tyr Arg Phe Arg Cys Leu Asn Val Ile Leu Trp Leu Gly  
 145 150 155 160

Phe Trp Ala Val Gln Leu Asn Val Cys Leu Ser Arg Ile Tyr Leu Ala  
 165 170 175

Ala His Phe Pro His Gln Val Val Ala Gly Val Leu Ser Gly Ile Ala  
 180 185 190

Val Ala Glu Thr Phe Ser His Ile His Ser Ile Tyr Asn Ala Ser Leu  
 195 200 205

Lys Lys Tyr Phe Leu Ile Thr Phe Phe Leu Phe Ser Phe Ala Ile Gly  
 210 215 220

Phe Tyr Leu Leu Leu Lys Gly Leu Gly Val Asp Leu Leu Trp Thr Leu  
 225 230 235 240

Glu Lys Ala Gln Arg Trp Cys Glu Gln Pro Glu Trp Val His Ile Asp  
 245 250 255

Thr Thr Pro Phe Ala Ser Leu Leu Lys Asn Leu Gly Thr Leu Phe Gly  
 260 265 270

Leu Gly Leu Ala Leu Asn Ser Ser Met Tyr Arg Glu Ser Cys Lys Gly  
 275 280 285

Lys Leu Ser Lys Trp Leu Pro Phe Arg Leu Ser Ser Ile Val Ala Ser  
 290 295 300

Leu Val Leu Leu His Val Phe Asp Ser Leu Lys Pro Pro Ser Gln Val  
 305 310 315 320

Glu Leu Val Phe Tyr Val Leu Ser Phe Cys Lys Ser Ala Val Val Pro

325

330

335

Leu Ala Ser Val Ser Val Ile Pro Tyr Cys Leu Ala Gln Val Leu Gly  
 340 345 350

Gln Pro His Lys Lys Ser Leu  
 355

<210> 61  
 <211> 1503  
 <212> DNA  
 <213> 智人

<400> 61  
 atggcctcca ggctgaccct gctgaccctc ctgctgctgc tgctggctgg ggatagagcc 60  
 tcctcaaate caaatgctac cagctccagc tcccaggatc cagagagttt gcaagacaga 120  
 ggogaaggga aggtogcaac aacagttatc tccaagatgc tattcgttga acccatcctg 180  
 gaggtttcca gcttgccgac aaccaactca acaaccaatt cagccaccaa aataacagct 240  
 aataccactg atgaaccac cacacaacc accacagagc ccaccaccca acccaccatc 300  
 caaccacccc aaccaactac ccagctocca acagattctc ctaccagcc cactactggg 360  
 tcctttctgc caggacctgt tactctctgc tctgacttgg agagtcatte aacagaggcc 420  
 gtgttggggg atgctttggg agattctctc ctgaagctct accacgcctt ctcagcaatg 480  
 aagaaggtgg agaccaacat ggccttttcc ccattcagca tcgccagcct ccttaccag 540  
 gtctgtctcg gggctgggga gaacacaaaa acaaacttgg agagcatcct ctcttaccoc 600  
 aaggacttca cctgtgtcca ccaggccctg aagggttca cgaccaaagg tgtcacctca 660  
 gtctctcaga tcttcacag cccagacctg gccataaggg acacctttgt gaatgcctct 720  
 cggacctgt acagcagcag ccccagagtc ctaagcaaca acagtgaagc caacttggag 780  
 ctcatcaaca cctgggtggc caagaacacc aacaacaaga tcagccggct gctagacagt 840  
 ctgccctccg ataccgcct tgcctctctc aatgctatct acctgagtgc caagtggaag 900  
 acaacatttg atcccaagaa aaccagaatg gaacccttcc acttcaaaaa ctcagttata 960  
 aaagtgccca tgatgaatag caagaagtac cctgtggccc atttcattga ccaaactttg 1020  
 aaagccaagg tggggcagct gcagctctcc cacaatctga gtttggatgat cctggtacc 1080  
 cagaacctga aacatcgtct tgaagacatg gaacaggctc tcagcccttc tgttttcaag 1140  
 gccatcatgg agaaactgga gatgtccaag ttccagccca ctctcctaac actaccccgc 1200

atcaaagtga cgaccagcca ggatatgctc tcaatcatgg agaaattgga attcttcgat 1260  
 tttttttatg accttaacct gtgtgggctg acagaggacc cagatcttca ggtttctgog 1320  
 atgcagcacc agacagtgtt ggaactgaca gagactgggg tggaggcggc tgcagcctcc 1380  
 gccatctctg tggcccgcac cctgctggtc tttgaagtgc agcagccctt cctcttcgtg 1440  
 ctctgggacc agcagcaciaa gttccctgtc ttcattggggc gagtatatga ccccagggcc 1500  
 tga 1503

<210> 62  
 <211> 500  
 <212> PRT  
 <213> 智人

<400> 62

Met Ala Ser Arg Leu Thr Leu Leu Thr Leu Leu Leu Leu Leu Leu Ala  
 1 5 10 15

Gly Asp Arg Ala Ser Ser Asn Pro Asn Ala Thr Ser Ser Ser Ser Gln  
 20 25 30

Asp Pro Glu Ser Leu Gln Asp Arg Gly Glu Gly Lys Val Ala Thr Thr  
 35 40 45

Val Ile Ser Lys Met Leu Phe Val Glu Pro Ile Leu Glu Val Ser Ser  
 50 55 60

Leu Pro Thr Thr Asn Ser Thr Thr Asn Ser Ala Thr Lys Ile Thr Ala  
 65 70 75 80

Asn Thr Thr Asp Glu Pro Thr Thr Gln Pro Thr Thr Glu Pro Thr Thr  
 85 90 95

Gln Pro Thr Ile Gln Pro Thr Gln Pro Thr Thr Gln Leu Pro Thr Asp  
 100 105 110

Ser Pro Thr Gln Pro Thr Thr Gly Ser Phe Cys Pro Gly Pro Val Thr  
 115 120 125

Leu Cys Ser Asp Leu Glu Ser His Ser Thr Glu Ala Val Leu Gly Asp  
 130 135 140

Ala Leu Val Asp Phe Ser Leu Lys Leu Tyr His Ala Phe Ser Ala Met



Lys Leu Glu Met Ser Lys Phe Gln Pro Thr Leu Leu Thr Leu Pro Arg  
385 390 395 400

Ile Lys Val Thr Thr Ser Gln Asp Met Leu Ser Ile Met Glu Lys Leu  
405 410 415

Glu Phe Phe Asp Phe Ser Tyr Asp Leu Asn Leu Cys Gly Leu Thr Glu  
420 425 430

Asp Pro Asp Leu Gln Val Ser Ala Met Gln His Gln Thr Val Leu Glu  
435 440 445

Leu Thr Glu Thr Gly Val Glu Ala Ala Ala Ala Ser Ala Ile Ser Val  
450 455 460

Ala Arg Thr Leu Leu Val Phe Glu Val Gln Gln Pro Phe Leu Phe Val  
465 470 475 480

Leu Trp Asp Gln Gln His Lys Phe Pro Val Phe Met Gly Arg Val Tyr  
485 490 495

Asp Pro Arg Ala  
500

<210> 63  
<211> 1140  
<212> DNA  
<213> 智人

<400> 63  
atgtcgcgca gtggaaccga tcctcagcaa cgccagcagg cgtcagaggc ggacgccgca 60  
gcagcaacct tccgggcaaa cgaccatcag catatccgct acaaccgcgt gcaggatgag 120  
tgggtgctgg tgtcagctca ccgcatgaag cggccctggc agggtaagt ggagccccag 180  
cttctgaaga cagtgccccg ccatgacct ctcaaccctc tgtgtcctgg ggccatccga 240  
gccaacggag aggtgaatcc ccagtaocat agcaccttc tgtttgacaa cgacttccca 300  
gctctgcagc ctgatgcccc cagtccagga cccagtgatc atcccccttt ccaagcaaag 360  
totgtctgag gagtctgtaa ggtcatgtgc ttccaccct ggtcggatgt aacgctgcca 420  
ctcatgtcgg tccctgagat cgggctggt gttgatgcat gggcctcagt cacagaggag 480  
ctgggtgccc agtaaccttg ggtgcagatc tttgaaaaca aaggtgcat gatgggctgt 540

ttaaccccc accccactg ccaggtatgg gccagcagtt tcctgccaga tattgccag 600  
 cgtgaggagc gatctcagca ggcctataag agtcagcatg gagagcccct gctaatggag 660  
 tacagccgcc aggagctact caggaaggaa cgtctgggcc taaccagtga gcactgggta 720  
 gtactgggcc ccttctgggc aacatggccc taccagacac tgctgctgcc cgtcggcat 780  
 gtgoggcggc tacctgagct gaccctgct gagcgtgatg atctagcctc catcatgaag 840  
 aagctcttga ccaagtatga caacctcttt gagacgtcct ttccctactc catgggctgg 900  
 catggggctc ccacaggatc agaggctggg gccaaactgga accattggca gctgcacgct 960  
 cactactacc ctccgctcct gcgctctgcc actgtccgga aattcatggt tggctacgaa 1020  
 atgottgctc aggctcagag ggacctcacc cctgagcagg ctgcagagag actaagggca 1080  
 cttctgagg ttcattacca cctggggcag aaggacagg agacagcaac catcgctga 1140

<210> 64  
 <211> 379  
 <212> PRT  
 <213> 智人

<400> 64

Met Ser Arg Ser Gly Thr Asp Pro Gln Gln Arg Gln Gln Ala Ser Glu  
 1 5 10 15

Ala Asp Ala Ala Ala Ala Thr Phe Arg Ala Asn Asp His Gln His Ile  
 20 25 30

Arg Tyr Asn Pro Leu Gln Asp Glu Trp Val Leu Val Ser Ala His Arg  
 35 40 45

Met Lys Arg Pro Trp Gln Gly Gln Val Glu Pro Gln Leu Leu Lys Thr  
 50 55 60

Val Pro Arg His Asp Pro Leu Asn Pro Leu Cys Pro Gly Ala Ile Arg  
 65 70 75 80

Ala Asn Gly Glu Val Asn Pro Gln Tyr Asp Ser Thr Phe Leu Phe Asp  
 85 90 95

Asn Asp Phe Pro Ala Leu Gln Pro Asp Ala Pro Ser Pro Gly Pro Ser  
 100 105 110

Asp His Pro Leu Phe Gln Ala Lys Ser Ala Arg Gly Val Cys Lys Val  
 115 120 125

Met Cys Phe His Pro Trp Ser Asp Val Thr Leu Pro Leu Met Ser Val  
 130 135 140

Pro Glu Ile Arg Ala Val Val Asp Ala Trp Ala Ser Val Thr Glu Glu  
 145 150 155 160

Leu Gly Ala Gln Tyr Pro Trp Val Gln Ile Phe Glu Asn Lys Gly Ala  
 165 170 175

Met Met Gly Cys Ser Asn Pro His Pro His Cys Gln Val Trp Ala Ser  
 180 185 190

Ser Phe Leu Pro Asp Ile Ala Gln Arg Glu Glu Arg Ser Gln Gln Ala  
 195 200 205

Tyr Lys Ser Gln His Gly Glu Pro Leu Leu Met Glu Tyr Ser Arg Gln  
 210 215 220

Glu Leu Leu Arg Lys Glu Arg Leu Val Leu Thr Ser Glu His Trp Leu  
 225 230 235 240

Val Leu Val Pro Phe Trp Ala Thr Trp Pro Tyr Gln Thr Leu Leu Leu  
 245 250 255

Pro Arg Arg His Val Arg Arg Leu Pro Glu Leu Thr Pro Ala Glu Arg  
 260 265 270

Asp Asp Leu Ala Ser Ile Met Lys Lys Leu Leu Thr Lys Tyr Asp Asn  
 275 280 285

Leu Phe Glu Thr Ser Phe Pro Tyr Ser Met Gly Trp His Gly Ala Pro  
 290 295 300

Thr Gly Ser Glu Ala Gly Ala Asn Trp Asn His Trp Gln Leu His Ala  
 305 310 315 320

His Tyr Tyr Pro Pro Leu Leu Arg Ser Ala Thr Val Arg Lys Phe Met  
 325 330 335

Val Gly Tyr Glu Met Leu Ala Gln Ala Gln Arg Asp Leu Thr Pro Glu

340

345

350

Gln Ala Ala Glu Arg Leu Arg Ala Leu Pro Glu Val His Tyr His Leu  
 355 360 365

Gly Gln Lys Asp Arg Glu Thr Ala Thr Ile Ala  
 370 375

<210> 65  
 <211> 1656  
 <212> DNA  
 <213> 智人

<400> 65  
 atgccttctg agacccccca ggcagaagtg gggcccacag gctgccccca cgcctcaggg 60  
 ccacactcgg cgaaggggag cctggagaag ggtccccag aggataagga agccaaggag 120  
 cccctgtgga tcgggccga tgctcogagc aggtgcaoct ggcagctggg cggcctgcc 180  
 tccgagtccc cacatcacca cactgccccg gcaaaatctc caaaaatctt gccagatatt 240  
 ctgaagaaaa tcggggacac ccctatggtc agaatcaaca agattgggaa gaagttcggc 300  
 ctgaagtgtg agctcttggc caagtgtgag ttcttcaacg cgggcgggag cgtgaaggac 360  
 cgcacacgcc tgcggatgat tgaggatgct gagcgcgacg ggacgctgaa gcccggggac 420  
 acgattatcg agccgacatc cgggaacacc gggatcgggc tggccctggc tgcggcagtg 480  
 aggggctatc gctgcatcat cgtgatgcca gagaagatga gctccgagaa ggtggacgtg 540  
 ctgcggggcac tgggggctga gattgtgagg acgcccacca atgccagggt cgactccccg 600  
 gagtccacacg tgggggtggc ctggcggctg aagaacgaaa tccccattc tcacatccta 660  
 gaccagtacc gcaacgccag caacccccctg gctcactacg acaccaccgc tgatgagatc 720  
 ctgcagcagt gtgatgggaa gctggacatg ctggtggctt cagtgggcac gggcggcacc 780  
 atcacgggca ttgccaggaa gctgaaggag aagtgtcctg gatgcaggat cattgggggtg 840  
 gatcccgaag ggtccatcct cgcagagccg gaggagctga accagacgga gcagacaacc 900  
 tacgaggttg aagggatcgg ctacgacttc atccccacgg tgctggacag gacggtggtg 960  
 gacaagtggg tcaagagcaa cgatgaggag gcgttcaoct ttgcccgcac gctgatcggc 1020  
 caagaggggc tgctgtgcgg tggcagtgct ggcagcacgg tggcgggtggc cgtgaaggcc 1080  
 gcgcaggagc tgcaggaggg ccagcgtgct gtggtcattc tgcccactc agtgcggaac 1140  
 tacatgacca agttcctgag cgacaggtgg atgctgcaga agggctttct gaaggaggag 1200

gacctcacgg agaagaagcc ctggtggtgg cacctccgtg ttcaggagct gggcctgtca 1260  
 gccccgctga ccgtgctccc gaccatcacc tgtgggcaca ccatcgagat cctccgggag 1320  
 aagggcttcg accaggcgcc cgtggtggat gaggcggggg taatcctggg aatggtgacg 1380  
 cttgggaaca tgctctcgtc cctgcttgcc gggaaggtgc agccgtcaga ccaagttggc 1440  
 aaagtcatct acaagcagtt caaacagatc cgcctcacgg acacgctggg caggctctcg 1500  
 cacatcctgg agatggacca cttcgccctg gtggtgcacg agcagatcca gtaccacagc 1560  
 accgggaagt ccagtcagcg gcagatggtg ttcgggggtg tcaccgccat tgacttgctg 1620  
 aacttcgtgg ccgcccagga gcgggaccag aagtga 1656

<210> 66  
 <211> 551  
 <212> PRT  
 <213> 智人

<400> 66

Met Pro Ser Glu Thr Pro Gln Ala Glu Val Gly Pro Thr Gly Cys Pro  
1 5 10 15

His Arg Ser Gly Pro His Ser Ala Lys Gly Ser Leu Glu Lys Gly Ser  
20 25 30

Pro Glu Asp Lys Glu Ala Lys Glu Pro Leu Trp Ile Arg Pro Asp Ala  
35 40 45

Pro Ser Arg Cys Thr Trp Gln Leu Gly Arg Pro Ala Ser Glu Ser Pro  
50 55 60

His His His Thr Ala Pro Ala Lys Ser Pro Lys Ile Leu Pro Asp Ile  
65 70 75 80

Leu Lys Lys Ile Gly Asp Thr Pro Met Val Arg Ile Asn Lys Ile Gly  
85 90 95

Lys Lys Phe Gly Leu Lys Cys Glu Leu Leu Ala Lys Cys Glu Phe Phe  
100 105 110

Asn Ala Gly Gly Ser Val Lys Asp Arg Ile Ser Leu Arg Met Ile Glu  
115 120 125

Asp Ala Glu Arg Asp Gly Thr Leu Lys Pro Gly Asp Thr Ile Ile Glu

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

Arg Cys Val Val Ile Leu Pro Asp Ser Val Arg Asn Tyr Met Thr Lys  
370 375 380

Phe Leu Ser Asp Arg Trp Met Leu Gln Lys Gly Phe Leu Lys Glu Glu  
385 390 395 400

Asp Leu Thr Glu Lys Lys Pro Trp Trp Trp His Leu Arg Val Gln Glu  
405 410 415

Leu Gly Leu Ser Ala Pro Leu Thr Val Leu Pro Thr Ile Thr Cys Gly  
420 425 430

His Thr Ile Glu Ile Leu Arg Glu Lys Gly Phe Asp Gln Ala Pro Val  
435 440 445

Val Asp Glu Ala Gly Val Ile Leu Gly Met Val Thr Leu Gly Asn Met  
450 455 460

Leu Ser Ser Leu Leu Ala Gly Lys Val Gln Pro Ser Asp Gln Val Gly  
465 470 475 480

Lys Val Ile Tyr Lys Gln Phe Lys Gln Ile Arg Leu Thr Asp Thr Leu  
485 490 495

Gly Arg Leu Ser His Ile Leu Glu Met Asp His Phe Ala Leu Val Val  
500 505 510

His Glu Gln Ile Gln Tyr His Ser Thr Gly Lys Ser Ser Gln Arg Gln  
515 520 525

Met Val Phe Gly Val Val Thr Ala Ile Asp Leu Leu Asn Phe Val Ala  
530 535 540

Ala Gln Glu Arg Asp Gln Lys  
545 550

<210> 67  
<211> 1428  
<212> DNA  
<213> 智人

<400> 67  
atggagagca aagccctgct cgtgctgact ctggccgtgt ggctccagag tctgaccgcc

60

tccgcggag gggtgccgc cgccgaccaa agaagagatt ttatcgacat cgaaagtaaa 120  
 tttgccctaa ggaccctga agacacagct gaggacactt gccacctcat tccggagta 180  
 gcagagtccg tggctacctg tcatttcaat cacagcagca aaaccttcat ggtgatccat 240  
 ggctggacgg taacaggaat gtatgagagt tgggtgccaa aacttgtggc cgccctgtac 300  
 aagagagaac cagactccaa tgtcattgtg gtggactggc tgtcacgggc tcaggagcat 360  
 taccagtgt ccgcgggcta caccaaaactg gtgggacagg atgtggcccg gtttatcaac 420  
 tggatggagg aggagtttaa ctaccctctg gacaatgtcc atctcttggg atacagcctt 480  
 ggagcccatg ctgctggcat tgcaggaagt ctgaccaata agaaagtcaa cagaattact 540  
 ggccctgatc cagctggacc taactttgag tatgcagaag ccccgagtgc tctttctcct 600  
 gatgatgcag atttttaga cgtcttacac acattcacca gagggtcccc tggtcgaagc 660  
 attggaatcc agaaaaccagt tgggcatggt gacatttacc cgaatggagg tacttttcag 720  
 ccaggatgta acattggaga agctatccgc gtgattgcag agagaggact tggagatgtg 780  
 gaccagctag tgaagtgctc ccacgagcgc tccattcadc tcttcatcga ctctctgttg 840  
 aatgaagaaa atccaagtaa ggcctacagg tgcagttoca aggaagcctt tgagaaaggg 900  
 ctctgcttga gtttagataa gaaccgctgc aacaatctgg gctatgagat caataaagtc 960  
 agagccaaaa gaagcagcaa aatgtacctg aagactcgtt ctcatatgcc ctacaaagtc 1020  
 ttccattacc aagtaaagat tcatttttct gggactgaga gtgaaacca taccatcag 1080  
 gcctttgaga tttctctgta tggcaccgtg gccgagagtg agaacatccc attcactctg 1140  
 cctgaagttt ccacaaataa gacctactcc ttctaattt acacagaggt agatattgga 1200  
 gaactactca tgttgaagct caaatggaag agtgattcat acttttagctg gtcagactgg 1260  
 tggagcagtc ccggcttcgc cattcagaag atcagagtaa aagcaggaga gactcagaaa 1320  
 aaggtgatct tctgttctag ggagaaagtg tctcatttgc agaaaggaaa ggcacctgcg 1380  
 gtatttgtga aatgcoatga caagtctctg aataagaagt caggctga 1428

<210> 68  
 <211> 475  
 <212> PRT  
 <213> 智人

<400> 68

Met Glu Ser Lys Ala Leu Leu Val Leu Thr Leu Ala Val Trp Leu Gln  
 1 5 10 15

Ser Leu Thr Ala Ser Arg Gly Gly Val Ala Ala Ala Asp Gln Arg Arg  
20 25 30

Asp Phe Ile Asp Ile Glu Ser Lys Phe Ala Leu Arg Thr Pro Glu Asp  
35 40 45

Thr Ala Glu Asp Thr Cys His Leu Ile Pro Gly Val Ala Glu Ser Val  
50 55 60

Ala Thr Cys His Phe Asn His Ser Ser Lys Thr Phe Met Val Ile His  
65 70 75 80

Gly Trp Thr Val Thr Gly Met Tyr Glu Ser Trp Val Pro Lys Leu Val  
85 90 95

Ala Ala Leu Tyr Lys Arg Glu Pro Asp Ser Asn Val Ile Val Val Asp  
100 105 110

Trp Leu Ser Arg Ala Gln Glu His Tyr Pro Val Ser Ala Gly Tyr Thr  
115 120 125

Lys Leu Val Gly Gln Asp Val Ala Arg Phe Ile Asn Trp Met Glu Glu  
130 135 140

Glu Phe Asn Tyr Pro Leu Asp Asn Val His Leu Leu Gly Tyr Ser Leu  
145 150 155 160

Gly Ala His Ala Ala Gly Ile Ala Gly Ser Leu Thr Asn Lys Lys Val  
165 170 175

Asn Arg Ile Thr Gly Leu Asp Pro Ala Gly Pro Asn Phe Glu Tyr Ala  
180 185 190

Glu Ala Pro Ser Arg Leu Ser Pro Asp Asp Ala Asp Phe Val Asp Val  
195 200 205

Leu His Thr Phe Thr Arg Gly Ser Pro Gly Arg Ser Ile Gly Ile Gln  
210 215 220

Lys Pro Val Gly His Val Asp Ile Tyr Pro Asn Gly Gly Thr Phe Gln  
225 230 235 240

