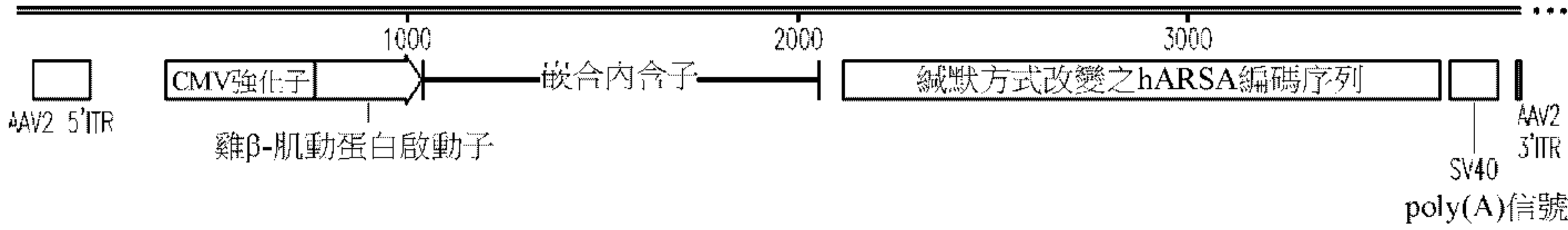
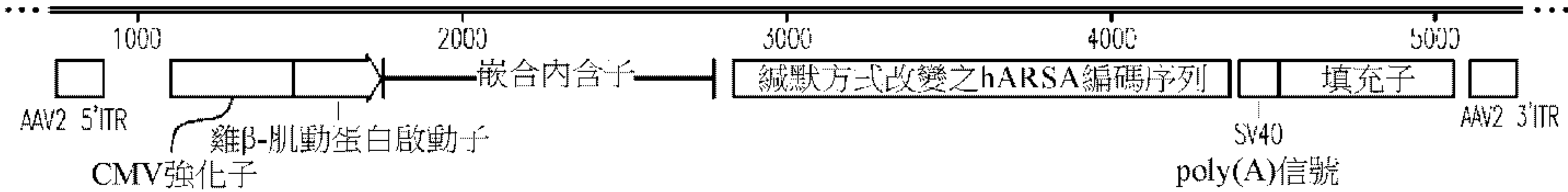




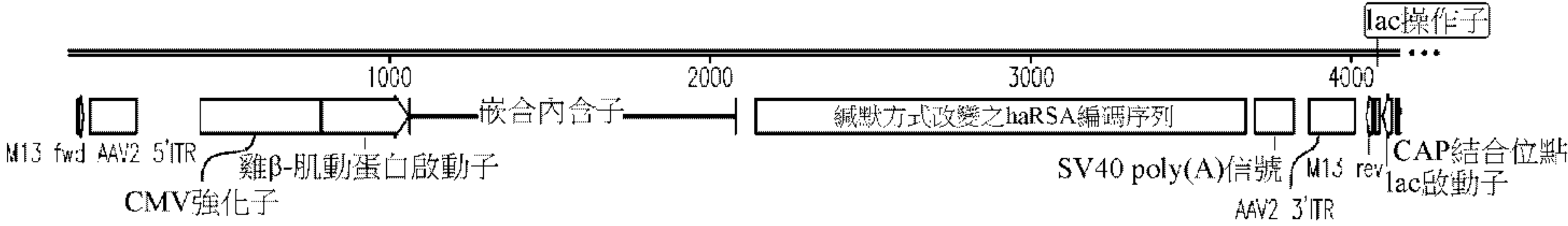
【圖 1A】



【圖1B】



【圖1C】



【圖1D】



202112807

【發明摘要】

【中文發明名稱】用於ARSA基因轉移之腺相關病毒組成物及其使用方法

【英文發明名稱】 ADENO-ASSOCIATED VIRUS COMPOSITIONS FOR ARSA
GENE TRANSFER AND METHODS OF USE THEREOF

【中文】

本文中提供一種腺相關病毒（AAV）組成物，其可在細胞中表現芳基硫酸酯酶 A（ARSA）多肽，藉此恢復 ARSA 基因功能。亦提供使用 AAV 組成物之方法，及用於製備 AAV 組成物之封裝系統。

【英文】

Provided herein are adeno-associated virus (AAV) compositions that can express an arylsulfatase A (ARSA) polypeptide in a cell, thereby restoring the ARSA gene function. Also provided are methods of using the AAV compositions, and packaging systems for making the AAV compositions.

【指定代表圖】 圖1

【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ARSA基因轉移之腺相關病毒組成物及其使用方法

【英文發明名稱】 ADENO-ASSOCIATED VIRUS COMPOSITIONS FOR ARSA
GENE TRANSFER AND METHODS OF USE THEREOF

【技術領域】

相關申請案之交叉參考

【0001】 本申請案主張 2019 年 6 月 10 日提交之美國臨時申請案第 62/859,539 號、2019 年 6 月 25 日提交之美國臨時申請案第 62/866,374 號、2019 年 10 月 15 日提交之美國臨時申請案第 62/915,523 號、2020 年 1 月 13 日提交之美國臨時申請案第 62/960,487 號、2020 年 3 月 10 日提交之美國臨時申請案第 62/987,858 號及 2020 年 4 月 16 日提交之美國臨時申請案第 63/010,970 號之權益，其各自以全文引用之方式併入本文中。

序列表

【0002】 本申請案含有序列表，其已以 ASCII 格式、以電子方式提交且以全文引用之方式併入本文中（該 ASCII 複本創建於 2020 年 4 月 16 日，名為「705151_HMW-030-6_ST25.txt」且大小為 295,995 個位元組）。

【先前技術】

【0003】 異染性白質失養症（MLD）為具有較高的未滿足之醫學需求之致死性溶酶體貯積病。此神經退化性疾病以三種形式（嬰兒晚期、青少年及成年人）存在且係由於溶酶體酶芳基硫酸酯酶-A（ARSA）不足引起。ARSA 位於稱為溶酶體之細胞結構中，其在溶酶體中有助於分解髓磷脂。此酶之不足可引起

腦部、脊髓及周邊器官中髓磷脂之大量積聚，引起髓鞘之嚴重損傷，髓鞘為神經纖維之主要保護層。髓鞘產生細胞中之髓磷脂積聚引起整個神經系統（包括腦部、脊髓以及連接腦部及脊髓與肌肉及偵測感覺（諸如觸覺、疼痛、熱及聲音）之感覺細胞之神經）中之白質之進行性損壞。因此，MLD 之特徵在於中樞神經系統且接著周邊神經系統之進行性軸突髓鞘脫失。此引起獲得性功能及/或技能之喪失、低張症、共濟失調、癲癇、失明、聽力喪失及過早地死亡。

【0004】 在患有異染性白質失養症之人中，白質損傷引起心智功能及運動技能（諸如行走能力）之進行性退化。受影響之個體亦出現四肢感覺喪失、失禁、癲癇、麻痺、無法說話、失明及聽力喪失。最終，此類個體喪失對其周圍環境之感知且變得無反應。儘管神經問題為異染性白質失養症之主要特徵，但已報導髓磷脂積聚對其他器官及組織之影響，最通常涉及膽囊。

【0005】 可藉由若干種治療來管理 MLD。舉例而言，用於減少 MLD 之病徵及症狀以及緩解相關疼痛之藥物。已證實造血幹細胞移植可藉由引入健康細胞以幫助置換患病細胞來延緩 MLD 之進程。其他治療包括用於促進肌肉及關節靈活性以及保持活動範圍之物理、職業及言語療法。然而，尚未能治癒 MLD。

【0006】 大部分患有 MLD 之個體具有芳基硫酸酯酶 A (ARSA) 基因之突變，且已鑑別超過 110 種引起 MLD 之不同的 ARSA 突變。每 100 人中便發現 1 人具有攜帶者突變，且在美國每 40,000 名活新生兒中便有 1 名或在全世界每 160,000 名活新生兒中便有 1 名受影響。

【0007】 基因療法提供治癒 MLD 之獨特機會。反轉錄病毒載體（包括慢病毒載體）能夠將核酸整合至宿主細胞基因組中，由於其對基因組之非靶向插入而產生安全性問題。舉例而言，存在載體破壞腫瘤抑制基因或活化致癌基因，從而引起惡性疾病之風險。實情為，在藉由用 γ 反轉錄病毒載體轉導 CD34⁺骨髓前驅體來進行之用於治療 X 性聯嚴重複合免疫缺乏症(SCID)之臨床試驗中，十

名患者中之四名產生白血病（Hacein-Bey-Abina 等人，《臨床研究雜誌（J Clin Invest.）》（2008）118(9):3132-42）。另一方面，非整合載體通常在活體內經歷表現量不足或表現持續時間不足。

【0008】因此，本領域中需要經改良之基因療法組成物及方法，其可有效及安全地恢復 MLD 患者中之 ARSA 基因功能。

【發明內容】

【0009】本文中提供可恢復細胞中之 ARSA 基因功能之腺相關病毒(AAV)組成物，及使用其治療與 ARSA 基因功能降低相關之疾病（例如 MLD）之方法。亦提供用於製備腺相關病毒組成物之封裝系統。

【0010】因此，在一個態樣中，本發明提供一種用於在細胞中表現芳基硫酸酯酶 A（ARSA）多肽之方法，該方法包含用重組腺相關病毒（rAAV）轉導細胞，該重組腺相關病毒包含：（a）AAV 衣殼，其包含 AAV 衣殼蛋白質（例如分枝系 F (Clade F)衣殼蛋白質）；及（b）轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序列之轉錄調節元件。

【0011】在某些具體實施例中，細胞為神經元及/或神經膠細胞。在某些具體實施例中，細胞為中樞神經系統及/或周邊神經系統之神經元及/或神經膠細胞。在某些具體實施例中，細胞為選自由以下組成之群組的中樞神經系統區域之細胞：脊髓、運動皮質、感覺皮質、海馬區、殼核、小腦（視情況，小腦核）及其任何組合。在某些具體實施例中，細胞為選自由以下組成之群組之細胞：運動神經元、星形膠質細胞、寡樹突神經膠細胞、中樞神經系統中之大腦皮質之細胞、周邊神經系統之感覺神經元、許旺細胞（Schwann cell）及其任何組合。在某些具體實施例中，細胞在哺乳動物個體中且以可有效轉導個體中之細胞之量向個體投與 AAV。

【0012】 在另一態樣中，本發明提供一種用於治療患有異染性白質失養症（MLD）之個體之方法，該方法包含向個體投與有效量之 rAAV，其包含：（a）AAV 衣殼，其包含 AAV 衣殼蛋白質（例如分枝系 F 衣殼蛋白質）；及（b）轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序列之轉錄調節元件。

【0013】 在某些具體實施例中，以緘默方式改變之 ARSA 編碼序列編碼 SEQ ID NO:23 中所示之胺基酸序列。在某些具體實施例中，以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO: 14、62 或 72 中所示之核苷酸序列。

【0014】 在某些具體實施例中，轉錄調節元件包含選自由以下組成之群組的元件中之一或多者：細胞巨大病毒（CMV）強化子元件、雞 β -肌動蛋白（CBA）啟動子、小型雞 β -肌動蛋白（SmCBA）啟動子、鈣調蛋白 1（CALM1）啟動子、蛋白脂質蛋白質 1（PLP1）啟動子、膠質原纖維酸性蛋白（GFAP）啟動子、突觸蛋白 2（SYN2）啟動子、金屬硫蛋白 3（MT3）啟動子及其任何組合。在某些具體實施例中，轉錄調節元件包含與選自由以下組成之群組的序列至少 90% 一致之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。在某些具體實施例中，轉錄調節元件包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。在某些具體實施例中，轉錄調節元件自 5'至 3'包含 SEQ ID NO: 58、25 及 32 中所示之核苷酸序列。在某些具體實施例中，轉錄調節元件包含 SEQ ID NO: 36 中所示之核苷酸序列。

【0015】 在某些具體實施例中，轉移基因組進一步包含位於以緘默方式改變之 ARSA 編碼序列 3'-聚腺苷酸化序列。在某些具體實施例中，聚腺苷酸化序列為外源聚腺苷酸化序列。在某些具體實施例中，外源聚腺苷酸化序列為 SV40 聚腺苷酸化序列。在某些具體實施例中，SV40 聚腺苷酸化序列包含 SEQ ID NO: 42 中所示之核苷酸序列。

【0016】 在某些具體實施例中，轉移基因組進一步包含填充序列。在某些具體實施例中，轉移基因組進一步包含位於以緘默方式改變之 ARSA 編碼序列 3' 之填充序列。在某些具體實施例中，填充序列位於聚腺苷酸化序列的 3'。

【0017】 在某些具體實施例中，轉移基因組包含選自由以下組成之群組之序列：SEQ ID NO: 41、44、46、65、67 及 75。

【0018】 在某些具體實施例中，轉移基因組進一步包含基因組的 5' 之 5' 反向末端重複（5' ITR）核苷酸序列及基因組的 3' 之 3' 反向末端重複（3' ITR）核苷酸序列。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95% 序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO: 19 具有至少 95% 序列一致性。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO: 26 具有至少 95% 序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO: 27 具有至少 95% 序列一致性。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95% 序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO: 57 具有至少 95% 序列一致性。

【0019】 在某些具體實施例中，轉移基因組包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 47、48、49、68、69 及 76。

【0020】 在某些具體實施例中，異染性白質失養症與芳基硫酸酯酶 A（ARSA）基因突變相關。在某些具體實施例中，個體為人類個體。

【0021】 在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 95% 序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中

對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0022】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0023】 在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列。

【0024】 在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、

5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0025】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16

之胺基酸 687 之胺基酸為 R；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0026】在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列。

【0027】在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 95% 序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 68 之胺基酸為 V；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質

中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0028】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 Y；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（f）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（g）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；（h）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（i）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0029】在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列。

【0030】在另一態樣中，本發明提供 rAAV，其包含：(a) AAV 衣殼，其包含 AAV 衣殼蛋白質（例如分枝系 F 衣殼蛋白質）；及 (b) 轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序列之轉錄調節元件。

【0031】在某些具體實施例中，以緘默方式改變之 ARSA 編碼序列編碼 SEQ ID NO: 23 中所示之胺基酸序列。在某些具體實施例中，以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO: 14 中所示之核苷酸序列。在某些具體實施例中，以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO: 62 或 72 中所示之核苷酸序列。

【0032】在某些具體實施例中，轉錄調節元件包含選自由以下組成之群組的元件中之一或多者：細胞巨大病毒 (CMV) 強化子元件、雞 β -肌動蛋白 (CBA) 啟動子、小型雞 β -肌動蛋白 (SmCBA) 啟動子、鈣調蛋白 1 (CALM1) 啟動子、蛋白脂質蛋白質 1 (PLP1) 啟動子、膠質原纖維酸性蛋白 (GFAP) 啟動子、突觸蛋白 2 (SYN2) 啟動子、金屬硫蛋白 3 (MT3) 啟動子及其任何組合。在某些具體實施例中，轉錄調節元件包含與選自由以下組成之群組的序列至少 90% 一致之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。在某些具體實施例中，轉錄調節元件包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。在某些具體實施例中，轉錄調節元件自 5' 至 3' 包含 SEQ ID NO: 58、25 及 32 中所示之核苷酸序列。在某些具體實施例中，轉錄調節元件包含 SEQ ID NO: 36 中所示之核苷酸序列。

【0033】在某些具體實施例中，轉移基因組進一步包含位於以緘默方式改變之 ARSA 編碼序列 3'-聚腺苷酸化序列。在某些具體實施例中，聚腺苷酸化序列為外源聚腺苷酸化序列。在某些具體實施例中，外源聚腺苷酸化序列為

SV40 聚腺苷酸化序列。在某些具體實施例中，SV40 聚腺苷酸化序列包含 SEQ ID NO: 42 中所示之核苷酸序列。

【0034】 在某些具體實施例中，轉移基因組包含選自由以下組成之群組之序列：SEQ ID NO: 41、44、46、65、67 及 75。

【0035】 在某些具體實施例中，轉移基因組進一步包含基因組的 5'之 5'反向末端重複（5' ITR）核苷酸序列及基因組的 3'之 3'反向末端重複（3' ITR）核苷酸序列。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95%序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO: 19 具有至少 95%序列一致性。

【0036】 在某些具體實施例中，轉移基因組包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 47、48、49、68、69 及 76。在某些具體實施例中，轉移基因組之核苷酸序列由選自由以下組成之群組之核苷酸序列組成：SEQ ID NO: 47、48、49、68、69 及 76。在某些具體實施例中，轉移基因組之核苷酸序列由 SEQ ID NO: 48 中所示之核苷酸序列組成。

【0037】 在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基

酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0038】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0039】 在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列。

【0040】 在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID

NO: 16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0041】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸

為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0042】 在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列。

【0043】 在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 68 之胺基酸為 V；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706

之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【0044】 在某些具體實施例中，（a）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；（b）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 Y；（c）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；（d）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；（e）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；（f）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；（g）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；（h）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或（i）衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【0045】 在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列。

【0046】 在另一態樣中，本發明提供一種醫藥組成物，其包含本文中所描

述之 rAAV。

【0047】 在另一態樣中，本發明提供一種聚核苷酸，其包含 SEQ ID NO: 14、62 及 72 中所示之核酸序列。

【0048】 在另一態樣中，本發明提供一種用於製備 rAAV 之封裝系統，其中該封裝系統包含 (a) 第一核苷酸序列，其編碼一或多種 AAV Rep 蛋白質；(b) 第二核苷酸序列，其編碼如技術方案 41 至 71 中任一項之 AAV 之衣殼蛋白質；及 (c) 第三核苷酸序列，其包含如技術方案 41 至 71 中任一項之 AAV 之 rAAV 基因組序列。

【0049】 在某些具體實施例中，封裝系統包含有包含第一核苷酸序列及第二核苷酸序列之第一載體，及包含第三核苷酸序列之第二載體。

【0050】 在某些具體實施例中，封裝系統進一步包含有包含一或多種輔助病毒基因之第四核苷酸序列。在某些具體實施例中，第四核苷酸序列包含於第三載體內。在某些具體實施例中，第四核苷酸序列包含一或多種來自選自由以下組成之群組的病毒之基因：腺病毒、疱疹病毒、痘瘡病毒及細胞巨大病毒 (CMV)。

【0051】 在某些具體實施例中，第一載體、第二載體及/或第三載體為質體。

【0052】 在另一態樣中，本發明提供一種用於 rAAV 之重組製備之方法，該方法包含在可產生 rAAV 之條件下將本文中所描述之封裝系統引入細胞中。

【0053】 在另一態樣中，本發明提供本文中所描述之 rAAV，其係用於如本文中所描述之用於在細胞中表現芳基硫酸酯酶 A (ARSA) 多肽之方法中。

【0054】 在另一態樣中，本發明提供本文中所描述之 rAAV，其係用於如本文中所描述之用於治療患有異染性白質失養症 (MLD) 之個體之方法中。

【圖式簡單說明】

【0055】圖 1A、1B、1C 及 1D 分別為 T-001、pHMI-5000、pHMI-5003 及 pHMI-hARSA1-TC-002 載體之載體圖。

【0056】圖 2A、2B 及 2C。圖 2A 為顯示在用媒劑對照物或封裝於 AAVHSC15 衣殼中之 pHMI-5000 處理之 ARSA (-/-) 小鼠中，使用抗 LAMP-1 抗體藉由免疫組織化學研究之來源於 LAMP-1 免疫反應性之總像素強度之定量之圖 (dWM：背部白質；vWM：腹部白質；及 vGM：腹部灰質)。圖 2B 為顯示隨時間推移，在對照組小鼠 (WT/Het) 及 ARSA (-/-) 小鼠之大腦中量測到的 C18:0 髓磷脂含量之圖。圖 2C 為顯示在用 4×10^{13} vg/kg 之劑量 (劑量-4) 之封裝於 AAVHSC15 衣殼中之 pHMI-hARSA1-TC-002 或媒劑對照物處理之 ARSA (-/-) 小鼠中的髓磷脂含量之變化 (以與年齡匹配野生型對照物相比之倍數形式) 之圖。圖 2D 為圖之集合，其顯示在用 4×10^{13} vg/kg 或 6×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物處理之 ARSA (-/-) 小鼠之前腦、中腦及後腦中的 C18:0 及 C18:1 髓磷脂異構體含量之變化 (以與年齡匹配野生型對照物相比之倍數形式)。圖 2E 為圖之集合，其顯示在用 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物處理之 ARSA (-/-) 小鼠之前腦、中腦及後腦中的 C18:0 及 C18:1 髓磷脂異構體含量之變化 (以與年齡匹配野生型對照物相比之倍數形式)。圖 2F 為圖之集合，其顯示在用 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物處理之 ARSA (-/-) 小鼠之前腦、中腦及後腦中的 C24:0 及 C24:1 髓磷脂異構體含量之變化 (以與年齡匹配野生型對照物相比之倍數形式)。圖 2G 為圖之集合，其顯示在用 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物處理之 ARSA (-/-) 小鼠之前腦、中腦及後腦中的總髓磷脂異構體含量之變化 (以與年齡匹配野生型對照物相比之倍數形式)。

【0057】圖 3A 及 3B。圖 3A 為顯示在對照組小鼠 (WT/Het) 及 ARSA (-/-)

小鼠中，在四週時量測之髓鞘及淋巴細胞蛋白質（MAL）mRNA 轉錄物之含量之圖。**圖 3B** 為顯示與年齡匹配野生型小鼠及用媒劑處理之 ARSA（-/-）小鼠相比，在用 $4e13$ vg/kg 之劑量（劑量-4）之封裝於 AAVHSC15 衣殼中之 pHMI-5000 處理之 ARSA（-/-）小鼠中偵測到的 MAL 轉錄物之含量之圖。**圖 3C** 為顯示在投與 $4e13$ vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物之後第 12 或 52 週，在野生型小鼠或 ARSA（-/-）小鼠中偵測到的 MAL 轉錄物複本數之圖。

【0058】圖 4 為顯示 ARSA（-/-）小鼠之大腦中的以每個經轉導之細胞計的載體基因組之數目與以每奈克 cDNA 計的 hARSA 之複本之數目之間的相關性之作圖。

【0059】圖 5 為顯示在靜脈內投與封裝於 AAV9 或 AAVHSC15 衣殼中之轉移載體 pHMI-5000（在每種情況下以 $2e13$ vg/kg 之劑量投與）之後，ARSA（-/-）小鼠之大腦中的以每個經轉導之細胞計的載體基因組之數目之圖。

【0060】圖 6 為顯示在靜脈內投與封裝於 AAV9 或 AAVHSC15 衣殼中且以所指示之劑量投與之轉移載體 pHMI-5000 之後，在 ARSA（-/-）小鼠之腦部中量測到的正常人類 ARSA 酶活性水準之百分比之圖。

【0061】圖 7 為顯示在靜脈內投與封裝於 AAV9 或 AAVHSC15 中之轉移載體 pHMI-5000（在每種情況下以 $4e13$ vg/kg 之劑量）之後，ARSA（-/-）小鼠之腦部中以每個細胞計的載體基因組之數目之圖。

【0062】圖 8 為顯示在靜脈內（IV）或鞘內（IT）投與封裝於 AAVHSC15 中之轉移載體 pHMI-5000 之後，後腦及中腦中之正常人類 ARSA 酶活性之百分比之圖。

【0063】圖 9A、9B、9C 及 9D。圖 9A 為顯示在以所指示之劑量向 ARSA（-/-）小鼠靜脈內投與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 之後，

在腦部中實現之正常 hARSA 活性之百分比之圖。**圖 9B** 為顯示在以所指示之劑量靜脈內投與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 之後，ARSA (-/-) 小鼠之大腦中之以每個細胞計的載體基因組之數目之圖。**圖 9C** 為顯示在給藥後 12 週之過程中，投與 4×10^{13} vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 的新生 ARSA (-/-) 小鼠中之 hARSA 酶活性水準之圖。**圖 9D** 為顯示投與 4×10^{13} vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 的成年 ARSA (-/-) 小鼠之大腦中之 ARSA 酶活性水準（經由 hARSA 轉錄物分析）之圖。**圖 9E** 為顯示在投與單次靜脈內 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA (-/-) 小鼠之大腦中，以每微克基因組 DNA 計的載體基因組之數目之圖。**圖 9F** 為顯示在投與單次靜脈內 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA (-/-) 小鼠之大腦中，以每奈克 RNA 計的 ARSA 轉錄物之複本之數目之圖。

【0064】 **圖 10A** 及 **10B** 分別為 TC-013.pHMIA2 及 TC-015.pKITR 載體之載體圖。

【0065】 **圖 11** 為顯示在投與轉移載體 pHMI-5000（CBA 啟動子）、TC-013.pHMIA2（CALM1 啟動子）及 TC-015.pKITR（smCBA 啟動子）（在每種情況下封裝於 AAVHSC15 衣殼中且以 4×10^{13} vg/kg 之劑量靜脈內投與）之小鼠 ARSA (-/-) 小鼠之大腦中，以每個細胞計的經轉導之病毒基因組之數目之圖。

【0066】 **圖 12** 為顯示在投與轉移載體 pHMI-5000（CBA 啟動子）及 TC-015.pKITR（smCBA 啟動子）（在每種情況下封裝於 AAVHSC15 衣殼中且以 4×10^{13} vg/kg 之劑量靜脈內投與）之小鼠 ARSA (-/-) 小鼠之大腦中偵測到的正常人類 ARSA 酶活性之百分比之圖。

【0067】 **圖 13** 為免疫墨點之照片，其顯示使用抗 hARSA 抗體之小鼠的大腦中 hARSA 之表現。向 ARSA (-/-) 小鼠投與轉移載體 pHMI-5000（CBA 啟動