Pro Gly Cys Asn Ile Gly Glu Ala Ile Arg Val Ile Ala Glu Arg Gly  
 245 250 255

Leu Gly Asp Val Asp Gln Leu Val Lys Cys Ser His Glu Arg Ser Ile  
 260 265 270

His Leu Phe Ile Asp Ser Leu Leu Asn Glu Glu Asn Pro Ser Lys Ala  
 275 280 285

Tyr Arg Cys Ser Ser Lys Glu Ala Phe Glu Lys Gly Leu Cys Leu Ser  
 290 295 300

Cys Arg Lys Asn Arg Cys Asn Asn Leu Gly Tyr Glu Ile Asn Lys Val  
 305 310 315 320

Arg Ala Lys Arg Ser Ser Lys Met Tyr Leu Lys Thr Arg Ser Gln Met  
 325 330 335

Pro Tyr Lys Val Phe His Tyr Gln Val Lys Ile His Phe Ser Gly Thr  
 340 345 350

Glu Ser Glu Thr His Thr Asn Gln Ala Phe Glu Ile Ser Leu Tyr Gly  
 355 360 365

Thr Val Ala Glu Ser Glu Asn Ile Pro Phe Thr Leu Pro Glu Val Ser  
 370 375 380

Thr Asn Lys Thr Tyr Ser Phe Leu Ile Tyr Thr Glu Val Asp Ile Gly  
 385 390 395 400

Glu Leu Leu Met Leu Lys Leu Lys Trp Lys Ser Asp Ser Tyr Phe Ser  
 405 410 415

Trp Ser Asp Trp Trp Ser Ser Pro Gly Phe Ala Ile Gln Lys Ile Arg  
 420 425 430

Val Lys Ala Gly Glu Thr Gln Lys Lys Val Ile Phe Cys Ser Arg Glu  
 435 440 445

Lys Val Ser His Leu Gln Lys Gly Lys Ala Pro Ala Val Phe Val Lys  
 450 455 460

Cys His Asp Lys Ser Leu Asn Lys Lys Ser Gly  
 465 470 475

<210> 69  
 <211> 1239  
 <212> DNA  
 <213> 智人

<400> 69  
 atgtccagca aaggctccgt gttctggcc tacagtggcg gcctggacac ctctgtcatc 60  
 ctctgtgtggc tgaaggaaca aggctatgac gtcattgcct atctggccaa cattggccag 120  
 aaggaagact tcgaggaagc caggaagaag gcactgaagc ttggggccaa aaaggtgttc 180  
 attgaggatg tcagcaggga gtttgtggag gagttcatct ggccggccat ccagtccagc 240  
 gcactgtatg aggaccgcta cctcctgggc acctctcttg ccaggccctg catcgcccgc 300  
 aaacaagtgg aaatcgccca gcgggagggg gccaagtatg tgtcccacgg cgccacagga 360  
 aaggggaacg atcaggtccg gtttgagctc agctgctact cactggcccc ccagataaag 420  
 gtcattgctc cctggaggat gcctgaattc tacaacoggt tcaagggcog caatgacctg 480  
 atggagtacg caaagcaaca cgggattccc atcccggtoa ctccaagaa cccgtggagc 540  
 atggatgaga acctcatgca catcagctac gaggctggaa tcctggagaa cccaagaac 600  
 caagcgcctc caggtctcta cacgaagacc caggaccag ccaaagcccc caacaccct 660  
 gacattctcg agatcgagtt caaaaaagg gtccctgtga aggtgacc aa cgtcaaggat 720  
 ggcaccacc accagacctc cttggagctc ttcatgtacc tgaacgaagt cgcgggcaag 780  
 catggcgtgg gcogtattga catcgtggag aaccgcttca ttggaatgaa gtcccagagt 840  
 atctacgaga cccagcagg caccatcctt taccatgctc atttagacat cgaggccttc 900  
 accatggacc ggggaagtgcg caaaatcaaa caaggcctgg gcttgaaatt tgctgagctg 960  
 gtgtataccg gtttctggca cagccctgag tgtgaatttg tccgccaactg catcgccaag 1020  
 tccaggagc gagtggaagg gaaagtgcag gtgtccgtcc tcaagggcca ggtgtacatc 1080  
 ctggccggg agtcccact gtctctctac aatgaggagc tggtgagcat gaacgtgcag 1140  
 ggtgattatg agccaactga tgccaccggg ttcatcaaca tcaattccct caggctgaag 1200  
 gaatatcatc gtctccagag caaggcact gccaaatag 1239

<210> 70  
 <211> 412  
 <212> PRT  
 <213> 智人

&lt;400&gt; 70

Met Ser Ser Lys Gly Ser Val Val Leu Ala Tyr Ser Gly Gly Leu Asp  
1 5 10 15

Thr Ser Cys Ile Leu Val Trp Leu Lys Glu Gln Gly Tyr Asp Val Ile  
20 25 30

Ala Tyr Leu Ala Asn Ile Gly Gln Lys Glu Asp Phe Glu Glu Ala Arg  
35 40 45

Lys Lys Ala Leu Lys Leu Gly Ala Lys Lys Val Phe Ile Glu Asp Val  
50 55 60

Ser Arg Glu Phe Val Glu Glu Phe Ile Trp Pro Ala Ile Gln Ser Ser  
65 70 75 80

Ala Leu Tyr Glu Asp Arg Tyr Leu Leu Gly Thr Ser Leu Ala Arg Pro  
85 90 95

Cys Ile Ala Arg Lys Gln Val Glu Ile Ala Gln Arg Glu Gly Ala Lys  
100 105 110

Tyr Val Ser His Gly Ala Thr Gly Lys Gly Asn Asp Gln Val Arg Phe  
115 120 125

Glu Leu Ser Cys Tyr Ser Leu Ala Pro Gln Ile Lys Val Ile Ala Pro  
130 135 140

Trp Arg Met Pro Glu Phe Tyr Asn Arg Phe Lys Gly Arg Asn Asp Leu  
145 150 155 160

Met Glu Tyr Ala Lys Gln His Gly Ile Pro Ile Pro Val Thr Pro Lys  
165 170 175

Asn Pro Trp Ser Met Asp Glu Asn Leu Met His Ile Ser Tyr Glu Ala  
180 185 190

Gly Ile Leu Glu Asn Pro Lys Asn Gln Ala Pro Pro Gly Leu Tyr Thr  
195 200 205

Lys Thr Gln Asp Pro Ala Lys Ala Pro Asn Thr Pro Asp Ile Leu Glu  
210 215 220

Ile Glu Phe Lys Lys Gly Val Pro Val Lys Val Thr Asn Val Lys Asp  
225 230 235 240

Gly Thr Thr His Gln Thr Ser Leu Glu Leu Phe Met Tyr Leu Asn Glu  
245 250 255

Val Ala Gly Lys His Gly Val Gly Arg Ile Asp Ile Val Glu Asn Arg  
260 265 270

Phe Ile Gly Met Lys Ser Arg Gly Ile Tyr Glu Thr Pro Ala Gly Thr  
275 280 285

Ile Leu Tyr His Ala His Leu Asp Ile Glu Ala Phe Thr Met Asp Arg  
290 295 300

Glu Val Arg Lys Ile Lys Gln Gly Leu Gly Leu Lys Phe Ala Glu Leu  
305 310 315 320

Val Tyr Thr Gly Phe Trp His Ser Pro Glu Cys Glu Phe Val Arg His  
325 330 335

Cys Ile Ala Lys Ser Gln Glu Arg Val Glu Gly Lys Val Gln Val Ser  
340 345 350

Val Leu Lys Gly Gln Val Tyr Ile Leu Gly Arg Glu Ser Pro Leu Ser  
355 360 365

Leu Tyr Asn Glu Glu Leu Val Ser Met Asn Val Gln Gly Asp Tyr Glu  
370 375 380

Pro Thr Asp Ala Thr Gly Phe Ile Asn Ile Asn Ser Leu Arg Leu Lys  
385 390 395 400

Glu Tyr His Arg Leu Gln Ser Lys Val Thr Ala Lys  
405 410

<210> 71  
<211> 1359  
<212> DNA  
<213> 智人

<400> 71  
atgtccactg cggtoctgga aaaccaggc ttgggcagga aactctctga ctttggacag

60

gaaacaagct atattgaaga caactgcaat caaaatgggtg ccatatcaact gatctttctca 120  
 ctcaaagaag aagttgggtgc attggccaaa gtattgcgct tatttgagga gaatgatgta 180  
 aacctgacct acattgaatc tagaccttct cgtttaaaga aagatgagta tgaatttttc 240  
 acccatttgg ataaaagtag cctgcctgct ctgacaaaaca tcatcaagat cttgaggcat 300  
 gacattgggtg ccaactgtcca tgagctttca cgagataaga agaaagacac agtgccttgg 360  
 ttccaagaa ccattcaaga gctggacaga tttgccaatc agattctcag ctatggagcg 420  
 gaactggatg ctgaccacctg tggttttaaa gatcctgtgt accgtgcaag acggaagcag 480  
 tttgctgaca ttgcctacaa ctaccgccat gggcagccca tccctcgagt ggaatacatg 540  
 gaggaagaaa agaaaacatg gggcacagtg ttcaagactc tgaagtcctt gtataaaacc 600  
 catgcttctg atgagtacaa tcacattttt ccacttcttg aaaagtaactg tggcttccat 660  
 gaagataaca ttccccagct ggaagacgtt tctcagttcc tgcagacttg cactggtttc 720  
 cgctccgac ctgtggctgg cctgctttcc tctcgggatt tcttgggtgg cctggccttc 780  
 cgagtcttcc actgcacaca gtacatcaga catggatoca agcccatgta tcccccgaa 840  
 cctgacatct gccatgagct gttgggacat gtgcccttgt tttcagatcg cagctttgcc 900  
 cagttttccc aggaaattgg ccttgctctt ctgggtgcac ctgatgaata cattgaaaag 960  
 ctgccacaa tttactgggt tactgtggag tttgggtctt gcaaacaagg agactccata 1020  
 aaggcatatg gtgctgggct cctgtcatcc tttggtgaat tacagtactg cttatcagag 1080  
 aagccaaagc ttotccccct ggagctggag aagacagcca tccaaaatta cactgtcacg 1140  
 gattccagc ccctctatta cgtggcagag agttttaatg atgccaagga gaaagtaagg 1200  
 aactttgctg ccacaatacc tggcccttc tcagttogct acgaccata cacccaaagg 1260  
 attgaggtct tggacaatac ccagcagctt aagattttgg ctgattccat taacagtgaa 1320  
 attggaatcc tttgcagtgc cctccagaaa ataaagtaa 1359

<210> 72  
 <211> 452  
 <212> PRT  
 <213> 智人  
  
 <400> 72

Met Ser Thr Ala Val Leu Glu Asn Pro Gly Leu Gly Arg Lys Leu Ser  
 1                    5                    10                    15

Asp Phe Gly Gln Glu Thr Ser Tyr Ile Glu Asp Asn Cys Asn Gln Asn  
20 25 30

Gly Ala Ile Ser Leu Ile Phe Ser Leu Lys Glu Glu Val Gly Ala Leu  
35 40 45

Ala Lys Val Leu Arg Leu Phe Glu Glu Asn Asp Val Asn Leu Thr His  
50 55 60

Ile Glu Ser Arg Pro Ser Arg Leu Lys Lys Asp Glu Tyr Glu Phe Phe  
65 70 75 80

Thr His Leu Asp Lys Arg Ser Leu Pro Ala Leu Thr Asn Ile Ile Lys  
85 90 95

Ile Leu Arg His Asp Ile Gly Ala Thr Val His Glu Leu Ser Arg Asp  
100 105 110

Lys Lys Lys Asp Thr Val Pro Trp Phe Pro Arg Thr Ile Gln Glu Leu  
115 120 125

Asp Arg Phe Ala Asn Gln Ile Leu Ser Tyr Gly Ala Glu Leu Asp Ala  
130 135 140

Asp His Pro Gly Phe Lys Asp Pro Val Tyr Arg Ala Arg Arg Lys Gln  
145 150 155 160

Phe Ala Asp Ile Ala Tyr Asn Tyr Arg His Gly Gln Pro Ile Pro Arg  
165 170 175

Val Glu Tyr Met Glu Glu Glu Lys Lys Thr Trp Gly Thr Val Phe Lys  
180 185 190

Thr Leu Lys Ser Leu Tyr Lys Thr His Ala Cys Tyr Glu Tyr Asn His  
195 200 205

Ile Phe Pro Leu Leu Glu Lys Tyr Cys Gly Phe His Glu Asp Asn Ile  
210 215 220

Pro Gln Leu Glu Asp Val Ser Gln Phe Leu Gln Thr Cys Thr Gly Phe  
225 230 235 240

Arg Leu Arg Pro Val Ala Gly Leu Leu Ser Ser Arg Asp Phe Leu Gly



<212> DNA  
 <213> 智人

<400> 73

```

atggggccct ggggctggaa attgcgctgg accgtcgcct tgctcctcgc cgcggcgggg      60
actgcagtgg gcgacagatg cgaagaaac gagttccagt gccaagacgg gaaatgcata      120
tctacaagt gggctcgcga tggcagcgt gagtgccagg atggctctga tgagtcccag      180
gagacgtgct tgtctgtcac ctgcaaatcc ggggacttca gctgtggggg cgtgtcaac      240
cgctgcattc ctcagttctg gaggtgcgat ggccaagtgg actgcgacaa cggctcagac      300
gagcaaggct gtcccccaa gacgtgctcc caggacgagt ttcgctgcca cगतgggaag      360
tgcattcttc ggcagttcgt ctgtgactca gaccgggact gcttggacgg ctcagacgag      420
gcctcctgcc cggtgctcac ctgtggtccc gccagcttcc agtgcaacag ctccacctgc      480
atcccccagc tgtgggcctg cgacaacgac cccgactgcg aagatggctc ggatgagtgg      540
ccgcagcgtc gtaggggtct ttacgtgttc caaggggaca gtagcccctg ctcggccttc      600
gagttccact gcctaagtgg cgagtgcata cactccagct ggcgctgtga tgggtggccc      660
gactgcaagg acaaatctga cgaggaaaac tgcgctgtgg ccacctgtcg ccctgacgaa      720
ttccagtgct ctgatggaaa ctgcatccat ggcagccggc agtgtgaccg ggaatatgac      780
tgcaaggaca tgagcgatga agttggctgc gttaatgtga cactctgcca gggacccaac      840
aagttcaagt gtcacagcgg cgaatgcata accctggaca aagtctgcaa catggctaga      900
gactgccggg actggtcaga tgaaccatac aaagagtgcg ggaccaacga atgcttggac      960
aacaacggcg gctgttccca cgtctgcaat gaccttaaga tcggctacga gtgcctgtgc     1020
ccgcacggct tccagctggt gggccagcga agatgcgaag atatcgatga gtgtcaggat     1080
ccgcacacct gcagccagct ctgcgtgaac ctggagggtg gctacaagtg ccagtgtgag     1140
gaaggcttcc agctggaccc ccacacgaag gcctgcaagg ctgtgggctc catcgcctac     1200
cttttttca ccaaccggca cgaggtcagg aagatgacgc tggaccggag cgagtacacc     1260
agcctcatcc ccaacctgag gaacgtggtc gctctggaca cggaggtggc cagcaataga     1320
atctactggt ctgacctgtc ccagagaatg atctgcagca cccagcttga cagagcccac     1380
ggcgtctctt cctatgacac cgtcatcagc agagacatcc aggccccga cgggctggct     1440
gtggactgga tccacagcaa catctactgg accgactctg tcctgggcac tgtctctgtt     1500
gcggatacca agggcgtgaa gaggaaaacg ttattcaggg agaacggctc caagccaagg     1560
gccatcgtgg tggatcctgt tcatggcttc atgtactgga ctgactgggg aactcccgcc     1620

```

aagatcaaga aagggggcct gaatggtgtg gacatctact cgctggtgac tgaaaacatt 1680  
cagtggccca atggcatcac cctagatctc ctcagtggcc gcctctactg ggttgactcc 1740  
aaacttcact ccatctcaag catcgatgtc aacgggggca accggaagac catcttggag 1800  
gatgaaaaga ggctggccca ccccttctcc ttggccgtct ttgaggacaa agtatttttg 1860  
acagatatca tcaacgaagc cattttcagt gccaaccgcc tcacaggttc cgatgtcaac 1920  
ttgttggctg aaaacctact gtccccagag gatatggttc tcttccacaa cctcaccag 1980  
ccaagaggag tgaactggtg tgagaggacc accctgagca atggcggctg ccagtatctg 2040  
tgctccctg ccccgagat caacccccac tcgccaagt ttacctgogc ctgcccggac 2100  
ggcatgctgc tggccaggga catgaggagc tgctcacag aggctgaggc tgcagtggcc 2160  
accaggaga catccaccgt caggctaaag gtcagctcca cagccgtaag gacacagcac 2220  
acaaccacc gacctgttcc cgacacctc cggctgctg gggccacccc tgggctcacc 2280  
acggtggaga tagtgacaat gtctcaccaa gctctggggc acgttgctgg cagaggaaat 2340  
gagaagaagc ccagtagcgt gagggctctg tcattgtcc tccccatcgt gctcctcgtc 2400  
ttcctttgcc tgggggtctt ccttctatgg aagaactggc ggcttaagaa catcaacagc 2460  
atcaactttg acaaccocgt ctatcagaag accacagagg atgaggcca catttgccac 2520  
aaccaggacg gctacagcta cccctogaga cagatggtea gtctggagga tgacgtggcg 2580  
tga 2583

<210> 74  
<211> 860  
<212> PRT  
<213> 智人

<400> 74

Met Gly Pro Trp Gly Trp Lys Leu Arg Trp Thr Val Ala Leu Leu Leu  
1 5 10 15

Ala Ala Ala Gly Thr Ala Val Gly Asp Arg Cys Glu Arg Asn Glu Phe  
20 25 30

Gln Cys Gln Asp Gly Lys Cys Ile Ser Tyr Lys Trp Val Cys Asp Gly  
35 40 45

Ser Ala Glu Cys Gln Asp Gly Ser Asp Glu Ser Gln Glu Thr Cys Leu  
50 55 60

Ser Val Thr Cys Lys Ser Gly Asp Phe Ser Cys Gly Gly Arg Val Asn  
65 70 75 80

Arg Cys Ile Pro Gln Phe Trp Arg Cys Asp Gly Gln Val Asp Cys Asp  
85 90 95

Asn Gly Ser Asp Glu Gln Gly Cys Pro Pro Lys Thr Cys Ser Gln Asp  
100 105 110

Glu Phe Arg Cys His Asp Gly Lys Cys Ile Ser Arg Gln Phe Val Cys  
115 120 125

Asp Ser Asp Arg Asp Cys Leu Asp Gly Ser Asp Glu Ala Ser Cys Pro  
130 135 140

Val Leu Thr Cys Gly Pro Ala Ser Phe Gln Cys Asn Ser Ser Thr Cys  
145 150 155 160

Ile Pro Gln Leu Trp Ala Cys Asp Asn Asp Pro Asp Cys Glu Asp Gly  
165 170 175

Ser Asp Glu Trp Pro Gln Arg Cys Arg Gly Leu Tyr Val Phe Gln Gly  
180 185 190

Asp Ser Ser Pro Cys Ser Ala Phe Glu Phe His Cys Leu Ser Gly Glu  
195 200 205

Cys Ile His Ser Ser Trp Arg Cys Asp Gly Gly Pro Asp Cys Lys Asp  
210 215 220

Lys Ser Asp Glu Glu Asn Cys Ala Val Ala Thr Cys Arg Pro Asp Glu  
225 230 235 240

Phe Gln Cys Ser Asp Gly Asn Cys Ile His Gly Ser Arg Gln Cys Asp  
245 250 255

Arg Glu Tyr Asp Cys Lys Asp Met Ser Asp Glu Val Gly Cys Val Asn  
260 265 270

Val Thr Leu Cys Glu Gly Pro Asn Lys Phe Lys Cys His Ser Gly Glu  
275 280 285

Cys Ile Thr Leu Asp Lys Val Cys Asn Met Ala Arg Asp Cys Arg Asp  
290 295 300

Trp Ser Asp Glu Pro Ile Lys Glu Cys Gly Thr Asn Glu Cys Leu Asp  
305 310 315 320

Asn Asn Gly Gly Cys Ser His Val Cys Asn Asp Leu Lys Ile Gly Tyr  
325 330 335

Glu Cys Leu Cys Pro Asp Gly Phe Gln Leu Val Ala Gln Arg Arg Cys  
340 345 350

Glu Asp Ile Asp Glu Cys Gln Asp Pro Asp Thr Cys Ser Gln Leu Cys  
355 360 365

Val Asn Leu Glu Gly Gly Tyr Lys Cys Gln Cys Glu Glu Gly Phe Gln  
370 375 380

Leu Asp Pro His Thr Lys Ala Cys Lys Ala Val Gly Ser Ile Ala Tyr  
385 390 395 400

Leu Phe Phe Thr Asn Arg His Glu Val Arg Lys Met Thr Leu Asp Arg  
405 410 415

Ser Glu Tyr Thr Ser Leu Ile Pro Asn Leu Arg Asn Val Val Ala Leu  
420 425 430

Asp Thr Glu Val Ala Ser Asn Arg Ile Tyr Trp Ser Asp Leu Ser Gln  
435 440 445

Arg Met Ile Cys Ser Thr Gln Leu Asp Arg Ala His Gly Val Ser Ser  
450 455 460

Tyr Asp Thr Val Ile Ser Arg Asp Ile Gln Ala Pro Asp Gly Leu Ala  
465 470 475 480

Val Asp Trp Ile His Ser Asn Ile Tyr Trp Thr Asp Ser Val Leu Gly  
485 490 495

Thr Val Ser Val Ala Asp Thr Lys Gly Val Lys Arg Lys Thr Leu Phe  
500 505 510

Arg Glu Asn Gly Ser Lys Pro Arg Ala Ile Val Val Asp Pro Val His  
515 520 525

Gly Phe Met Tyr Trp Thr Asp Trp Gly Thr Pro Ala Lys Ile Lys Lys  
530 535 540

Gly Gly Leu Asn Gly Val Asp Ile Tyr Ser Leu Val Thr Glu Asn Ile  
545 550 555 560

Gln Trp Pro Asn Gly Ile Thr Leu Asp Leu Leu Ser Gly Arg Leu Tyr  
565 570 575

Trp Val Asp Ser Lys Leu His Ser Ile Ser Ser Ile Asp Val Asn Gly  
580 585 590

Gly Asn Arg Lys Thr Ile Leu Glu Asp Glu Lys Arg Leu Ala His Pro  
595 600 605

Phe Ser Leu Ala Val Phe Glu Asp Lys Val Phe Trp Thr Asp Ile Ile  
610 615 620

Asn Glu Ala Ile Phe Ser Ala Asn Arg Leu Thr Gly Ser Asp Val Asn  
625 630 635 640

Leu Leu Ala Glu Asn Leu Leu Ser Pro Glu Asp Met Val Leu Phe His  
645 650 655

Asn Leu Thr Gln Pro Arg Gly Val Asn Trp Cys Glu Arg Thr Thr Leu  
660 665 670

Ser Asn Gly Gly Cys Gln Tyr Leu Cys Leu Pro Ala Pro Gln Ile Asn  
675 680 685

Pro His Ser Pro Lys Phe Thr Cys Ala Cys Pro Asp Gly Met Leu Leu  
690 695 700

Ala Arg Asp Met Arg Ser Cys Leu Thr Glu Ala Glu Ala Ala Val Ala  
705 710 715 720

Thr Gln Glu Thr Ser Thr Val Arg Leu Lys Val Ser Ser Thr Ala Val  
725 730 735

Arg Thr Gln His Thr Thr Thr Arg Pro Val Pro Asp Thr Ser Arg Leu



tcctgtgagt ggccctcttta tatgtggccc ttccaagaagc ccaattatac agaaatccga 660  
 cagtactgca atcactggcg aaattttgct gacattgatg attcctggaa aagtataaag 720  
 agtatcttgg actggacatc ttttaaccag gagagaattg ttgatgttgc tggaccaggg 780  
 ggttggaatg acccagatat gttagtgatt ggcaactttg gcctcagctg gaatcagcaa 840  
 gtaactcaga tggccctctg ggctatcatg gctgctcctt tattcatgtc taatgacctc 900  
 cgacacatca gccctcaagc caaagctctc cttcaggata aggacgtaat tgccatcaat 960  
 caggaccctt tgggcaagca agggtagcagg cttagacagg gagacaactt tgaagtgtgg 1020  
 gaacgacctc tctcaggctt agcctgggct gtagctatga taaaccggca ggagattggg 1080  
 ggacctcgct cttataccat cgcagttgct tccttgggta aaggagtggc ctgtaatcct 1140  
 gcctgcttca tcacacagct cctccctgtg aaaaggaagc tagggttcta tgaatggact 1200  
 tcaaggttaa gaagtcacat aaatcccaca ggcactgttt tgcttcagct agaaaatata 1260  
 atgcagatgt cattaaaaga cttactttaa 1290

<210> 76  
 <211> 429  
 <212> PRT  
 <213> 智人

<400> 76

Met Gln Leu Arg Asn Pro Glu Leu His Leu Gly Cys Ala Leu Ala Leu  
1 5 10 15

Arg Phe Leu Ala Leu Val Ser Trp Asp Ile Pro Gly Ala Arg Ala Leu  
20 25 30

Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu  
35 40 45

Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys Ile  
50 55 60

Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu Gly  
65 70 75 80

Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met  
85 90 95

Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg

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

Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val Ala  
 340 345 350

Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile Ala  
 355 360 365

Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile  
 370 375 380

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

Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu Gln  
 405 410 415

Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu  
 420 425

<210> 77  
 <211> 1257  
 <212> DNA  
 <213> 智人

<400> 77  
 atgccgtcct ctgtctcgtg gggcatcctc ctgctggcag gcctgtgctg cctggtcctt 60  
 gtctccctgg ctgaggatcc ccaggagat gctgcccaga agacagatac atcccaccat 120  
 gatcaggatc acccaacctt caacaagatc accccaacc tggctgagtt cgccttcagc 180  
 ctataccgcc agctggcaca ccagtccaac agcaccaata tcttcttctc cccagtgagc 240  
 atcgctacag cctttgcaat gctctccctg gggaccaagg ctgacactca cgatgaaatc 300  
 ctggagggcc tgaatttcaa cctcacggag attccggagg ctgagatcca tgaaggcttc 360  
 caggaactcc tccgtaccct caaccagcca gacagccagc tccagctgac caccggcaat 420  
 ggocctgttc tcagcgaggg cctgaagcta gtggataagt ttttggagga tgttaaaaag 480  
 ttgtaccact cagaagcctt cactgtcaac ttcggggaca ccgaagaggc caagaaacag 540  
 atcaacgatt acgtggagaa gggtactcaa gggaaaattg tggatttggc caaggagctt 600  
 gacagagaca cagtttttgc tctggtgaat tacatcttct ttaaaggcaa atgggagaga 660  
 ccctttgaag tcaaggacac cgaggaagag gacttccaag tggaccaggt gaccaccgtg 720  
 aaggtgccta tgatgaagcg tttagcatg tttaacatcc agcactgtaa gaagctgtcc 780

agctgggtgc tgctgatgaa atacctgggc aatgccaccg ccatcttctt cctgcctgat 840  
 gaggggaaac tacagcacct ggaaaatgaa ctaccccacg atatcatcac caagttcctg 900  
 gaaaatgaag acagaaggtc tgccagctta catttaccca aactgtccat tactggaacc 960  
 tatgatctga agagcgtcct gggtcaactg ggcatcacta aggtcttcag caatggggct 1020  
 gacctctccg gggtcacaga ggaggcacc ctagaagctct ccaaggccgt gcataaggct 1080  
 gtgctgacca tcgacgagaa agggactgaa gctgctgggg ccatgttttt agaggccata 1140  
 cccatgtcta tccccccga ggtcaagttc aacaaccct ttgtcttctt aatgattgaa 1200  
 caaaatacca agtctcccct cttcatggga aaagtgggta atcccaccca aaaataa 1257

<210> 78  
 <211> 418  
 <212> PRT  
 <213> 智人

<400> 78

Met Pro Ser Ser Val Ser Trp Gly Ile Leu Leu Leu Ala Gly Leu Cys  
1 5 10 15

Cys Leu Val Pro Val Ser Leu Ala Glu Asp Pro Gln Gly Asp Ala Ala  
20 25 30

Gln Lys Thr Asp Thr Ser His His Asp Gln Asp His Pro Thr Phe Asn  
35 40 45

Lys Ile Thr Pro Asn Leu Ala Glu Phe Ala Phe Ser Leu Tyr Arg Gln  
50 55 60

Leu Ala His Gln Ser Asn Ser Thr Asn Ile Phe Phe Ser Pro Val Ser  
65 70 75 80

Ile Ala Thr Ala Phe Ala Met Leu Ser Leu Gly Thr Lys Ala Asp Thr  
85 90 95

His Asp Glu Ile Leu Glu Gly Leu Asn Phe Asn Leu Thr Glu Ile Pro  
100 105 110

Glu Ala Gln Ile His Glu Gly Phe Gln Glu Leu Leu Arg Thr Leu Asn  
115 120 125

Gln Pro Asp Ser Gln Leu Gln Leu Thr Thr Gly Asn Gly Leu Phe Leu  
 130 135 140

Ser Glu Gly Leu Lys Leu Val Asp Lys Phe Leu Glu Asp Val Lys Lys  
 145 150 155 160

Leu Tyr His Ser Glu Ala Phe Thr Val Asn Phe Gly Asp Thr Glu Glu  
 165 170 175

Ala Lys Lys Gln Ile Asn Asp Tyr Val Glu Lys Gly Thr Gln Gly Lys  
 180 185 190

Ile Val Asp Leu Val Lys Glu Leu Asp Arg Asp Thr Val Phe Ala Leu  
 195 200 205

Val Asn Tyr Ile Phe Phe Lys Gly Lys Trp Glu Arg Pro Phe Glu Val  
 210 215 220

Lys Asp Thr Glu Glu Glu Asp Phe His Val Asp Gln Val Thr Thr Val  
 225 230 235 240

Lys Val Pro Met Met Lys Arg Leu Gly Met Phe Asn Ile Gln His Cys  
 245 250 255

Lys Lys Leu Ser Ser Trp Val Leu Leu Met Lys Tyr Leu Gly Asn Ala  
 260 265 270

Thr Ala Ile Phe Phe Leu Pro Asp Glu Gly Lys Leu Gln His Leu Glu  
 275 280 285

Asn Glu Leu Thr His Asp Ile Ile Thr Lys Phe Leu Glu Asn Glu Asp  
 290 295 300

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

Tyr Asp Leu Lys Ser Val Leu Gly Gln Leu Gly Ile Thr Lys Val Phe  
 325 330 335

Ser Asn Gly Ala Asp Leu Ser Gly Val Thr Glu Glu Ala Pro Leu Lys  
 340 345 350

Leu Ser Lys Ala Val His Lys Ala Val Leu Thr Ile Asp Glu Lys Gly

355

360

365

Thr Glu Ala Ala Gly Ala Met Phe Leu Glu Ala Ile Pro Met Ser Ile  
 370 375 380

Pro Pro Glu Val Lys Phe Asn Lys Pro Phe Val Phe Leu Met Ile Glu  
 385 390 395 400

Gln Asn Thr Lys Ser Pro Leu Phe Met Gly Lys Val Val Asn Pro Thr  
 405 410 415

Gln Lys

<210> 79  
 <211> 2859  
 <212> DNA  
 <213> 智人

<400> 79  
 atgggagtga ggcacccgcc ctgctccac cggctcctgg cegtctgogc cctcgtgtcc 60  
 ttggcaaccg ctgcaactct ggggcacatc ctactccatg atttctgtct ggttccccga 120  
 gagctgagtg gctcctcccc agtccctggag gagactcacc cagctcacca gcagggagcc 180  
 agcagaccag ggcocccggga tgcccaggca cccccggcc gtcccagagc agtgcccaca 240  
 cagtgcgacg tccccccaa cagccgcttc gattgcgccc ctgacaaggc catcaccag 300  
 gaacagtgcg aggcccgagg ctgttgctac atccctgcaa agcaggggct gcagggagcc 360  
 cagatggggc agccctggtg cttcttccca ccagctacc ccagctacaa gctggagaac 420  
 ctgagctcct ctgaaatggg ctacacggcc accctgacc gtaccacccc caccttcttc 480  
 cccaaggaca tcctgaccct gggctggac gtgatgatgg agactgagaa ccgcctccac 540  
 ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttgagagc ccgcatgtc 600  
 cacagccggg caccgtcccc actctacagc gtggagtctt ccgaggagcc cttcgggggtg 660  
 atcgtgcgcc ggcagctgga cggcccgctg ctgctgaaca cgacggtggc gccctgttc 720  
 tttgoggacc agttccttca gctgtccacc tcgctgcctt cgcagtatat cacaggcctc 780  
 gccgagcacc tcagtcccct gatgtcagc accagctgga ccaggatcac cctgtggaac 840  
 cgggaccttg cgcacacgcc cgggtggaac ctctacgggt ctacccttt ctacctggcg 900  
 ctggaggacg gcgggtcggc acacggggtg ttctgctaa acagcaatgc catggatgtg 960

gtcctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggatcct ggatgtctac	1020
atcttcctgg gccagagacc caagagcgtg gtgcagcagt acctggacgt tgtgggatac	1080
ccgttcatgc cgccatactg gggcctgggc ttccacctgt gccgctgggg ctactcctcc	1140
accgctatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc	1200
cagtggaaac acctggacta catggactcc cggagggact tcacgttcaa caaggatggc	1260
ttccgggact tcccgccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg	1320
atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag	1380
ggtctgcgga ggggggtttt catcaccaac gagaccggcc agccgctgat tgggaaggta	1440
tggcccgggt ccactgcctt ccccgacttc accaaccoca cagccctggc ctggtgggag	1500
gacatggtgg ctgagttcca tgaccaggtg cccttcgacg gcatgtggat tgacatgaac	1560
gagccttcca acttcatcag gggctctgag gacggctgcc ccaacaatga gctggagaac	1620
ccaccctacg tgccctgggt ggttgggggg acctccagg cggccaccat ctgtgcctcc	1680
agccaccagt ttctctccac aactacaac ctgcacaacc tctacggcct gaccgaagcc	1740
atgcctccc acagggcgct ggtgaaggct cgggggacac gccatttgt gatctcccgc	1800
togaccttg ctggccacgg ccgatacggc ggccactgga cgggggacgt gtggagctcc	1860
tgggagcagc tcgcctctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct	1920
ctggtcgggg ccgacgtctg cggcttctg ggcaaacct cagaggagct gtgtgtgcgc	1980
tggaccagc tgggggcctt ctacccttc atgcggaacc acaacagcct gctcagtctg	2040
cccaggagc cgtacagctt cagcgagccg gccagcagg ccatgaggaa ggccctcacc	2100
ctgcgctacg cactcctccc ccacctctac aactgttcc accaggccca cgtcgcgggg	2160
gagaccgtgg cccggcccct cttcctggag ttccccaaagg actctagcac ctggactgtg	2220
gaccaccagc tcctgtgggg ggaggccctg ctcatcacc cagtgtcca ggccgggaag	2280
gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccagta	2340
gaggcccttg gcagcctccc acccccacct gcagctccc gtgagccagc catccacagc	2400
gaggggcagt ggggtgacgt gccggcccc ctggacacca tcaacgtcca cctccgggct	2460
gggtacatca tccccctgca gggccctggc ctcaacaacca cagagtcccg ccagcagccc	2520
atggccctgg ctgtggccct gaccaagggt ggggaggccc gaggggagct gttctgggac	2580
gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat cttcctggcc	2640
aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgagggagc tggcctgcag	2700

ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggtcct ctccaacggt 2760  
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcttggacat ctgtgtctcg 2820  
ctgttgatgg gagagcagtt tctcgtcagc tgggttag 2859

<210> 80  
<211> 952  
<212> PRT  
<213> 智人

<400> 80

Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys  
1 5 10 15

Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu Gly His Ile Leu Leu  
20 25 30

His Asp Phe Leu Leu Val Pro Arg Glu Leu Ser Gly Ser Ser Pro Val  
35 40 45

Leu Glu Glu Thr His Pro Ala His Gln Gln Gly Ala Ser Arg Pro Gly  
50 55 60

Pro Arg Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr  
65 70 75 80

Gln Cys Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys  
85 90 95

Ala Ile Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro  
100 105 110

Ala Lys Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe  
115 120 125

Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser  
130 135 140

Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe  
145 150 155 160

Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu  
165 170 175

Asn Arg Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu  
180 185 190

Val Pro Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu  
195 200 205

Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg  
210 215 220

Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe  
225 230 235 240

Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr  
245 250 255

Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser  
260 265 270

Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly  
275 280 285

Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly  
290 295 300

Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val  
305 310 315 320

Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile  
325 330 335

Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln  
340 345 350

Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly  
355 360 365

Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr  
370 375 380

Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val  
385 390 395 400

Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe  
405 410 415

Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His  
420 425 430

Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser  
435 440 445

Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg  
450 455 460

Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val  
465 470 475 480

Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu  
485 490 495

Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe  
500 505 510

Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly  
515 520 525

Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val  
530 535 540

Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser  
545 550 555 560

Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly  
565 570 575

Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly  
580 585 590

Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg  
595 600 605

Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu  
610 615 620

Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro  
625 630 635 640

Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu  
645 650 655

Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg  
660 665 670

Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser  
675 680 685

Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala  
690 695 700

Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly  
705 710 715 720

Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser  
725 730 735

Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile  
740 745 750

Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro  
755 760 765

Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Val Glu Ala Leu Gly  
770 775 780

Ser Leu Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser  
785 790 795 800

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

His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr  
820 825 830

Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr  
835 840 845

Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser

850						855											860
Leu	Glu	Val	Leu	Glu	Arg	Gly	Ala	Tyr	Thr	Gln	Val	Ile	Phe	Leu	Ala		
865					870					875					880		
Arg	Asn	Asn	Thr	Ile	Val	Asn	Glu	Leu	Val	Arg	Val	Thr	Ser	Glu	Gly		
				885					890					895			
Ala	Gly	Leu	Gln	Leu	Gln	Lys	Val	Thr	Val	Leu	Gly	Val	Ala	Thr	Ala		
			900					905					910				
Pro	Gln	Gln	Val	Leu	Ser	Asn	Gly	Val	Pro	Val	Ser	Asn	Phe	Thr	Tyr		
			915				920					925					
Ser	Pro	Asp	Thr	Lys	Val	Leu	Asp	Ile	Cys	Val	Ser	Leu	Leu	Met	Gly		
	930					935					940						
Glu	Gln	Phe	Leu	Val	Ser	Trp	Cys										
945					950												

## 【發明申請專利範圍】

【請求項 1】一種用於治療遺傳疾病之系統，該系統包含：

(a)基因編輯載體，其包含編碼靶向 PCSK9 基因之核酸酶的核酸序列；及

(b)供體載體，其包含轉基因匣，該轉基因匣包含編碼轉基因之核酸序列及指導該轉基因在標靶細胞中之表現的調控序列，該供體載體進一步包含該轉基因匣之同源定向重組 (HDR) 臂 5' 及 3'，其中該轉基因不是 PCSK9。

【請求項 2】如請求項 1 之系統，其進一步包含調控序列，該調控序列指導該核酸酶在包含 PCSK9 基因之標靶細胞中的表現。

【請求項 3】如請求項 1 或 2 之系統，其中該核酸酶靶向 PCSK9 外顯子 7。

【請求項 4】如請求項 1 至 3 中任一項之系統，其中該核酸酶為對於 PCSK9 特異性的巨型核酸酶。

【請求項 5】如請求項 4 之系統，其中該巨型核酸酶為 ARCUS 巨型核酸酶。

【請求項 6】如請求項 1 或 2 之系統，其中該基因編輯載體包含編碼兩側有核定位訊號之 Cas9 的序列。

【請求項 7】如請求項 6 之系統，其中該基因編輯載體進一步包含含有至少 20 個核苷酸的 sgRNA，其與該 PCSK9 基因中的標靶位點特異性結合，該標靶位點位於被該 Cas9 特異性識別的前間隔序列相鄰基序 (PAM) 的 5' 處。

【請求項 8】如請求項 6 之系統，其中該供體載體進一步包含含有至少 20 個核苷酸種子區的 sgRNA，其中該 sgRNA 特異性結合該 PCSK9 基因中的標靶位點，該標靶位點位於被該 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處。

【請求項 9】如請求項 6 至 8 中任一項之系統，其進一步包含 RNA 聚合酶啟動子。

【請求項 10】如請求項 9 之系統，其中該 RNA 聚合酶啟動子為 U6 啟動子。

【請求項 11】如請求項 10 之系統，其中該 U6 啟動子位於該 sgRNA 的 5'。

【請求項 12】如請求項 7 至 11 中任一項之系統，其中該種子區與該標靶位點序列 100%互補。

【請求項 13】如請求項 7 至 11 中任一項之系統，其中該種子區與該標靶位點序列小於 100%互補。

【請求項 14】如請求項 1 至 13 中任一項之系統，其中該轉基因為 OTC、PKU、CTLN1 或 LDLR。

【請求項 15】如請求項 1 至 14 中任一項之系統，其中該供體載體及基因編輯載體中的至少一者為腺相關病毒(AAV)載體，且該 AAV 載體包含 AAV 5' ITR 及 AAV 3' ITR。

【請求項 16】如請求項 15 之系統，其中(a)的基因編輯 AAV 載體與(b)的供體 AAV 載體的比率使得(b)的供體 AAV 載體超過(a)的基因編輯載體。

【請求項 17】一種用於治療遺傳疾病之系統，該系統包含：

(a)基因編輯 AAV，其包含 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5' ITR、編碼巨型核酸酶之序列及 3' ITR，該巨型核酸酶在調控序列的控制下靶向 PCSK9，該調控序列指導該巨型核酸酶在包含 PCSK9 基因的標靶細胞中的表現；及

(b)供體 AAV 載體，其包含 AAV 衣殼及第二載體基因體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、及 3' ITR。

【請求項 18】一種用於治療遺傳疾病之系統，該系統包含：

(a)基因編輯 AAV，其包含 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5' ITR、5'核定位訊號(NLS)、編碼 Cas9 之序列及指導 saCas9 在包含 PCSK9 基因之標靶細胞中之表現的調控序列、3' NLS 及 3' ITR；及

(b)供體 AAV 載體，其包含 AAV 衣殼及第二載體基因體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3' HDR 臂、U6 啟動子、sgRNA 及 3' ITR，該 sgRNA 包含至少 20 個核苷酸，其特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處。

【請求項 19】一種用於治療遺傳疾病之系統，該系統包含：

(a)基因編輯 AAV 載體，其包含 AAV 衣殼及第一載體基因體，該第一載體基因體包含 5' ITR、U6 啟動子、sgRNA、5'核定位訊號(NLS)、編碼 Cas9 的序列及指導該 Cas9 在包含 PCSK9 基因之標靶細胞中之表現的調控序列、3' NLS 及 3' ITR，該 sgRNA 包含至少 20 個核苷酸，其特異性結合該 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處；及

(b)供體 AAV 載體，其包含 AAV 衣殼及第二載體基因體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3' HDR 臂及 3' ITR。

【請求項 20】一種用於治療遺傳疾病之系統，該系統包含：

(a)基因編輯載體，其包含：

(i)脂質奈米顆粒；

(ii)包含至少 20 個核苷酸之 sgRNA，其特異性結合 PCSK9 基因中的標靶位點，該標靶位點位於被 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處；

(iii)mRNA，其包含 5'核定位訊號(NLS)、編碼 Cas9 之序列、3' NLS；及

(b)供體 AAV 載體，其包含 AAV 衣殼及第二載體基因體，該第二載體基因體包含：5'ITR、5'同源定向重組(HDR)臂、轉基因及指導該轉基因在標靶細胞中之表現的調控序列、3' HDR 臂及 3' ITR。

【請求項 21】如請求項 17 至 19 中任一項之系統，其中(a)之該基因編輯 AAV 載體及(b)之該供體 AAV 載體具有相同 AAV 衣殼。

【請求項 22】如請求項 21 之系統，其中該 AAV 衣殼選自 AAV8、AAV9、rh10、AAV6.2、AAV3B、hu37、rh79 及 rh64。

【請求項 23】如請求項 6 至 18 或 18 至 22 中任一項之系統，其中 Cas9 選自金黃色葡萄球菌 Cas9 或釀膿鏈球菌 Cas9。

【請求項 24】如請求項 2 至 19 中任一項之系統，其中該核酸酶處於組織特異性啟動子的控制之下。

【請求項 25】如請求項 2 至 19 中任一項之系統，其中該核酸酶處於組成型啟動子的控制之下。

【請求項 26】如請求項 24 之系統，其中該核酸酶處於肝特異性啟動子、可選擇地人類甲狀腺素結合球蛋白(TBG)啟動子、或雜交肝啟動子(HLP)的控制之下。

【請求項 27】一種如請求項 1 至 26 中任一項之系統之用途，其用於治療人類之病症。

【請求項 28】一種下列(a)及(b)用於治療新生兒受試者之肝臟代謝性病症之用途，

(a) 基因編輯 AAV 載體，其包含編碼核酸酶之序列及指導該核酸酶在包含 PCSK9 基因之標靶細胞中之表現的調控序列；及

(b) 供體 AAV 載體，其包含轉基因及指導該轉基因在標靶細胞中之表現的調控序列，該供體載體進一步包含該轉基因匣之同源定向重組 (HDR) 臂 5' 及 3'。