子) 及 TC-015.pKITR (smCBA 啟動子)，在每種情況下封裝於 AAVHSC15 衣殼中且分別以 4×10^{13} vg/kg 及 8×10^{13} vg/kg 之劑量靜脈內投與 (對於每種載體，n=5 隻小鼠)。

【0068】圖 14 為轉移載體 pHMI-5004 之載體圖。

【0069】圖 15 為轉移載體 pHMI-5005 之載體圖。

【0070】圖 16 為顯示用所指示之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5005 處理或用媒劑對照物處理之非人類靈長類動物中之丙胺酸轉胺酶 (ALT) 含量之圖。

【0071】圖 17 為顯示投與封裝於 AAVHSC15 衣殼中之 pHMI-5005 的非人類靈長類動物之中樞神經系統 (CNS) 及腦脊髓液 (CSF) 中之 ARSA 活性之圖。

【實施方式】

【0072】本文中提供可恢復細胞中之 ARSA 基因功能之腺相關病毒 (AAV) 組成物，及使用其治療與 ARSA 基因功能降低相關之疾病 (例如 MLD) 之方法。亦提供用於製備腺相關病毒組成物之封裝系統。

I. 定義

【0073】如本文中所使用，術語「複製缺陷型腺相關病毒」係指包含缺乏 Rep 及 Cap 基因之基因組之 AAV。

【0074】如本文中所使用，術語「ARSA 基因」係指芳基硫酸酯酶 A 基因。由國家生物技術資訊中心 (National Center for Biotechnology Information; NCBI) 基因 ID 410 鑑別人類 ARSA 基因。ARSA mRNA 之例示性核苷酸序列提供為 SEQ ID NO: 14。ARSA 多肽之例示性胺基酸序列提供為 SEQ ID NO: 23。

【0075】如本文中所使用，術語「轉移基因組」係指包含可操作地連接至外源轉錄調節元件之編碼序列之重組 AAV 基因組，該外源轉錄調節元件在轉移

基因組被引入細胞中時介導編碼序列之表現。在某些具體實施例中，轉移基因組未整合在細胞之染色體 DNA 中。本領域中熟習此項技術者將瞭解，轉移基因組中之包含可操作地連接至 ARSA 編碼序列之轉錄調節元件之部分相對於 ARSA 編碼序列之轉錄方向可為有義或反義定向。

【0076】如本文中所使用，術語「分枝系 F 衣殼蛋白質」係指 AAV VP1、VP2 或 VP3 衣殼蛋白質，其分別與本文中之 SEQ ID NO: 1 之胺基酸 1-736、138-736 及 203-736 中示之 VP1、VP2 或 VP3 胺基酸序列具有至少 90%一致性。

【0077】如本文中所使用，兩個核苷酸序列之間或兩個胺基酸序列之間的「一致性百分比」係藉由將所比對之序列對之間的匹配數目乘以 100 且除以所比對之區域之長度（包括內部間隙(gaps)）來計算。一致性評分僅對完美匹配進行計數，且不考慮胺基酸彼此之間的相似性程度。長度中僅包括內部間隙，而不包括序列末端處之間隙。

【0078】如本文中所使用，術語「與 ARSA 基因突變相關之疾病或病症」係指由 ARSA 基因之突變引起、由 ARSA 基因之突變加劇或在遺傳學上與 ARSA 基因之突變相關之任何疾病或病症。在某些具體實施例中，與 ARSA 基因突變相關之疾病或病症為異染性白質失養症（MLD）。

【0079】如本文中所使用，術語「編碼序列」係指互補 DNA（cDNA）中之編碼多肽之部分，在起始密碼子處開始且在終止密碼子處終止。基因可由於替代性剪接、替代性轉譯起始及群體內之變化而具有一或多個編碼序列。編碼序列可為野生型或密碼子改變型。例示性野生型 ARSA 編碼序列示於 SEQ ID NO: 24 中。

【0080】如本文中所使用，術語「以緘默方式改變」係指在不改變由編碼序列或經填充子插入之編碼序列編碼之多肽之胺基酸序列之情況下，基因之編碼序列或經填充子插入之編碼序列之變化（例如藉由核苷酸取代）。此類緘默

變化之有利之處在於，其可提高編碼序列之轉譯效率，及/或在編碼序列轉導至細胞中時防止與內源基因之相應序列之重組。

【0081】 在本發明中，相對於起始密碼子之第一核苷酸指定 ARSA 基因中之核苷酸位置。起始密碼子之第一核苷酸為位置 1；位於起始密碼子之第一核苷酸 5'之核苷酸具有負數；位於起始密碼子之第一核苷酸 3'之核苷酸具有正數。人類 ARSA 基因之例示性核苷酸 1 為 NCBI 參考序列之核苷酸 374：NG_009260.2（區域：5028-10426），且人類 ARSA 基因之例示性核苷酸 3 為 NCBI 參考序列之核苷酸 376：NG_009260.2（區域：5028-10426）。與起始密碼子之 5'相鄰之核苷酸為核苷酸-1。

【0082】 在本發明中，相對於涵蓋起始密碼子之第一核苷酸之外顯子指定 ARSA 基因中之外顯子及內含子，該第一核苷酸為 NCBI 參考序列之核苷酸 374：NG_009260.2（區域：5028-10426）。涵蓋起始密碼子之第一核苷酸之外顯子為外顯子 1。外顯子 3'至外顯子 1 為 5'至 3'：外顯子 2、外顯子 3 等。內含子 3'至外顯子 1 為 5'至 3'：內含子 1、內含子 2 等。因此，ARSA 基因自 5'至 3'包含：外顯子 1、內含子 1、外顯子 2、內含子 2、外顯子 3 等。人類 ARSA 基因之例示性外顯子 1 為 NCBI 參考序列之核苷酸 374-597：NG_009260.2（區域：5028-10426）。人類 ARSA 基因之例示性內含子 1 為 NCBI 參考序列之核苷酸 598-746：NG_009260.2（區域：5028-10426）。

【0083】 如本文中所使用，術語「轉錄調節元件」或「TRE」係指順式作用核苷酸序列，例如 DNA 序列，其藉由 RNA 聚合酶調節（例如控制、增加或減少）可操作地連接之核苷酸序列之轉錄以形成 RNA 分子。TRE 依靠一或多種反式作用分子（諸如轉錄因子）來調節轉錄。因此，一個 TRE 在其與不同的反式作用分子接觸時（例如當其位於不同類型的細胞中時）可以不同方式調節轉錄。TRE 可包含一或多種啟動子元件及/或強化子元件。本領域中熟習此項技術

者將瞭解，基因中之啟動子及強化子元件可在位置上靠近，且術語「啟動子」可指包含啟動子元件及強化子元件之序列。因此，術語「啟動子」不排除序列中之強化子元件。啟動子及強化子元件無需來源於相同基因或物種，且每個啟動子或強化子元件之序列可與基因組中之相應內源序列一致或實質上一致。

【0084】如本文中所使用，術語「可操作地連接」用於描述 TRE 與待轉錄之編碼序列之間的連接。通常，使基因表現處於包含一或多種啟動子及/或強化子元件之 TRE 之控制下。若編碼序列之轉錄由 TRE 控制或影響，則編碼序列「可操作地連接」至 TRE。TRE 之啟動子及強化子元件可自編碼序列呈任何定向及/或距離，只要獲得所需轉錄活性即可。在某些具體實施例中，TRE 位於編碼序列之上游。

【0085】如本文中所使用，術語「核糖體跳躍元件(ribosomal skipping element)」係指編碼短肽序列之核苷酸序列，其能夠造成由一個 mRNA 分子之轉譯產生兩個肽鏈。在某些具體實施例中，核糖體跳躍元件編碼包含 $X_1X_2EX_3NPGP$ 之共有基元(motif)之肽，其中 X_1 為 D 或 G， X_2 為 V 或 I，且 X_3 為任何胺基酸 (SEQ ID NO:34)。在某些具體實施例中，核糖體跳躍元件編碼明脈扁刺蛾 (*Thosea asigna*) 病毒 2A 肽 (T2A)、豬捷申病毒 (porcine teschovirus) -1 2A 肽 (P2A)、口蹄疫病毒 2A 肽 (F2A)、馬鼻炎 A 病毒 2A 肽 (E2A)、細胞質多角體病毒 2A 肽 (BmCPV 2A) 或家蠶 (*B. mori*) 2A 肽之軟化病病毒 (BmIFV 2A)。T2A 肽及 P2A 肽之例示性胺基酸序列分別示於 SEQ ID NO:37 及 38 中。T2A 元件及 P2A 元件之例示性核苷酸序列分別示於 SEQ ID NO:66 及 63 中。在某些具體實施例中，核糖體跳躍元件編碼在 N 端進一步包含 Gly-Ser-Gly 之序列之肽，視情況其中 Gly-Ser-Gly 之序列由 GGCAGCGGA 之核苷酸序列編碼。儘管不希望受理論約束，但假設核糖體跳躍元件藉由以下發揮功能：終止第一肽鏈之轉譯及再起始第二肽鏈之轉譯；或由核糖體跳躍元件編碼之肽序列

中之肽鍵藉由經編碼之肽之內源蛋白酶活性進行之裂解，或環境（例如胞溶質）中之另一種蛋白酶。

【0086】如本文中所使用，術語「核糖體跳躍肽」係指由核糖體跳躍元件編碼之肽。

【0087】如本文中所使用，術語「聚腺苷酸化序列」係指在轉錄至 RNA 中時組成聚腺苷酸化信號序列之 DNA 序列。聚腺苷酸化序列可為原生的（例如來自 ARSA 基因）或外源的。外源聚腺苷酸化序列可為哺乳動物或病毒聚腺苷酸化序列（例如 SV40 聚腺苷酸化序列）。

【0088】如本文中所使用，「外源聚腺苷酸化序列」係指與 ARSA 基因（例如人類 ARSA 基因）之內源聚腺苷酸化序列不一致或實質上一致之聚腺苷酸化序列。在某些具體實施例中，外源聚腺苷酸化序列為相同物種（例如人類）中之非 ARSA 基因之聚腺苷酸化序列。在某些具體實施例中，外源聚腺苷酸化序列為不同物種（例如病毒）之聚腺苷酸化序列。

【0089】如本文中所使用，在向個體投與 AAV 之情形下，術語「有效量」係指實現所需預防或治療作用之 AAV 之量。

II. 腺相關病毒組成物

【0090】在一個態樣中，本文中提供新穎重組 AAV（例如複製缺陷型 AAV）組成物，其適用於在具有降低的或以其他方式缺乏 ARSA 基因功能之細胞中表現 ARSA 多肽。在某些具體實施例中，本文中所揭示之 rAAV 包含：AAV 衣殼，其包含衣殼蛋白質（例如分枝系 F 衣殼蛋白質）；及轉移基因組，其包含可操作地連接至 ARSA 編碼序列（例如以緘默方式改變之 ARSA 編碼序列）之轉錄調節元件，從而實現經 AAV 轉導之細胞中的 ARSA 之染色體外表現。

【0091】來自本領域中已知的任何衣殼之衣殼蛋白質可用於本文中所揭示之 rAAV 組成物中，包括（但不限於）來自 AAV1、AAV2、AAV3、AAV4、AAV5、

AAV6、AAV7、AAV8 或 AAV9 血清型之衣殼蛋白質。舉例而言，在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列，其中：衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為

H，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C。在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO:2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列。

【0092】舉例而言，在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO:2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO:2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列，其中：衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺

基酸為 A；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C。在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列。

【0093】舉例而言，在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列。在某些具體實施例中，衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列，其中：衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 2 之胺基酸為 T；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 65 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 68 之胺基酸為 V；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 77 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 119 之胺基酸為 L；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 151 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 160 之胺基酸為 D；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 206 之胺基酸為 C；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為 H；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 312 之胺基酸為 Q；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺基酸為 A；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 468 之胺基酸為 S；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 590 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G 或 Y；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R；衣殼蛋白質中對應於 SEQ ID

NO:16 之胺基酸 690 之胺基酸為 K；衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C；或衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 2 之胺基酸為 T，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 312 之胺基酸為 Q。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 65 之胺基酸為 I，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 Y。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 77 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 690 之胺基酸為 K。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 119 之胺基酸為 L，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 468 之胺基酸為 S。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 626 之胺基酸為 G，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 718 之胺基酸為 G。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 296 之胺基酸為 H，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 464 之胺基酸為 N，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 681 之胺基酸為 M。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 687 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 346 之胺基酸為 A，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R。在某些具體實施例中，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 501 之胺基酸為 I，衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 505 之胺基酸為 R，且衣殼蛋白質中對應於 SEQ ID NO:16 之胺基酸 706 之胺基酸為 C。在某些具體實施例中，衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列。

【0094】 在某些具體實施例中，AAV 衣殼包含以下中之兩者或更多者：(a) 包含 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含：(a) 具有由 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 組成之胺基酸序列之衣殼蛋白質；(b) 具有由 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 組成之胺基酸序列之衣殼蛋白質；及 (c) 具有由 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 組成之胺基酸序列之衣殼蛋白質。

【0095】 在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含與 SEQ ID NO:8 之胺基酸序列 203-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；(b) 包含與 SEQ ID NO:8 之胺基酸序列 138-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；及 (c) 包含與 SEQ ID NO:8 之胺基酸序列 1-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含 SEQ ID NO:8 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:8 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:8 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體

實施例中，AAV 衣殼包含以下中之兩者或更多者：(a) 包含 SEQ ID NO:8 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:8 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:8 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含：(a) 具有由 SEQ ID NO:8 之胺基酸 203-736 組成之胺基酸序列之衣殼蛋白質；(b) 具有由 SEQ ID NO:8 之胺基酸 138-736 組成之胺基酸序列之衣殼蛋白質；及 (c) 具有由 SEQ ID NO:8 之胺基酸 1-736 組成之胺基酸序列之衣殼蛋白質。

【0096】 在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含與 SEQ ID NO:11 之胺基酸序列 203-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；(b) 包含與 SEQ ID NO:11 之胺基酸序列 138-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；及 (c) 包含與 SEQ ID NO:11 之胺基酸序列 1-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含 SEQ ID NO:11 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:11 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:11 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之兩者或更多者：(a) 包含 SEQ ID NO:11 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:11 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:11 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含：(a) 具有

由 SEQ ID NO:11 之胺基酸 203-736 組成之胺基酸序列之衣殼蛋白質；(b) 具有由 SEQ ID NO:11 之胺基酸 138-736 組成之胺基酸序列之衣殼蛋白質；及 (c) 具有由 SEQ ID NO:11 之胺基酸 1-736 組成之胺基酸序列之衣殼蛋白質。

【0097】 在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含與 SEQ ID NO:13 之胺基酸序列 203-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；(b) 包含與 SEQ ID NO:13 之胺基酸序列 138-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；及 (c) 包含與 SEQ ID NO:13 之胺基酸序列 1-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含 SEQ ID NO:13 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:13 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:13 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之兩者或更多者：(a) 包含 SEQ ID NO:13 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:13 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:13 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含：(a) 具有由 SEQ ID NO:13 之胺基酸 203-736 組成之胺基酸序列之衣殼蛋白質；(b) 具有由 SEQ ID NO:13 之胺基酸 138-736 組成之胺基酸序列之衣殼蛋白質；及 (c) 具有由 SEQ ID NO:13 之胺基酸 1-736 組成之胺基酸序列之衣殼蛋白質。

【0098】 在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a)

包含與 SEQ ID NO:16 之胺基酸序列 203-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；(b) 包含與 SEQ ID NO:16 之胺基酸序列 138-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質；及 (c) 包含與 SEQ ID NO:16 之胺基酸序列 1-736 具有至少 80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之一或多者：(a) 包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:16 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含以下中之兩者或更多者：(a) 包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列之衣殼蛋白質；(b) 包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列之衣殼蛋白質；及 (c) 包含 SEQ ID NO:16 之胺基酸 1-736 之胺基酸序列之衣殼蛋白質。在某些具體實施例中，AAV 衣殼包含：(a) 具有由 SEQ ID NO:16 之胺基酸 203-736 組成之胺基酸序列之衣殼蛋白質；(b) 具有由 SEQ ID NO:16 之胺基酸 138-736 組成之胺基酸序列之衣殼蛋白質；及 (c) 具有由 SEQ ID NO:16 之胺基酸 1-736 組成之胺基酸序列之衣殼蛋白質。

【0099】 適用於本文中所揭示之 AAV 組成物中之轉移基因組通常包含可操作地連接至 ARSA 編碼序列之轉錄調節元件 (TRE)。在某些具體實施例中，轉移基因組包含 TRE 及 ARSA 編碼序列的 5'之 5'反向末端重複 (5' ITR) 核苷酸序列，及 TRE 及 ARSA 編碼序列的 3'之 3'反向末端重複 (3' ITR) 核苷酸序列。

【0100】 在某些具體實施例中，ARSA 編碼序列包含 ARSA 基因之全部或

實質上全部編碼序列。在某些具體實施例中，轉移基因組包含編碼 SEQ ID NO:23 之核苷酸序列且可視情況進一步包含位於 ARSA 編碼序列 3'-外源聚腺苷酸化序列。在某些具體實施例中，編碼 SEQ ID NO:23 之核苷酸序列為野生型的（例如具有 SEQ ID NO:24 中所示之序列）。在某些具體實施例中，編碼 SEQ ID NO:23 之核苷酸序列係以緘默方式改變（例如具有 SEQ ID NO:14、62 或 72 中所示之序列）。

【0101】 在某些具體實施例中，ARSA 編碼序列編碼包含 ARSA 蛋白質之全部或實質上全部胺基酸序列之多肽。在某些具體實施例中，ARSA 編碼序列編碼野生型 ARSA 蛋白質（例如人類 ARSA 蛋白質）之胺基酸序列。在某些具體實施例中，ARSA 編碼序列編碼突變型 ARSA 蛋白質（例如人類 ARSA 蛋白質）之胺基酸序列，其中突變型 ARSA 多肽為野生型 ARSA 多肽之功能等效物，亦即，可充當野生型 ARSA 多肽。在某些具體實施例中，功能上等效的 ARSA 多肽進一步包含至少一種未在野生型 ARSA 多肽中發現之特徵，例如抵抗蛋白質降解之能力。

【0102】 在某些具體實施例中，適用於本文中所揭示之 AAV 組成物中之轉移基因組通常包含轉錄調節元件（TRE），其可操作地連接至編碼 ARSA 及/或 SUMF1 之編碼序列。硫酸酯酶修飾因子 1（SUMF1）基因編碼一種酶，該酶藉由使受質硫酸酯酶中之半胱胺酸殘基氧化成活性位點 3-側氧基丙胺酸殘基（其亦稱為 C- α -甲醯基甘胺酸）來催化硫酸酯之水解。與 SUMF1 相關之疾病包括多種硫酸酯酶不足及異染性白質失養症。

【0103】 在某些具體實施例中，SUMF1 編碼序列包含 SUMF1 基因之全部或實質上全部編碼序列。在某些具體實施例中，轉移基因組包含編碼 SEQ ID NO:29 之核苷酸序列且可視情況進一步包含位於 SUMF1 編碼序列 3'-外源聚腺苷酸化序列。在某些具體實施例中，編碼 SEQ ID NO:29 之核苷酸序列為野生型

(例如具有 SEQ ID NO:64 中所示之序列)。在某些具體實施例中，編碼 SEQ ID NO:29 之核苷酸序列係以緘默方式改變。

【0104】 在某些具體實施例中，SUMF1 編碼序列編碼包含 SUMF1 蛋白質之全部或實質上全部胺基酸序列之多肽。在某些具體實施例中，SUMF1 編碼序列編碼野生型 SUMF1 蛋白質(例如人類 SUMF1 蛋白質(hSUMF1))之胺基酸序列。在某些具體實施例中，SUMF1 編碼序列編碼突變型 SUMF1 蛋白質(例如人類 SUMF1 蛋白質)之胺基酸序列，其中突變型 SUMF1 多肽為野生型 SUMF1 多肽之功能等效物，亦即，可充當野生型 SUMF1 多肽。在某些具體實施例中，功能上等效的 SUMF1 多肽進一步包含至少一種未在野生型 SUMF1 多肽中發現之特徵，例如抵抗蛋白質降解之能力。

【0105】 在某些具體實施例中，轉移基因組經設計以表現 hARSA 及 hSUMF1 且包含核苷酸序列，該核苷酸序列包含編碼 hARSA 之第一編碼序列及編碼 hSUMF1 之第二編碼序列。在某些具體實施例中，編碼 hARSA 之第一編碼序列及編碼 hSUMF1 之第二編碼序列係由核糖體跳躍元件分離。可使用本領域中已知的任何核糖體跳躍元件，例如本文中其他地方描述之核糖體跳躍元件。在某些具體實施例中，包含編碼 hARSA 之第一編碼序列及編碼 hSUMF1 之第二編碼序列之核苷酸序列包含 SEQ ID NO:30 中所示之核苷酸序列。

【0106】 在某些具體實施例中，適用於本文中所揭示之 AAV 組成物中之轉移基因組通常包含轉錄調節元件(TRE)，其可操作地連接至編碼 ARSA 及/或 SapB 之編碼序列。鞘脂激活蛋白原(PSAP)基因編碼高保守性前原蛋白(preproprotein)，該前原蛋白經蛋白水解處理以產生四種主要裂解產物，包括鞘脂激活蛋白 A、B、C 及 D。前驅蛋白質之每個結構域之長度為約 80 個胺基酸殘基，且具有幾乎相同的半胱胺酸殘基及糖基化位點之佈置。鞘脂激活蛋白 A-D 主要位於溶酶體隔室，其在溶酶體隔室中促進具有短寡醣基團之鞘醣脂之分解

代謝。前驅蛋白質以分泌蛋白質及整合膜蛋白質兩者之形式存在且具有神經營養活性。此基因中之突變與高歇氏病（Gaucher disease）及異染性白質失養症相關。已證實鞘脂激活蛋白 B（SapB）可刺激半乳糖-腦甘脂硫酸酯由 ARSA 進行之水解、GM1 神經節苷脂由 β -半乳糖苷酶進行之水解及球形三醯神經醯胺由 α -半乳糖 A 進行之水解。已證實 SapB 與鞘脂水解酶之受質形成溶解複合物。

【0107】在某些具體實施例中，SapB 編碼序列包含 SapB 基因之全部或實質上全部編碼序列。在某些具體實施例中，轉移基因組包含編碼 SEQ ID NO:33 之核苷酸序列且可視情況進一步包含位於 SapB 編碼序列 3'-外源聚腺苷酸化序列。在某些具體實施例中，編碼 SEQ ID NO:33 之核苷酸序列為野生型（例如具有 SEQ ID NO:73 中所示之序列）。在某些具體實施例中，編碼 SEQ ID NO:33 之核苷酸序列係以緘默方式改變。

【0108】在某些具體實施例中，SapB 編碼序列編碼包含 SapB 蛋白質之全部或實質上全部胺基酸序列之多肽。在某些具體實施例中，SapB 編碼序列編碼野生型 SapB 蛋白質（例如人類 SapB 蛋白質（hSapB））之胺基酸序列。在某些具體實施例中，SapB 編碼序列編碼突變型 SapB 蛋白質（例如人類 SapB 蛋白質）之胺基酸序列，其中突變型 SapB 多肽為野生型 SapB 多肽之功能等效物，亦即，可充當野生型 SapB 多肽。在某些具體實施例中，功能上等效的 SapB 多肽進一步包含至少一種未在野生型 SapB 多肽中發現之特徵，例如抵抗蛋白質降解之能力。

【0109】在某些具體實施例中，轉移基因組經設計以表現 hARSA 及 hSapB 且包含核苷酸序列，該核苷酸序列包含編碼 hARSA 之第一編碼序列及編碼 hSapB 之第二編碼序列。在某些具體實施例中，編碼 hARSA 之第一編碼序列及編碼 hSapB 之第二編碼序列係由核糖體跳躍元件分離。可使用本領域中已知的任何核糖體跳躍元件，例如本文中其他地方描述之核糖體跳躍元件。在某些具

體實施例中，包含編碼 hARSA 之第一編碼序列及編碼 hSapB 之第二編碼序列之核苷酸序列包含 SEQ ID NO:74 中所示之核苷酸序列。

【0110】 轉移基因組可用於表現任何哺乳動物細胞（例如人類細胞）中之 ARSA、SUMF1 及/或 SapB。因此，TRE 可在任何哺乳動物細胞（例如人類細胞）中具有活性。在某些具體實施例中，TRE 在廣泛的人類細胞中具有活性。此類 TRE 可包含組成性啟動子及/或強化子元件，包括細胞巨大病毒（CMV）啟動子/強化子（例如包含與 SEQ ID NO:58 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、SV40 啟動子、雞 β 肌動蛋白（CBA）啟動子（例如包含與 SEQ ID NO:59 或 25 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、smCBA 啟動子（例如包含與 SEQ ID NO:55 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、人類延長因子 1 α （EF1 α ）啟動子（例如包含與 SEQ ID NO:40 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、包含轉錄因子結合位點之小鼠微小病毒（MVM）內含子（例如包含與 SEQ ID NO:35 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、人類磷酸甘油酸激酶（PGK1）啟動子、人類泛蛋白 C（Ubc）啟動子、人類 β 肌動蛋白啟動子、人類神經元特異性烯醇酶（ENO2）啟動子、人類 β -葡糖苷酸酶（GUSB）啟動子、兔 β -球蛋白元件（例如包含與 SEQ ID NO:60 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）、人類鈣調蛋白 1（CALM1）啟動子（例如包含與 SEQ ID NO:54 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）及/或人類甲基-CpG 結合蛋白質 2（MeCP2）啟動子。此等 TRE 中之任一者可以任何順序組合以驅動有效轉錄。舉例而言，轉移基因組可包含 CMV 強化子、CBA