【請求項 29】如請求項 28 之用途，其中 (a) 之該基因編輯 AAV 載體及 (b) 之該供體載體經由相同的途徑實質上被同時遞送。

【請求項 30】如請求項 28 或 29 之用途，其中 (a) 之該基因編輯 AAV 載體以約  $2 \times 10^{11}$  GC/mL 至約  $2 \times 10^{12}$  GC/mL 之濃度懸浮於注射用媒劑中。

【請求項 31】如請求項 28 或 29 之用途，其中 (a) 之該 AAV 靶向載體以約  $2 \times 10^{12}$  GC/mL 至約  $1 \times 10^{13}$  GC/mL 之濃度懸浮於注射用媒劑中。

【請求項 32】如請求項 28 至 31 中任一項之用途，其中該肝臟代謝性病症為鳥胺酸胺甲醯基轉移酶。

【請求項 33】如請求項 28 至 31 中任一項之用途，其中該肝臟代謝性病症為 OTC、FH、1 型瓜胺酸血症 (CTLN1) 或苯丙酮尿症。

【請求項 34】一種用於治療遺傳疾病之系統，該系統包含：

(a) 脂質奈米顆粒 (LNP)，其包含編碼核酸酶之 mRNA 序列；及

(b)供體 AAV 載體，其包含轉基因及指導其在標靶細胞中之表現的調控序列，該供體載體進一步包含該轉基因之同源定向重組(HDR)臂 5'及 3'。

【請求項 35】如請求項 34 之系統，其中該核酸酶靶向該 PCSK9 基因。

【請求項 36】如請求項 34 之系統，其中該核酸酶靶向 PCSK9 外顯子 7。

【請求項 37】如請求項 34 之系統，其中該核酸酶為對於 PCSK9 特異性的巨型核酸酶。

【請求項 38】如請求項 37 之系統，其中該巨型核酸酶為 ARCUS 巨型核酸酶。

【請求項 39】如請求項 34 之系統，其中該核酸酶為 Cas9 核酸酶，且其中該 LNP 包含 sgRNA。

【請求項 40】如請求項 39 之系統，其中該 Cas9 核酸酶的兩側是核定位訊號。

【請求項 41】如請求項 39 或 40 之系統，其中該 sgRNA 包含至少 20 個核苷酸，其特異性結合該 PCSK9 基因中的標靶位點，該標靶位點位於被該 Cas9 特異性識別的前間隔序列相鄰基序(PAM)的 5'處。

【請求項 42】如請求項 34 至 41 中任一項之系統，其進一步包含 RNA 聚合酶啟動子。

【請求項 43】如請求項 42 之系統，其中該 RNA 聚合酶啟動子為 U6 啟動子。

【請求項 44】如請求項 43 之系統，其中該 U6 啟動子位於該 sgRNA 之 5'處。

【請求項 45】如請求項 34 至 44 中任一項之系統，其中該 sgRNA 與該標靶位點序列 100%互補。

【請求項 46】如請求項 34 至 44 中任一項之系統，其中該 sgRNA 與該標靶位點序列小於 100%互補。

【請求項 47】如請求項 34 至 46 中任一項之系統，其中該轉基因為肝表現的基因。

【請求項 48】如請求項 34 至 47 中任一項之系統，其中該轉基因選自 OTC、PKU、CTLN1 及 FH。

【請求項 49】一種用於治療遺傳疾病之系統，該系統包含：

(a) 基因編輯載體，其包含編碼核酸酶的核酸序列；及

(b) 供體載體，其包含編碼用於從 PCSK9 基因座表現之外源產物的核酸序列，其中插入的核酸序列並不編碼 PCSK9，

其中該系統進一步包含指導該核酸酶特異性地靶向天然 PCSK9 基因座的序列；且

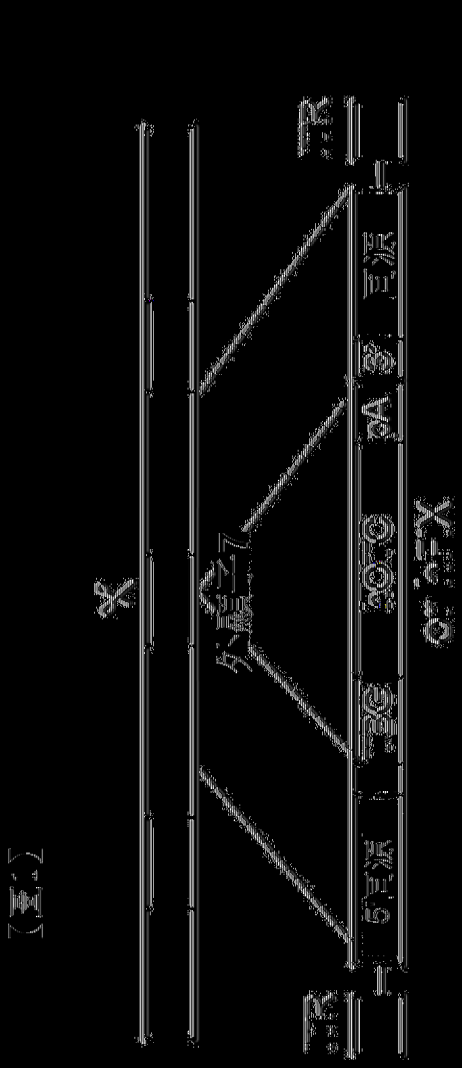
其中在以雙重載體系統投藥後，標靶細胞中的天然 PCSK9 可選擇地消融或減少。

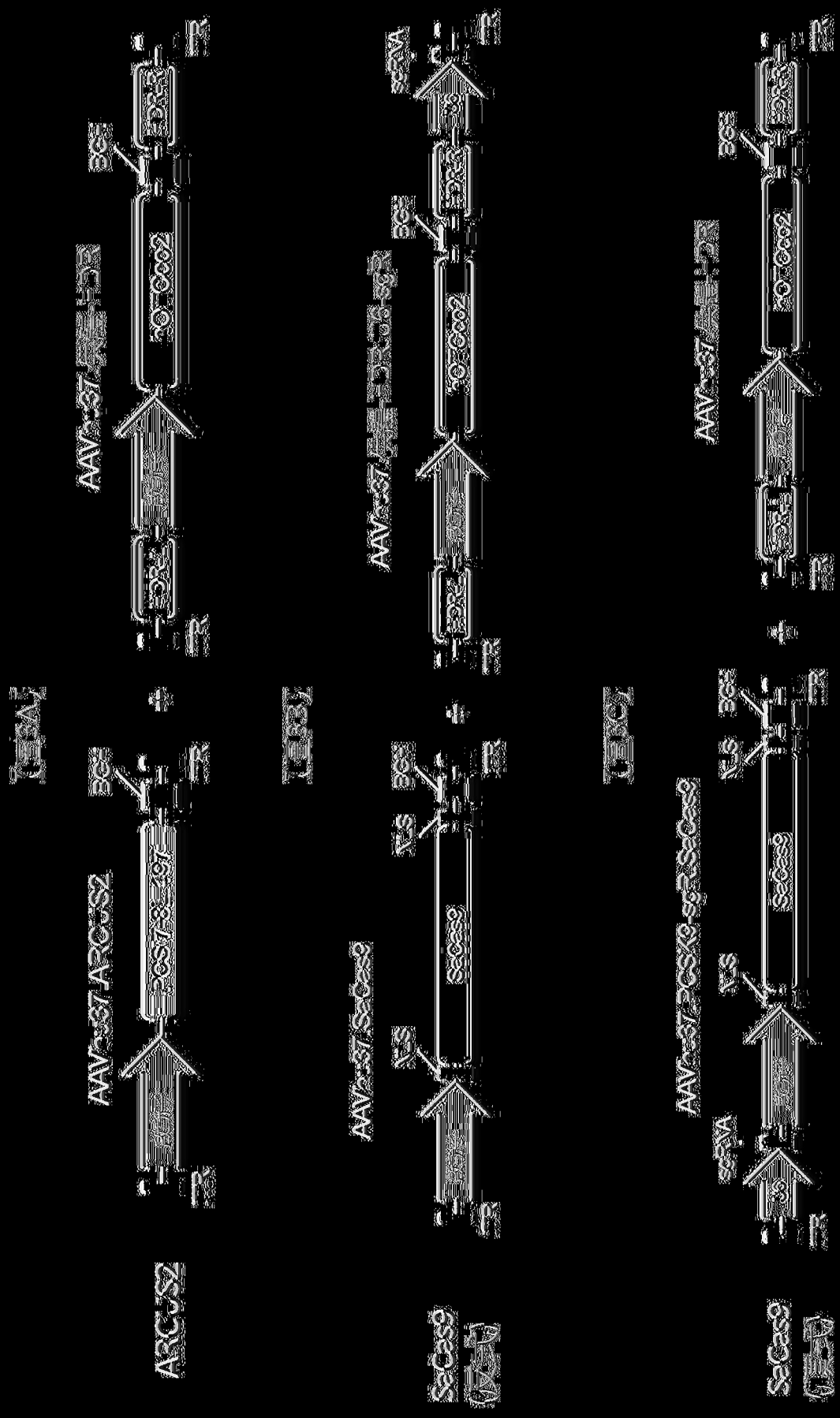
【請求項 50】一種表現匣，其包含 SEQ ID NO: 17 之工程化編碼序列或與其共享至少 90%同一性的序列。

【請求項 51】如請求項 50 之表現匣，其進一步包含 AAV 5' ITR 及 3' ITR。

【請求項 52】一種 AAV 載體，其包含如請求項 50 或 51 之表現匣。

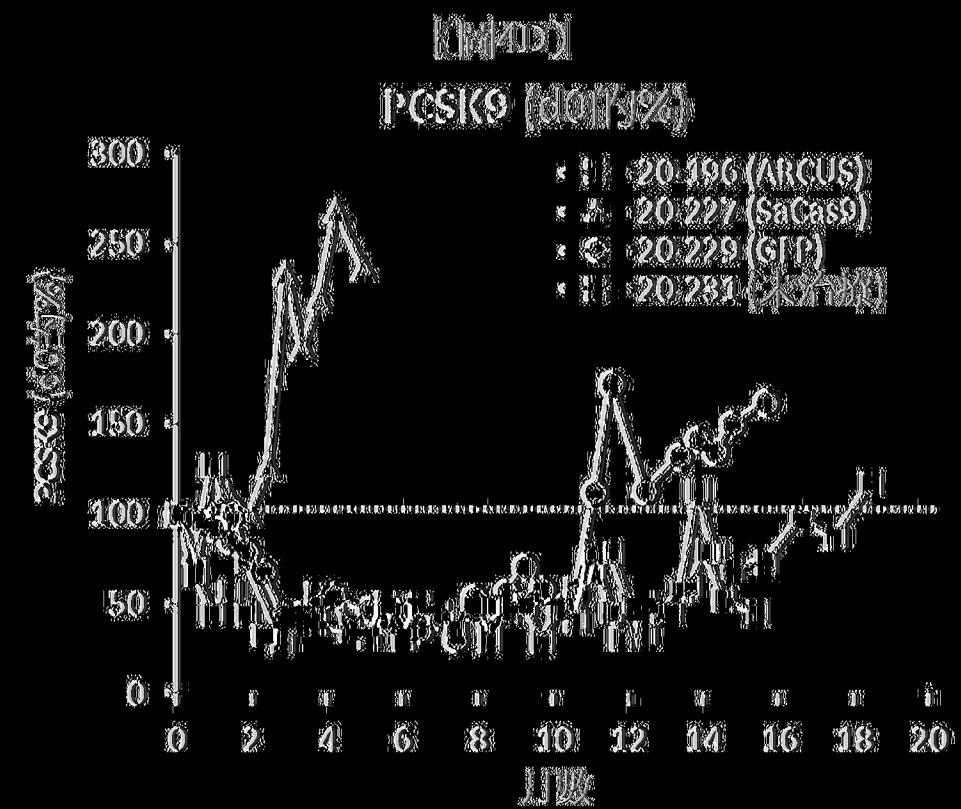
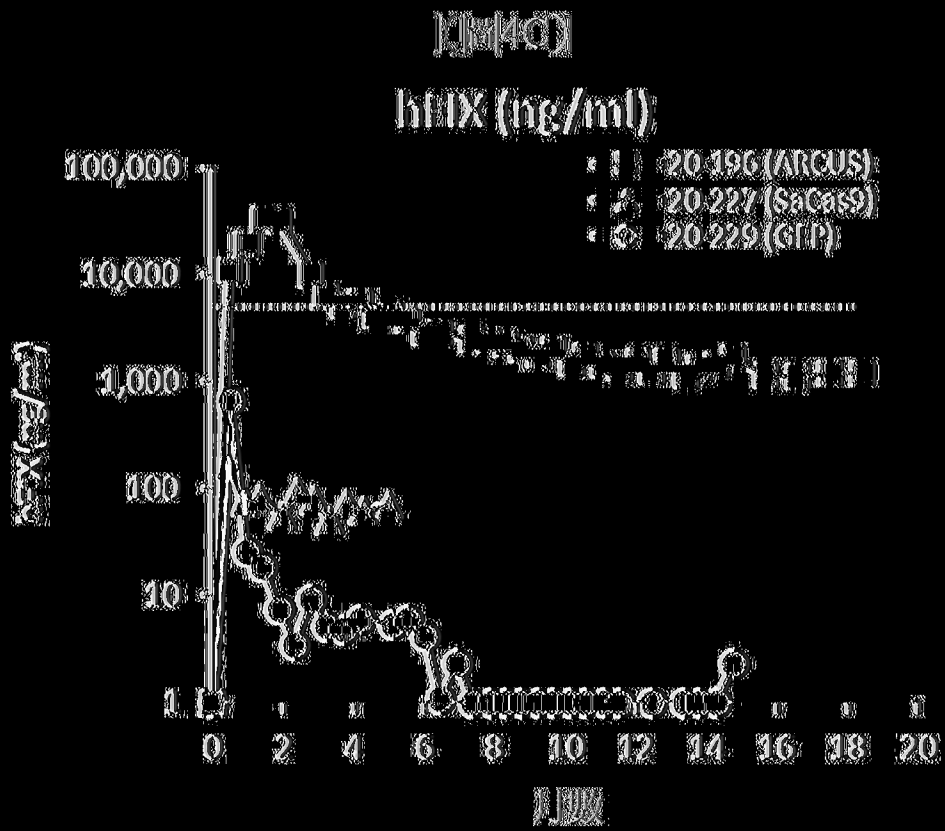
(發明圖式)

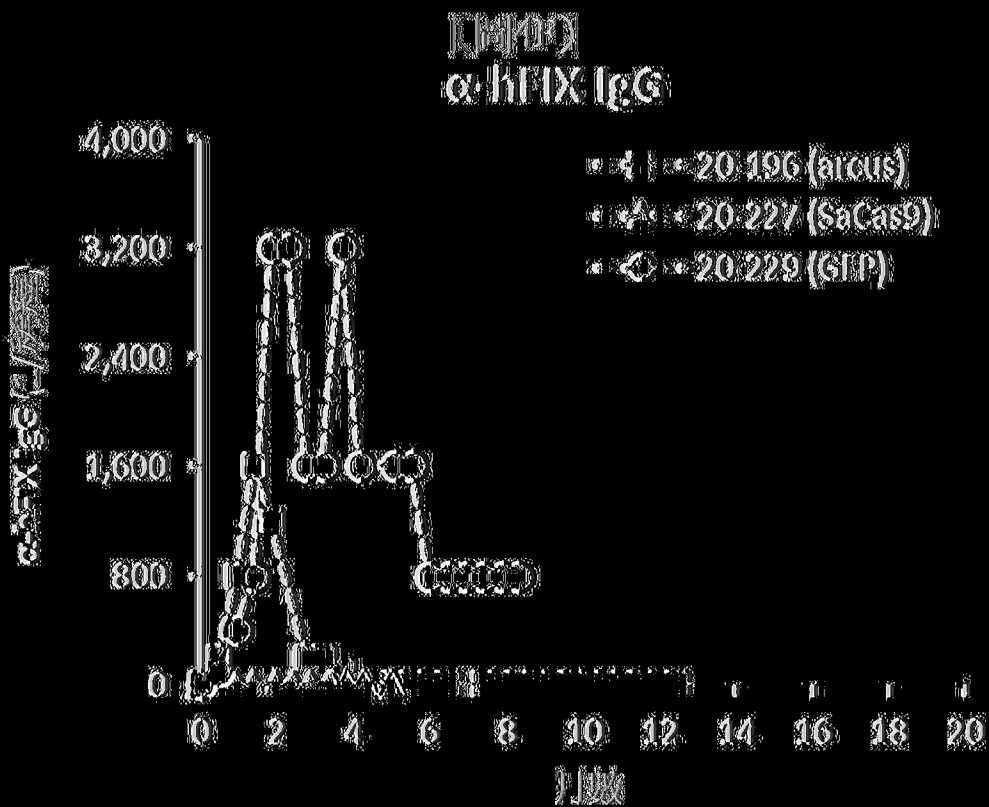
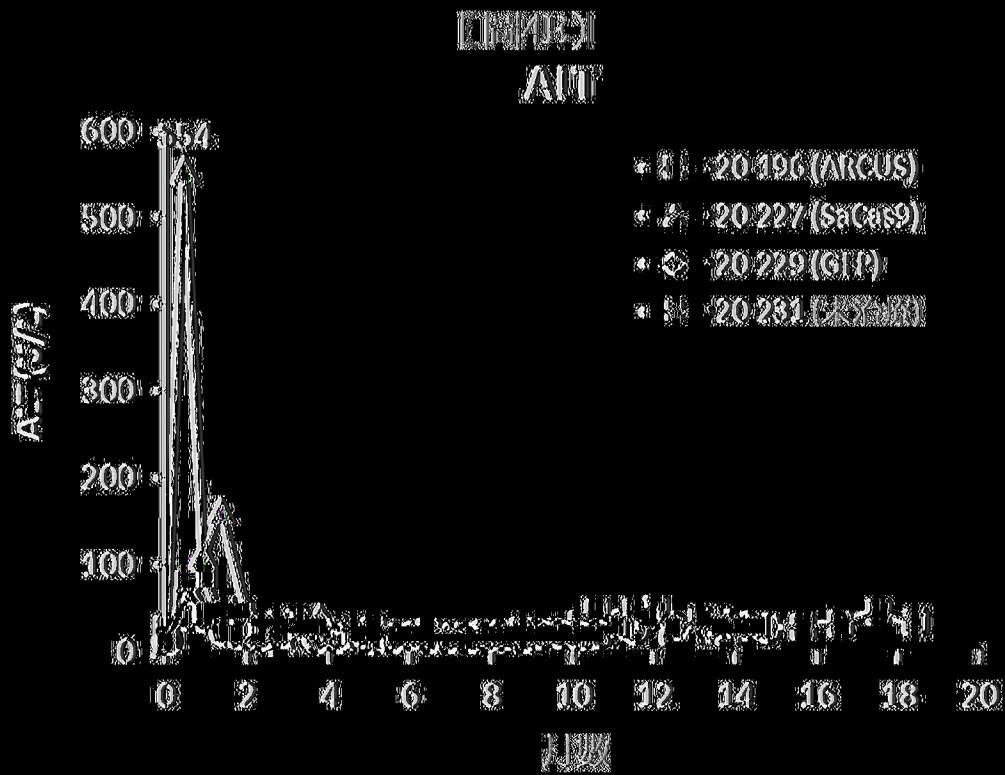


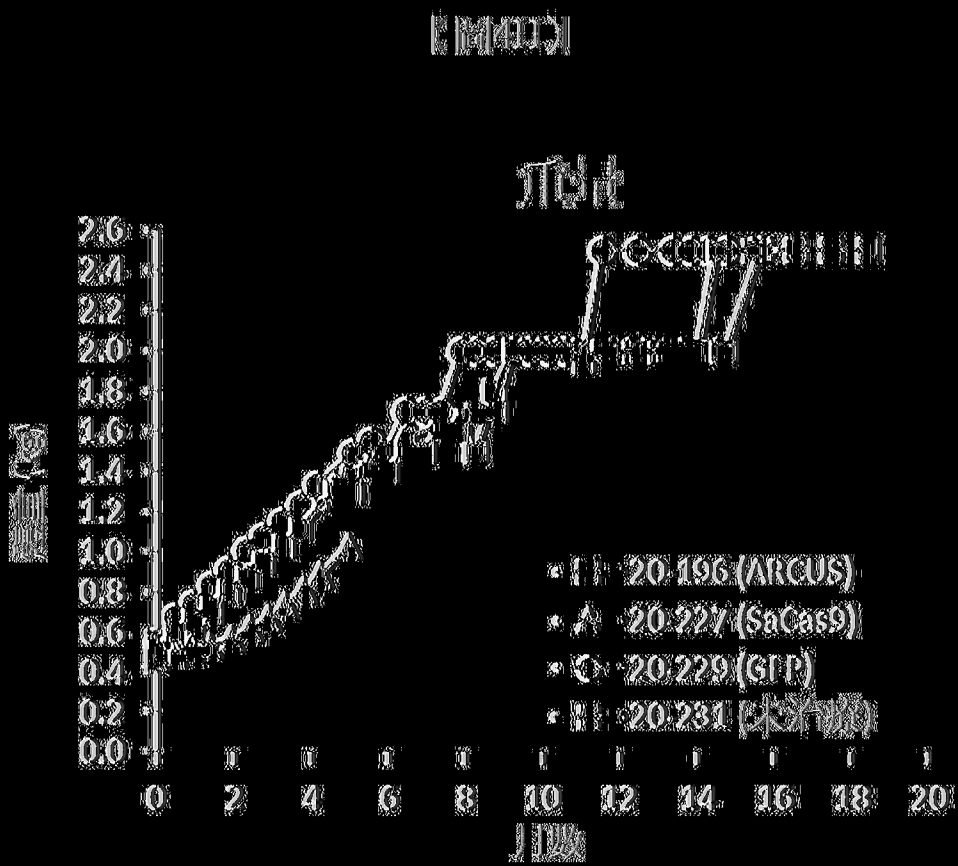
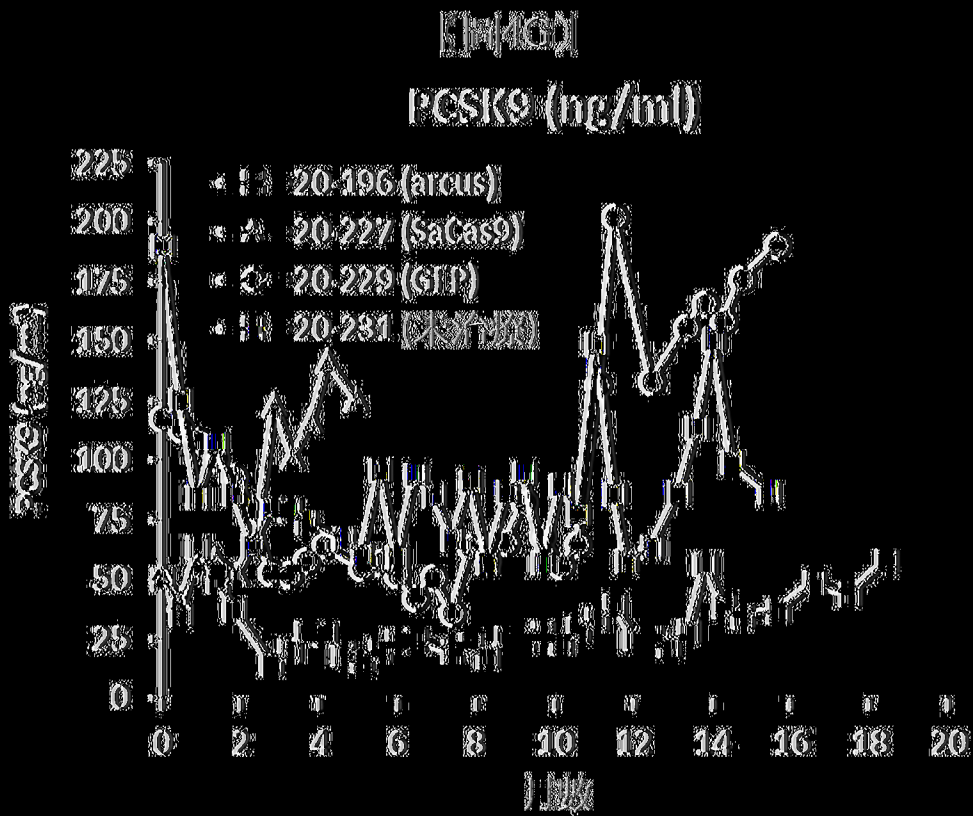