啟動子及來自兔 β -球蛋白基因之外顯子 3 之剪接受體，統稱為 CAG 啟動子（例如包含與 SEQ ID NO:28 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）。舉例而言，轉移基因組可包含 CMV 強化子及 CBA 啟動子之雜交物，接著為剪接供體及剪接受體，統稱為 CASI 啟動子區域（例如包含與 SEQ ID NO:63 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）。

【0111】或者，TRE 可為組織特異性 TRE，亦即，其在特定組織及/或器官中具有活性。組織特異性 TRE 包含一或多種組織特異性啟動子及/或強化子元件，及視情況選用之一或多種組成性啟動子及/或強化子元件。本領域中熟習此項技術者將理解，可藉由本領域中熟知之方法自組織中特異性表現之基因分離組織特異性啟動子及/或強化子元件。

【0112】在某些具體實施例中，TRE 為腦部特異性（例如神經元特異性、神經膠細胞特異性、星形膠質細胞特異性、寡樹突神經膠細胞特異性、微神經膠質細胞特異性及/或中樞神經系統特異性）。例示性腦部特異性 TRE 可包含一或多個來自（但不限於）人類膠質原纖維酸性蛋白（GFAP）啟動子、人類突觸蛋白 1（SYN1）啟動子、人類突觸蛋白 2（SYN2）啟動子、人類金屬硫蛋白 3（MT3）啟動子及/或人類蛋白脂質蛋白質 1（PLP1）啟動子之元件。更多的腦部特異性啟動子元件揭示於 WO 2016/100575A1 中，其以全文引用之方式併入本文中。

【0113】在某些具體實施例中，轉移基因組包含兩個或更多個 TRE，視情況包含以上揭示之 TRE 中之至少一者。本領域中熟習此項技術者將理解，此等 TRE 中之任一者可以任何順序組合，且組成性 TRE 及組織特異性 TRE 之組合可驅動有效及組織特異性轉錄。

【0114】在某些具體實施例中，轉移載體進一步包含非編碼填充序列（例

如包含與 SEQ ID NO:39 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%一致之核苷酸序列）。非編碼填充序列可用於將載體之尺寸保持在適合的限制內以用於有效 DNA 封裝，且因此可用於提高 DNA 封裝之功效。本領域中熟習此項技術者將認識到，填充序列之性質可對載體之功能具有影響，且因此將選擇最適合的填充序列來使用。

【0115】 在某些具體實施例中，轉移載體進一步包含相對於 ARSA 編碼序列位於 5'或插入 ARSA 編碼序列中之內含子。此類內含子可增加轉基因表現，例如藉由減少轉錄沉默及增強自細胞核至細胞質之 mRNA 輸出。在某些具體實施例中，轉移基因組自 5'至 3'包含：非編碼外顯子、內含子及 ARSA 編碼序列。在某些具體實施例中，將內含子序列插入 ARSA 編碼序列中，視情況其中在連接兩個原生外顯子之核苷酸間鍵處插入內含子。在某些具體實施例中，在連接原生外顯子 1 與外顯子 2 之核苷酸間鍵處插入內含子。

【0116】 內含子可包含 ARSA 基因之原生內含子序列、來自不同物種之內含子序列或來自相同物種之不同基因及/或合成內含子序列。本領域中熟悉此項技術者將瞭解，合成內含子序列可經設計以藉由引入本領域中已知的任何共有剪接基元來介導 RNA 剪接（例如 Sibley 等人, (2016) 《自然綜述遺傳學（Nature Reviews Genetics）》，17, 407-21，其以全文引用之方式併入本文中）。例示性內含子序列提供於 Lu 等人 (2013) 《分子療法(Molecular Therapy)》21(5): 954-63 及 Lu 等人 (2017) 《人類基因療法（Hum. Gene Ther.）》28(1): 125-34 中，其以全文引用之方式併入本文中。在某些具體實施例中，轉移基因組包含 SV40 內含子（例如包含與 SEQ ID NO:31 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之核苷酸序列）或小鼠微小病毒（MVM）內含子（例如包含與 SEQ ID NO:35 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之核苷酸序列）。在某些具體實施例中，轉移基因組包含 SV40

內含子(例如包含 SEQ ID NO:31 中所示之核苷酸序列)或小鼠微小病毒(MVM)內含子(例如包含 SEQ ID NO:35 中所示之核苷酸序列)。在某些具體實施例中，轉移基因組包含嵌合內含子序列，其包含雞及兔序列之組合，部分包含未經轉錄之雞 ACTB (cACTB) 啟動子、全部 cACTB 外顯子 1、部分 cACTB 內含子 1、部分兔 HBB2 (rHBB2) 內含子 2 及部分 rHBB2 外顯子 3 (例如 SEQ ID NO:32)。在某些具體實施例中，轉移基因組包含嵌合內含子序列(例如包含與 SEQ ID NO:32 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之核苷酸序列)。在某些具體實施例中，轉移基因組包含嵌合內含子序列(例如包含 SEQ ID NO:32 中所示之核苷酸序列)。

【0117】 在某些具體實施例中，轉移基因組包含 TRE，其包含 CMV 強化子、CBA 啟動子及嵌合內含子序列(例如包含與 SEQ ID NO:36 至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之核苷酸序列)。在某些具體實施例中，轉移基因組包含有包含 SEQ ID NO:36 之 TRE。

【0118】 在某些具體實施例中，本文中所揭示之轉移基因組進一步包含轉錄終止子(例如聚腺苷酸化序列)。在某些具體實施例中，轉錄終止子位於 ARSA 編碼序列的 3'。轉錄終止子可為任何有效終止轉錄之序列且本領域中熟習此項技術者將理解，可自任何表現於需要 ARSA 編碼序列之轉錄的細胞中之基因分離此類序列。在某些具體實施例中，轉錄終止子包含聚腺苷酸化序列。在某些具體實施例中，聚腺苷酸化序列與人類 ARSA 基因之內源聚腺苷酸化序列一致或實質上一致。在某些具體實施例中，聚腺苷酸化序列為外源聚腺苷酸化序列。在某些具體實施例中，聚腺苷酸化序列為 SV40 聚腺苷酸化序列(例如包含 SEQ ID NO:31、42、43 或 45 中所示之核苷酸序列，或與其互補之核苷酸序列)。在某些具體實施例中，聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列。

【0119】 在某些具體實施例中，轉移基因組自 5'至 3'包含：TRE、ARSA

編碼序列及聚腺苷酸化序列。在某些具體實施例中，TRE 與 SEQ ID NO:25、32、36、54、55 及/或 58 中之任一者具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；ARSA 編碼序列與 SEQ ID NO:14、24、62 或 72 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；及/或聚腺苷酸化序列與 SEQ ID NO:42、43 及 45 中之任一者具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性。

【0120】 在某些具體實施例中，TRE 包含 SEQ ID NO:36 中所示之序列；ARSA 編碼序列包含 SEQ ID NO:14 中所示之序列；及/或聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列。在某些具體實施例中，TRE 自 5'至 3'包含 SEQ ID NO:58 中所示之序列、SEQ ID NO:25 中所示之序列及 SEQ ID NO:32 中所示之序列。

【0121】 在某些具體實施例中，TRE 包含 SEQ ID NO:54 中所示之序列；ARSA 編碼序列包含 SEQ ID NO:62 中所示之序列；及/或聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列。在某些具體實施例中，TRE 包含 SEQ ID NO:55 中所示之序列；ARSA 編碼序列包含 SEQ ID NO:62 中所示之序列；及/或聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列。

【0122】 在某些具體實施例中，TRE 包含 SEQ ID NO:36 中所示之序列；ARSA 編碼序列包含 SEQ ID NO:72 中所示之序列；及/或聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列。在某些具體實施例中，TRE 自 5'至 3'包含 SEQ ID NO:58 中所示之序列、SEQ ID NO:25 中所示之序列及 SEQ ID NO:32 中所示之序列。

【0123】 在某些具體實施例中，轉移基因組進一步包含 hSUMF1 編碼序列。在某些具體實施例中，轉移基因組自 5'至 3'包含：TRE、ARSA 編碼序列、2A 元件及 hSUMF1 編碼序列。在某些具體實施例中，TRE 與 SEQ ID NO:25、32、

36、54、55 及/或 58 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；ARSA 編碼序列與 SEQ ID NO:62 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；2A 元件與 SEQ ID NO:63 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；及 hSUMF1 序列與 SEQ ID NO:64 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性。在某些具體實施例中，進一步包含 hSUMF1 編碼序列之轉移基因組自 5'至 3'包含：包含 SEQ ID NO:54 或 55 中所示之序列之 TRE、包含 SEQ ID NO:62 中所示之序列之 hARSA 編碼序列、包含 SEQ ID NO:63 中所示之序列之 2A 元件及包含 SEQ ID NO:64 中所示之序列之 hSUMF1 編碼序列。在某些具體實施例中，hARSA-2A-hSUMF1 編碼序列包含 SEQ ID NO:30 中所示之序列。

【0124】 在某些具體實施例中，轉移基因組進一步包含 hSapB 編碼序列。在某些具體實施例中，轉移基因組自 5'至 3'包含：TRE、ARSA 編碼序列、2A 元件及 hSapB 編碼序列。在某些具體實施例中，TRE 與 SEQ ID NO: 25、32、36、54、55 及/或 58 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；ARSA 編碼序列與 SEQ ID NO:72 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；2A 元件與 SEQ ID NO:63 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；及 hSapB 序列與 SEQ ID NO:73 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性。在某些具體實施例中，進一步包含 hSapB 編碼序列之轉移基因組自 5'至 3'包含：包含 SEQ ID NO:36 中所示之序列之 TRE、包含 SEQ ID NO:72 中所示之序列之 hARSA 編碼序列、包含 SEQ ID NO:63 中所示之序列之 2A 元件及包含 SEQ ID NO:74 中所示之序列之 hSapB 編碼序列。在某些具體實施例中，hARSA-2A-hSapB 編碼序列包含 SEQ ID NO:74

中所示之序列。

【0125】 在某些具體實施例中，轉移基因組包含與 SEQ ID NO: 41、44、46、65、67 或 75 至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之序列。在某些具體實施例中，轉移基因組包含 SEQ ID NO:41、44、46、65、67 或 75 中所示之核苷酸序列。在某些具體實施例中，轉移基因組之核苷酸序列由 SEQ ID NO: 41、44、46、65、67 或 75 中所示之核苷酸序列組成。在某些具體實施例中，轉移基因組包含 SEQ ID NO:44 中所示之核苷酸序列。在某些具體實施例中，轉移基因組之核苷酸序列由 SEQ ID NO:44 中所示之核苷酸序列組成。

【0126】 在某些具體實施例中，本文中所揭示之轉移基因組進一步包含 TRE 的 5'之 5'反向末端重複（5' ITR）核苷酸序列，及 ARSA 編碼序列的 3'之 3'反向末端重複（3' ITR）核苷酸序列。來自任何 AAV 血清型或其變異體之 ITR 序列皆可用於本文中所揭示之轉移基因組中。5'及 3' ITR 可來自相同血清型之 AAV 或不同血清型之 AAV。用於本文中所揭示之轉移基因組中之例示性 ITR 示於本文中之 SEQ ID NO: 18-21、26 及 27 中。

【0127】 在某些具體實施例中，5' ITR 或 3' ITR 係來自 AAV2。在某些具體實施例中，5' ITR 及 3' ITR 皆來自 AAV2。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO:18 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性，或 3' ITR 核苷酸序列與 SEQ ID NO:19 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO:18 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO:19 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致

性。在某些具體實施例中，轉移基因組包含 SEQ ID NO: 41、44、46、65、67 或 75 中之任一者中所示之核苷酸序列，5' ITR 核苷酸序列具有 SEQ ID NO:18 之序列，且 3' ITR 核苷酸序列具有 SEQ ID NO:19 之序列。

【0128】 在某些具體實施例中，5' ITR 或 3' ITR 係來自 AAV5。在某些具體實施例中，5' ITR 及 3' ITR 皆來自 AAV5。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO:20 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性，或 3' ITR 核苷酸序列與 SEQ ID NO:21 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性。在某些具體實施例中，5' ITR 核苷酸序列與 SEQ ID NO:20 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性，且 3' ITR 核苷酸序列與 SEQ ID NO:21 具有至少 90%（例如至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%）序列一致性。在某些具體實施例中，轉移基因組包含 SEQ ID NO:46-50 中之任一者中所示之核苷酸序列、具有 SEQ ID NO:20 之序列之 5' ITR 核苷酸序列及具有 SEQ ID NO:21 之序列之 3' ITR 核苷酸序列。

【0129】 在某些具體實施例中，5' ITR 核苷酸序列及 3' ITR 核苷酸序列彼此實質上互補（例如除在 5'或 3' ITR 中之 1、2、3、4 或 5 個核苷酸位置處之失配以外彼此互補）。

【0130】 在某些具體實施例中，5' ITR 或 3' ITR 經修飾以減少或消除由 Rep 蛋白質進行之解析（「不可解析的 ITR」）。在某些具體實施例中，不可解析的 ITR 包含末端解析位點之核苷酸序列中之插入、缺失或取代。此類修飾使得能夠在被感染之細胞中複製轉移基因組之後形成 AAV 之自互補、雙股 DNA 基因組。例示性不可解析的 ITR 序列為本領域中已知的（參見例如美國專利案第 7,790,154 號及第 9,783,824 號中提供之序列，其以全文引用之方式併入本文中）。

在某些具體實施例中，5' ITR 包含與 SEQ ID NO:26 至少 95%、96%、97%、98% 或 99%一致之核苷酸序列。在某些具體實施例中，5' ITR 由與 SEQ ID NO:26 至少 95%、96%、97%、98%或 99%一致之核苷酸序列組成。在某些具體實施例中，5' ITR 由 SEQ ID NO:26 中所示之核苷酸序列組成。在某些具體實施例中，3' ITR 包含與 SEQ ID NO:27 至少 95%、96%、97%、98%或 99%一致之核苷酸序列。在某些具體實施例中，5' ITR 由與 SEQ ID NO:27 至少 95%、96%、97%、98%或 99%一致之核苷酸序列組成。在某些具體實施例中，3' ITR 由 SEQ ID NO:27 中所示之核苷酸序列組成。在某些具體實施例中，5' ITR 由 SEQ ID NO:26 中所示之核苷酸序列組成，且 3' ITR 由 SEQ ID NO:27 中所示之核苷酸序列組成。在某些具體實施例中，5' ITR 由 SEQ ID NO:26 中所示之核苷酸序列組成，且 3' ITR 由 SEQ ID NO:19 中所示之核苷酸序列組成。

【0131】 在某些具體實施例中，3' ITR 由來源於野生型 AAV2 基因組序列之其他核苷酸序列側接。在某些具體實施例中，3' ITR 由源自於與野生型 AAV2 ITR 相鄰之野生型 AAV2 序列之另一個 37 bp 序列側接。參見例如 Savy 等人,《人類基因療法方法 (*Human Gene Therapy Methods*) 》(2017) 28(5): 277-289 (其以全文引用之方式併入本文中)。在某些具體實施例中，該另一個 37 bp 序列位於 3' ITR 內部。在某些具體實施例中，37 bp 序列由 SEQ ID NO:56 中所示之序列組成。在某些具體實施例中，3' ITR 包含與 SEQ ID NO:57 至少 95%、96%、97%、98%或 99%一致之核苷酸序列。在某些具體實施例中，3' ITR 包含 SEQ ID NO:57 中所示之核苷酸序列。在某些具體實施例中，3' ITR 之核苷酸序列由與 SEQ ID NO:57 至少 95%、96%、97%、98%或 99%一致之核苷酸序列組成。在某些具體實施例中，3' ITR 之核苷酸序列由 SEQ ID NO:57 中所示之核苷酸序列組成。

【0132】 在某些具體實施例中，轉移基因組自 5'至 3'包含：5' ITR；內部元件，其自 5'至 3'包含：TRE、視情況選用之非編碼外顯子及內含子、ARSA 編碼

序列及聚腺苷酸化序列，如本文中所揭示；不可解析的 ITR；與內部元件互補之核苷酸序列；及 3' ITR。此類轉移基因組可在感染之後及複製之前形成 AAV 之自互補、雙股 DNA 基因組。

【0133】 在某些具體實施例中，轉移基因組自 5'至 3'包含：5' ITR、TRE、ARSA 編碼序列、聚腺苷酸化序列及 3' ITR。在某些具體實施例中，5' ITR 與 SEQ ID:18、20 或 26 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99%序列一致性；TRE 與 SEQ ID NO:25、32、36、54、55 及/或 58 中之任一者具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；ARSA 編碼序列與 SEQ ID NO:14、24、62 或 72 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；聚腺苷酸化序列與 SEQ ID NO:42、43 及 45 中之任一者具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性；及/或 3' ITR 與 SEQ ID:19、21、27 或 57 具有至少 90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%序列一致性。在某些具體實施例中，5' ITR 包含選自由 SEQ ID NO:18、20 及 26 組成之群組之核苷酸序列或由其組成；TRE 包含選自由以下組成之群組之核苷酸序列：SEQ ID NO:25、32、36、54、55 及/或 58；ARSA 編碼序列包含 SEQ ID NO:14、24、62 或 72 中所示之序列；聚腺苷酸化序列包含選自由 SEQ ID NO:42、43 及 45 組成之群組之核苷酸序列；及/或 3' ITR 包含選自由 SEQ ID NO:19、21、27 或 57 組成之群組之核苷酸序列或由其組成。

【0134】 在某些具體實施例中，5' ITR 包含 SEQ ID NO:18 中所示之序列或由其組成；TRE 包含 SEQ ID NO:36 中所示之序列；ARSA 編碼序列包含 SEQ ID NO:14、24、62 或 72 中所示之序列；聚腺苷酸化序列包含 SEQ ID NO:42 中所示之序列；及/或 3' ITR 包含 SEQ ID NO:19 中所示之序列或由其組成。

【0135】 在某些具體實施例中，轉移基因組包含與 SEQ ID NO:47、48、49、

68、69 或 76 至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%一致之序列。在某些具體實施例中，轉移基因組包含 SEQ ID NO:47、48、49、68、69 或 76 中所示之核苷酸序列。在某些具體實施例中，轉移基因組之核苷酸序列由 SEQ ID NO:47、48、49、68、69 或 76 中所示之核苷酸序列組成。在某些具體實施例中，轉移基因組包含 SEQ ID NO:48 中所示之核苷酸序列。在某些具體實施例中，轉移基因組之核苷酸序列由 SEQ ID NO:48 中所示之核苷酸序列組成。

【0136】 在某些具體實施例中，rAAV 包含：（a）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列，及轉移基因組，其自 5'至 3'包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）；（b）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列，及轉移基因組，其自 5'至 3'包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）；及/或（c）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸序列，及轉移基因組，其自 5'至 3'包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動

子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）。

【0137】 在某些具體實施例中，rAAV 包含：（a）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列；（b）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列；及/或（c）AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列。

【0138】 在另一態樣中，本文中提供聚核苷酸，其包含與 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中所示之核酸序列至少 80%（例如至少 85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%）一致之核酸序列。在某些具體實施例中，聚核苷酸包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中所示之核酸序列。在某些具體實施例中，聚核苷酸由 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中所示之核酸序列組成。在某些具體實施例中，聚核苷酸包含 SEQ ID NO:44 或 48 中所示之核酸序列。在某些具體實施例中，聚核苷酸由 SEQ ID NO:44 或 48 中所示之核酸序列組成。

【0139】 本文中亦提供聚核苷酸，其包含與 SEQ ID NO:14、62 或 72 中所示之核酸序列至少 80%（例如至少 85%、90%、91%、92%、93%、94%、95%、

96%、97%、98%或 99%) 一致之核酸序列。在某些具體實施例中，聚核苷酸包含 SEQ ID NO:14、62 或 72 中所示之核酸序列。在某些具體實施例中，聚核苷酸由 SEQ ID NO:14、62 或 72 中所示之核酸序列組成。在某些具體實施例中，聚核苷酸包含 SEQ ID NO:14 中所示之核酸序列。在某些具體實施例中，聚核苷酸由 SEQ ID NO:14 中所示之核酸序列組成。

【0140】 在另一態樣中，本發明提供一種醫藥組成物，其包含如本文中所揭示之 AAV 及醫藥學上可接受之賦形劑、佐劑、稀釋劑、媒劑或載體或其組合。「醫藥學上可接受之載劑」包括任何在與組成物之活性成分組合時使該成分保留生物活性且不引起破壞性生理學反應（諸如不希望的免疫反應）之物質。醫藥學上可接受之載劑包括水、磷酸鹽緩衝生理食鹽水、乳液（諸如油/水乳液）及濕潤劑。包含此類載劑之組成物係藉由熟知的習知方法方法調配，諸如《雷明頓氏醫藥科學（Remington's Pharmaceutical Sciences）》，最新版, Mack Publishing Co., Easton Pa. 18042, USA；A. Gennaro (2000) 《雷明頓：藥劑學之科學及實踐（Remington: The Science and Practice of Pharmacy）》，第 20 版, Lippincott, Williams, & Wilkins；《醫藥劑型及藥物遞送系統（Pharmaceutical Dosage Forms and Drug Delivery Systems）》(1999) H. C. Ansel 等人, 第 7 版, Lippincott, Williams, & Wilkins；及《醫藥賦形劑手冊（Handbook of Pharmaceutical Excipients）》(2000) A. H. Kibbe 等人, 第 3 版, Amer. Pharmaceutical Assoc 中所闡述之方法。

【0141】 在另一態樣中，本發明提供聚核苷酸，其包含編碼人類 ARSA 蛋白質或其片段之編碼序列，其中該編碼序列已以緘默方式改變以與野生型人類 ARSA 基因具有小於 100%（例如小於 95%、90%、85%、80%、75%、70%、65%、60%、55%或 50%）一致性。在某些具體實施例中，聚核苷酸包含 SEQ ID NO:14、62 或 72 中所示之序列。在某些具體實施例中，聚核苷酸由 SEQ ID NO:14、62 或 72 中所示之序列組成。聚核苷酸可包含 DNA、RNA、經修飾之 DNA、經修

飾之 RNA 或其組合。在某些具體實施例中，聚核苷酸為表現載體。

III. 使用方法

【0142】 在另一態樣中，本發明提供用於在細胞中表現 ARSA 多肽之方法。該等方法通常包含用如本文中所揭示之 rAAV 轉導細胞。此類方法在恢復 ARSA 表現方面為高效的。因此，在某些具體實施例中，本文中所揭示之方法涉及用如本文中所揭示之 rAAV 轉導細胞。

【0143】 本文中所揭示之方法可應用於任何在 ARSA 基因中具有突變之細胞。本領域中熟習此項技術者將瞭解，對需要活性內源 ARSA 之細胞尤其感興趣。因此，在某些具體實施例中，將該等方法應用於任何損失內源 ARSA 活性之細胞。在某些具體實施例中，將該方法應用於神經元及/或神經膠細胞。在某些具體實施例中，尤其感興趣的是需要活性內源 ARSA 之神經元及/或神經膠細胞。在某些具體實施例中，將該方法應用於中樞神經系統之細胞及/或周邊神經系統之細胞。在某些具體實施例中，尤其感興趣的是需要活性內源 ARSA 之中樞神經系統及/或周邊神經系統之細胞。在某些具體實施例中，尤其感興趣的是前腦、中腦、後腦、脊髓及其任何組合中之細胞。在某些具體實施例中，尤其感興趣的是選自由以下組成之群組之中樞神經系統區域之細胞：脊髓、運動皮質、感覺皮質、丘腦、海馬區、殼核、小腦（例如小腦核）及其任何組合。在某些具體實施例中，尤其感興趣的是腦中之腦橋及髓質、脊髓之上升束及其任何組合中之細胞。在某些具體實施例中，尤其感興趣的是選自由以下組成之群組之中樞神經系統區域中的需要活性內源 ARSA 之細胞：脊髓、運動皮質、感覺皮質、丘腦、海馬區、殼核、小腦（例如小腦核）及其任何組合。在某些具體實施例中，尤其感興趣的是中樞神經系統（CNS）中之運動神經元及星形膠質細胞分佈、CNS 中之寡樹突神經膠質細胞（上升纖維）、CNS 中之大腦皮質之細胞群體以及周邊神經系統（PNS）中之感覺神經元。在某些具體實施例中，

尤其感興趣的是寡樹突神經膠質細胞，諸如脊髓之背束中之細胞。在某些具體實施例中，尤其感興趣的是中樞神經系統中之神經膠質分佈，包括（但不限於）星形膠質細胞、寡樹突神經膠質細胞、許旺細胞及其任何組合。在某些具體實施例中，尤其感興趣的是運動神經元、星形膠質細胞、寡樹突神經膠質細胞、中樞神經系統中之大腦皮質之細胞、周邊神經系統中之感覺神經元、周邊神經系統中之神經膠細胞（例如許旺細胞）及其任何組合。

【0144】 本文中所揭示之方法可活體外進行以用於研究目的或可離體或活體內進行以用於治療目的。

【0145】 在某些具體實施例中，待轉導之細胞在哺乳動物個體中且以可有效轉導個體中之細胞之量向個體投與 AAV。因此，在某些具體實施例中，本發明提供一種用於治療患有與 ARSA 基因突變相關之疾病或病症的個體之方法，該方法通常包含向個體投與有效量之如本文中所揭示之 rAAV。個體可為具有 ARSA 突變之人類個體、非人類靈長類動物個體（例如食蟹獼猴）或啮齒動物個體（例如小鼠）。可使用本文中所揭示之方法治療任何與 ARSA 基因突變相關之疾病或病症。適合的疾病或病症包括（但不限於）異染性白質失養症。