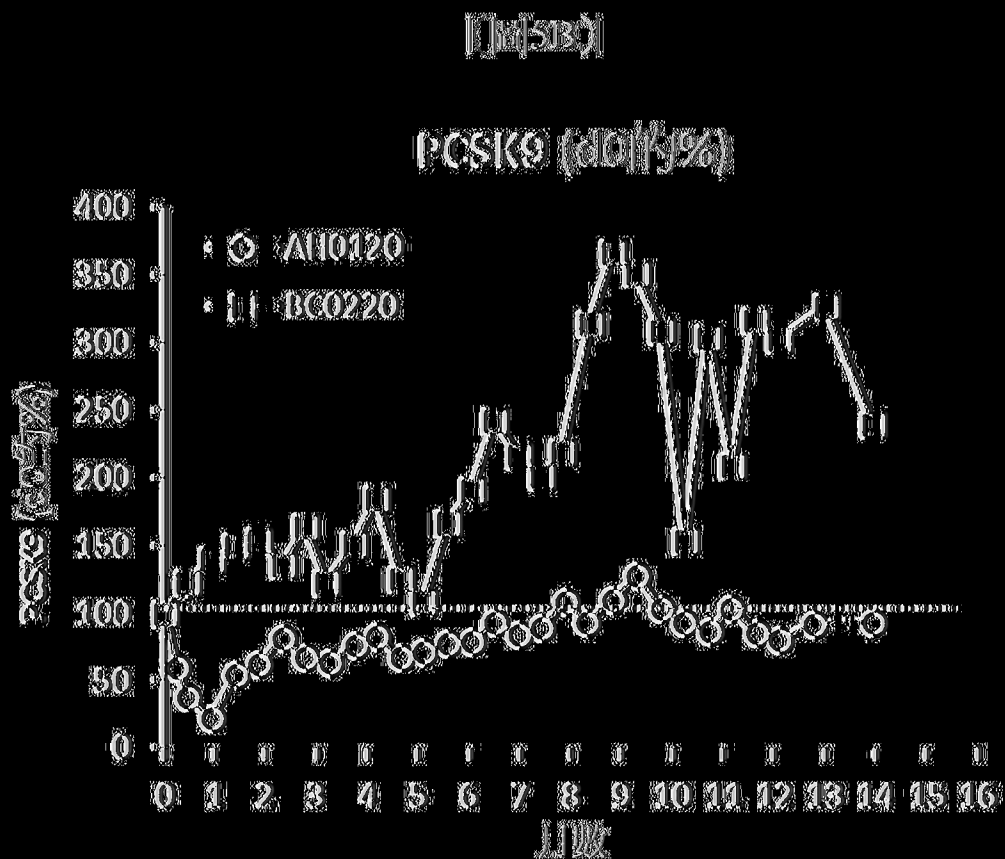
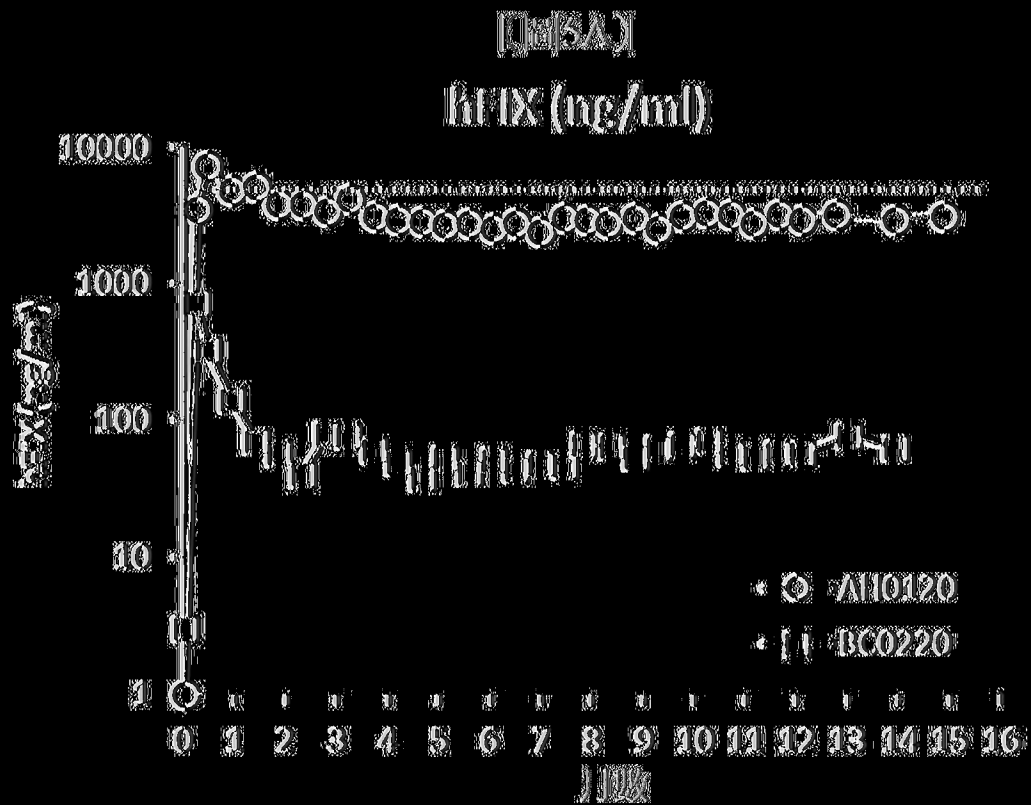






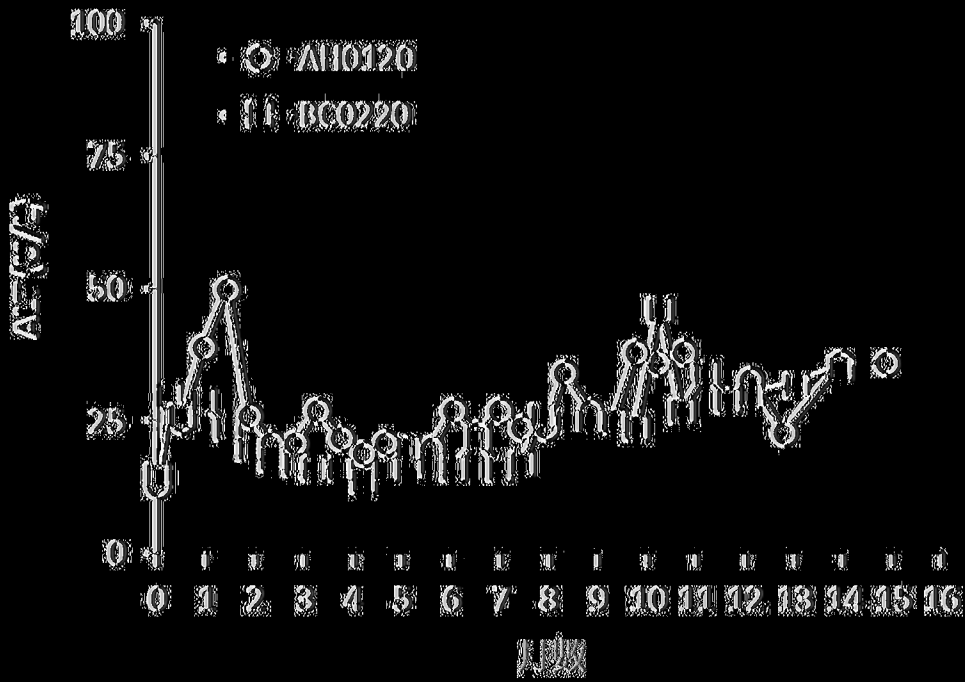






KIM5C3

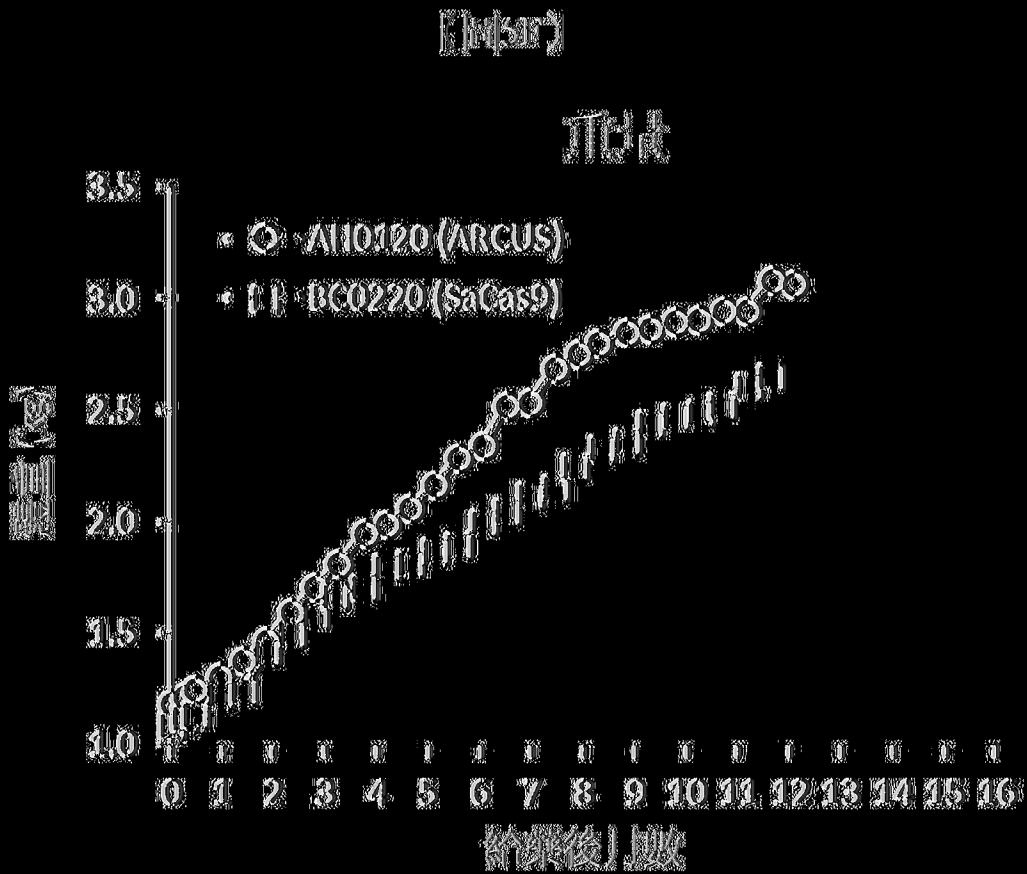
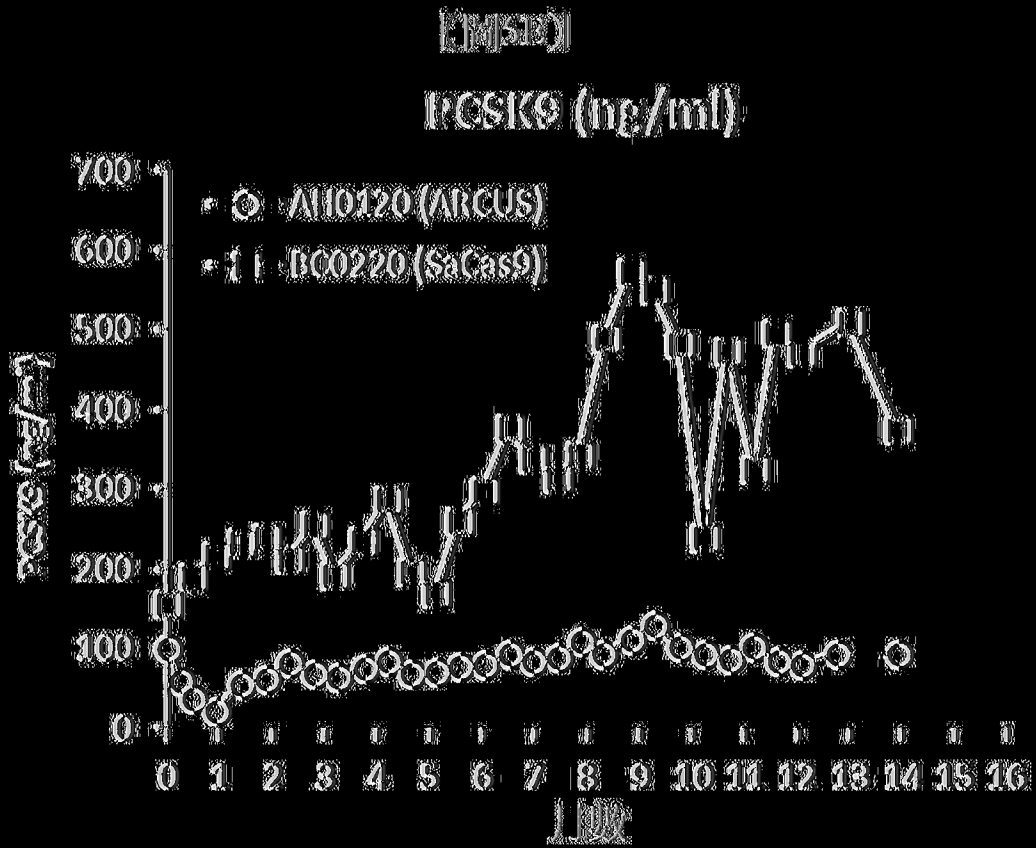
ALT



KIM5T3C3

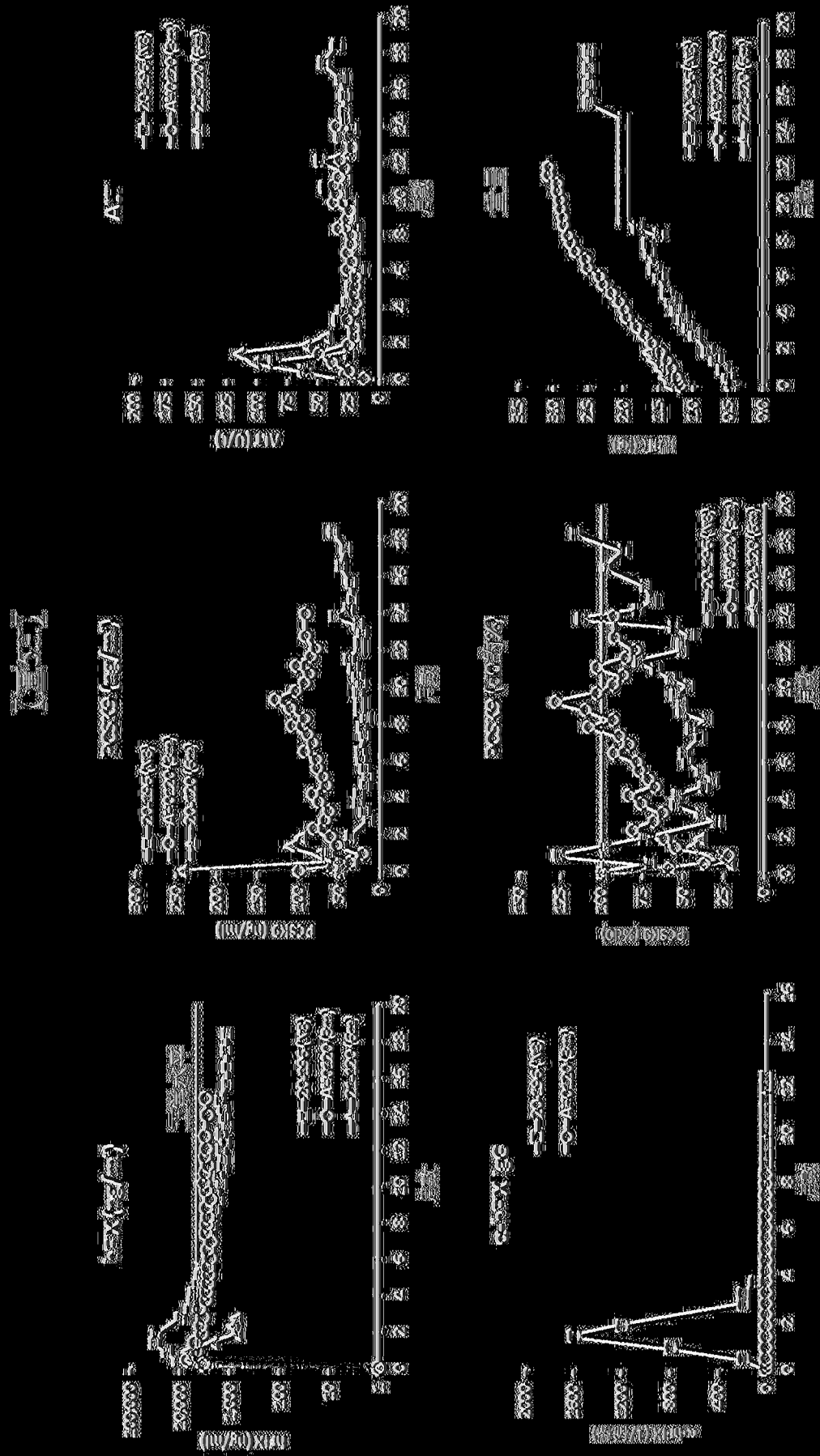
α-hMIX IgG





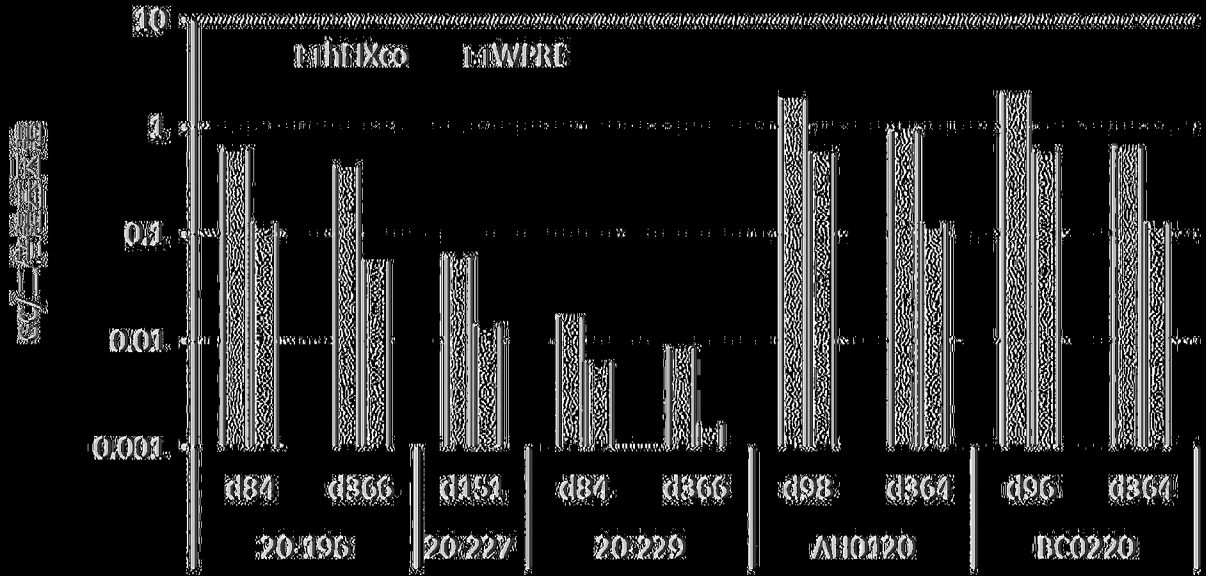
TEST

試料番号	試験機	試験片形状	試験片寸法	試験片重量	試験片面積	試験片厚さ	試験片容積	試験片密度	試験片吸水率	試験片吸水率標準偏差	試験片吸水率標準偏差率
20-220	ARGUS	2	37.749 (37)	1.239 (350)	2.23	0.20	0.447	1.239	2.1	0.03	2.1%
20-220	ARGUS	3	37.749 (37)	1.239 (350)	2.13	0.12	0.252	1.239	2.1	0.03	2.1%
20-220	ARGUS	3	37.749 (37)	1.239 (350)	2.07	0.08	0.166	1.239	2.1	0.03	2.1%
20-220	ARGUS	3	37.749 (37)	1.239 (350)	2.05	0.05	0.103	1.239	2.1	0.03	2.1%
20-220	ARGUS	3	37.749 (37)	1.239 (350)	2.00	0.00	0.000	1.239	2.1	0.03	2.1%
20-227	SECSO	2	30 (35)	88 (354)	0.8	0.02	0.016	88	0.2	0.01	0.2%
20-227	SECSO	3	30 (35)	88 (354)	1.6	0.02	0.032	88	0.2	0.01	0.2%
20-227	SECSO	3	30 (35)	88 (354)	0.07	0.00	0.000	88	0.2	0.01	0.2%
20-229	SECSO	2	37.749 (37)	1.239 (350)	0.1	0.00	0.000	1.239	0.0	0.00	0.0%
20-229	SECSO	3	37.749 (37)	1.239 (350)	0.0	0.00	0.000	1.239	0.0	0.00	0.0%



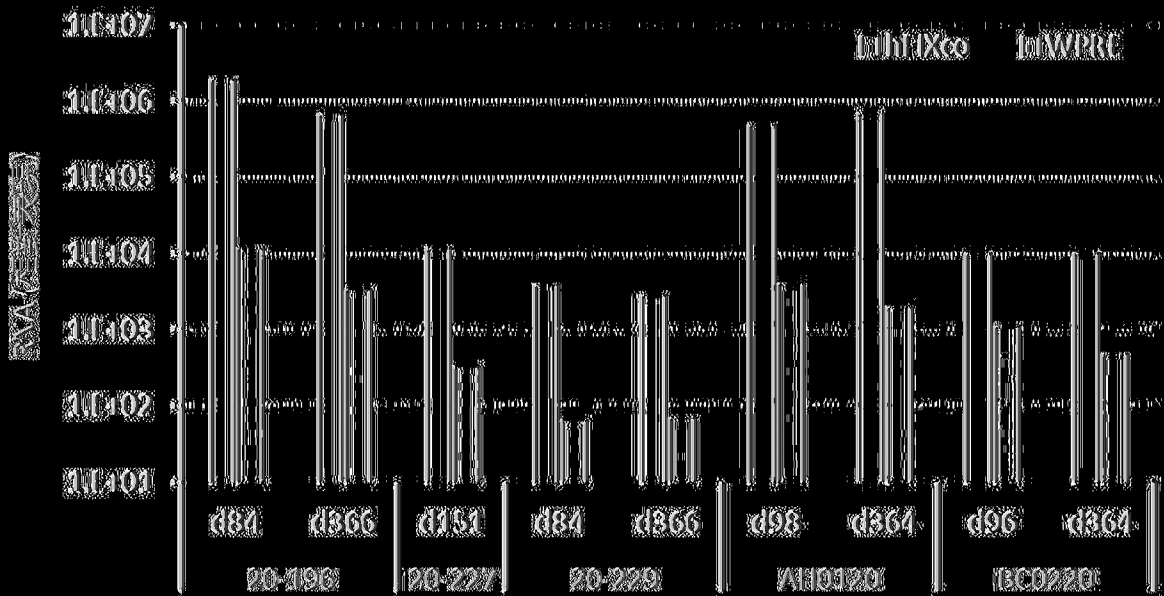
【圖表】

【圖表】之【圖表】



【圖表】

【圖表】中之【圖表】



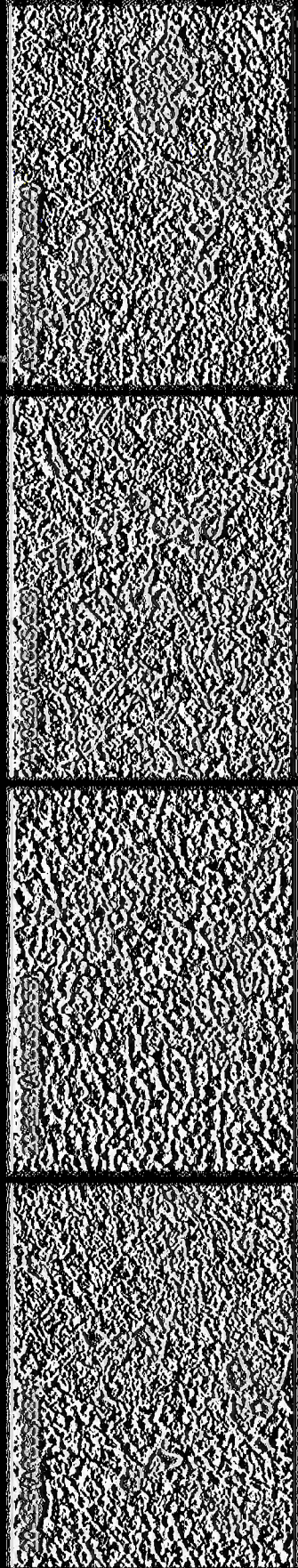
1107

行旅中定时的列车

1108

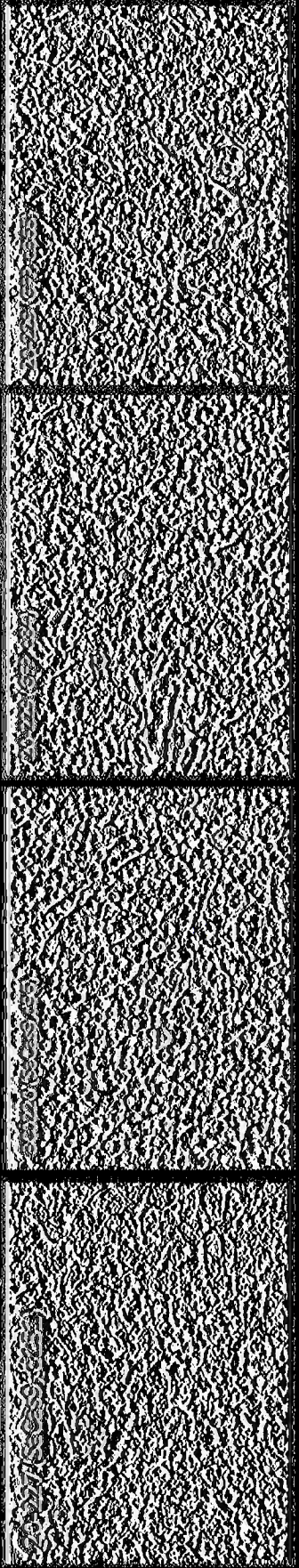
1109

1110



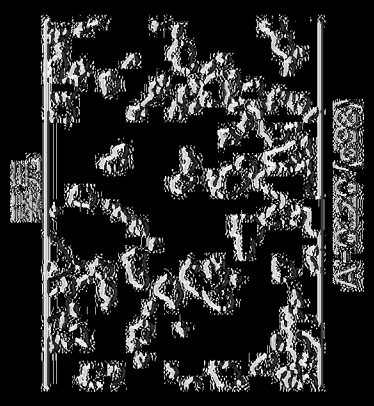
1111

1112

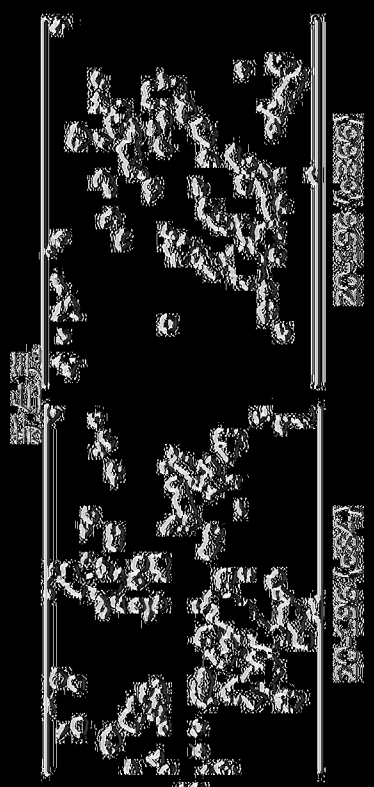


ARCUS

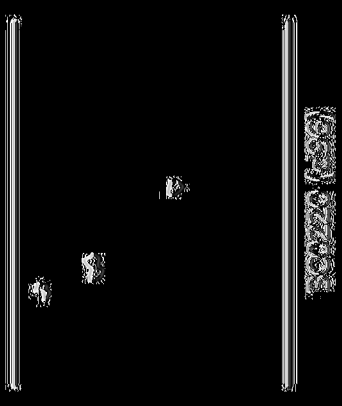
ARCUS



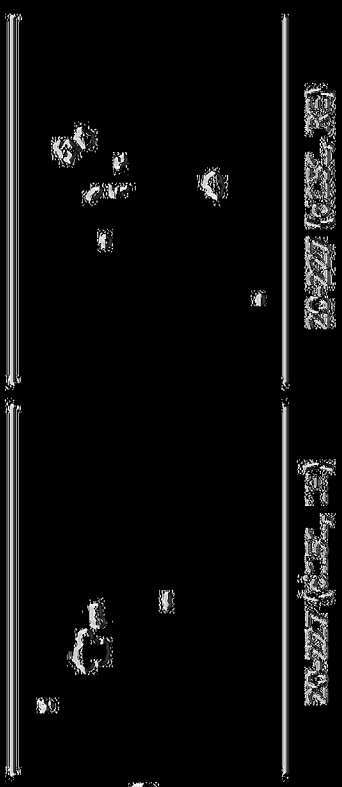
ARCUS



ARCUS



ARCUS



ARCUS

ARCUS



ARCUS

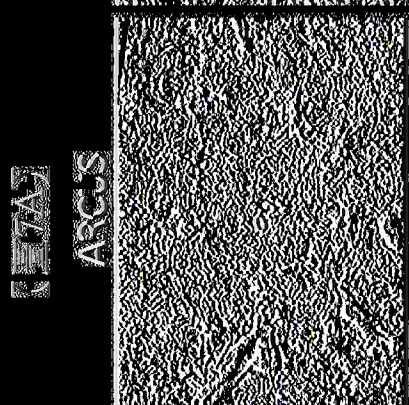
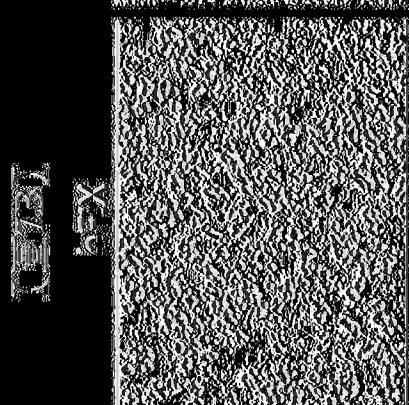
ARCUS

ARCUS

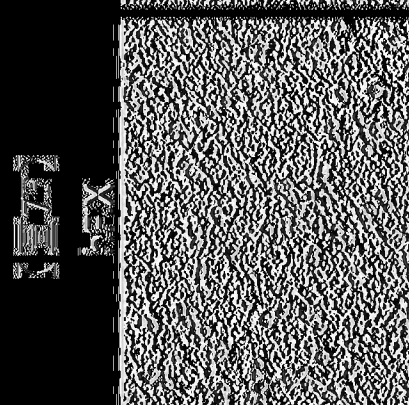
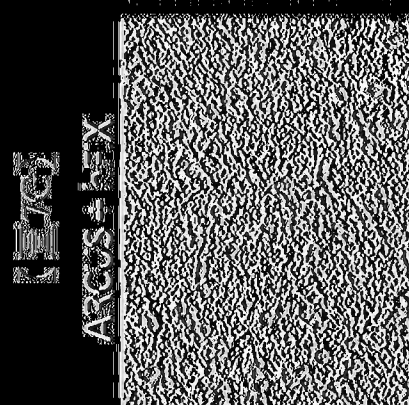
ARCUS

ARCUS

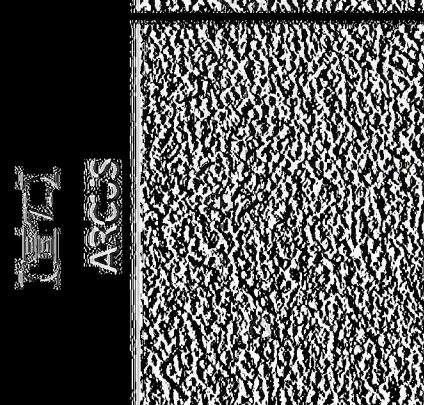
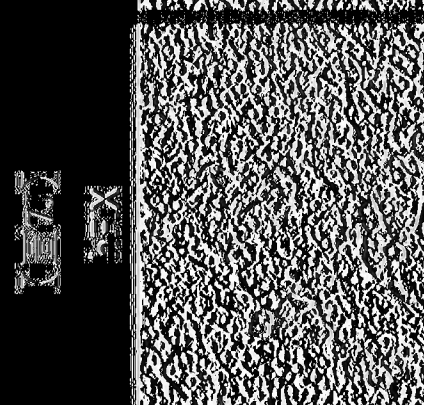




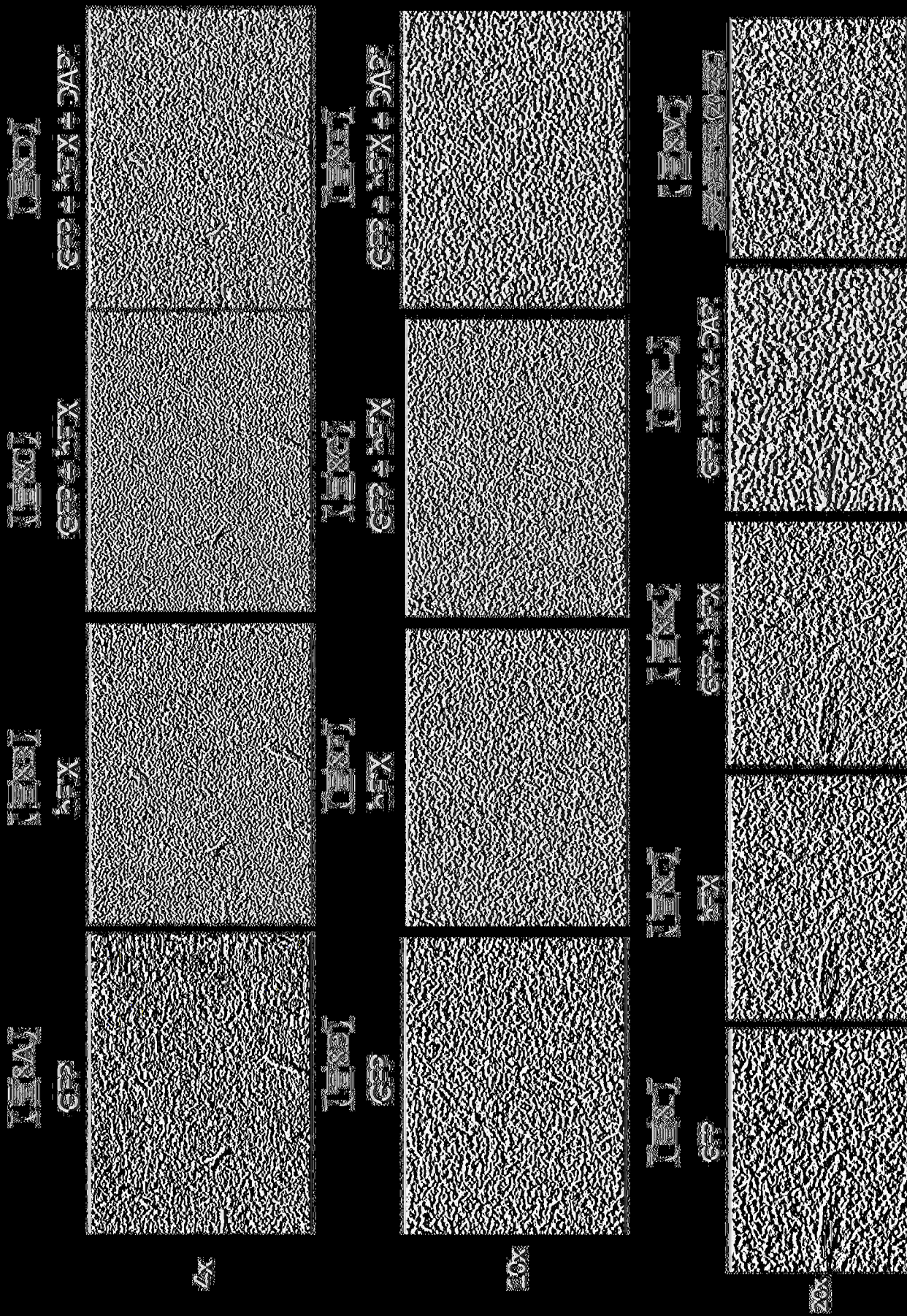
4x



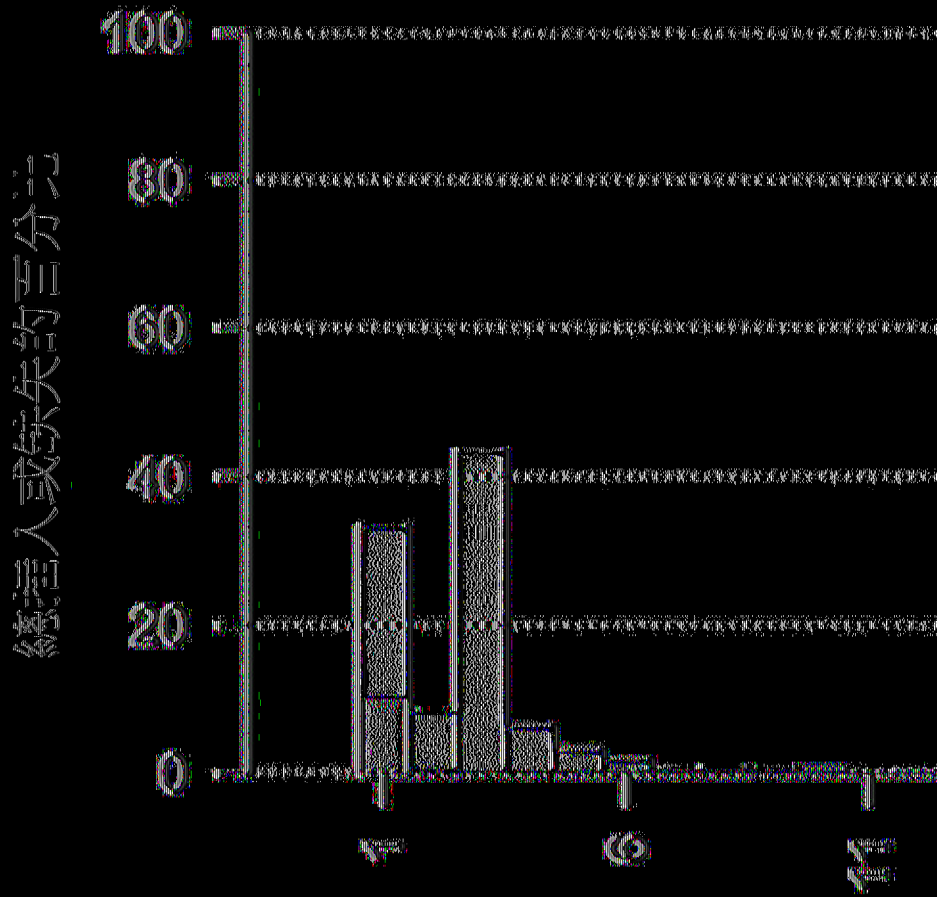
20x



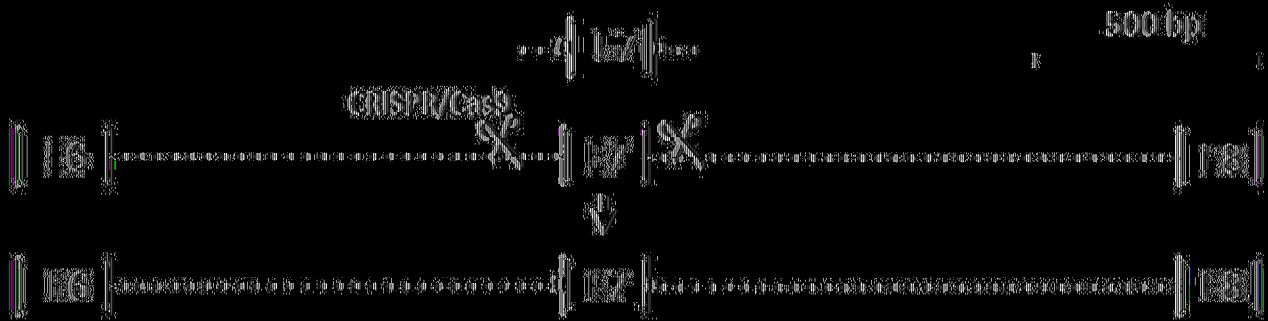
20x



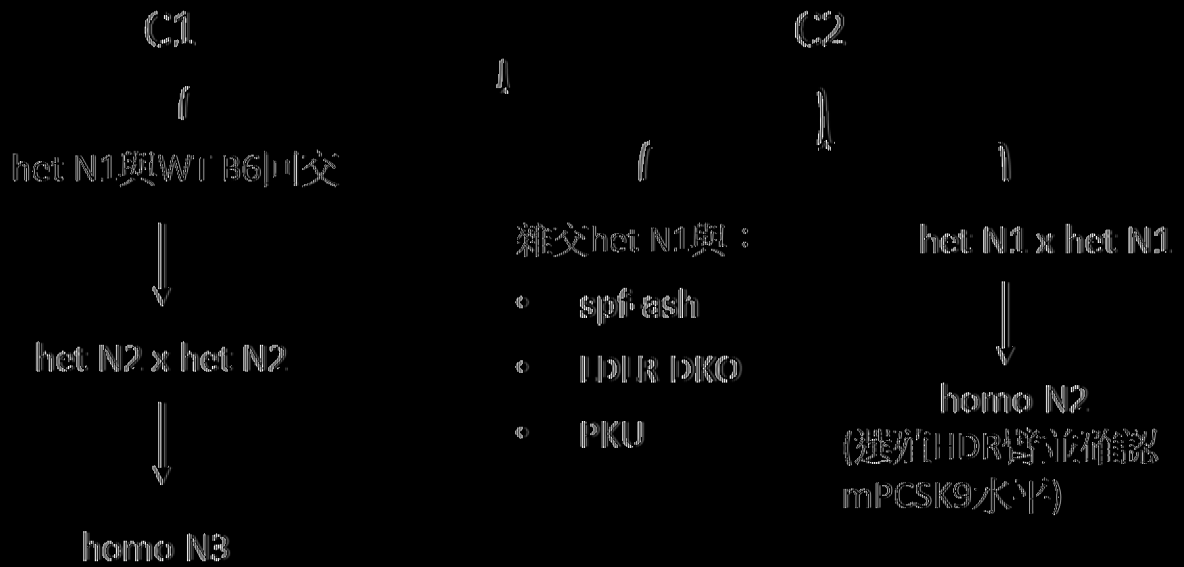
(附9)



(圖)10A)



(圖)10B)



[表1A]

工区	高脚子 (核種號碼)	土質區	標柱	新物 ID	位置	採取時 台数 (天)	歩AAV 3A3 万圓 (1日あたり)	台数 (天)	採集率 (%)	採集子数 (個)	OT位點 編號	
												採集率 (%)
長	-	AAV6-79	G1	21-000	東	5	50	5	0.4	0.6	0.6	0
				21-006	西	1	50	5	2.0	0.7	0.05	0.0
				21-005	東	7	50	5	2.3	1.5	1.5	0.0
				21-012	西	25	1000	0	0	0	0.5	0.0
				21-013	西	35	400	0	0	0	0.1	0.0
短	T96	AAV6-79	G4	21-020	東	11	1000	0	0	0.1	0.0	
				21-021	西	11	1000	0	0	0.1	0.0	
				21-022	東	7	50	0	0	0.1	0.0	
				21-023	西	7	50	0	0	0.1	0.0	
				21-024	東	9	25	0	0	0.1	0.0	
長	F113	AAV6-79	G7	21-053	西	10	1000	1.9	2.6	0.9	0.0	
				21-055	西	9	25	1.6	2.6	0.7	0.1	
				21-056	東	9	100	4.5	4.0	4.5	0.0	
				21-071	西	9	50	3.7	2.9	2.9	2.0	
				21-097	東	6	25	1.6	1.6	1.6	1.6	
	T96	AAV6-79	G6	21-073	東	11	1000	3.0	3.5	3.5	0.0	