【0146】 在某些具體實施例中，前述方法使用包含以下之 rAAV：(a) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列，及轉移基因組，其自 5' 至 3' 包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）；(b) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列，及轉移基因組，其自 5' 至 3'

包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）；及/或 (c) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸序列，及轉移基因組，其自 5' 至 3' 包含以下基因元件：5' ITR 元件（例如 SEQ ID NO:18 之 5' ITR）、強化子元件（例如 SEQ ID NO:58 之強化子元件）、啟動子序列（例如 SEQ ID NO:25 之啟動子序列）、嵌合內含子序列（例如 SEQ ID NO:32 之嵌合內含子序列）、以緘默方式改變之人類 ARSA 編碼序列（例如 SEQ ID NO:14 之 hARSA 編碼序列）、SV40 聚腺苷酸化序列（例如 SEQ ID NO:42 之 SV40 聚腺苷酸化序列）及 3' ITR 元件（例如 SEQ ID NO:19 之 3' ITR）。

【0147】 在某些具體實施例中，前述方法使用包含以下之 rAAV：(a) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 203-736 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列；(b) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸 138-736 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列；及/或 (c) AAV 衣殼蛋白質，其包含 SEQ ID NO:16 之胺基酸序列，及轉移基因組，其包含 SEQ ID NO:41、44、46、47、48、49、65、67、68、69、75 或 76 中之任一者中所示之核苷酸序列。

【0148】 本文中所揭示之方法的有利之處尤其在於其能夠活體內及活體外在細胞中以高效率表現 ARSA 蛋白質。在某些具體實施例中，ARSA 蛋白質

之表現量為在 ARSA 基因中不具有突變之相同類型的細胞中之內源 ARSA 蛋白質之表現量之至少 10%、20%、30%、40%、50%、60%、70%、80%、90% 或 100%。在某些具體實施例中，ARSA 蛋白質之表現量比在 ARSA 基因中不具有突變之相同類型的細胞中之內源 ARSA 蛋白質之表現量高至少 1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、2、3、4、5、6、7、8、9 或 10 倍。可使用任何用於測定 ARSA 蛋白質之表現量之方法，包括(但不限於)ELISA、西方墨點法(Western blotting)、免疫染色及質譜。

【0149】 在某些具體實施例中，用本文中所揭示之 AAV 組成物轉導細胞可如本文所提供來進行或藉由本領域中一般熟習此項技術者已知的任何轉導方法來進行。在某些具體實施例中，可使細胞在以下感染倍率(MOI)下與 AAV 接觸：50,000；100,000；150,000；200,000；250,000；300,000；350,000；400,000；450,000；或 500,000，或任何提供細胞之最佳轉導之 MOI。

【0150】 可藉由任何適合的途徑向個體投與本文中所揭示之 AAV 組成物，包括(但不限於)靜脈內、鞘內、腹膜內、皮下、肌肉內、鼻內、局部或皮內途徑。在某些具體實施例中，組成物經調配以用於經由靜脈內注射或皮下注射來投藥。

IV. AAV 封裝系統

【0151】 在另一態樣中，本發明提供用於本文中所揭示之重組腺相關病毒(rAAV)之重組製備之封裝系統。此類封裝系統通常包含：第一核苷酸，其編碼一或多種 AAV Rep 蛋白質；第二核苷酸，其編碼如本文中所揭示之 AAV 中之任一者之衣殼蛋白質；及第三核苷酸序列，其包含如本文中所揭示之 rAAV 基因組中之任一者，其中封裝系統在細胞中具有操作性以用於將轉移基因組包裹在衣殼中以形成 AAV。

【0152】 在某些具體實施例中，封裝系統包含第一載體，其包含編碼一或

多種 AAV Rep 蛋白質之第一核苷酸序列及編碼 AAV 衣殼蛋白質之第二核苷酸序列，及第二載體，其包含有包含 rAAV 基因組之第三核苷酸序列。如在如本文中所描述之封裝系統之情形中所使用，「載體」係指一種核酸分子，其為用於將核酸引入細胞之媒劑（例如質體、病毒、黏質體、人工染色體等）。

【0153】 任何 AAV Rep 蛋白質皆可用於本文中所揭示之封裝系統中。在封裝系統之某些具體實施例中，Rep 核苷酸序列編碼 AAV2 Rep 蛋白質。適合的 AAV2 Rep 蛋白質包括（但不限於）Rep 78/68 或 Rep 68/52。在封裝系統之某些具體實施例中，編碼 AAV2 Rep 蛋白質之核苷酸序列包含編碼與 SEQ ID NO:22 之 AAV2 Rep 胺基酸序列具有最小序列一致性百分比之蛋白質之核苷酸序列，其中跨越 AAV2 Rep 蛋白質之胺基酸序列之長度，最小序列一致性百分比為至少 70%（例如至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 98%、至少 99%或 100%）。在封裝系統之某些具體實施例中，AAV2 Rep 蛋白質具有 SEQ ID NO:22 中所示之胺基酸序列。

【0154】 在封裝系統之某些具體實施例中，封裝系統進一步包含第四核苷酸序列，其包含一或多種輔助病毒基因。在封裝系統之某些具體實施例中，封裝系統進一步包含第三載體，例如輔助病毒載體，其包含有包含一或多種輔助病毒基因之第四核苷酸序列。第三載體可為獨立的第三載體、與第一載體整合在一起或與第二載體整合在一起。

【0155】 在封裝系統之某些具體實施例中，輔助病毒係選自由以下組成之群組：腺病毒、疱疹病毒（包括單純疱疹病毒（HSV））、痘病毒（諸如痘瘡病毒）、細胞巨大病毒（CMV）及桿狀病毒。在封裝系統之某些具體實施例中，其中輔助病毒為腺病毒，腺病毒基因組包含一或多種選自由以下組成之群組之腺病毒 RNA 基因：E1、E2、E4 及 VA。在封裝系統之某些具體實施例中，其中輔助病毒為 HSV，HSV 基因組包含選自由以下組成之群組之 HSV 基因中之一或

多者：UL5/8/52、ICPO、ICP4、ICP22 及 UL30/UL42。

【0156】 在封裝系統之某些具體實施例中，第一、第二及/或第三載體包含於一或多種質體內。在某些具體實施例中，第一載體及第三載體包含於第一質體內。在某些具體實施例中，第二載體及第三載體包含於第二質體內。

【0157】 在封裝系統之某些具體實施例中，第一、第二及/或第三載體包含於一或多種重組輔助病毒內。在某些具體實施例中，第一載體及第三載體包含於重組輔助病毒內。在某些具體實施例中，第二載體及第三載體包含於重組輔助病毒內。

【0158】 在另一態樣中，本發明提供一種用於如本文中所描述之 AAV 之重組製備之方法，其中該方法包含在可操作以將 rAAV 基因組包裹於衣殼中以形成如本文中所描述之 rAAV 之條件下，用如本文中所描述之封裝系統轉染或轉導細胞。用於 rAAV 之重組製備之例示性方法包括短暫轉染（例如用一或多種含有如本文中所描述之第一及第二以及視情況選用之第三載體之轉染質體）、病毒感染（例如用一或多種重組輔助病毒，諸如腺病毒、痘病毒（諸如痘瘡病毒）、疱疹病毒（包括 HSV、細胞巨大病毒或桿狀病毒，其含有如本文中所描述之第一及第二以及視情況選用之第三載體）及穩定的生產者細胞株轉染或感染（例如用穩定的生產者細胞，諸如哺乳動物或昆蟲細胞，其含有如本文中所描述之編碼一或多種 AAV Rep 蛋白質之 Rep 核苷酸序列及/或編碼一或多種衣殼蛋白質之 Cap 核苷酸序列，且其中以質體或重組輔助病毒形式遞送如本文中所描述之轉移基因組）。

【0159】 因此，本發明提供用於製備重組 AAV（rAAV）之封裝系統，其中該封裝系統包含第一核苷酸序列，其編碼一或多種 AAV Rep 蛋白質；第二核苷酸序列，其編碼本文中所描述之 AAV 中之任一者之衣殼蛋白質；第三核苷酸序列，其包含本文中所描述之 AAV 中之任一者之 rAAV 基因組序列；及視情況

之第四核苷酸序列，其包含一或多種輔助病毒基因。

V. 實例

【0160】 本文中所揭示之重組 AAV 載體活體外及活體內介導高效基因轉移。以下實例說明使用如本文中所揭示之基於 AAV 之載體進行之 ARSA 基因(其在某些人類疾病(諸如異染性白質失養症)中突變)之表現之有效恢復。提供此等實例作為說明，而非作為限制。

實例 1：人類 ARSA 轉移載體

【0161】 此實例提供用於經載體轉導之細胞(例如人類細胞或小鼠細胞)中之人類 ARSA (hARSA) 之表現之人類 ARSA 轉移載體 T-001、pHMI-5000、pHMI-5003 及 pHMI-hARSA1-TC-002。

a) T-001

【0162】 如圖 1A 中所示，ARSA 轉移載體 TC-001 自 5'至 3'包含以下基因元件：5' ITR 元件、轉錄調節元件(其包含 CMV 強化子元件、雞 β -肌動蛋白啟動子及嵌合內含子序列)；野生型人類 ARSA 編碼序列；SV40 聚腺苷酸化序列；及 3' ITR 元件。此等元件之序列闡述於表 1 中。此載體能夠在經載體轉導之細胞(例如人類細胞或小鼠細胞)中表現人類 ARSA 蛋白質。

b) pHMI-5000

【0163】 如圖 1B 中所示，ARSA 轉移載體 pHMI-5000 自 5'至 3'包含以下基因元件：5' ITR 元件；轉錄調節元件，其包含 CMV 強化子元件、雞 β -肌動蛋白啟動子及嵌合內含子序列；以緘默方式改變之人類 ARSA 編碼序列；SV40 聚腺苷酸化序列；及 3' ITR 元件。此等元件之序列闡述於表 1 中。此載體能夠在經載體轉導之細胞(例如人類細胞或小鼠細胞)中表現人類 ARSA 蛋白質。

c) pHMI-5003

【0164】 如圖 1C 中所示，ARSA 轉移載體 pHMI-5003 自 5'至 3'包含以下

基因元件：5' ITR 元件；轉錄調節元件，其包含 CMV 強化子元件、雞 β-肌動蛋白啟動子及嵌合內含子序列；以緘默方式改變之人類 ARSA 編碼序列；SV40 聚腺苷酸化序列；非編碼填充序列，及 3' ITR 元件。此等元件之序列闡述於表 1 中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 蛋白質。

d) pHMI-hARSA1-TC-002

【0165】如圖 1D 中所示，ARSA 轉移載體 pHMI-hARSA1-TC-002 自 5'至 3'包含與 pHMI-5000 相同之基因元件。此等元件之序列闡述於表 1 中。pHMI-hARSA1-TC-002 與 pHMI-5000 之間的差異位於載體主鏈序列中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 蛋白質。

表 1：人類 ARSA 轉移載體 T-001、pHMI-5000、pHMI-5003 及 pHMI-hARSA1-TC-002 中之基因元件

基因元件	T-001	pHMI-5000	pHMI-5003	pHMI-hARSA1-TC-002
	SEQ ID NO:			
5' ITR 元件	18	18	18	18
強化子元件	58	58	58	58
啟動子序列	25	25	25	25
內含子序列	32	32	32	32
轉錄調節元件	36	36	36	36
人類 ARSA 編碼序列	24	14	14	14
SV40 聚腺苷酸化序列	42	42	42	42
填充序列	N/A	N/A	39	N/A
3' ITR 元件	19	19	19	19
轉移基因組 (自啟動子至聚腺苷酸化序	41	44	46	44

基因元件	T-001	pHMI-5000	pHMI-5003	pHMI-hARSA1-TC-002
	SEQ ID NO:			
列)				
轉移基因組 (自 5' ITR 至 3' ITR)	47	48	49	48
完全載體序列	50	51	52	53

【0166】 本文中所揭示之載體可封裝於 AAV 衣殼中，諸如（但不限於）AAVHSC5、AAVHSC7、AAVHSC15 或 AAVHSC17 衣殼。可將經封裝之病毒顆粒投與野生型動物或 ARSA 缺失型動物。

實例 2：ARSA (-/-) 小鼠模型中之 ARSA 基因轉移

【0167】 為了研究小鼠中 ARSA 基因轉移之作用，產生 ARSA (-/-) 小鼠模型。ARSA (-/-) 小鼠模型為藉由將新黴素卡匣插入小鼠 ARSA 基因之外顯子 4 而產生之 ARSA 基因剔除小鼠（參見 Hess 等人,《美國國家科學院院刊 (*Proc. Natl. Acad. Sci. U.S.A.*) 》 1996, 93(25):14821-14826，其以全文引用之方式併入本文中）。與人類相比，ARSA (-/-) 小鼠發展類似的但程度較輕的異染性白質失養症 (MLD)。ARSA (-/-) 小鼠未顯示廣泛的髓鞘脫失之證據。

【0168】 可使用各種生物標記研究 MLD。舉例而言，可量測腦中髓磷脂之含量。已報導寡樹突神經膠細胞 (C24:0) 及神經元 (C18:0) 髓磷脂之增加，其中積聚隨動物年齡而增加。可量測髓鞘及淋巴細胞蛋白質 (MAL) mRNA 轉錄物之含量。MAL 由寡樹突神經膠質細胞及許旺細胞表現，使神經膠質-軸突接合區穩定且與 MLD 之病變有關。已報導 ARSA (-/-) 小鼠中之 MAL 轉錄物之含量降低。溶酶體相關膜蛋白質 (LAMP-1) 為另一種可用於研究 MLD 之生物標記。已使用抗 LAMP-1 抗體，藉由對 ARSA (-/-) 及野生型小鼠中之脊髓組織進行免疫組織化學來研究 LAMP-1 免疫反應性，證實 ARSA (-/-) 小鼠中之 LAMP-1 免疫反應性提高。圖 2A 顯示來源於藉由對來自 ARSA (-/-) 小鼠之脊

髓組織進行免疫組織化學 (IHC) 而研究的 LAMP-1 免疫反應性之總像素強度之定量。在用媒劑對照物或封裝於 AAVHSC15 衣殼中之 pHMI-5000 處理的 ARSA (-/-) 小鼠中使用抗 LAMP-1 抗體進行 IHC。如圖 2A 中所示，在給藥 (4e13 vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000) 後第 12 週時，偵測到與投與媒劑對照物之 ARSA (-/-) 動物相比顯著降低之 LAMP-1 含量。

【0169】 將腦組織稱重且在 Precellys 珠粒均質器中，在 250 μ L 水中均質化，且移出均質物之 10 μ L 等分試樣以用於 Pierce BCA 蛋白質分析法定量。將 760 μ L 乙腈將添加至各均質物中且將混合物第二次均質化。均質物在 14,000 x g 下離心 15 分鐘且移出離心機澄清之上清液且在 75% 乙腈中稀釋 5 倍以用於 RapidFire-MS 分析。使用 C19:0 髓磷脂 (Matreya 目錄號 1888) 作為內標且用 Sciex API4000 三重四極質譜儀以 MRM 模式與 C18:0、C18:1、C24:0 及 C24:1 髓磷脂一起監測。各樣品與 8 種不同濃度之 C19:0 髓磷脂 IS 一起注射 8 次以產生各樣品之獨特標準曲線，使用其計算各分析物之濃度。圖 2B 顯示隨時間推移，對照組小鼠 (WT/Het) 及 ARSA (-/-) 小鼠之大腦中之 C18:0 髓磷脂含量。對照組為野生型動物 (ARSA (+/+)) 與雜合動物 (ARSA (+/-)) 之混合。如圖 2B 中所示，ARSA (-/-) 小鼠之腦中之 C18:0 髓磷脂含量隨時間推移而增加，而對照組小鼠之腦中之 C18:0 髓磷脂含量隨時間推移而保持基本不變。由兩隻對照組小鼠及兩隻 ARSA (-/-) 小鼠之分析產生圖 2B 中之資料。為了研究 ARSA 缺失型小鼠中 ARSA 基因遞送對髓磷脂積聚之作用，用 4e13 vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-hARSA1-TC-002 處理 ARSA (-/-) 小鼠 (圖 2C)。如圖 2C 中所示，與用媒劑對照物處理之 ARSA (-/-) 小鼠相比，在給藥後七個月時觀測到經處理之 ARSA (-/-) 小鼠中之腦部髓磷脂含量之顯著降低。

【0170】 在用 4e13 vg/kg 及 6e13 vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 或媒劑對照物處理後七個月時測定 ARSA (-/-) 小鼠之前腦、中腦及

後腦中之 C18:0 及 C18:1 髓磷脂同功異型物含量（圖 2D）。髓磷脂同功異型物含量呈現為與相同年齡之野生型對照動物相比之倍數。如圖 2D 中所示，與用煤劑對照物處理之 ARSA(-/-)小鼠相比，在給藥後七個月時觀測到經處理之 ARSA(-/-)小鼠之所有三個腦部區域中之腦部髓磷脂含量之顯著降低。所使用之方法及材料與上述相同。使用非配對 T 測試（unpaired T-test）分析資料。

【0171】在用 4×10^{13} vg/kg 之封裝於 AAVHS15 衣殼中之 pHMI-5000 或煤劑對照物處理後第 52 週測定 ARSA(-/-)小鼠之前腦、中腦及後腦中之 C18:0 及 C18:1 髓磷脂同功異型物含量（圖 2E）、C24:0 及 C24:1 髓磷脂同功異型物含量（圖 2F）及總髓磷脂同功異型物含量（圖 2G）。所使用之方法及材料與上述相同。使用非配對 T 測試分析資料。

【0172】圖 3A 顯示對照組小鼠（WT/Het）及 ARSA(-/-)小鼠中，在四週時 MAL 轉錄物之含量。對照組為野生型動物(ARSA(+/+))與雜合動物(ARSA(+/-))之混合。藉由 Trizol 提取，接著進行 Qiagen RNEasy 管柱純化來製備小鼠總 RNA。RNA 用作使用 ThermoFisher High Capacity cDNA 套組進行之 cDNA 合成之模板以產生轉錄物。使用液滴式數位 PCR 及對小鼠髓鞘及淋巴細胞蛋白質(MAL)具有特異性之引子/探針集合評估 MAL 轉錄物，其中針對小鼠 HPRT1 標準化複本數。如所顯示，在四週時，與雜合小鼠相比，ARSA(-/-)小鼠中之 MAL 轉錄物之含量降低。由五隻對照組小鼠及六隻 ARSA(-/-)小鼠之分析產生圖 3 中之資料。為了研究 ARSA 缺失型小鼠中 ARSA 基因遞送對 MAL 轉錄物含量之影響，用 4×10^{13} vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 處理 ARSA(-/-)小鼠（圖 3B）。如圖 3B 中所示，與野生型小鼠及用煤劑處理之 ARSA(-/-)小鼠相比，在給藥後三個月時觀測到經處理之 ARSA(-/-)小鼠中之 MAL 轉錄物含量之顯著增加。

【0173】測定用 4×10^{13} vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 處

理的 ARSA (-/-) 小鼠中之 MAL 轉錄物複本數之水準 (圖 3C)。圖 3C 顯示在野生型小鼠中或在被投與媒劑對照物或 4e13 vg/kg 之封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA (-/-) 小鼠中，在給藥後第 12 或 52 週時偵測到的 MAL 轉錄物之複本數。所使用之方法及材料與上述相同。使用非配對 T 測試分析資料。在圖 3C 中，各動物組之間的統計顯著性如下：12 週媒劑對比經處理之動物， $p=0.0012$ ；12 週處理對比野生型動物， $p<0.0001$ ；52 週媒劑對比經處理之動物， $p=0.0004$ ；及 52 週處理對比野生型動物，不顯著。

【0174】為了研究是否可獲得 hARSA 活性之治療水準，向 ARSA (-/-) 小鼠投與封裝於 AAV9 衣殼中之轉移載體 T-001 (參見 PCT 公開案第 WO2002/052052 號，其以全文引用之方式併入本文中)。自未經處理之對照性 ARSA (-/-) 小鼠及被投與封裝於 AAV9 衣殼中之轉移載體 T-001 之 ARSA (-/-) 小鼠獲得之腦薄片之抗 ARSA 免疫反應性表明，在 2e13 載體基因組/公斤體重 (vg/kg) 之劑量下獲得達到治療水準(10%)之 hARSA 酶活性。自經處理之 ARSA (-/-) 小鼠獲得之腦薄片之抗 ARSA 免疫反應性亦顯示腦中之 ARSA 酶活性之劑量依賴性增加。

實例 3：ARSA (-/-) 小鼠模型中之 ARSA 基因轉移

【0175】此實例提供與人類 ARSA 轉移載體 pHMI-5000 之用途相關之實驗資料。如本文中所描述，轉移載體 pHMI-5000 包含以緘默方式改變之人類 ARSA 編碼序列，證實其呈現顯著改良之 ARSA 蛋白質之表現。

【0176】圖 4 為顯示腦部中以每個經轉導之細胞計的載體基因組之數目與以每奈克 cDNA 計的 hARSA 之複本數目之間的相關性之作圖。使用來自 Qiagen 之 QIAamp Fast DNA Tissue 套組製備小鼠基因組 DNA。藉由液滴式數位 PCR 及對經密碼子最佳化之人類 ARSA 載體基因組之編碼區具有特異性之引子/探針集合測定 VG 計數，其中針對內源小鼠基因組序列進行標準化。如本文中所描述

製備小鼠總 RNA 且使用液滴式數位 PCR 及用於測定 VG 計數之相同引子/探針集合評估 ARSA 轉錄物，其中針對小鼠 GUSB 標準化複本數。如所顯示，對於使用封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 轉導之細胞，所偵測到的以每個經轉導之細胞計的載體基因組之數目與以每奈克 cDNA 計的 hARSA 之複本數目強相關 ($R^2=0.9332$)。

【0177】發現在 AAVHSC15 與 AAV9 衣殼介導之遞送之間的比較中，AAVHSC15 在腦部中顯著優於 AAV9。圖 5 顯示在 $2e13$ vg/kg 之封裝於 AAV9 或 AAVHSC15 衣殼中之轉移載體 pHMI-5000 的劑量下，腦部中之以每個經轉導之細胞計的載體基因組之數目。如所顯示，與 AAV9 衣殼相比，當轉移載體 pHMI-5000 封裝於 AAVHSC15 衣殼中時，觀測到高十倍的以每個細胞計的載體基因組計數。圖 6 顯示對於以所指示之劑量投與之封裝於 AAV9 或 AAVHSC15 衣殼中之轉移載體 pHMI-5000，所量測的正常人類 ARSA 酶活性水準之百分比。圖 7 顯示在以 $4e13$ vg/kg 投與封裝於 AAV9 或 AAVHSC15 中之轉移載體 pHMI-5000 之小鼠中，以每個經轉導之腦細胞計的載體基因組之數目。

【0178】與封裝於 AAV9 衣殼中之 pHMI-5000 相比，封裝於 AAVHSC15 衣殼中之 pHMI-5000 顯示更強且更廣泛的腦部及脊髓表現概況。抗 ARSA 免疫反應性實驗表明，在 $3e13$ vg/kg 之劑量下，在各種情況下，與靜脈內投與封裝於 AAV9 衣殼中之 pHMI-5000 之小鼠相比，在靜脈內投與封裝於 AAVHSC15 衣殼中之 pHMI-5000 之小鼠的腦薄片上偵測到顯著更高的量。

【0179】為了評估投藥途徑對腦部中 hARSA 之生物分佈之影響，分別在 $4e13$ vg/kg 及 $4e12$ vg/kg 之劑量下經由靜脈內 (IV) 及鞘內 (IT) 途徑投與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000。在向 ARSA (-/-) 小鼠靜脈內投與封裝於 AAVHSC15 中之 pHMI-5000 之後，在重要的中樞神經系統區域中存在抗 ARSA 免疫反應性。廣泛偵測到抗小鼠 ARSA (mARSA) 或人類 ARSA

(hARSA)，包括（但不限於）運動及感覺皮質、海馬區（CA3 區域）、殼核及小腦。在封裝於 AAVHSC15 中之轉移載體 pHMI-5000 之 IV 或 IT 投藥之後，後腦及中腦中之正常人類 ARSA 酶活性之百分比之定量顯示於圖 8 中。

【0180】在持續 4 週以 $4e13$ vg/kg 投與封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA (-/-) 小鼠中，在腦部之重要生理學區域以及中樞神經系統 (CNS) 之整個後尾軸中偵測到 hARSA 之生物相關分佈。使用抗 hARSA 抗體偵測 hARSA，且在脊髓、運動皮質、丘腦、海馬區及小腦核中偵測到。亦在以下中偵測到 hARSA：CNS 中之運動神經元及星形膠質細胞分佈；CNS 中之寡樹突神經膠質細胞（在上升纖維中大量偵測到）；CNS 中之大腦皮質之細胞群；以及周邊神經系統 (PNS) 中之感覺神經元及許旺細胞。最早可在處理後第 2 週便偵測到類似的生物學分佈。

【0181】在以 $2e13$ vg/kg 投與封裝於 AAVHSC15 衣殼中之 pHMI-5000 之小鼠中，觀測到與在以 $4e13$ vg/kg 或更高的劑量進行投藥之小鼠中所發現相同的組織學分佈。在此等實驗中，以溶酶體特有之點狀模式在細胞質中偵測到 hARSA。

【0182】如圖 9A 及 9B 中所示，在給藥後第 4 週，經處理之 ARSA (-/-) 小鼠之大腦中的人類 ARSA 酶活性之生理學水準得到恢復。使用來自 ARSA (-/-) 小鼠之腦溶解物評估 hARSA 酶活性。劑量範圍發現研究表明，hARSA 酶活性與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 的靜脈內投藥之劑量相關。在經處理之動物中，但未在媒劑對照性動物中偵測到酶活性。對於所測試之劑量，酶活性水準（約 40-145%）顯著高於治療目標（約 10-15%），如先前在臨床上測定（參見 Patil 及 Maegawa, 《藥物設計及發展及療法 (*Drug Des. Devel. Ther.*) 》2013, 7:729-745）。圖 9A 顯示藉由以所指示之劑量向 ARSA (-/-) 小鼠投與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 而實現的正常 hARSA

活性之百分比。如所顯示，獲得 hARSA 活性之劑量依賴性反應。圖 9B 顯示以所指示之劑量投與封裝於 AAVHSC15 衣殼中之轉移載體 pHMI-5000 的 ARSA (-/-) 小鼠之腦部中以每個細胞計的載體基因組之數目。對於 1×10^{13} vg/kg、 4×10^{13} vg/kg 及 6×10^{13} vg/kg 劑量， $n=5$ 隻小鼠。對於 2×10^{13} vg/kg 劑量， $n=4$ 隻小鼠。所有小鼠為 5 週齡且皆為雄性。在圖 9C 中，使用比色芳基硫酸酯酶 A 特異性分析法評估 ARSA 酶活性，該分析法量測由可溶性受質對硝基兒茶酚-硫酸酯 (pNCS) 引起之硫酸酯之裂解。使用芳基硫酸酯酶 A 特異性免疫沈澱步驟消除由競爭酶引起之硫酸酯之非特異性裂解。藉由各兩位正常人類男性及女性之額葉皮質中的 ARSA 酶活性之分析來測定腦部中之正常人類 ARSA 酶活性。人類額葉皮質樣品係自 BioiVT 購得且在每個 ARSA 酶活性分析盤上與測試樣品一起一式三份地操作。資料表示為以每小時每毫克蛋白質計的脫硫酸化 pNCS 之平均量 (ng) 之百分比。圖 9C 顯示最早在處理後第 1 週且直至處理後第 12 週，單次靜脈內 4×10^{13} vg/kg 劑量之封裝於 AAVHSC15 衣殼中之 pHMI-5000 引起在新生 ARSA (-/-) 小鼠之大腦中偵測到 hARSA 酶活性，其水準超過既定人類治療目標 (10-15%) (如虛線所指示)。在給藥後第 1、2、3、4 及 12 週收集材料。對於每個時間點， $n=6$ 隻小鼠，3 隻雄性及 3 隻雌性，8 週齡。

【0183】在圖 9D 中，藉由 Trizol 提取，接著進行 Qiagen RNEasy 管柱純化來製備小鼠總 RNA。RNA 用作使用 ThermoFisher High Capacity cDNA 套組進行之 cDNA 合成之模板以產生轉錄物。使用液滴式數位 PCR 及對經密碼子最佳化之人類 ARSA 轉錄物具有特異性之引子/探針集合評估 ARSA 轉錄物，其中針對小鼠 GUSB 標準化複本數。圖 9D 顯示最早在處理後第 1 週，單次靜脈內 4×10^{13} vg/kg 劑量之封裝於 AAVHSC15 衣殼中之 pHMI-5000 引起在成年 ARSA (-/-) 小鼠之大腦中偵測到正常水準之 hARSA 酶活性 (經由 hARSA 轉錄物分析)。在給藥後第 2 至 3 週觀測到 hARSA 酶活性之峰值水準，接著穩定的平線區持續

至處理後第 52 週（以超過既定人類治療目標（10-15%）之水準）。在給藥後第 1、2、3、4、8、12、26 及 52 週收集材料。圖 9E 顯示在投與單次靜脈內 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA(-/-)小鼠之大腦中，以每微克基因組 DNA 計的載體基因組之數目。在給藥後第 1、2、3、8、12、26 及 52 週收集材料。圖 9F 顯示在投與單次靜脈內 4×10^{13} vg/kg 之劑量的封裝於 AAVHSC15 衣殼中之 pHMI-5000 之 ARSA(-/-)小鼠之大腦中，以每奈克 RNA 計的 ARSA 轉錄物之複本數目。在給藥後第 4、8、12、26 及 52 週收集材料。

實例 4：人類 ARSA 轉移載體

【0184】此實例提供用於在經載體轉導之細胞（例如人類細胞或小鼠細胞）中的 hARSA 之表現之人類 ARSA 轉移載體 TC-013.pHMIA2 及 TC-015.pKITR。除表現 hARSA 以外，此等載體經設計以亦表現人類 SUMF1。hARSA 及 hSUMF1 之編碼序列由 2A 元件分離。在某些具體實施例中，核糖體跳躍元件（例如 2A 元件）編碼在 N 端進一步包含 Gly-Ser-Gly 之序列之肽，視情況其中 Gly-Ser-Gly 之序列由 GGCAGCGGA 之核苷酸序列編碼。儘管不希望受理論約束，但假設核糖體跳躍元件藉由以下發揮功能：終止第一肽鏈之轉譯及再起始第二肽鏈之轉譯；或由核糖體跳躍元件編碼之肽序列中之肽鍵藉由經編碼之肽之內源蛋白酶活性進行之裂解，或環境（例如胞溶質）中之另一種蛋白酶。

a) TC-013.pHMIA2

【0185】如圖 10A 中所示，ARSA 轉移載體 TC-013.pHMIA2 自 5'至 3'包含以下基因元件：5' ITR 元件、包含 CALM1 啟動子之轉錄調節元件；以緘默方式改變之人類 ARSA 編碼序列；2A 元件；以緘默方式改變之人類 SUMF1 編碼序列；及 3' ITR 元件。此等元件之序列闡述於表 2 中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 蛋白質及人類 SUMF1

蛋白質。

b) TC-015.pKITR

【0186】 如圖 10B 中所示，ARSA 轉移載體 TC-015.pKITR 自 5'至 3'包含以下基因元件：5' ITR 元件、包含 smCBA 啟動子之轉錄調節元件；以緘默方式改變之人類 ARSA 編碼序列；2A 元件；以緘默方式改變之人類 SUMF1 編碼序列；及 3' ITR 元件。此等元件之序列闡述於表 2 中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 蛋白質及人類 SUMF1 蛋白質。

表 2：人類 ARSA 轉移載體 TC-013.pHMIA2 及 TC-015.pKITR 中之基因元件

基因元件	TC-013.pHMIA2	TC-015.pKITR
	SEQ ID NO:	
5' ITR 元件	18	18
啟動子序列	54	55
轉錄調節元件	54	55
人類 ARSA 編碼序列	62	62
2A 元件	63	63
人類 SUMF1 編碼序列	64	64
hARSA-2A-hSUMF1 序列	30	30
3' ITR 元件	19	19
轉移基因組（自啟動子至 SUMF1 編碼序列）	65	67
轉移基因組（自 5' ITR 至 3' ITR）	68	69
完全載體序列	70	71

【0187】 本文中所揭示之載體可封裝於 AAV 衣殼中，諸如（但不限於）AAVHSC5、AAVHSC7、AAVHSC15 或 AAVHSC17 衣殼。可將經封裝之病毒顆粒投與野生型動物或 ARSA 缺失型動物。

【0188】 為了評估腦部中啟動子對 hARSA 表現之影響，將轉移載體

pHMI-5000、TC-013.pHMIA2 及 TC-015.pKITR 封裝於 AAVHSC15 衣殼中且靜脈內投與 ARSA (-/-) 小鼠。在以 4×10^{13} vg/kg 之劑量投與 pHMI-5000 載體（雞 β -肌動蛋白（CBA）啟動子）及以 8×10^{13} vg/kg 之劑量投與 TC-015.pKITR（smCBA 啟動子）之情況下在腦部中偵測到 hARSA 表現及酶活性，其中以每個細胞計數計的病毒基因組類似。與所測試之其他啟動子相比，CBA 啟動子在最低劑量下引起 hARSA 之最高表現。圖 11 顯示 pHMI-5000(CBA 啟動子)、TC-013.pHMIA2 (CALM1 啟動子) 及 TC-015.pKITR (smCBA 啟動子) 之以每個細胞計的經轉導之病毒基因組之數目，在每種情況下封裝於 AAVHSC15 衣殼中且以 4×10^{13} vg/kg 之劑量投與（對於每種載體， $n=5$ 隻小鼠）。圖 12 顯示所偵測之 pHMI-5000(CBA 啟動子) 及 TC-015.pKITR (smCBA 啟動子) 之正常人類 ARSA 酶活性之百分比，在每種情況下封裝於 AAVHSC15 衣殼中且以 4×10^{13} vg/kg 之劑量投與（對於每種載體， $n=5$ 隻小鼠）。圖 13 顯示對於封裝於 AAVHSC15 衣殼中且以 4×10^{13} vg/kg 之劑量投與之 pHMI-5000(CBA 啟動子) 及封裝於 AAVHSC15 衣殼中且以 8×10^{13} vg/kg 之劑量投與之 TC-015.pKITR (smCBA 啟動子)，在西方墨點法中，可在使用抗 hARSA 抗體之小鼠之大腦中偵測到 hARSA 之表現（對於每種載體， $n=5$ 隻小鼠）。

實例 5：人類 ARSA 轉移載體

【0189】此實例提供用於在經載體轉導之細胞（例如人類細胞或小鼠細胞）中的 hARSA 之表現之人類 ARSA 轉移載體 pHMI-5004。除表現 hARSA 以外，此載體經設計以亦表現人類鞘脂激活蛋白 B（SapB）。hARSA 及 SapB 之編碼序列由 2A 元件分離。

【0190】如圖 14 中所示，ARSA 轉移載體 pHMI-5004 自 5' 至 3' 包含以下基因元件：5' ITR 元件；轉錄調節元件，其包含 CMV 強化子元件、雞 β -肌動蛋白啟動子及嵌合內含子序列；以緘默方式改變之人類 ARSA 編碼序列；2A 元件；

野生型人類 SapB 編碼序列；及 3' ITR 元件。此等元件之序列闡述於表 3 中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 及/或 SapB 蛋白質。

表 3：人類 ARSA 轉移載體 pHMI-5004 中之基因元件

基因元件	SEQ ID NO:
5' ITR 元件	18
強化子元件	58
啟動子序列	25
內含子序列	32
轉錄調節元件	36
人類 ARSA 編碼序列	72
2A 元件	63
人類 SapB 編碼序列	73
hARSA-2A-hSapB 序列	74
SV40 聚腺苷酸化序列	42
3' ITR 元件	19
轉移基因組（自啟動子至聚腺苷酸化序列）	75
轉移基因組（自 5' ITR 至 3' ITR）	76
完全載體序列	77

實例 6：非人類靈長類動物中之 ARSA 基因轉移

【0191】 為了研究非人類靈長類動物中單次劑量之 AAVHSC 介導之 ARSA 基因遞送之作用，根據表 4 及 5 中所闡述之實驗設計向六隻雄性未經處理之幼年食蟹獼猴給藥。

表 4：用於非人類靈長類動物研究之實驗設計

組	路徑	動物/ 組	給藥 天數	劑量（vg/kg）	體積 （mL/kg）	濃度 （vg/mL）	屍檢
1	IV	2 隻雄性	1	0	5.0	-	第 28/29 天
2	IV	2 隻雄性	1	4e13	5.0	1.2e13	
3	CM	2 隻雄性	1	靜脈內劑量之 約 10%，基於動 物體重以固定	0.5 mL	儲備溶液 為 1.98 vg/mL	

				劑量形式提供 (約 6e12 vg/kg)			
--	--	--	--	--------------------------	--	--	--

表 5：用於非人類靈長類動物研究之實驗設計

動物	體重 (kg)	治療	路徑	劑量 (vg/kg)	Vg/動物
18C42	1.38	媒劑	IV	0	0
18C17	1.55	媒劑	IV	0	0
18C21	1.28	AAVHSC15-pHMI-5005	IV	4e13	5.12e13
18C27	1.28	AAVHSC15-pHMI-5005	IV	4e13	5.12e13
18C13	1.9	AAVHSC15-pHMI-5005	CM	4e12	7.6e12
18C7	1.74	AAVHSC15-pHMI-5005	CM	4e12	6.96e12

【0192】 如圖 15 中所示，ARSA 轉移載體 pHMI-5005 自 5'至 3'包含以下基因元件：5' ITR 元件；轉錄調節元件，其包含 CMV 強化子元件、雞 β-肌動蛋白啟動子及嵌合內含子序列；以緘默方式改變之人類 ARSA 編碼序列；V5 標籤；及 3' ITR 元件。此等元件之序列闡述於表 6 中。此載體能夠在經載體轉導之細胞（例如人類細胞或小鼠細胞）中表現人類 ARSA 蛋白質。

表 6：人類 ARSA 轉移載體 pHMI-5005 中之基因元件

基因元件	SEQ ID NO:
5' ITR 元件	18
強化子元件	58
啟動子序列	25
內含子序列	32
轉錄調節元件	36
人類 ARSA 編碼序列	14
V5 標籤	78
SV40 聚腺苷酸化序列	42
3' ITR 元件	19
轉移基因組（自啟動子至聚腺苷酸化序列）	79
轉移基因組（自 5' ITR 至 3' ITR）	80
完全載體序列	81

【0193】 pHMI-5005 為經 V5 標記之 ARSA 轉移載體。根據表 4 及 5 中所

闡述之實驗設計向非人類靈長類動物（NHP）投與封裝於 AAVHSC15 衣殼中之 pHMI-5005。在第 0 天，經由經頭部/隱靜脈進行 1-2 分鐘緩慢推注靜脈內注射(IV) 或直接注射至大池（CM）中來進行投藥。每天兩次針對死亡及瀕死跡象來進行存活率檢驗。在每天上午以及在給藥當天在完成給藥之後（15 分鐘）及給藥後 4 小時進行臨床觀測。在即將給藥之前以及在給藥後第 1、2 及 4 週獲得用於血液學及臨床化學之血液。在第 28 及 29 天之屍檢時，在腦脊髓液（CSF）及血液收集之後，向動物灌注 1.0 L 低溫生理食鹽水以移除血球。在屍檢時收集腦部、肝臟、脊髓（子宮頸及腰）、子宮頸及腰背根神經節（DRG）、三叉神經節、腎臟、坐骨神經、周邊淋巴結、脾、心臟、肺及睪丸。

【0194】對於生物分析，在即將給藥之前以及在第 1、2 及 4 週收集血清以用於 V5 Elisa（0.5 mL 全血，處理成血清/分成兩個等分試樣）。在給藥前收集 0.5 mL CSF（來自第 3 組 CM 給藥組動物）及在屍檢時收集 1-2 mL（對於所有動物）。在屍檢之前，自全血收集 15 mL 周邊血液單核細胞（PBMC）。

【0195】圖 16 顯示在投與封裝於 AAVHSC15 衣殼中之 pHMI-5005 的 NHP 中，丙胺酸轉胺酶（ALT）之含量上升。升高之 ALT 在給藥後第 14 天恢復至基線水準。

【0196】在給藥後第 28 及 29 天處死接受單次靜脈內劑量之 4×10^{13} vg/kg 之封裝於 AAVHSC15 中之 pHMI-5005 之 NHP（第 2 組動物）。在經處死之第 2 組動物之中樞神經系統（CNS）及腦脊髓液（CSF）中偵測人類 ARSA 酶活性水準（圖 17）。如圖 17 中所示，偵測到高於治療閾值（野生型人類腦部含量之 15%）之 hARSA 活性水準，如由虛線指示。動物 18C27（第 2 組）之 CNS 及周邊神經系統（PNS）中之免疫螢光染色證實存在 hARSA（經由 V5 標籤偵測）且存在於包括背根神經節、脊椎運動神經元及小腦之特定區域中。

【0197】本發明之範疇不受本文中所描述之特定實施例限制。實際上，根

第 70 頁，共 71 頁(發明說明書)

據先前描述及附圖，除所描述之修改以外，本發明之各種修改對本領域中熟習此項技術者而言將變得顯而易見。此類修改意欲屬於隨附申請專利範圍之範疇內。

【0198】 本文中所引用之所有參考文獻（例如公開案或專利案或專利申請案）皆以全文引用之方式且出於所有目的併入本文中，其併入程度如同各個別參考文獻（例如公開案或專利案或專利申請案）具體且個別地指示為出於所有目的以全文引用之方式併入一般。其他實施例屬於以下申請專利範圍內。

【符號說明】

【0199】 無

序列表

<110> 美商霍蒙拉奇醫藥公司(HOMOLOGY MEDICINES, INC.)

<120> 用於ARSA基因轉移之腺相關病毒組成物及其使用方法

<130> 705151: HMW-030-6

<160> 81

<170> PatentIn version 3.5

<210> 1

<211> 736

<212> PRT

<213> 腺相關AAV9

<400> 1

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 Gln Pro
 20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
 85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 2
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 2

Met Thr 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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Gln Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 3
<211> 736

<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 3

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Gly Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Gly Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 4
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 4

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 Gln Pro
20 25 30

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

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

Ile 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Tyr Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 5
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 5

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 6
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 6

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

Asn Leu Gly Arg Ala Val Leu Gln Ala Lys Lys Arg Leu 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 Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	
				245					250					255		
Tyr	Lys	Gln	Ile	Ser	Asn	Ser	Thr	Ser	Gly	Gly	Ser	Ser	Asn	Asp	Asn	
			260						265					270		
Ala	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	
		275						280				285				
Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	
	290						295				300					
Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile	
305					310					315					320	
Gln	Val	Lys	Glu	Val	Thr	Asp	Asn	Asn	Gly	Val	Lys	Thr	Ile	Ala	Asn	
				325					330					335		
Asn	Leu	Thr	Ser	Thr	Val	Gln	Val	Phe	Thr	Asp	Ser	Asp	Tyr	Gln	Leu	
			340					345					350			
Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Glu	Gly	Cys	Leu	Pro	Pro	Phe	Pro	
	355						360					365				
Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asp	
	370						375				380					
Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	
385					390					395					400	
Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Ser	Tyr	Glu	
			405						410					415		
Phe	Glu	Asn	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	
			420						425				430			
Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser	
	435						440				445					

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Ser Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 7
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 7

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Arg Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 8
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 8

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 Gln Pro
20 25 30

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

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

Val Asn Ala Val 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 9
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 9

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 10
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 10

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Cys Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 11
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 11

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 Gln Pro
20 25 30

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

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

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

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

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Lys Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 12
<211> 736
<212> PRT
<213> 人工序列

<220>

202112807

<223> AAV分離物

<400> 12

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro His Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Asn
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Met Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 13
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 13

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 14
<211> 1527
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 14
atgtctatgg gggctcctcg ctccctgctg ctggcactgg ccgccgggct ggctgtcgca 60

agaccaccta atatcgctct gattttttgca gacgatctgg gatacggcga cctgggatgc 120

tatggccacc caagctccac cacacccaac ctggaccagc tggcagcagg aggcctgcgg 180

ttcaccgact tctacgtgcc agtgagcctg tgcacccccct ccagagccgc cctgctgaca 240

ggcaggctgc cagtgcgcat gggcatgtat cctggcgtgc tggtgccatc tagcaggggc 300

ggcctgccac	tggaggaggt	gaccgtggca	gaggtgctgg	cagccagagg	ctacctgaca	360
ggaatggccg	gcaagtggca	cctgggagtg	ggaccagagg	gagccttcct	gccccctcac	420
cagggettcc	accggtttct	gggcatecct	tattctcacg	accagggccc	atgccagaac	480
ctgacctgtt	ttccaccagc	aacaccatgc	gacggaggat	gtgatcaggg	cctggtgcca	540
atcccactgc	tggcaaatct	gagcgtggag	gcacagcctc	catggctgcc	tggcctggag	600
gcaagataca	tggccttcgc	ccacgacctg	atggcagatg	cacagcggca	ggatagacct	660
ttctttctgt	actatgcctc	ccaccacacc	cactatccac	agttcagcgg	ccagtccttt	720
gccgagaggt	ccggaagggg	accattcggc	gactctctga	tggagctgga	tgccgccgtg	780
ggcacctga	tgacagcaat	cggcgacctg	ggcctgctgg	aggagacact	ggtcactctc	840
accgccgata	acggccctga	gacaatgcgg	atgtctagag	gcggatgcag	cggcctgctg	900
agatgtggca	agggaaccac	atacgaggga	ggcgtgcgcg	agcctgccct	ggcattttgg	960
ccaggacaca	tcgcacctgg	agtgaccac	gagctggcct	cctctctgga	cctgctgcca	1020
acactggccg	ccctggcagg	agcacctctg	ccaaatgtga	ccctggacgg	cttcgatctg	1080
agcccactgc	tgctgggaac	cggcaagtcc	cctaggcagt	ctctgttctt	ttacccctcc	1140
tatcctgatg	aggtgcgggg	cgtgtttgcc	gtgagaaccg	gcaagtacaa	ggcccacttc	1200
tttacacagg	gctctgcccc	cagcgacacc	acagcagatc	cagcatgcca	cgccagctcc	1260
tctctgaccg	cacacgagcc	acctctgctg	tacgacctgt	ccaaggatcc	cggcgagaa	1320
tataatctgc	tgggaggagt	ggcaggagca	accctgagg	tgctgcaggc	cctgaagcag	1380
ctgcagctgc	tgaaggcaca	gctggacgca	gcagtgacat	tcggcccaag	ccaggtggcc	1440
agaggcgagg	atcccgccct	gcagatctgt	tgccaccccg	gctgcacccc	aagacctgcc	1500
tgttgccatt	gccccgaccc	acacgcc				1527

- <210> 15
- <211> 736
- <212> PRT
- <213> 人工序列

<220>

202112807

<223> AAV分離物

<400> 15

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Arg Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 16
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 16

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala	Pro	Val	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	Ser
210						215					220				
Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Gln	Trp	Leu	Gly	Asp	Arg	Val	Ile
225					230					235					240
Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu
				245					250					255	
Tyr	Lys	Gln	Ile	Ser	Asn	Ser	Thr	Ser	Gly	Gly	Ser	Ser	Asn	Asp	Asn
			260						265					270	
Ala	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg
		275						280					285		
Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn
	290						295				300				
Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile
305						310				315					320
Gln	Val	Lys	Glu	Val	Thr	Asp	Asn	Asn	Gly	Val	Lys	Thr	Ile	Ala	Asn
				325					330						335
Asn	Leu	Thr	Ser	Thr	Val	Gln	Val	Phe	Ala	Asp	Ser	Asp	Tyr	Gln	Leu
			340						345					350	
Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Glu	Gly	Cys	Leu	Pro	Pro	Phe	Pro
		355					360						365		
Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asp
		370					375					380			
Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe
385						390				395					400
Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Ser	Tyr	Glu
				405					410					415	

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 17
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> AAV分離物

<400> 17

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 Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly 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 Lys Tyr Asn His Ala
85 90 95

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

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

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Ile Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

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

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Cys Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 18
<211> 145
<212> DNA
<213> 人工序列

<220>
<223> AAV2 5' ITR

<400> 18
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gacgcgcgag agagggagtg 120

gccaaactcca tcactagggg ttct 145

<210> 19
<211> 145
<212> DNA
<213> 人工序列

<220>
<223> AAV2 3' ITR

<400> 19
aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccc ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaa 145

<210> 20
<211> 167
<212> DNA
<213> 人工序列

<220>
<223> AAV5 5' ITR

<400> 20
ctctccccc tgtcgcgttc gctcgctcgc tggctcgttt gggggggtgg cagctcaaag 60

agctgccaga cgacggccct ctggccgtcg ccccccaaa cgagccagcg agcgagcgaa 120

cgcgacaggg gggagagtgc cacactctca agcaaggggg ttttgta 167

<210> 21
<211> 167
<212> DNA
<213> 人工序列

<220>
<223> AAV5 3' ITR

<400> 21
tacaaaacct ccttgcttga gagtgtggca ctctccccc tgtcgcgttc gctcgctcgc 60

tggctcgttt gggggggtgg cagctcaaag agctgccaga cgacggccct ctggccgtcg 120

ccccccaaa cgagccagcg agcgagcgaa cgcgacaggg gggagag 167

<210> 22
<211> 621
<212> PRT
<213> 人工序列

<220>
<223> AAV2 Rep

<400> 22

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

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245 250 255

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

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

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290 295 300

Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305 310 315 320

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355 360 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370 375 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385 390 395 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405 410 415

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
420 425 430

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435 440 445

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
450 455 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
465 470 475 480

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
485 490 495

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500 505 510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515 520 525

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530 535 540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545 550 555 560

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

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580 585 590

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595 600 605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610 615 620