圖 1

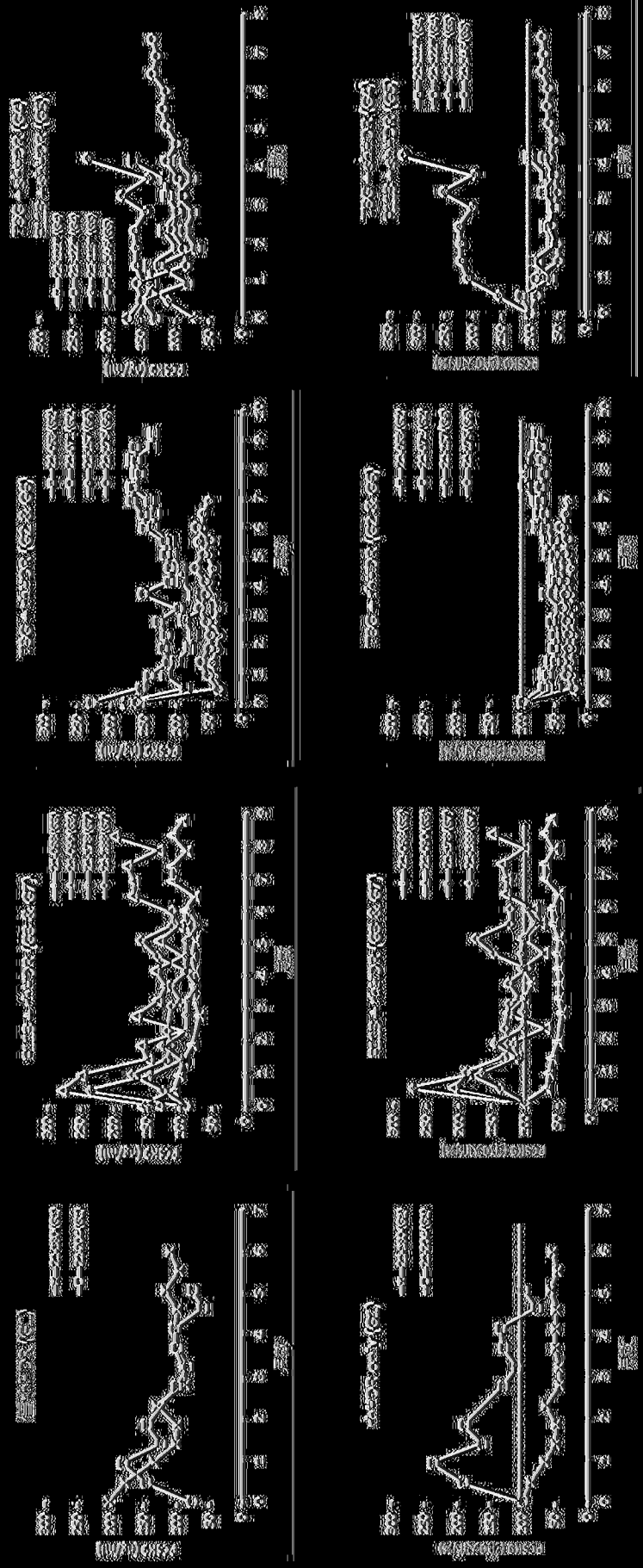
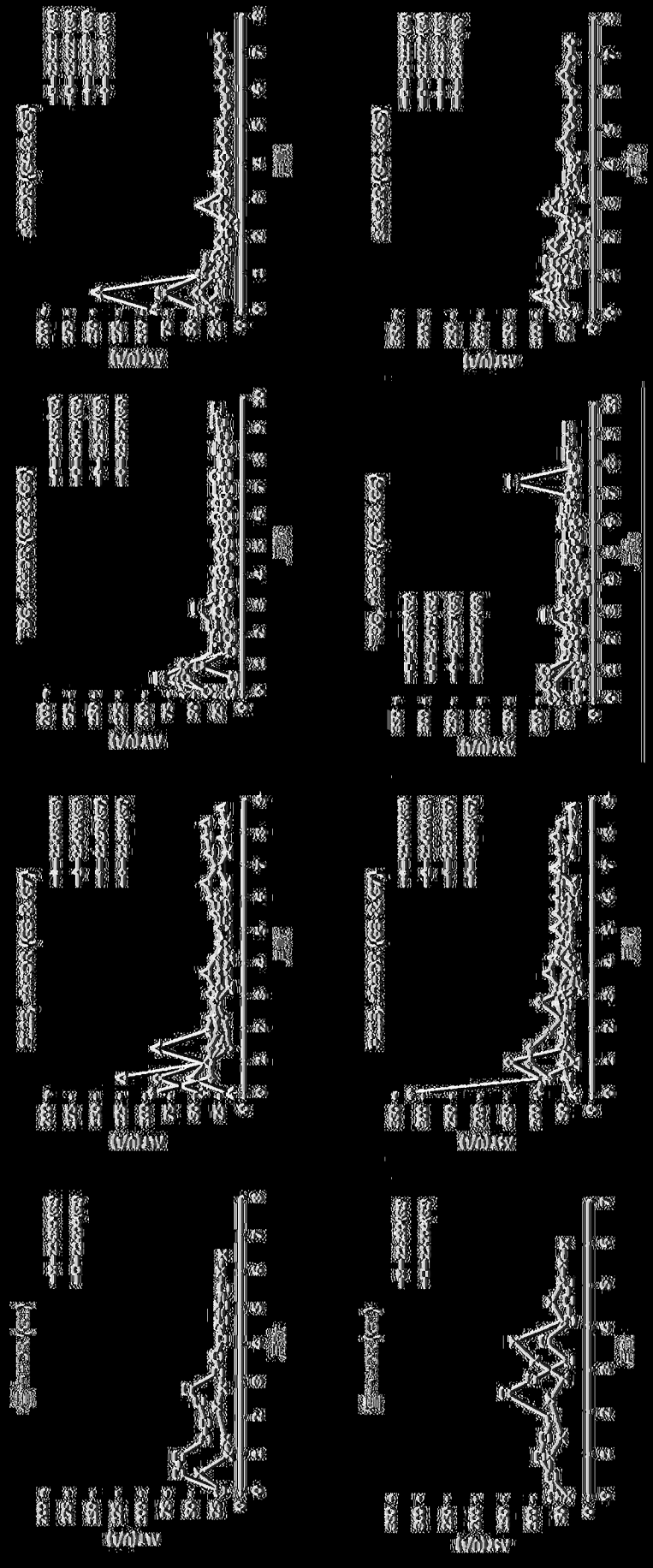


圖 10



# 【資料】

合計 14.4%

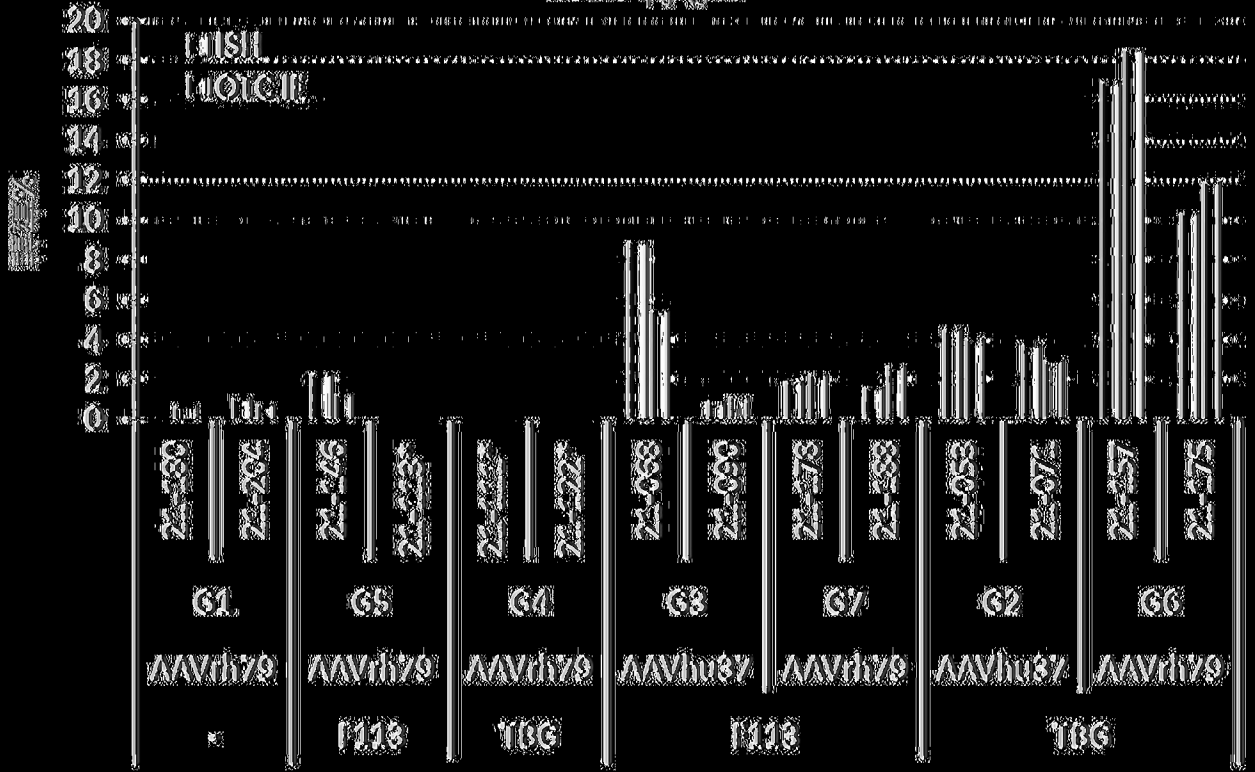
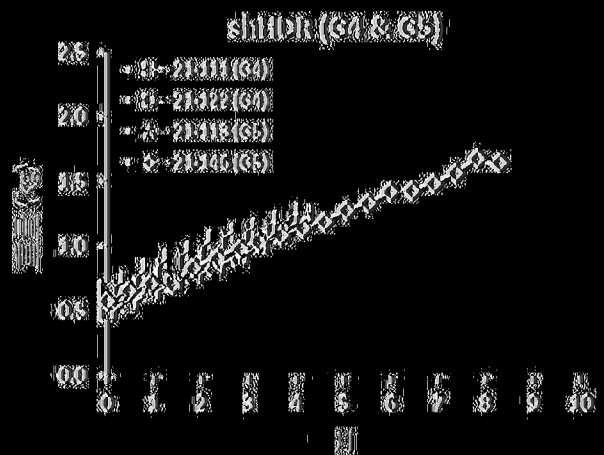
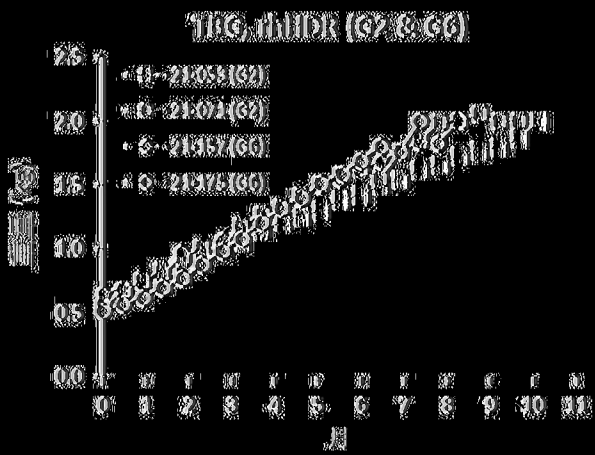
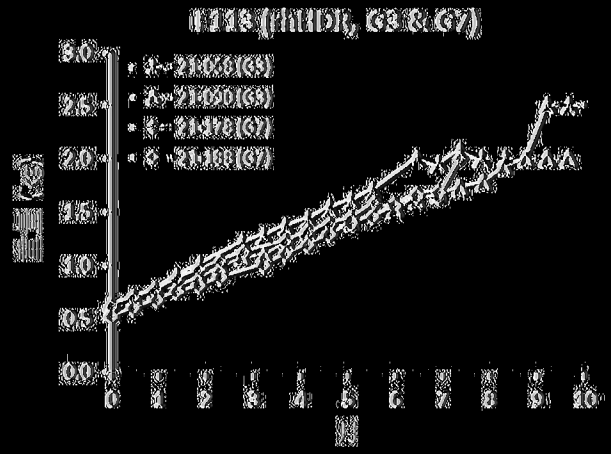
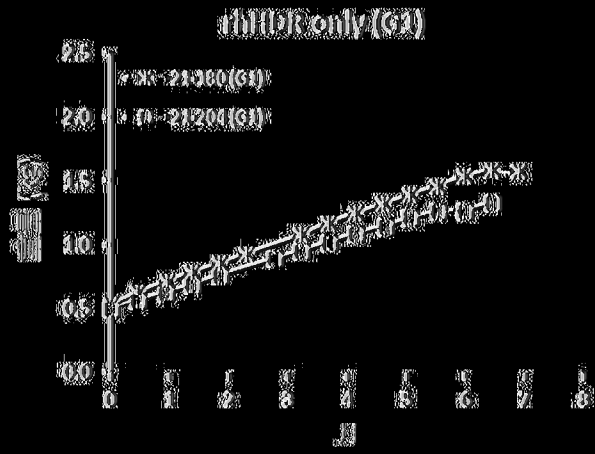
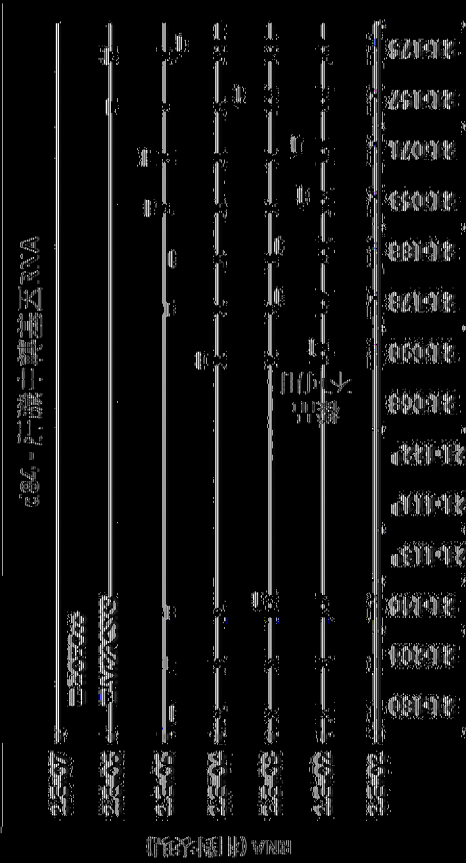


圖 10 試驗結果



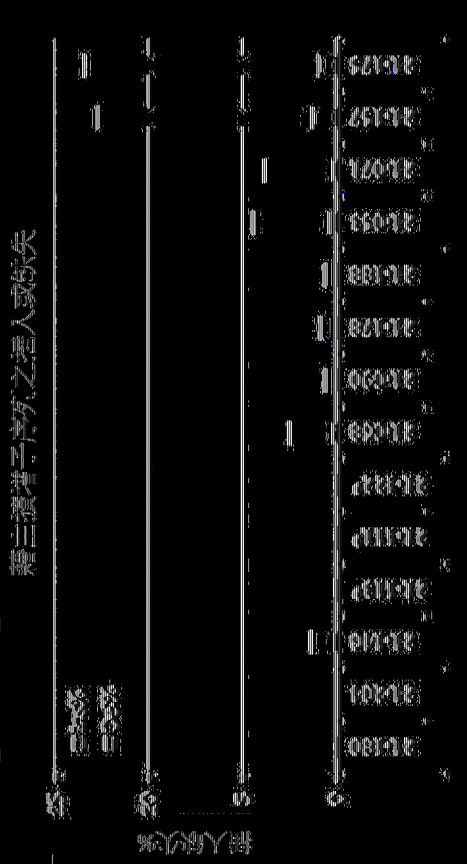
(圖 16)



(圖 17)



(圖 18)



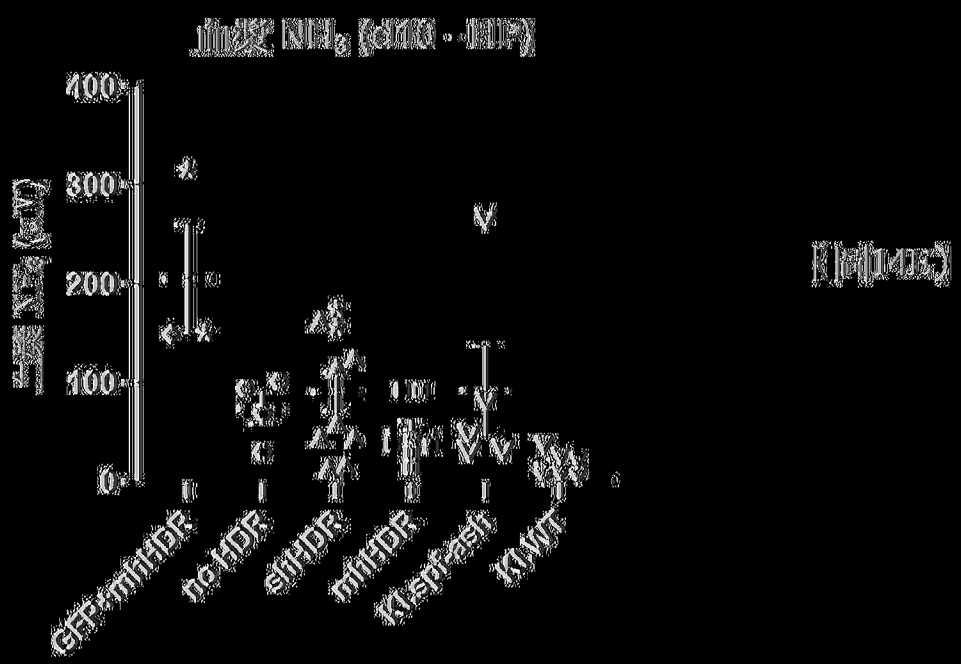
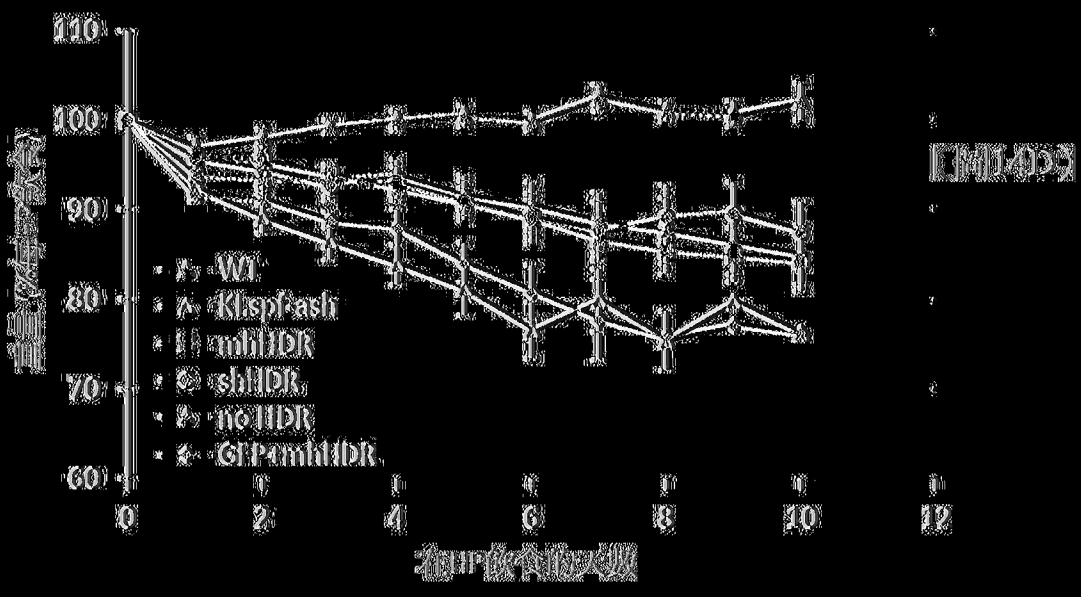
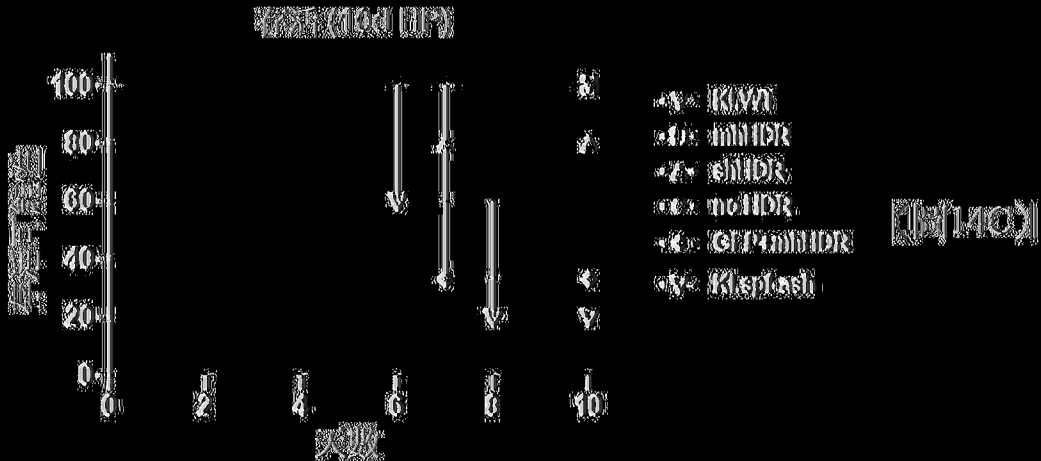