<210> 23
<211> 509
<212> PRT
<213> 智人

<400> 23

Met Ser Met Gly Ala Pro Arg Ser Leu Leu Leu Ala Leu Ala Ala Gly
1 5 10 15

Leu Ala Val Ala Arg Pro Pro Asn Ile Val Leu Ile Phe Ala Asp Asp
20 25 30

Leu Gly Tyr Gly Asp Leu Gly Cys Tyr Gly His Pro Ser Ser Thr Thr
35 40 45

Pro Asn Leu Asp Gln Leu Ala Ala Gly Gly Leu Arg Phe Thr Asp Phe
50 55 60

Tyr Val Pro Val Ser Leu Cys Thr Pro Ser Arg Ala Ala Leu Leu Thr
65 70 75 80

Gly Arg Leu Pro Val Arg Met Gly Met Tyr Pro Gly Val Leu Val Pro
85 90 95

Ser Ser Arg Gly Gly Leu Pro Leu Glu Glu Val Thr Val Ala Glu Val
100 105 110

Leu Ala Ala Arg Gly Tyr Leu Thr Gly Met Ala Gly Lys Trp His Leu
115 120 125

Gly Val Gly Pro Glu Gly Ala Phe Leu Pro Pro His Gln Gly Phe His
130 135 140

Arg Phe Leu Gly Ile Pro Tyr Ser His Asp Gln Gly Pro Cys Gln Asn

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

Val Thr Leu Asp Gly Phe Asp Leu Ser Pro Leu Leu Leu Gly Thr Gly
355 360 365

Lys Ser Pro Arg Gln Ser Leu Phe Phe Tyr Pro Ser Tyr Pro Asp Glu
370 375 380

Val Arg Gly Val Phe Ala Val Arg Thr Gly Lys Tyr Lys Ala His Phe
385 390 395 400

Phe Thr Gln Gly Ser Ala His Ser Asp Thr Thr Ala Asp Pro Ala Cys
405 410 415

His Ala Ser Ser Ser Leu Thr Ala His Glu Pro Pro Leu Leu Tyr Asp
420 425 430

Leu Ser Lys Asp Pro Gly Glu Asn Tyr Asn Leu Leu Gly Gly Val Ala
435 440 445

Gly Ala Thr Pro Glu Val Leu Gln Ala Leu Lys Gln Leu Gln Leu Leu
450 455 460

Lys Ala Gln Leu Asp Ala Ala Val Thr Phe Gly Pro Ser Gln Val Ala
465 470 475 480

Arg Gly Glu Asp Pro Ala Leu Gln Ile Cys Cys His Pro Gly Cys Thr
485 490 495

Pro Arg Pro Ala Cys Cys His Cys Pro Asp Pro His Ala
500 505

<210> 24
<211> 1527
<212> DNA
<213> 智人

<400> 24
atgtccatgg gggcaccgcg gtcctcctc ctggccctgg ctgctggcct ggccgttgcc 60

cgtcgcceca acatcgtgct gatctttgcc gacgacctcg gctatgggga cctgggctgc 120

tatgggcacc ccagctctac cactcccaac ctggaccagc tggcggcggg agggctgcgg	180
ttcacagact tctacgtgcc tgtgtctctg tgcacaccct ctagggccgc cctcctgacc	240
ggccggctcc cggttcggat gggcatgtac cctggcgtec tggtgcccag ctcccggggg	300
ggcctgcccc tggaggaggt gaccgtggcc gaagtcctgg ctgcccgagg ctacctaca	360
ggaatggccg gcaagtggca ccttggggtg gggcctgagg gggccttcct gccccccat	420
cagggttcc atcgatttct aggcateccg tactcccacg accagggcc ctgccagaac	480
ctgacctgt tcccgccggc cactccttgc gacggtggct gtgaccaggg cctggtcctc	540
atcccactgt tggccaacct gtccgtggag gcgcagcccc cctggctgcc cggactagag	600
gcccgtaca tggctttcgc ccatgacctc atggccgacg cccagcgcca ggatcgcccc	660
ttcttctgt actatgcctc tcaccacacc cactaccctc agttcagtgg gcagagcttt	720
gcagagcgtt caggccgcgg gccatttggg gactccctga tggagctgga tgcagctgtg	780
gggaccctga tgacagccat aggggacctg gggctgcttg aagagacgct ggtcatcttc	840
actgcagaca atggacctga gaccatgcgt atgtcccag gcggtgctc cggctctctg	900
cgggtgtggaa agggaacgac ctacgagggc ggtgtccgag agcctgcctt ggccttctgg	960
ccaggtcata tcgctcccgg cgtgaccac gagctggcca gctccctgga cctgctgcct	1020
accctggcag ccctggctgg ggccccactg cccaatgtca ccttggatgg ctttgacctc	1080
agccccctgc tgctgggcac aggcaagagc cctcggcagt ctctcttctt ctacccgtec	1140
taccagacg aggtccgtgg ggTTTTtGct gtgcggactg gaaagtacaa ggctcacttc	1200
ttcaccagg gctctgcca cagtatacc actgcagacc ctgcctgcca cgcctccagc	1260
tctctgactg ctcatgagcc cccgtgctc tatgacctgt ccaaggacct tggtgagaac	1320
tacaacctgc tggggggtgt ggccggggcc accccagagg tgctgcaagc cctgaaacag	1380
cttcagctgc tcaaggccca gttagacgca gctgtgacct tcggccccag ccaggtggcc	1440
cggggcgagg accccgccct gcagatctgc tgtcatcctg gctgcacccc ccgccagct	1500
tgctgccatt gccagatcc ccatgcc	1527

<210> 25
<211> 278
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 25
tcgaggtgag ccccacgttc tgcttcactc tccccatctc cccccctcc ccacccccaa 60

ttttgtattt atttatTTTT taattatTTT gtgcagcgat gggggcgggg gggggggggg 120

ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg 180

cggcggcagc caatcagagc ggcgcgctcc gaaagtTTcc ttttatggcg aggcggcggc 240

ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcg 278

<210> 26
<211> 106
<212> DNA
<213> 人工序列

<220>
<223> 5' ITR

<400> 26
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgctc ggcgaccttt 60

ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtgg 106

<210> 27
<211> 143
<212> DNA
<213> 人工序列

<220>
<223> 3' ITR

<400> 27
aggaaccctt agtgatggag ttggccactc cctctctgcg cgtcgcctcg ctactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccc ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gcc 143

<210> 28
<211> 1873
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 28	
gatcttcaat attggccatt agccatatta ttcat ttggtt atatagcata aatcaatatt	60
ggctattggc cattgcatac gttgtatcta tatcataata tgtacattta tattggctca	120
tgtccaatat gaccgccatg ttggcattga ttattgacta gttattaata gtaatcaatt	180
acgggggtcat tagttcatag cccatatatg gagttccgcg ttacataact tacggtaa	240
ggcccgccctg gctgaccgcc caacgacccc cgcccattga cgtcaataat gacgtatgtt	300
cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggagta tttacggtaa	360
actgcccact tggcagtaca tcaagtgtat catatgccaa gtccgccccc tattgacgtc	420
aatgacggta aatggccccg ctggcattat gcccagtaca tgaccttacg ggactttcct	480
acttggcagt acatctacgt attagtcatc gctattacca tggtcgaggt gagccccacg	540
ttctgettca ctctccccat ctcccccccc tccccacccc caatttttgta tttatttatt	600
ttttaattat tttgtgcagc gatggggggcg ggggggggggg gggggcgcgcg gccaggcggg	660
gcggggcgggg gcgagggggcg gggcgggggcg aggcggagag gtgcggcggc agccaatcag	720
agcggcgcgcg tccgaaagtt tccttttatg gcgaggcggc ggcggcgggcg gccctataaa	780
aagcgaagcg cgcggcgggc gggagtcgct gcgacgtgc cttcgccccg tgccccgctc	840
cgcgcgccgc tcgcgccgc cgcgccggct ctgactgacc gcgttactcc cacaggtgag	900
cgggcgggac ggcccttctc ctccgggctg taattagcgc ttggtttaat gacggcttgt	960
ttcttttctg tggctgcgtg aaagccttga ggggctccgg gagggccctt tgtgcggggg	1020
ggagcggctc ggggggtgcg tgcgtgtgtg tgtgcgtggg gagcgccgcg tgcggcccgc	1080
gctgccccgc ggctgtgagc gctgcgggcg cggcgcgggg ctttgtgcgc tccgcagtgt	1140
gcgcgagggg agcgcggccg ggggcggtgc cccgcggtgc ggggggggct gcgaggggaa	1200
caaaggctgc gtgcggggtg tgtgcgtggg ggggtgagca gggggtgtgg gcgcggcggt	1260

cgggctgtaa cccccccctg cccccccctc cccgagttgc tgagcacggc ccggttcgg	1320
gtgcggggct ccgtacgggg cgtggcgcgg ggctcgccgt gccgggcggg gggtggcggc	1380
aggtgggggt gccgggcggg gcggggccgc ctcgggccgg ggagggtcg ggggaggggc	1440
gcggcggccc ccggagcgcc ggcggtgtc gaggcgcggc gagccgcagc cattgccttt	1500
tatggtaatc gtgcgagagg gcgcagggac ttcctttgtc ccaaatctgt gcggagccga	1560
aatctgggag gcgccgcgc accccctcta gcgggcgcgg ggccaagcgg tgcggcgccg	1620
gcaggaagga aatgggcggg gagggccttc gtgcgtcgcc gcgccgccgt ccccttctcc	1680
ctctccagcc tcggggctgt ccgcgggggg acggtgcct tcggggggga cggggcaggg	1740
cggggttcgg ctcttggcgt gtgaccggcg gctctagagc ctctgctaac catgttcatg	1800
ccctcttctt ttccctacag ctctgggca acgtgctggt tattgtgctg tctcatcatt	1860
ttggcaaaga att	1873

<210> 29
<211> 374
<212> PRT
<213> 智人

<400> 29

Met	Ala	Ala	Pro	Ala	Leu	Gly	Leu	Val	Cys	Gly	Arg	Cys	Pro	Glu	Leu
1			5					10						15	

Gly	Leu	Val	Leu	Leu	Leu	Leu	Leu	Leu	Ser	Leu	Leu	Cys	Gly	Ala	Ala
			20					25					30		

Gly	Ser	Gln	Glu	Ala	Gly	Thr	Gly	Ala	Gly	Ala	Gly	Ser	Leu	Ala	Gly
		35					40					45			

Ser	Cys	Gly	Cys	Gly	Thr	Pro	Gln	Arg	Pro	Gly	Ala	His	Gly	Ser	Ser
	50						55				60				

Ala	Ala	Ala	His	Arg	Tyr	Ser	Arg	Glu	Ala	Asn	Ala	Pro	Gly	Pro	Val
65					70					75					80

Pro	Gly	Glu	Arg	Gln	Leu	Ala	His	Ser	Lys	Met	Val	Pro	Ile	Pro	Ala	
				85					90					95		
Gly	Val	Phe	Thr	Met	Gly	Thr	Asp	Asp	Pro	Gln	Ile	Lys	Gln	Asp	Gly	
			100					105					110			
Glu	Ala	Pro	Ala	Arg	Arg	Val	Thr	Ile	Asp	Ala	Phe	Tyr	Met	Asp	Ala	
		115					120					125				
Tyr	Glu	Val	Ser	Asn	Thr	Glu	Phe	Glu	Lys	Phe	Val	Asn	Ser	Thr	Gly	
	130					135					140					
Tyr	Leu	Thr	Glu	Ala	Glu	Lys	Phe	Gly	Asp	Ser	Phe	Val	Phe	Glu	Gly	
145					150				155						160	
Met	Leu	Ser	Glu	Gln	Val	Lys	Thr	Asn	Ile	Gln	Gln	Ala	Val	Ala	Ala	
			165					170					175			
Ala	Pro	Trp	Trp	Leu	Pro	Val	Lys	Gly	Ala	Asn	Trp	Arg	His	Pro	Glu	
		180						185				190				
Gly	Pro	Asp	Ser	Thr	Ile	Leu	His	Arg	Pro	Asp	His	Pro	Val	Leu	His	
	195					200					205					
Val	Ser	Trp	Asn	Asp	Ala	Val	Ala	Tyr	Cys	Thr	Trp	Ala	Gly	Lys	Arg	
	210					215				220						
Leu	Pro	Thr	Glu	Ala	Glu	Trp	Glu	Tyr	Ser	Cys	Arg	Gly	Gly	Leu	His	
225				230				235						240		
Asn	Arg	Leu	Phe	Pro	Trp	Gly	Asn	Lys	Leu	Gln	Pro	Lys	Gly	Gln	His	
			245					250					255			
Tyr	Ala	Asn	Ile	Trp	Gln	Gly	Glu	Phe	Pro	Val	Thr	Asn	Thr	Gly	Glu	
		260				265					270					
Asp	Gly	Phe	Gln	Gly	Thr	Ala	Pro	Val	Asp	Ala	Phe	Pro	Pro	Asn	Gly	

275

280

285

Tyr Gly Leu Tyr Asn Ile Val Gly Asn Ala Trp Glu Trp Thr Ser Asp
290 295 300

Trp Trp Thr Val His His Ser Val Glu Glu Thr Leu Asn Pro Lys Gly
305 310 315 320

Pro Pro Ser Gly Lys Asp Arg Val Lys Lys Gly Gly Ser Tyr Met Cys
325 330 335

His Arg Ser Tyr Cys Tyr Arg Tyr Arg Cys Ala Ala Arg Ser Gln Asn
340 345 350

Thr Pro Asp Ser Ser Ala Ser Asn Leu Gly Phe Arg Cys Ala Ala Asp
355 360 365

Arg Leu Pro Thr Met Asp
370

<210> 30
<211> 2718
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 30
atgagcatgg gcgccccag aagcctgtta cttgcttttag ctgctggcct tgcagtggca 60

aggcccccta acatcgtgct gatctttgca gatgacttgg gatatgggga tcttggttgt 120

tatggccacc catcaagcac aactcccaat ctggatcagt tggctgcagg aggtctgagg 180

tttacagact tttatgttcc agtctccctg tgcactcctt ctcgggctgc cctgcttact 240

gggaggctcc ctgtgagaat gggtatgtac cctggagtgt tggteccatc cagcagggga 300

gggctgcccc tggaagaggt gacagtggca gaggtgctgg cagcacgagg ctatctgact 360

ggcatggcag gcaagtggca cctgggtgta gggccagagg gtgctttcct gcctcccat 420

cagggctttc ataggtttct gggaatccca tactctcatg accaaggacc ctgccagaac	480
ctcacctgtt tccccctgc aacaccatgt gatgggggct gtgatcaagg tctggttcct	540
ataccactgc ttgctaatct ttcagtggaa gctcaaccac cctggctgcc tggcttggag	600
gctagataca tggccttcgc acatgatctg atggcagatg cccagagaca agataggcct	660
ttcttctctt actatgcata tcaccacacc cactatcctc agttctcagg ccaatcattt	720
gctgagcgta gtggcagggg ccattttggg gacagtttga tggaaactgga tgccgcagtt	780
ggtaccctca tgacagcaat aggggactta ggtttgctgg aggaaacatt ggtaattttc	840
acagctgata atggccctga gacaatgaga atgtctaggg gaggctgctc tggctctctg	900
aggtgtggta aagggactac atatgagggg ggagtgaggg aaccagctct tgcccttttg	960
ccaggtcaca tagccccctgg agttacacat gaactagctt ctccctgga cttgcttctt	1020
acactggcag ccctggcagg tgccccctc cctaattgtaa ctttagatgg atttgacctc	1080
tctccactac ttttagggac agggaaaagt ccaaggcagt cttattctt ctatccttc	1140
taccagatg aggtgagggg tgtttttgcc gtgaggactg ggaaatacaa agctcatttt	1200
tttaccagg gatcagctca ttcagacacc acagctgac ctgcctgtca tgccagcagt	1260
agcttgacag cacatgagcc tcccttactg tatgacctga gcaaggaccc aggggagAAC	1320
tataacctgc ttgggggggt tgctggggcc accccagaag tgcttcaggc actaaagcag	1380
ctgcaactgc ttaaagcaca gttggatgct gcagtgacct ttggcccttc ccaggtggcc	1440
agaggcgagg atcccgccct gcagatctgc tgccaccag gctgcacacc cagacctgcc	1500
tgctgtcact gccccgacc acacgccggc agcggagcta ctaacttcag cctgctgaag	1560
caggctggag acgtggagga gaacctgga cctatggctg cccagccct ggggctggtg	1620
tgtggcagat gccctgagct gggcctggtg ctgcttctcc tgetgctgag cctcctgtgt	1680
ggtgctgctg gctctcagga agcagggaca ggagcaggag caggttctct ggctggctca	1740
tgcggttggt ggacccccca gaggccaggg gctcatgggt cctctgcagc tgcccacagg	1800
tactcaaggg aagcaaatgc ccctggcccc gtacctgggg aaaggcaact tgetcactcc	1860
aagatggttc ctatccctgc aggagttttt actatgggaa ctgatgacce tcagatcaag	1920

caggatggtg aagcaccagc taggagagtc acaattgatg ctttctatat ggatgcctat	1980
gaagtgtcaa acacagaatt tgagaaattt gtaaacagca ctggatacct tacagaggct	2040
gagaaatttg gtgacagttt tgtttttgaa ggcatgctaa gtgagcaggt gaagaccaat	2100
atccaacagg cagtggctgc agccccctgg tggctgcctg ttaaaggagc caattggaga	2160
caccagagg gaccagactc aactatcctc cacaggcctg accaccctgt gctgcatgtg	2220
tcctggaatg atgcagtggc atactgcacc tgggctggga aaaggttacc aacagaggca	2280
gaatgggagt attcctgccg gggcggactg cacaacagac tgttccccctg gggcaataag	2340
ctgcaaccta aaggacagca ttatgccaat atttggcagg gagagtcccc agtcacaaac	2400
actggtgagg atggcttcca gggaactgcc cctgtggatg ctttcccacc caatggctat	2460
gggttgtaca atatagttag gaatgcctgg gagtggactt ctgactggtg gacggtecat	2520
cacagtgtgg aagagacact gaacccaaag gggccccctc caggcaagga cagagtcaag	2580
aaaggtggct cttatatgtg tcacagaagc tattgctaca gatataggtg tgctgcaaga	2640
agtcagaaca cccctgacag ctcagctagc aatctgggat ttagatgtgc agcagataga	2700
ctccccacca tggactga	2718

<210> 31
<211> 93
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 31	
ctctaaggta aatataaaat ttttaagtgt ataatgtgtt aaactactga ttctaattgt	60
ttctctcttt tagattccaa cctttggaac tga	93

<210> 32
<211> 1017
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400>	32	
ggagtcgctg	cgcgctgcct	tcgccccgtg ccccgcctccg ccgccgcctc gcgccgcccg 60
ccccggctct	gactgaccgc	gttactccca caggtgagcg ggccgggacgg cccttctcct 120
ccgggctgta	attagcgctt	ggtttaatga cggcttgttt cttttctgtg gctgcgtgaa 180
agccttgagg	ggctccggga	gggccctttg tgcgggggga gcggctcggg gggtgcgctgc 240
gtgtgtgtgt	gcgtggggag	cgccgcgtgc ggctccgcgc tgcccggcgg ctgtgagcgc 300
tgcgggcgcg	gcgcggggct	ttgtgcgctc cgcagtgtgc gcgaggggag cgcggccggg 360
ggcggtgccc	cgcggtgcgg	ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg 420
tgcgtggggg	ggtgagcagg	gggtgtgggc gcgtcggtcg ggctgcaacc cccctgcac 480
ccccctcccc	gagttgctga	gcacggcccc gcttcgggtg cggggctccg tacggggcgt 540
ggcgcggggc	tcgccgtgcc	gggcgggggg tggcggcagg tgggggtgcc gggcggggcg 600
gggccgcctc	gggccgggga	gggctcgggg gaggggcgcg gcggcccccg gagcgccggc 660
ggctgtcgag	gcgcggcgag	ccgcagccat tgccttttat ggtaatcgtg cgagagggcg 720
cagggacttc	ctttgtccca	aatctgtgcg gagccgaaat ctgggaggcg ccgccgcacc 780
ccctctagcg	ggcgcggggc	gaagcgtgct ggccccggca ggaaggaaat gggcggggag 840
ggccttcgtg	cgtcgccgcg	ccgccgtccc cttctccctc tccagcctcg gggctgtccg 900
cggggggacg	gctgccttcg	ggggggacgg ggcagggcgg ggttcggctt ctggcgtgtg 960
accggcggct	ctagagcctc	tgctaacctt gttcatgcct tctttttttt cctacag 1017

<210> 33
<211> 79
<212> PRT
<213> 智人

<400>	33	
Gly Asp Val Cys Gln Asp Cys Ile Gln Met Val Thr Asp Ile Gln Thr		
1	5	10 15
Ala Val Arg Thr Asn Ser Thr Phe Val Gln Ala Leu Val Glu His Val		
20	25	30

Lys Glu Glu Cys Asp Arg Leu Gly Pro Gly Met Ala Asp Ile Cys Lys
35 40 45

Asn Tyr Ile Ser Gln Tyr Ser Glu Ile Ala Ile Gln Met Met Met His
50 55 60

Met Gln Pro Lys Glu Ile Cys Ala Leu Val Gly Phe Cys Asp Glu
65 70 75

<210> 34
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成聚核苷酸

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> Xaa爲D或G

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa爲V或I

<220>
<221> MISC_FEATURE
<222> (4)..(4)
<223> Xaa爲任何胺基酸

<400> 34

Xaa Xaa Glu Xaa Asn Pro Gly Pro
1 5

<210> 35
<211> 92
<212> DNA
<213> 人工序列

<220>

<223> 合成聚核苷酸		
<400> 35		
aagaggtaag ggtttaaggg atggttggtt ggtggggtat taatgtttaa ttacctggag	60	
cacctgcctg aaateacttt ttttcaggtt gg	92	
<210> 36		
<211> 1676		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成聚核苷酸		
<400> 36		
ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60	
catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgceca	120	
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180	
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240	
aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggcccgctt	300	
ggcatttatgc ccagtacatg accttacggg acttttcctac ttggcagtac atctacgtat	360	
tagtcategc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct	420	
cccccccctc cccacccccca attttgtatt tatttatttt ttaattattt tgtgcagcga	480	
tgggggcggg ggggggggggg gggcgcgcg caggcggggc ggggcggggc gaggggcggg	540	
gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600	
cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg	660	
gagtcgctgc gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc	720	
cccggtcttg actgaccgcg ttactcccac aggtgagcgg gcgggacggc ccttctcctc	780	
cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa	840	
gccttgaggg gctccgggag ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg	900	
tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgct gcccggcggc tgtgagcgt	960	
gcgggcgcgg cgcggggctt tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg	1020	

gcggtgcccc	gcggtgcggg	gggggctgcg	aggggaacaa	aggctgcgtg	cggggtgtgt	1080
gcgtgggggg	gtgagcaggg	ggtgtgggcg	cgtcggtcgg	gctgcaaccc	ccccgcacc	1140
ccccccccg	agttgctgag	caegggcccg	cttcgggtgc	ggggctccgt	acggggcggtg	1200
gcgcggggct	cgccgtgccg	ggcgggggggt	ggcggcaggt	gggggtgccg	ggcggggcg	1260
ggccgcctcg	ggccggggag	ggctcggggg	aggggcgcgg	cggcccccg	agcgccggcg	1320
gctgtcgagg	cgcggcgagc	cgcagccatt	gccttttatg	gtaatcgtgc	gagagggcgc	1380
agggacttcc	tttgtcccaa	atctgtgcgg	agccgaaatc	tgggaggcgc	cgccgcaccc	1440
cctctagcgg	gcgcggggcg	aagcgggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	1500
gccttcgtgc	gtcgccgcgc	cgccgtcccc	ttctccctct	ccagcctcgg	ggctgtccgc	1560
ggggggacgg	ctgccttcgg	gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	1620
cgggcggctc	tagagcctct	gctaaccatg	ttcatgccct	cttctttttc	ctacag	1676

<210> 37
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> T2A肽

<400> 37

Glu	Gly	Arg	Gly	Ser	Leu	Leu	Thr	Cys	Gly	Asp	Val	Glu	Glu	Asn	Pro
1				5					10					15	

<210> 38
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> P2A肽

<400> 38

Ala	Thr	Asn	Phe	Ser	Leu	Leu	Lys	Gln	Ala	Gly	Asp	Val	Glu	Glu	Asn
1				5					10					15	

<210>	39	
<211>	540	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成聚核苷酸	
<400>	39	
cctgcaggct caccagtgtt tgtgactggg aactctccct gccaaatatt ggcataatgc	60	
tgtccttttag gttgcagctt attgccccag gggaacagtc tgttgtgcag tccaccccgg	120	
caggaatact cccattctgc ctctgttggg aaccttttcc cagcccagggt gcagtatgcc	180	
actgcatcat tccaggacac atgcagcaca gggtggtcag gcctgtggag gatagttgag	240	
tctggtcctt ctgggtgtct ccaattggct cttttaacag gcagccacca gggggctgca	300	
gccactgcct gttggatatt ggtcttcacc tgctcactta gcatgccttc aaaaacaaaa	360	
ctgtcaccaa atttctcagc ctctgtaagg tatccagtgc tgtttacaaa tttctcaat	420	
tctgtgtttg acacttcata ggcatccata tagaaggcat caattgtgac tctcctagct	480	
ggtgcttcac catcctgctt gatctgaggg tcatcagttc ccatagtaaa aactcctgca	540	