7:40Z

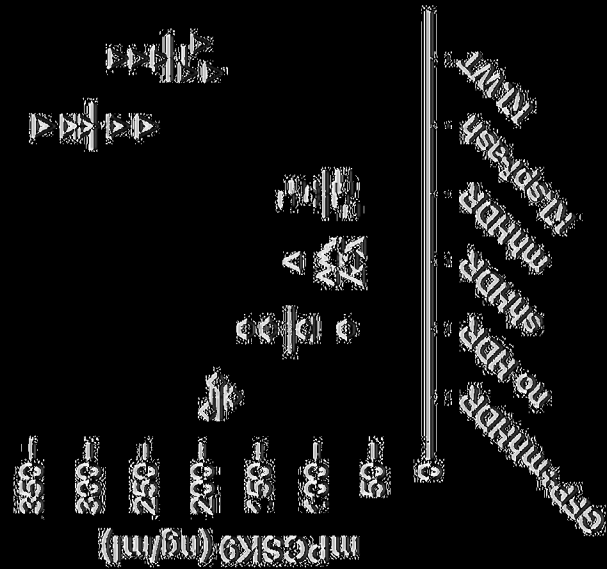
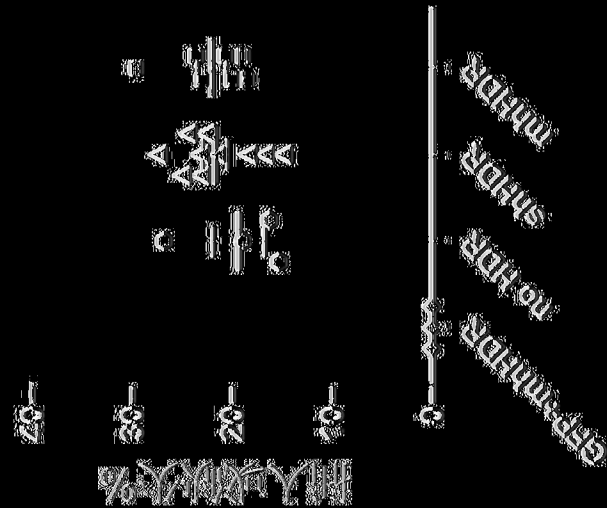
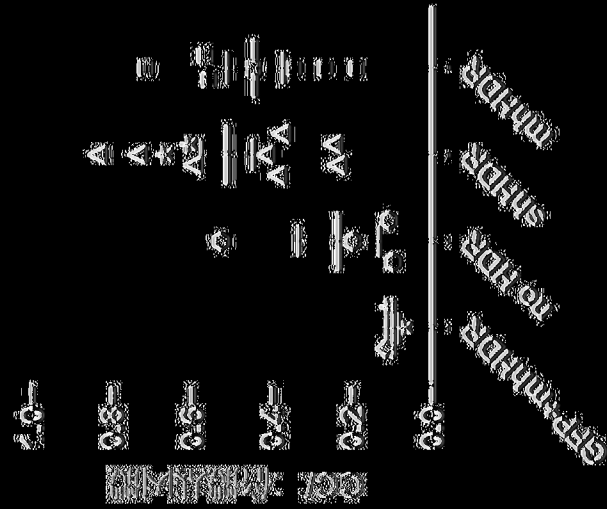
7:40Z

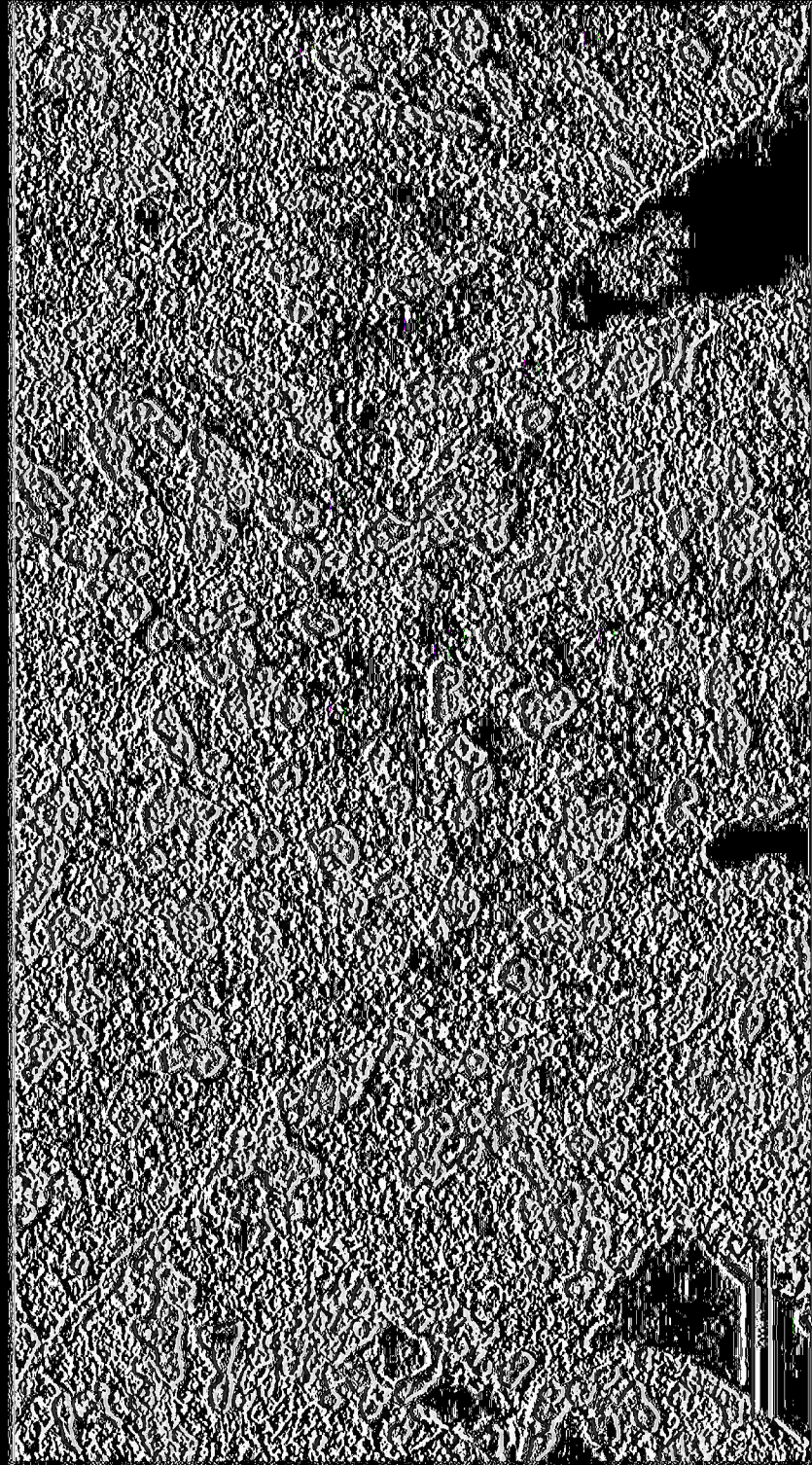
7:47Z

000000 (659)

無人観測 (076)

mPSC (679)

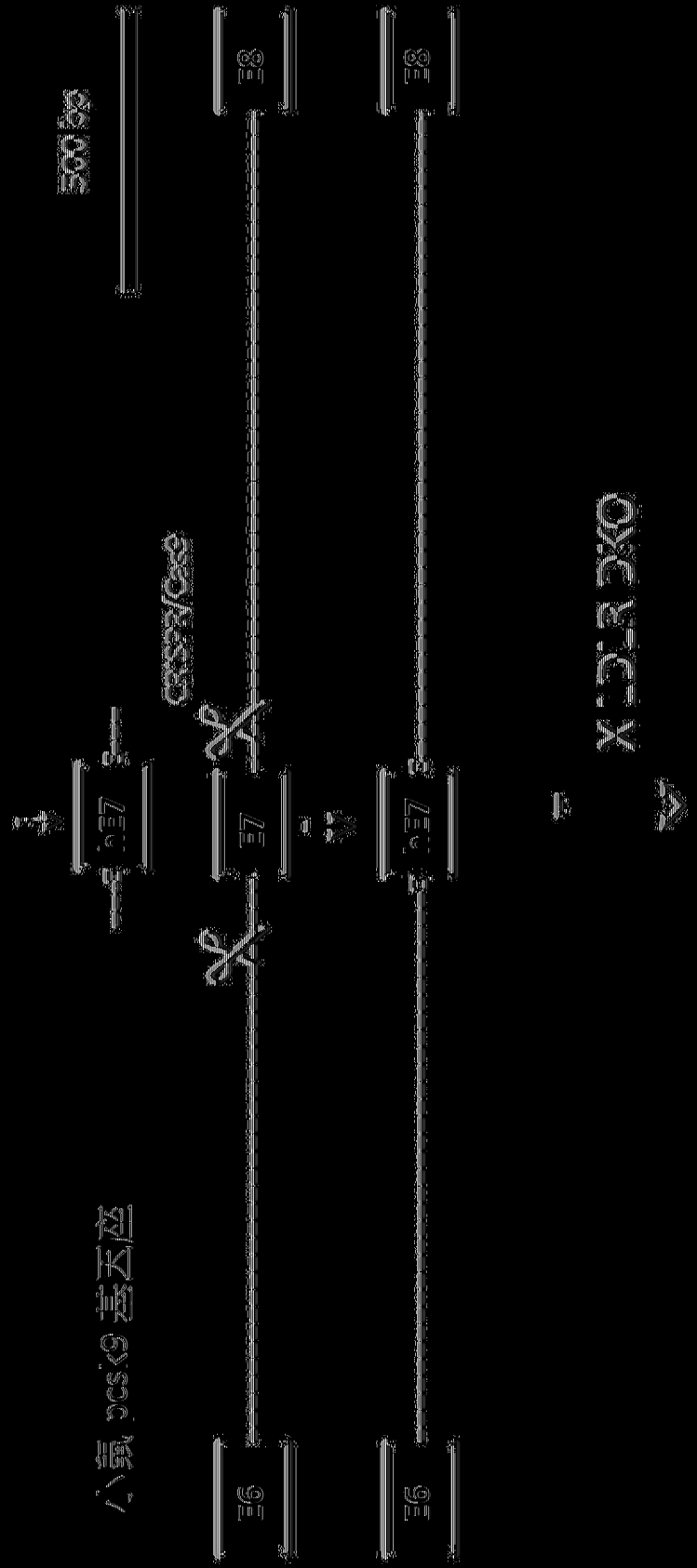




1111

[圖 5]

依據：265 bp 的人類 PCSK9 序列 (40 bp + 184 bp (hE7) + 41 bp)  
圖為小氣泡全序列的 HDR 變

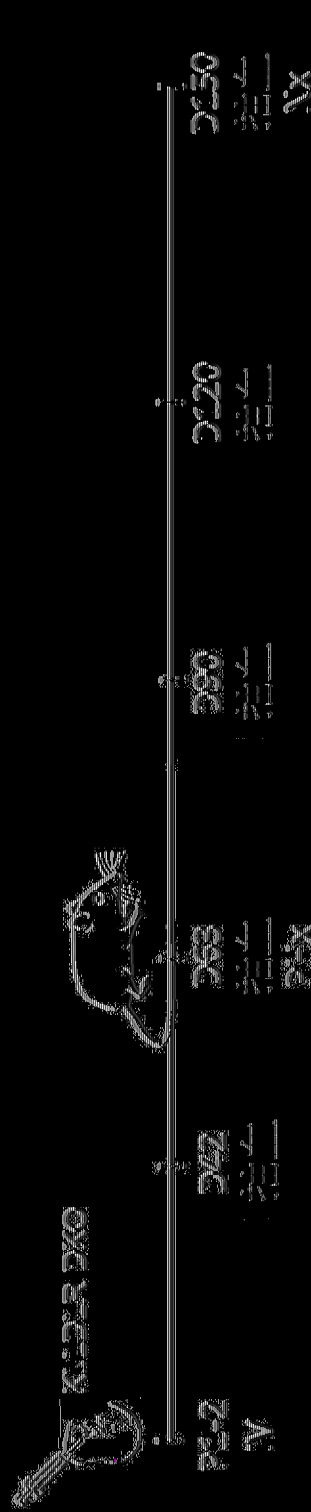


hE7<sup>KI</sup>/hE7<sup>KI</sup>-csg/csg-apobec1/apobec1 (hE7 疾病模型)



(圖 7)

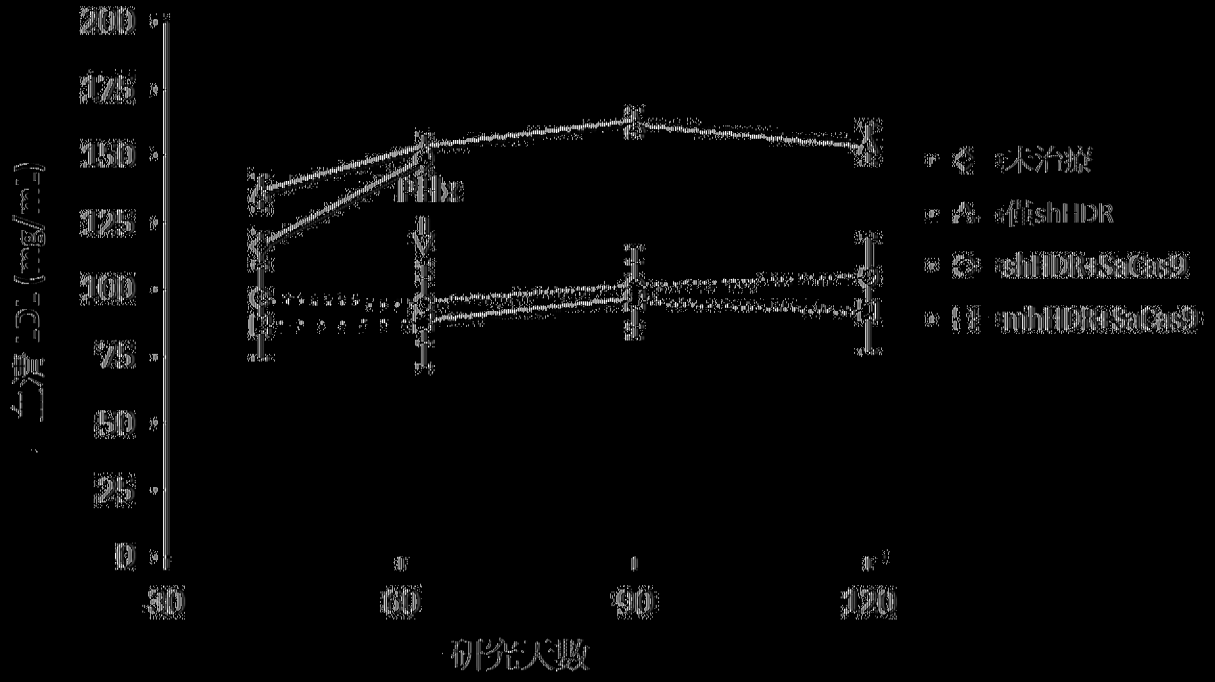
WS448-研究設計



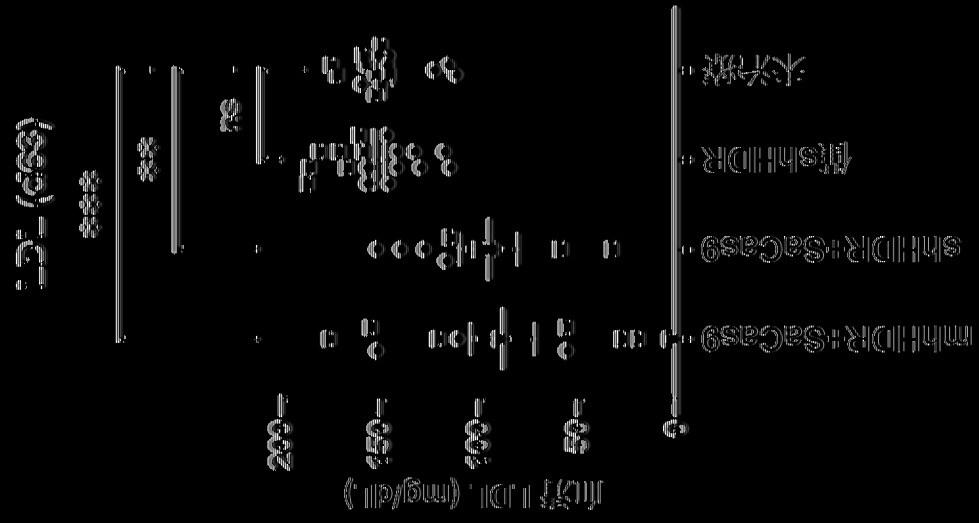
組別	巨噬菌載體 ( $10^8$ EC/kg)	核酸酶 ( $10^{-6}$ mg/ml P-Se-Cas9)	巨噬菌/核酸酶比例
1	mDR-2021	Se-Cas9	3:1
2	sDR-2021	Se-Cas9	3:1
3	sDR-2021	7BS	NA
4	NA	NA	NA
5	NA	NA	NA

(表18A)

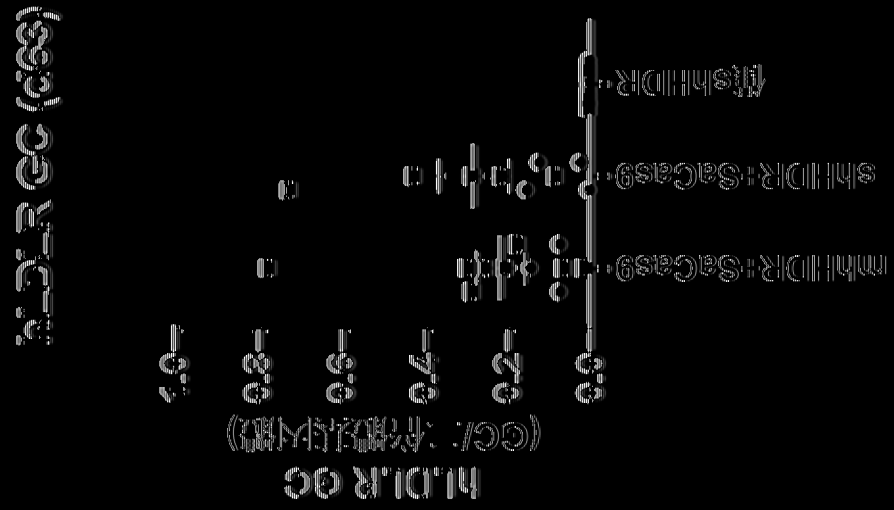
(b)



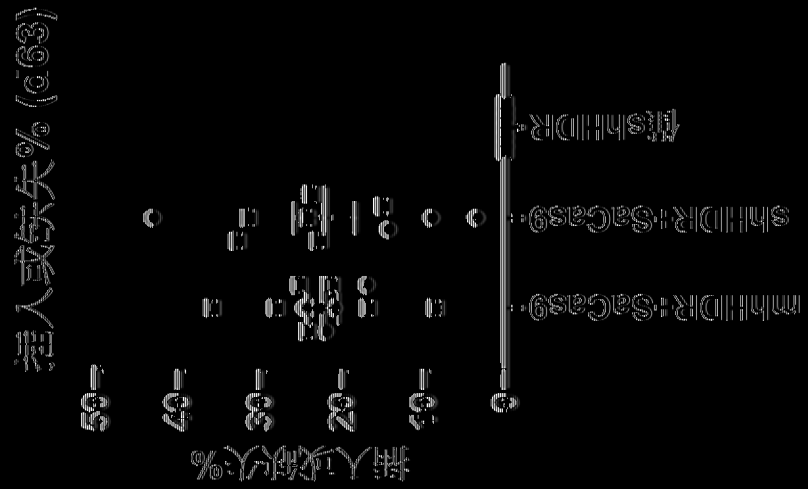
【表8】



【表8】



【表8】



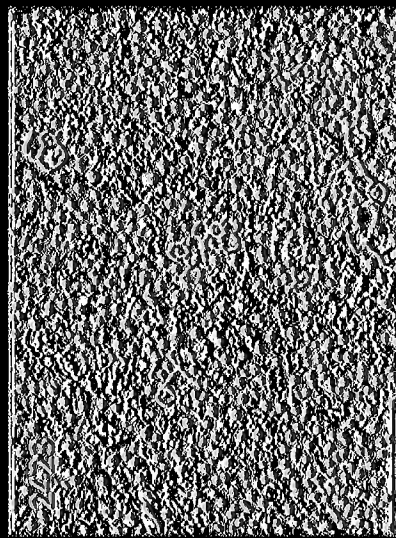
1529

第63天來自部分巨藻羽藻類之片鱗

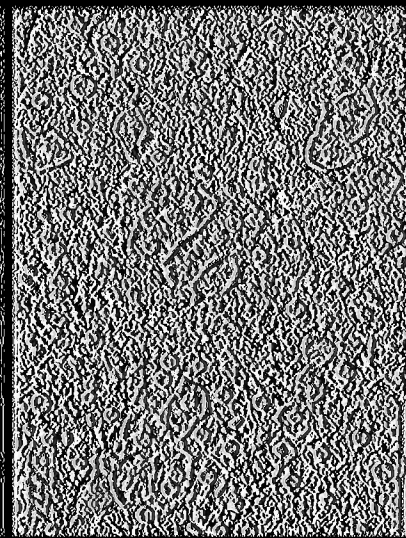
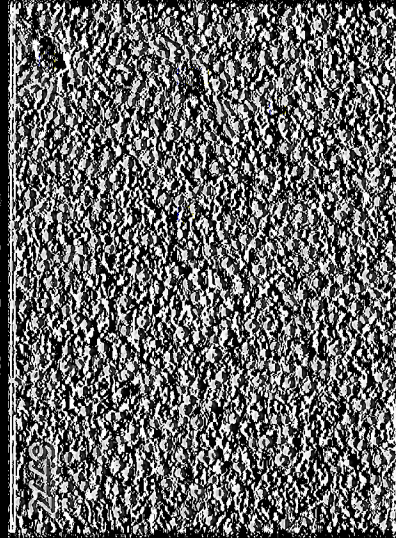
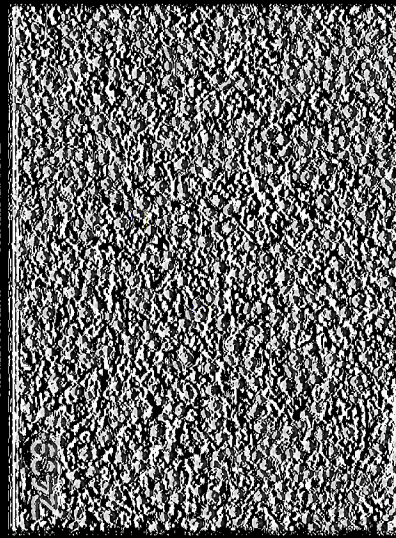
SeCae9

SeCae9

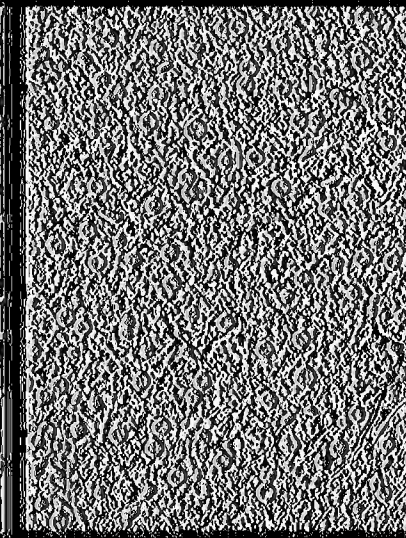
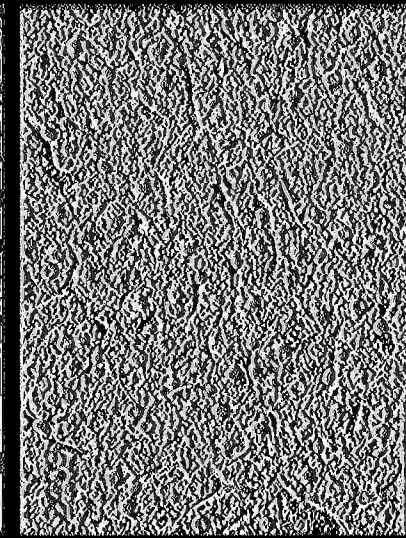
SeCae9



20x



40x



第63天來自部分巨藻羽藻類之片鱗