<210>	40	
<211>	1168	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成聚核苷酸	
<400>	40	
cgtgaggctc cggtgcccgt cagtgggcag agcgcacatc gccacagtc cccgagaagt	60	
tggggggagg ggtcggcaat tgaaccggtg cctagagaag gtggcgcggg gtaaactggg	120	
aaagtgatgt cgtgtactgg ctccgccttt ttcccagagg tgggggagaa ccgtatataa	180	
gtgcagtagt cgccgtgaac gttctttttc gcaacggggt tgccgccaga acacaggtaa	240	
gtgccgtgtg tggttcccgc gggcctggcc tctttacggg ttatggccct tgcgtgcctt	300	
gaattacttc cacctggctc cagtacgtga ttcttgatcc cgagctggag ccaggggctg	360	

gccttgcgct ttaggagccc cttegcctcg tgcttgagtt gaggcctggc ctgggcgctg	420
gggccgccgc gtgcgaatct ggtggcacct tcgcgcctgt ctcgctgctt tcgataagtc	480
tctagccatt taaaattttt gatgacctgc tgcgacgctt tttttctggc aagatagtct	540
tgtaaatgcg ggccaggatc tgcacactgg tatttcggtt tttggggccg cgggcggcga	600
cggggccccgt gcgtcccagc gcacatgttc ggcgaggcgg ggcctgcgag cgcggccacc	660
gagaatcgga cgggggtagt ctcaagctgg ccggcctgct ctggtgcctg gcctcgcgcc	720
gccgtgtatc gccccgccct gggcggcaag gctggcccgg tcggcaccag ttgcgtgagc	780
ggaaagatgg ccgcttcccc gccctgctcc agggggctca aaatggagga cgcggcgctc	840
gggagagcgg gcgggtgagt cacccacaca aaggaaaggg gcctttccgt cctcagccgt	900
cgtttcatgt gactccacgg agtaccgggc gccgtccagg cacctcgatt agttctggag	960
cttttggagt acgtcgtctt taggttgggg ggaggggttt tatgcgatgg agtttcecca	1020
cactgagtgg gtggagactg aagttaggcc agcttggcac ttgatgtaat tctccttgga	1080
atttgccctt tttgagtttg gatcttggtt cattctcaag cctcagacag tggttcaaag	1140
tttttttctt ccatttcagg tgtcgtga	1168

<210> 41
<211> 3416
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 41 ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgccca	120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggccccct	300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360

tagtcacgc	tattaccatg	gtcgaggtga	gccccacgtt	ctgcttcact	ctccccatct	420
cccccccctc	cccacccccca	attttgtatt	tatttatttt	ttaattattt	tgtgcagcga	480
tgggggcggg	ggggggggggg	gggcgcgcgc	caggcggggc	ggggcggggc	gaggggcggg	540
gcggggcgag	gcggagaggt	gcggcggcag	ccaatcagag	cggcgcgctc	cgaaagtttc	600
cttttatggc	gaggcggcgg	cggcggcggc	cctataaaaa	gcgaagcgcg	cggcgggcgg	660
gagtcgctgc	gcgtgcctt	cgccccgtgc	cccgctccgc	cgcgcctcg	cgcgcctcgc	720
cccggctctg	actgaccgcg	ttactcccac	aggtgagcgg	gcgggacggc	ccttctcctc	780
cgggctgtaa	ttagcgcttg	gtttaatgac	ggcttgtttc	ttttctgtgg	ctgcgtgaaa	840
gccttgaggg	gctccgggag	ggccctttgt	gcggggggag	cggctcgggg	ggtgcgtgcg	900
tgtgtgtgtg	cgtggggagc	gccgcgtgcg	gctccgcgct	gcccggcggc	tgtgagcgt	960
gcgggcgcgg	cgcggggctt	tgtgcgctcc	gcagtgtgcg	cgaggggagc	gcggccgggg	1020
gcggtgcccc	gcggtgcggg	gggggctgcg	aggggaacaa	aggctgcgtg	cggggtgtgt	1080
gcgtgggggg	gtgagcaggg	ggtgtgggcg	cgtcggtcgg	gctgcaacct	cccctgcacc	1140
cccctccccg	agttgctgag	cacggccccg	cttcgggtgc	ggggctccgt	acggggcgtg	1200
gcgcggggct	cgccgtgccg	ggcgggggggt	ggcggcaggt	gggggtgccg	ggcggggcgg	1260
ggccgcctcg	ggccggggag	ggctcggggg	aggggcgcgg	cggcccccg	agcgcgggcg	1320
gctgtcgagg	cgcggcgagc	cgcagccatt	gccttttatg	gtaatcgtgc	gagagggcgc	1380
agggacttcc	tttgtcccaa	atctgtgcgg	agccgaaatc	tgggaggcgc	cgcgcaccc	1440
cctctagcgg	gcgcggggcg	aagcggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	1500
gccttcgtgc	gtcgccgcgc	cgcgctcccc	ttctccctct	ccagcctcgg	ggctgtccgc	1560
ggggggacgg	ctgccttcgg	gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	1620
cgggcggctc	tagagcctct	gctaaccatg	ttcatgccct	cttctttttc	ctacagctcc	1680
tgggcaacgt	gctggttatt	gtgctgtctc	atcattttgg	caaagaattc	cgccaccatg	1740
tccatggggg	caccgcggtc	cctcctcctg	gccctggctg	ctggcctggc	cgttgcccgt	1800
cgcccaaca	tcgtgctgat	ctttgccgac	gacctcggct	atggggacct	gggctgctat	1860

gggcacccca gctctaccac tcccaacctg gaccagctgg cggcgggagg gctgcggttc	1920
acagacttct acgtgcctgt gtctctgtgc acacctcta gggccgccct cctgaccggc	1980
cggctccccg ttcggatggg catgtacctt ggcgtcctgg tgcccagctc ccggggggggc	2040
ctgcccctgg aggaggtgac cgtggccgaa gtccctggctg cccgaggcta cctcacagga	2100
atggccggca agtggcacct tggggtgggg cctgaggggg ctttcctgcc ccccatcag	2160
ggcttccatc gatttctagg catcccgta cccacgacc agggcccttg ccagaacctg	2220
acctgcttcc cgccggccac tccttgcgac ggtggctgtg accagggcct ggtcccatc	2280
ccactgttgg ccaacctgtc cgtggaggcg cagccccctt ggctgcccgg actagaggcc	2340
cgtacatgg ctttcgccc tgacctcatg gccgacgcc agcgccagga tcgccccttc	2400
ttcctgtact atgcctctca ccacaccac taccctcagt tcagtgggca gagctttgca	2460
gagcgttcag gccgcggggc atttggggac tccttgatgg agctggatgc agctgtgggg	2520
acctgatga cagccatagg ggacctgggg ctgcttgaag agacgctggt catcttcact	2580
gcagacaatg gacctgagac catgcgtatg tcccgaggcg gctgctccgg tctcttgcg	2640
tgtggaaagg gaacgacct caggggcggt gtccgagagc ctgccttggc cttctggcca	2700
ggtcatatcg ctcccggcgt gaccacgag ctggccagct ccctggacct gctgcctacc	2760
ctggcagccc tggttggggc cccactgccc aatgtcacct tgcatggctt tgacctcagc	2820
ccctgctgc tgggcacagg caagagccct cggcagtctc tcttcttcta ccgctcctac	2880
ccagacgagg tccgtggggt ttttgcctgt cggactggaa agtacaaggc tcacttcttc	2940
accagggct ctgcccacag tgataccact gcagaccctg cctgccacgc ctccagctct	3000
ctgactgctc atgagcccc gctgctctat gacctgtcca aggacctgg tgagaactac	3060
aacctgctgg ggggtgtggc cggggccacc ccagaggtgc tgcaagccct gaaacagctt	3120
cagctgctca agggccagtt agacgcagct gtgaccttcg gcccagcca ggtggcccgg	3180
ggcgaggacc ccgccctgca gatctgctgt catcctggct gcacccccg ccagcttgc	3240
tgccattgcc cagatcccca tgcctgagat tctagagtcg agccgcggac tagtaacttg	3300
tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa	3360

gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt atctta	3416
<210> 42	
<211> 122	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成聚核苷酸	
<400> 42	
aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca	60
aataaagcat ttttttcaact gcattctagt tgtggtttgt ccaaactcat caatgtatct	120
ta	122
<210> 43	
<211> 133	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成聚核苷酸	
<400> 43	
tgctttatatt gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat	60
aaacaagtta acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg	120
gaggtttttt aaa	133
<210> 44	
<211> 3416	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成聚核苷酸	
<400> 44	
ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gtccgcggtt acataactta cggtaaattg cccgcctggc tgaccgccca	120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggccccct	300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360
tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct	420
cccccccctc cccacccccca attttgtatt tatttatttt ttaattattt tgtgcagcga	480
tggggggcggg ggggggggggg gggcgcgcg caggcggggc ggggcggggc gaggggcggg	540
gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg	660
gagtcgctgc gcgctgcctt cgcctcgtgc cccgctccgc cgccgcctcg cgccgcccgc	720
cccggctctg actgaccgcg ttactcccac aggtgagcgg gcgggacggc ccttctcctc	780
cgggctgtaa ttagcgcttg gttaaagac ggcttgtttc tttctgtgg ctgcgtgaaa	840
gccttgaggg gctccgggag ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg	900
tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgct gcccggcggc tgtgagcgt	960
gcgggcgcgg cgcggggctt tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg	1020
gcggtgcccc gcggtgcggg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt	1080
gcgtgggggg gtgagcaggg ggtgtgggcg cgtcggtcgg gctgcaacct cccctgcacc	1140
ccccccccg agttgctgag cacggccccg ctfcgggtgc ggggctccgt acggggcgtg	1200
gcgcggggct cgccgtgccg ggcggggggt ggcggcaggt gggggtgccg ggcggggcgg	1260
ggccgcctcg ggccggggag ggctcggggg aggggcgcgg cgccccccgg agcgccggcg	1320
gctgtcgagg cgcggcgagc cgcagccatt gccttttatg gtaatcgtgc gagagggcgc	1380
agggacttcc tttgtcccaa atctgtgcgg agccgaaatc tgggaggcgc cgccgcaccc	1440
cctctagcgg gcgcggggcg aagcgggtgcg gcgccggcag gaaggaaatg ggcggggagg	1500
gccttcgtgc gtcgccgcgc cgccgtcccc ttctccctct ccagcctcgg ggcgtgccgc	1560
ggggggacgg ctgccttcgg gggggacggg gcagggcggg gtteggcttc tggcgtgtga	1620
ccggcggtc tagagcctct gctaaccatg ttcatgccct cttctttttc ctacagctcc	1680

tgggcaacgt	gctggttatt	gtgctgtctc	atcattttgg	caaagaattc	cgccaccatg	1740
tctatggggg	ctcctcgctc	cctgctgctg	gcactggccg	ccgggctggc	tgtcgcaaga	1800
ccacctaata	tcgtcctgat	ttttgcagac	gatctgggat	acggcgacct	gggatgctat	1860
ggccacccaa	gtccaccac	acccaacctg	gaccagctgg	cagcaggagg	cctgcggttc	1920
accgacttct	acgtgccagt	gagcctgtgc	accccccca	gagccgccct	gctgacaggc	1980
aggctgccag	tgcgcatggg	catgtatcct	ggcgtgctgg	tgccatctag	caggggcggc	2040
ctgccactgg	aggaggtgac	cgtggcagag	gtgctggcag	ccagaggcta	cctgacagga	2100
atggccggca	agtggcacct	gggagtggga	ccagagggag	ccttcctgcc	ccctcaccag	2160
ggcttcacc	ggtttctggg	catcccttat	tctcacgacc	agggcccatg	ccagaacctg	2220
acctgttttc	caccagcaac	accatgcgac	ggaggatgtg	atcagggcct	ggtgccaatc	2280
ccactgctgg	caaatctgag	cgtggaggca	cagcctccat	ggctgcctgg	cctggaggca	2340
agatacatgg	ccttcgcccc	cgacctgatg	gcagatgcac	agcggcagga	tagacctttc	2400
tttctgtact	atgcctccca	ccacacccac	tatccacagt	tcagcggcca	gtcctttgcc	2460
gagaggtccg	gaaggggacc	attcggcgac	tctctgatgg	agctggatgc	cgccgtgggc	2520
accctgatga	cagcaatcgg	cgacctgggc	ctgctggagg	agacactggc	catcttcacc	2580
gccgataacg	gccctgagac	aatgcggatg	tctagaggcg	gatgcagcgg	cctgctgaga	2640
tgtggcaagg	gaaccacata	cgagggaggc	gtgcgcgagc	ctgccctggc	attttggcca	2700
ggacacatcg	cacctggagt	gaccacgag	ctggcctcct	ctctggacct	gctgccaaca	2760
ctggccgccc	tggcaggagc	acctctgcca	aatgtgacct	tggacggctt	cgatctgagc	2820
ccactgctgc	tgggaaccgg	caagtccct	aggcagtctc	tgttctttta	ccccctctat	2880
cctgatgagg	tgcggggcgt	gtttgccgtg	agaaccggca	agtacaaggc	ccacttcttt	2940
acacagggct	ctgcccacag	cgacaccaca	gcagatccag	catgccacgc	cagctcctct	3000
ctgaccgcac	acgagccacc	tctgctgtac	gacctgtcca	aggatcccgg	cgagaactat	3060
aatctgctgg	gaggagtggc	aggagcaacc	cctgaggtgc	tgcaggccct	gaagcagctg	3120
cagctgctga	aggcacagct	ggacgcagca	gtgacattcg	gccaagcca	ggtggccaga	3180

ggcgaggatc ccgccctgca gatctgttgc caccocggct gcaccccaag acctgcctgt	3240
tgccattgcc ccgaccacaca cgcctaagat tctagagtcg agccgcggac tagtaacttg	3300
tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa	3360
gcattttttt cactgcattc tagttgttgt ttgtccaaac tcatcaatgt atctta	3416

<210> 45
<211> 198
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 45 gatccagaca tgataagata cattgatgag ttgggacaaa ccacaactag aatgcagtga	60
aaaaaatgct ttatttgtga aatttgtgat gctattgcct tatttgtaac cattataagc	120
tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag	180
gtgtgggagg ttttttaa	198

<210> 46
<211> 3416
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 46 ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgcca	120
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggcccgct	300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360
tagtcategc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct	420

cccccccctc	cccacccccca	attttgtatt	tattttatttt	ttaattattt	tgtgcagcga	480
tgggggcggg	ggggggggggg	gggcgcgcgc	caggcggggc	ggggcggggc	gaggggcggg	540
gcggggcgag	gcggagaggt	gcggcggcag	ccaatcagag	cggcgcgctc	cgaaagtttc	600
cttttatggc	gaggcggcgg	cggcggcggc	cctataaaaa	gcgaagcgcg	cggcgggcgg	660
gagtcgctgc	gcgtgcctt	cgccccgtgc	cccgtccgc	cgcgcctcg	cgccgccgc	720
cccggctctg	actgaccgcg	ttactccac	aggtgagcgg	gcgggacggc	ccttctcctc	780
cgggctgtaa	ttagcgcttg	gtttaatgac	ggcttgtttc	ttttctgtgg	ctgcgtgaaa	840
gccttgaggg	gctccgggag	ggccctttgt	gcggggggag	cggctcgggg	ggtgcgtgcg	900
tgtgtgtgtg	cgtggggagc	gccgcgtgcg	gctccgcgct	gcccggcggc	tgtgagcgt	960
gcgggcgcgg	cgcggggctt	tgtgcgctcc	gcagtgtgcg	cgaggggagc	gcggccgggg	1020
gcggtgcccc	gcggtgcggg	gggggctgcg	aggggaacaa	aggctgcgtg	cggggtgtgt	1080
gcgtgggggg	gtgagcaggg	ggtgtgggcg	cgtcggtcgg	gctgcaacc	ccccgcacc	1140
ccccccccg	agttgctgag	cacggccccg	cttcgggtgc	ggggctccgt	acggggcgtg	1200
gcgcgggggt	cgccgtgccg	ggcgggggggt	ggcggcaggt	gggggtgccg	ggcggggcgg	1260
ggccgcctcg	ggccggggag	ggctcggggg	aggggcgcgg	cggcccccg	agcgccggcg	1320
gctgtcgagg	cgcggcgagc	cgcagccatt	gccttttatg	gtaatcgtgc	gagagggcgc	1380
agggacttcc	tttgtcccaa	atctgtgcgg	agccgaaatc	tgggaggcgc	cgccgcaccc	1440
cctctagcgg	gcgcggggcg	aagcgggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	1500
gccttcgtgc	gtcgccgcgc	cgcgctcccc	ttctccctct	ccagcctcgg	ggctgtccgc	1560
ggggggacgg	ctgccttcgg	gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	1620
ccggcggctc	tagagcctct	gctaaccatg	ttcatgcctt	cttctttttc	ctacagctcc	1680
tgggcaacgt	gctggttatt	gtgctgtctc	atcattttgg	caaagaattc	cgccaccatg	1740
tctatggggg	ctcctcgctc	cctgctgctg	gactggccg	ccgggctggc	tgtcgcaaga	1800
ccacctaaata	tcgtcctgat	ttttgcagac	gatctgggat	acggcgacct	gggatgctat	1860
ggccacccaa	gtccaccac	acccaacctg	gaccagctgg	cagcaggagg	cctgcggttc	1920

accgacttct	acgtgccagt	gagcctgtgc	acccccctcca	gagccgccct	gctgacaggc	1980
aggctgccag	tgcgcatggg	catgtatcct	ggcgtgctgg	tgccatctag	caggggcggc	2040
ctgccactgg	aggaggtgac	cgtggcagag	gtgctggcag	ccagaggcta	cctgacagga	2100
atggccggca	agtggcacct	gggagtggga	ccagagggag	ccttcctgcc	ccctcaccag	2160
ggcttcacc	ggtttctggg	catcccttat	tctcacgacc	agggcccatg	ccagaacctg	2220
acctgttttc	caccagcaac	accatgcgac	ggaggatgtg	atcagggcct	ggtgccaatc	2280
ccactgctgg	caaatctgag	cgtggaggca	cagcctccat	ggctgcctgg	cctggaggca	2340
agatacatgg	ccttcgcccc	cgacctgatg	gcagatgcac	agcggcagga	tagacctttc	2400
tttctgtact	atgcctcccc	ccacaccac	tatccacagt	tcagcggcca	gtcctttgcc	2460
gagaggtccg	gaaggggacc	attcggcgac	tctctgatgg	agctggatgc	cgccgtgggc	2520
accctgatga	cagcaatcgg	cgacctgggc	ctgctggagg	agacactggg	catcttcacc	2580
gccgataacg	gccctgagac	aatgcggatg	tctagaggcg	gatgcagcgg	cctgctgaga	2640
tgtggcaagg	gaaccacata	cgagggaggc	gtgcgcgagc	ctgccctggc	attttggcca	2700
ggacacatcg	cacctggagt	gacccacgag	ctggcctcct	ctctggacct	gctgccaaca	2760
ctggccgccc	tggcaggagc	acctctgcca	aatgtgacct	tggacggctt	cgatctgagc	2820
ccactgctgc	tgggaaccgg	caagtcccc	aggcagtcct	tgttctttta	ccccctctat	2880
cctgatgagg	tgcggggcgt	gtttgccgtg	agaaccggca	agtacaaggc	ccacttcttt	2940
acacagggct	ctgcccacag	cgacaccaca	gcagatccag	catgccacgc	cagctcctct	3000
ctgaccgcac	acgagccacc	tctgctgtac	gacctgtcca	aggatccccg	cgagaactat	3060
aatctgctgg	gaggagtggc	aggagcaacc	cctgaggtgc	tgcaggccct	gaagcagctg	3120
cagctgctga	aggcacagct	ggacgcagca	gtgacattcg	gcccagcca	ggtggccaga	3180
ggcgaggatc	ccgccctgca	gatctgttgc	caccccggct	gcacccaag	acctgcctgt	3240
tgccattgcc	ccgaccacac	cgcctaagat	tctagagtcg	agccgcggac	tagtaacttg	3300
tttattgcag	cttataatgg	ttacaaataa	agcaatagca	tcacaaattt	cacaaataaa	3360
gcattttttt	cactgcattc	tagttgtggg	ttgtccaaac	tcatcaatgt	atctta	3416

<210> 47
<211> 3949
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 47	
ttggccactc cctctctgcg cgctcgtcgc ctcaactgagg ccggggcgacc aaaggtcggc	60
cgacgccccg gctttgcccc ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tcactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag	180
ggttagggag gtctgcaga tcttcaatat tggccattag ccatattatt cattggttat	240
atagcataaa tcaatattgg ctattggcca ttgcatacgt tgtatctata tcataaatatg	300
tacatttata ttggctcatg tccaatatga ccgccatgtt ggcatlgatt attgactagt	360
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt	420
acataactta cggtaaattg ccgcctggc tgaccgcca acgacccccg cccattgacg	480
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	540
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt	600
ccgcccccta ttgacgtcaa tgacggtaaa tggcccgctt ggcatltatgc ccagtacatg	660
accttacggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	720
gtcgaggtga gccccacgtt ctgcttcaact ctccccatct cccccccctc cccaccccca	780
attttgtatt tattttatit ttaattatit tgtgcagcga tgggggagggg ggggggggggg	840
gggcgcgcgc caggcggggc ggggcggggc gaggggcggg gcggggcgag gcggagaggt	900
gcggcggcag ccaatcagag cggcgcgctc cgaaagtitt cttttatggc gaggcggcgg	960
cggcggcggc cctataaaaa gcgaagcgcg cggcggggcgg gagtgcgtgc gcgctgcctt	1020
cgccccgtgc cccgctccgc cgccgcctcg cgccgccccg cccggctctg actgaccgcg	1080
ttactcccac aggtgagcgg gcgggacggc ctttctctc cgggctgtaa ttagcgcttg	1140
gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa gccttgaggg gctccgggag	1200

ggccctttgt	gcggggggag	cggctcgggg	ggtgcgtgcg	tgtgtgtgtg	cgtggggagc	1260
gccgcgtgcg	gctccgcgct	gcccggcggc	tgtgagcgct	gcgggcgcgg	cgcggggctt	1320
tgtgcgctcc	gcagtgtgcg	cgaggggagc	gcggccgggg	gcggtgcccc	gcggtgcggg	1380
gggggctgcg	aggggaacaa	aggctgcgtg	cggggtgtgt	gcgtgggggg	gtgagcaggg	1440
ggtgtgggcg	cgtcggtcgg	gctgcaaccc	ccccgcacc	ccccccccg	agttgctgag	1500
cacggccccg	cttcgggtgc	ggggctccgt	acggggcgtg	gcgcgggggt	cgccgtgccg	1560
ggcgggggggt	ggcggcaggt	gggggtgccg	ggcggggcgg	ggccgcctcg	ggccggggag	1620
ggctcggggg	aggggcgcgg	cggcccccg	agcgcgcggc	gctgtcgagg	cgcggcgagc	1680
cgcagccatt	gccttttatg	gtaatcgtgc	gagagggcgc	agggacttcc	tttgtcccaa	1740
atctgtgcgg	agccgaaatc	tgggaggcgc	cgccgcaccc	cctctagcgg	gcgcggggcg	1800
aagcggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	gccttcgtgc	gtcgccgcgc	1860
cgccgtcccc	ttctccctct	ccagcctcgg	ggctgtccgc	ggggggacgg	ctgccttcgg	1920
gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	ccggcggctc	tagagcctct	1980
gctaaccatg	ttcatgcctt	cttctttttc	ctacagctcc	tgggcaacgt	gctggttatt	2040
gtgctgtctc	atcattttgg	caaagaattc	cgccaccatg	tccatggggg	caccgcggtc	2100
cctcctcctg	gccctggctg	ctggcctggc	cgttgcccgt	ccgcccacaa	tcgtgctgat	2160
ctttgccgac	gacctcggct	atggggacct	gggctgctat	gggcacccca	gctctaccac	2220
tcccaacctg	gaccagctgg	cggcgggagg	gctgcggttc	acagacttct	acgtgcctgt	2280
gtctctgtgc	acacctctca	gggccgccct	cctgaccggc	cggctccccg	ttcggatggg	2340
catgtaccct	ggcgtcctgg	tgcccagctc	ccgggggggg	ctgcccctgg	aggaggtgac	2400
cgtggccgaa	gtcctggctg	cccgaggcta	cctcacagga	atggccggca	agtggcacct	2460
tggggtgggg	cctgaggggg	ccttcctgcc	ccccatcag	ggttccatc	gatttctagg	2520
catcccgtac	tcccacgacc	agggcccttg	ccagaacctg	acctgcttcc	cgccggccac	2580
tccttgcgac	ggtggctgtg	accagggcct	ggtccccatc	ccactgttgg	ccaacctgtc	2640
cgtggaggcg	cagccccctt	ggctgcccgg	actagaggcc	cgtacatgg	ctttcgccca	2700

tgacctcatg gccgacgccc agcgccagga tcgcccccttc ttccctgtact atgcctctca	2760
ccacacccac taccctcagt tcagtgggca gagcttttgc gagcgttcag gccgcgggcc	2820
at ttggggac tccctgatgg agctggatgc agctgtgggg accctgatga cagccatagg	2880
ggacctgggg ctgcttgaag agacgctggt catcttcact gcagacaatg gacctgagac	2940
catgcgtatg tcccgaggcg gctgctccgg tctcttgcgg tgtggaaagg gaacgaccta	3000
cgagggcggt gtccgagagc ctgcccttggc cttcttgcca ggtcatatcg ctcccggcgt	3060
gaccacgag ctggccagct ccctggacct gctgcctacc ctggcagccc tggctggggc	3120
cccactgccc aatgtcacct tggatggctt tgacctcage cccctgctgc tgggcacagg	3180
caagagccct cggcagtctc tcttcttcta cccgtcctac ccagacgagg tccgtgggggt	3240
ttttgctgtg cggactggaa agtacaaggc tcacttcttc acccagggct ctgcccacag	3300
tgataccact gcagacctg cctgccacgc ctccagctct ctgactgctc atgagceccc	3360
gctgctctat gacctgtcca aggacctgg tgagaactac aacctgctgg ggggtgtggc	3420
cggggccacc ccagaggtgc tgcaagccct gaaacagctt cagctgctca aggcccagtt	3480
agacgcagct gtgaccttcg gccccagcca ggtggccccg ggcgaggacc ccgccctgca	3540
gatctgctgt catcctggct gcaccccccg ccagcttgc tgccattgcc cagatcccca	3600
tgcctgagat tctagagtcg agccgcggac tagtaacttg ttatttgcag cttataatgg	3660
ttacaaataa agcaatagca tcacaaatit cacaaataaa gcattttttt cactgcattc	3720
tagttgtggt ttgtccaaac tcatcaatgt atcttaggtc tagatacgta gataagtagc	3780
atggcgggtt aatcattaac tacaaggaac ccttagtgat ggagttggcc actccctctc	3840
tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt cgcgccgacgc ccgggctttg	3900
cccgggcggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaa	3949

<210> 48
<211> 3949
<212> DNA
<213> 人工序列

<220>

<223> 合成聚核苷酸

<400> 48

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc	60
cgacgccccg gctttgcccc ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tcactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag	180
ggttagggag gtcctgcaga tcttcaatat tggccattag ccatattatt cattggttat	240
atagcataaa tcaatattgg ctattggcca ttgcatacgt tgtatctata tcataatatg	300
tacatttata ttggctcatg tccaatatga ccgccatgtt ggcatlgatt attgactagt	360
tattaatagt aatcaattac ggggtcatta gtccatagcc catatatgga gttccgcgtt	420
acataactta cggtaaattg ccgcctggc tgaccgcca acgacccccg ccattgacg	480
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	540
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt	600
ccgcccccta ttgacgtcaa tgacggtaaa tggcccgctt ggcatatatc ccagtacatg	660
accttacggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	720
gtcgagggtga gccccacgtt ctgcttcaact ctccccatct cccccccctc cccaccccc	780
attttgtatt tatttatfff ttaattatff tgtgcagcga tgggggcggg gggggggggg	840
gggcgcgcgc caggcggggc ggggcggggc gaggggcggg gcggggcgag gcggagaggt	900
gcggcggcag ccaatcagag cggcgcgctc cgaaagtffc cttttatggc gaggcggcgg	960
cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg gagtcgctgc gcgctgcctt	1020
cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc cccggctctg actgaccgcg	1080
ttactcccac aggtgagcgg gcgggacggc ctttctcctc cgggctgtaa ttagcgcttg	1140
gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa gccttgaggg gctccgggag	1200
ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg tgtgtgtgtg cgtggggagc	1260
gccgcgtgcg gctccgcgt gcccggcggc tgtgagcgt gcgggcgcgg cgcggggctt	1320
tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg gcggtgcccc gcggtgcggg	1380
gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt gcgtgggggg gtgagcaggg	1440

ggtgtgggcg	cgtcggtcgg	gctgcaaccc	ccccgcacc	ccccccccg	agttgctgag	1500
cacggccccg	cttcgggtgc	ggggctccgt	acggggcgig	gcgcggggct	cgccgtgccg	1560
ggcggggggg	ggcggcaggt	gggggtgccg	ggcggggcgg	ggccgcctcg	ggccggggag	1620
ggctcggggg	aggggcgcgg	cggcccccg	agcggcggcg	gctgtcgagg	cgcggcgagc	1680
cgcagccatt	gccttttatg	gtaatcgtgc	gagagggcgc	agggacttcc	tttgtcccaa	1740
atctgtgcgg	agccgaaatc	tgggaggcgc	cggcgcaccc	cctctagcgg	gcgcggggcg	1800
aagcggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	gccttcgtgc	gtcgccgcgc	1860
cgccgtcccc	ttctccctct	ccagcctcgg	ggctgtccgc	ggggggacgg	ctgccttcgg	1920
gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	ccggcggctc	tagagcctct	1980
gctaaccatg	ttcatgcctt	cttctttttc	ctacagctcc	tgggcaacgt	gctggttatt	2040
gtgctgtctc	atcatttttg	caaagaattc	cgccaccatg	tctatggggg	ctcctcgctc	2100
cctgctgctg	gcactggccg	ccgggctggc	tgtcgcaaga	ccacctata	tcgtcctgat	2160
ttttgcagac	gatctgggat	acggcgacct	gggatgctat	ggccacccaa	gtccaccac	2220
accaacctg	gaccagctgg	cagcaggagg	cctgcggttc	accgacttct	acgtgccagt	2280
gagcctgtgc	acccctcca	gagccgccct	gctgacaggc	aggctgccag	tgcgcatggg	2340
catgtatcct	ggcgtgctgg	tgccatctag	caggggcggc	ctgccactgg	aggaggtgac	2400
cgtggcagag	gtgctggcag	ccagaggcta	cctgacagga	atggccggca	agtggcacct	2460
gggagtggga	ccagagggag	ccttcctgcc	ccctcaccag	ggcttcacc	ggtttctggg	2520
catcccttat	tctcacgacc	agggcccatg	ccagaacctg	acctgttttc	caccageaac	2580
accatgcgac	ggaggatgtg	atcagggcct	ggtgccaatc	ccactgctgg	caaatctgag	2640
cgtggaggca	cagcctccat	ggctgcctgg	cctggaggca	agatacatgg	ccttcgcca	2700
cgacctgatg	gcagatgcac	agcggcagga	tagacctttc	tttctgtact	atgcctccca	2760
ccacaccac	tatccacagt	tcagcggcca	gtcctttgcc	gagagggtccg	gaaggggacc	2820
attcggcgac	tctctgatgg	agctggatgc	cgccgtgggc	accctgatga	cagcaatcgg	2880
cgacctgggc	ctgctggagg	agacactggg	catcttcacc	gccgataacg	gccctgagac	2940

aatgcggatg tctagaggcg gatgcagcgg cctgctgaga tgtggcaagg gaaccacata	3000
cgagggaggc gtgcgcgagc ctgccctggc attttggcca ggacacatcg cacctggagt	3060
gaccaacgag ctggcctcct ctctggacct gctgccaaca ctggccgccc tggcaggagc	3120
acctctgcca aatgtgacct tggacggctt cgatctgagc ccactgctgc tgggaaccgg	3180
caagtcacct aggcagtctc tgttctttta cccctcctat cctgatgagg tgcggggcgt	3240
gtttgccgtg agaaccggca agtacaaggc ccacttcttt acacagggct ctgcccacag	3300
cgacaccaca gcagatccag catgccacgc cagctcctct ctgaccgcac acgagccacc	3360
tctgctgtac gacctgtcca aggatcccgg cgagaactat aatctgctgg gaggagtggc	3420
aggagcaacc cctgaggctg tgcaggccct gaagcagctg cagctgctga aggcacagct	3480
ggacgcagca gtgacattcg gccaagcca ggtggccaga ggcgaggatc ccgccctgca	3540
gatctgttgc caccceggct gcacccaag acctgcctgt tgccattgcc ccgaccaca	3600
cgcctaagat tctagagtcg agccgcggac tagtaacttg ttattgcag cttataatgg	3660
ttacaaataa agcaatagca tcacaaatit cacaaataaa gcattttttt cactgcattc	3720
tagttgtggt ttgtccaaac tcatcaatgt atcttaggtc tagatacgta gataagtagc	3780
atggcggggtt aatcattaac tacaaggaac ccctagtgat ggagttggcc actccctctc	3840
tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt cgcgccgacgc ccgggctttg	3900
cccgggcggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaa	3949

<210> 49
<211> 4500
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 49	
ttggccaactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc	60
cgacgccccg gctttgcccc ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tcactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag	180

ggttagggag gtcctgcata tgcggccgcg atcttcaata ttggccatta gccatattat	240
tcattggtta tatagcataa atcaatattg gctattggcc attgcatacg ttgtatctat	300
atcataatat gtacatttat attggctcat gtccaatatg accgccatgt tggcattgat	360
tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc ccatatatgg	420
agttccgcgt tacataactt acggtaaattg gcccgcttg ctgaccgcc aacgaccccc	480
gcccatlgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt	540
gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc	600
atatgccaaag tccgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg	660
cccagtacat gaccttacgg gactttccta cttggcagta catctacgta ttagtcatcg	720
ctattaccat ggtcgaggtg agccccacgt tctgcttcac tctccccatc tccccccct	780
ccccaccccc aattttgtat ttatttatit tttaattatt ttgtgcagcg atggggggcgg	840
ggggggggggg gggggcgcgcg ccaggcgggg cggggcgggg cgagggggcgg ggcggggcgga	900
ggcggagagg tgcggcgga gccaatcaga gcggcgcgct ccgaaagttt ccttttatgg	960
cgaggcgggcg gcggcgggcg ccctataaaa agcgaagcgc gcggcgggcg ggagtcgctg	1020
cgcgctgcct tcgccccgtg ccccgctccg ccgccgcctc gcgccggccg ccccggtct	1080
gactgaccgc gttactccca caggtgagcg ggcgggacgg cccttctcct ccgggctgta	1140
attagcgctt ggtttaatga cggcttgttt cttttctgtg gctgcgtgaa agccttgagg	1200
ggctccggga gggccctttg tgcgggggga gcggctcggg gggtgcgctgc gtgtgtgtgt	1260
gcgtggggag cgccgcgtgc ggctccgcgc tgcccggcgg ctgtgagcgc tgcgggcgcg	1320
gcgcggggct ttgtgcgctc cgcagtgtgc gcgaggggag cgcggccggg ggcggtgcc	1380
cgcggtgcgg ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg tgcgtggggg	1440
ggtgagcagg ggggtgtgggc gcgtcggtcg ggctgcaacc cccctgcac cccctcccc	1500
gagttgctga gcacggcccc gcttcgggtg cggggctccg tacggggcgt ggcgcggggc	1560
tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg gggccgcctc	1620
gggccgggga gggctcgggg gaggggcgcg gcggcccccg gagcgccggc ggctgtcgag	1680

gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagagggcg cagggacttc	1740
ctttgtccca aatctgtgcg gagccgaaat ctgggaggcg ccgccgcacc cctctagcg	1800
ggcgcggggc gaagcggcgc ggcgccggca ggaaggaaat gggcggggag ggccttcgtg	1860
cgtcgccgcg ccgccgtccc cttctccctc tccagcctcg gggctgtccg cggggggacg	1920
gctgccttcg ggggggacgg ggcagggcgg ggttcggcct ctggcgtgtg accggcggct	1980
ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagctc ctgggcaacg	2040
tgctggttat tgtgctgtct catcattttg gcaaagaatt ccgccaccat gtctatgggg	2100
gctcctcgct ccttgcctgc ggcactggcc gccgggctgg ctgtcgcaag accacctaat	2160
atcgtcctga tttttgcaga cgatctggga tacggcgacc tgggatgcta tggccacca	2220
agctccacca cacccaacct ggaccagctg gcagcaggag gcctgcggtt caccgacttc	2280
tacgtgccag tgagcctgtg caccctctcc agagccgccc tgcctgacagg caggctgcca	2340
gtgcgcatgg gcatgtatcc tggcgtgctg gtgccatcta gcaggggcgg cctgccactg	2400
gaggaggtga ccgtggcaga ggtgctggca gccagaggct acctgacagg aatggccggc	2460
aagtggcacc tgggagtggg accagaggga gccttcctgc cccctcacca gggcttcac	2520
cggtttctgg gcatccctta ttctcacgac cagggcccat gccagaacct gacctgtttt	2580
ccaccagcaa caccatgcga cggaggatgt gatcagggcc tggtgccaat cccactgctg	2640
gcaaatctga gcgtggaggc acagcctcca tggctgcctg gcctggaggc aagatacatg	2700
gccttcgccc acgacctgat ggcagatgca cagcggcagg atagaccttt ctttctgtac	2760
tatgcctccc accacacca ctatccacag ttcagcggcc agtcctttgc cgagaggctc	2820
ggaaggggac cattcggcga ctctctgatg gagctggatg ccgccgtggg caccctgatg	2880
acagcaatcg gcgacctggg cctgctggag gagacactgg tcattcttcac cgccgataac	2940
ggccctgaga caatgcggat gtctagaggc ggatgcagcg gcctgctgag atgtggcaag	3000
ggaaccacat acgagggagg cgtgcgcgag cctgccctgg cattttggcc aggacacatc	3060
gcacctggag tgaccacga gctggcctcc tctctggacc tgctgccaac actggccgcc	3120
ctggcaggag cacctctgcc aaatgtgacc ctggacggct tcgatctgag cccactgctg	3180

ctgggaaccg gcaagtcgcc taggcagtct ctgttctttt acccctccta tcctgatgag	3240
gtgcggggcg tgtttgccgt gagaaccggc aagtacaagg cccacttctt tacacagggc	3300
tctgcccaca gcgacaccac agcagatcca gcatgccacg ccagctcctc tctgaccgca	3360
cacgagccac ctctgctgta cgacctgtcc aaggatcccc gcgagaacta taatctgctg	3420
ggaggagtgg caggagcaac ccctgagggt ctgcaggccc tgaagcagct gcagctgctg	3480
aaggcacagc tggacgcagc agtgacattc ggccccagcc aggtggccag aggcgaggat	3540
cccgccctgc agatctgttg ccaccccggc tgcaccccaa gacctgcctg ttgccattgc	3600
cccgaccac acgcctaaga ttctagagtc gagccgcgga ctagtaactt gtttattgca	3660
gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt	3720
tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttacct gcaggctcac	3780
cagtgtttgt gactgggaac tctccctgcc aaatattggc ataatgctgt cctttagggt	3840
gcagcttatt gccccagggg aacagtctgt tgtgcagtc accccggcag gaatactccc	3900
attctgcctc tgttggtaac cttttcccag ccaggtgca gtatgccact gcatcattcc	3960
aggacacatg cagcacaggg tggtcaggcc tgtggaggat agttgagtct ggtccctctg	4020
ggtgtctcca attggctcct ttaacaggca gccaccaggg ggctgcagcc actgcctgtt	4080
ggatattggg cttcacctgc tcacttagca tgccttcaaa aacaaaactg tcaccaaatt	4140
tctcagcctc tgtaaggat ccagtctgt ttacaaattt ctcaaattct gtgtttgaca	4200
cttcataggc atccatatag aaggcatcaa ttgtgactct cctagctggg gcttcacat	4260
cctgcttgat ctgagggtca tcagttccca tagtaaaaac tcctgcaggt ctagatacgt	4320
agataagtag catggcgggt taatcattaa ctacaaggaa ccctagtgga tggagttggc	4380
cactccctct ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgcg	4440
cccgggcttt gcccgggcgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa	4500

<210> 50
<211> 6612
<212> DNA
<213> 人工序列

<220>

<223> 合成聚核苷酸

<400> 50

cgccagggtt ttcccagtcg cgacgttgta aaacgacggc cagtgccaaag cttgcatgcc	60
tgcaatttggc cactccctct ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg	120
tcgcccgcg cccgggcttt gcccgggcgg cctcagtgag cgagcgagcg cgcagagagg	180
gagtggccaa ctccatcact aggggttcct ggaggggtgg agtcgtgacg tgaattacgt	240
catagggtta gggaggtcct gcagatcttc aatattggcc attagccata ttattcattg	300
gttatatagc ataaatcaat attggctatt ggccattgca tacgttgtat ctatatcata	360
atatgtacat ttatatggc tcatgtccaa tatgaccgcc atgttggcat tgattattga	420
ctagttatta atagtaatca attacgggtt cattagttca tagcccatat atggagttec	480
gcgttacata acttacggta aatggcccg cctggctgacc gccaacgcac ccccgcccat	540
tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc	600
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	660
caagtcgcgc ccctattgac gtcaatgacg gtaaattggc cgcttggeat tatgcccagt	720
acatgacctt acgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	780
ccatggtcga ggtgagcccc acgttctgct tcactctccc catctcccc ccctccccac	840
ccccaatttt gtatttatTT attttttaat tattttgtgc agcgatgggg gcgggggggg	900
gggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga	960
gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttccTTTT atggcgaggc	1020
ggcggcgggc gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgcgt	1080
gccttcgccc cgtgccccgc tccgccgcg cctcgcgccg cccgccccgg ctctgactga	1140
ccgcgttact cccacaggtg agcgggcggg acggcccttc tcctccgggc tgtaattagc	1200
gcttggttta atgacggctt gtttcttttc tgtggctgcg tgaaagcctt gaggggctcc	1260
gggaggggccc tttgtgcggg gggagcggt cgggggggtgc gtgcgtgtgt gtgtgcgtgg	1320
ggagcgccgc gtgcggctcc gcgctgccc gcggctgtga gcgctgcggg cgcggcgcg	1380

ggctttgtgc gctccgcagt gtgcgcgagg ggagcgcggc cgggggcggg gccccgcggg	1440
gcgggggggg ctgcgagggg aacaaaggct gcgtgcgggg tgtgtgcgtg ggggggtgag	1500
cagggggtgt gggcgcgtcg gtgggctgc aacccccct gcacccccct ccccgagtgt	1560
ctgagcacgg cccggcttcg ggtgcggggc tccgtacggg gcgtggcgcg gggctcgcg	1620
tgccgggcgg ggggtggcgg caggtggggg tgccgggcgg ggcggggccg cctcgggccg	1680
gggagggctc gggggagggg cgcggcggcc cccggagcgc cggcggctgt cgaggcgcgg	1740
cgagccgcag ccattgcctt ttatggtaat cgtgcgagag ggcgcaggga cttcctttgt	1800
cccaaatctg tgcggagccg aaatctggga ggcgccgccg cccccctct agcgggcgcg	1860
gggcgaagcg gtgcggcgcc ggcaggaagg aaatgggcgg ggagggcctt cgtgcgtcgc	1920
cgcgccgccg tccccctct cctctccagc ctgggggctg tccgcggggg gacggctgcc	1980
ttcggggggg acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag	2040
cctctgctaa ccatgttcat gccttcttct ttttctaca gctcctgggc aacgtgctgg	2100
ttattgtgct gtctcatcat tttggcaaag aattccgcc aatgtccat gggggcaccg	2160
cggtcctcc tcctggccct ggctgctggc ctggccgttg cccgtccgcc caacatcgtg	2220
ctgatctttg ccgacgacct cggctatggg gacctgggct gctatgggca cccagctct	2280
accactccca acctggacca gctggcggcg ggagggtgc ggttcacaga cttctacgtg	2340
cctgtgtctc tgtcacacc ctctagggcc gccctcctga ccggccggct cccggttcgg	2400
atgggcatgt accctggcgt cctggtgcc agctcccggg ggggcctgcc cctggaggag	2460
gtgaccgtgg ccgaagtcct ggctgcccga ggctacctca caggaatggc cggcaagtgg	2520
caccttgggg tggggcctga gggggccttc ctgcccccc atcagggtt ccatcgattt	2580
ctaggcatcc cgtactccca cgaccagggc ccctgccaga acctgacctg ctteccgccg	2640
gccactcctt gcgacggtgg ctgtgaccag ggcttggtcc ccataccact gttggccaac	2700
ctgtccgtgg aggcgcagcc cccctggctg cccggactag aggcccgcta catggctttc	2760
gcccattgacc tcatggccga cgcccagcgc caggatgcc ctttcttct gtactatgcc	2820
tctcaccaca cccactaccc tcagttcagt gggcagagct ttgcagagcg ttcaggccgc	2880

gggccatttg	gggactccct	gatggagctg	gatgcagctg	tggggaccct	gatgacagcc	2940
ataggggacc	tggggctgct	tgaagagacg	ctggtcattct	tcactgcaga	caatggacct	3000
gagaccatgc	gtatgtcccg	aggcggctgc	tccggtctct	tgcggtgtgg	aaagggaacg	3060
acctacgagg	gcggtgtccg	agagcctgcc	ttggccttct	ggccaggtea	tatcgctccc	3120
ggcgtgacct	acgagctggc	cagctccctg	gacctgctgc	ctaccctggc	agccctggct	3180
ggggcccccac	tgcccaatgt	caccttggat	ggctttgacc	tcagcccccct	gctgctgggc	3240
acaggcaaga	gccctcggca	gtctctcttc	ttctaccctg	cctaccacga	cgaggctccg	3300
gggggtttttg	ctgtgcggac	tggaaagtac	aaggctcact	tcttcaccca	gggctctgcc	3360
cacagtgata	ccactgcaga	ccctgcctgc	cacgcctcca	gctctctgac	tgtcatgag	3420
cccccgctgc	tctatgacct	gtccaaggac	cctggtgaga	actacaacct	gctgggggggt	3480
gtggccggggg	ccaccccaga	ggtgctgcaa	gccctgaaac	agcttcagct	gtcaaggcc	3540
cagttagacg	cagctgtgac	cttcggcccc	agccagggtg	cccggggcga	ggacccccgc	3600
ctgcagatct	gctgtcatec	tggctgcacc	ccccgcccag	cttgttgcca	ttgcccagat	3660
cccatgcct	gagattctag	agtcgagccg	cggactagta	acttgtttat	tgcagcttat	3720
aatggttaca	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	tttttactg	3780
cattctagtt	gtggtttgtc	caaactcacc	aatgtatctt	aggtctagat	acgtagataa	3840
gtagcatggc	gggttaatca	ttaactacaa	ggaaccacct	gtgatggagt	tggccactcc	3900
ctctctgcgc	gctcgtctgc	tcactgaggc	cgggcgacca	aaggctgccc	gacgcccggg	3960
ctttgccccg	gcggcctcag	tgagcgagcg	agcgcgcaga	gagggagtgg	ccaaagatcc	4020
ccgggtaccg	agctcgaatt	cgtaatcatg	tcatagctgt	ttcctgtgtg	aaattgttat	4080
ccgctcacaa	ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	ctgggggtgcc	4140
taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	tgcccgcttt	ccagtcggga	4200
aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgcgt	4260
attggcgaa	ttttgctgag	ttgaaggatc	agatcacgca	tcttccccgac	aacgcagacc	4320
gttccgtggc	aaagcaaaag	ttcaaaatca	gtaaccgtca	gtgccgataa	gttcaaagtt	4380

aaacctggtg ttgataccaa cattgaaacg ctgatcgaaa acgcgctgaa aaacgctgct	4440
gaatgtgcga gcttcttccg cttectcgct cactgactcg ctgcgctcgg tcgttcggct	4500
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga	4560
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc	4620
cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg	4680
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg	4740
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt	4800
tctcccttcg ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcggc	4860
gtaggtcggt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg	4920
cgccttatcc ggtaactatc gtcttgagtc caacccggtc agacacgact tatcgccact	4980
ggcagcagcc actggtaaca ggattagcag agcgaggatc gtaggcggtg ctacagagtt	5040
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggtc tctgcgctct	5100
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac	5160
cgcctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc	5220
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg atccgtcgag	5280
aggtctgcct cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc	5340
cagccagaaa gtgagggagc cacggttgat gagagctttg ttgtaggtgg accagttggt	5400
gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg	5460
atccittcaac tcagcaaaag ttcgatttat tcaacaaagc cacgttggtg ctcaaaatct	5520
ctgatgttac attgcacaag ataaaaatat atcatcatga acaataaaac tgtctgctta	5580
cataaacagt aatacaaggg gtgttatgag ccatattcaa cgggaaacgt cttgctcgaa	5640
gccgcgatta aattccaaca tggatgctga tttatatggg tataaatggg ctgcgcgataa	5700
tgtcgggcaa tcaggtgcga caatctatcg attgtatggg aagcccgatg cgccagagtt	5760
gtttctgaaa catggcaaag gtagcgttgc caatgatgtt acagatgaga tggtcagact	5820
aaactggctg acggaattta tgcctcttcc gaccatcaag cattttatcc gtactcctga	5880

tgatgcatgg ttactcacca ctgcgatccc cgggaaaaca gcattccagg tattagaaga	5940
atatacctgat tcaggtgaaa atattgttga tgcgctggca gtgttcctgc gccggttgca	6000
ttcgattcct gtttgtaatt gtctttttaa cagcgatcgc gtatttcgtc tcgctcaggc	6060
gcaatcacga atgaataacg gtttggttga tgcgagtgat tttgatgacg agcgtaatgg	6120
ctggcctggt gaacaagtct ggaaagaaat gcataagctt ttgccattct caccggattc	6180
agtcgtcact catggtgatt tctcacttga taaccttatt tttgacgagg ggaaattaat	6240
aggttgtatt gatgttgga c gagtcggaat cgcagaccga taccaggatc ttgccatcct	6300
atggaactgc ctcggtgagt tttctccttc attacagaaa cggttttttc aaaaatatgg	6360
tattgataat cctgatatga ataaattgca gtttcatttg atgctcgatg agtttttcta	6420
atcagaattg gttaattggt tgtaacactg gcagagcatt acgctgactt gacgggacgg	6480
cggttttggt gaataaatcg cattcgccat tcaggctgcg caactgttgg gaagggcgat	6540
cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat	6600
taagttgggt aa	6612

- <210> 51
- <211> 5792
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 51	
gtcaggtggc acttttcggg gaaatgtggc atgcctgcat ttggccactc cctctctgcg	60
cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg	120
ggcggcctca gtgagcgagc gagcgcgcag agagggagtg gccaaactcca tctactagggg	180
ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag ggttagggag gtccctgcaga	240
tcttcaatat tggccattag ccatattatt cattggttat atagcataaa tcaatattgg	300
ctattggcca ttgcatacgt tgtatctata tcataatatg tacattttata ttggctcatg	360
tccaatatga ccgccatggt ggcatlgatt attgactagt tattaatagt aatcaattac	420

ggggtcatta gttcatagcc catafatgga gttccgcgtt acataactta cggtaaattgg	480
cccgcttggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc	540
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac	600
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt ccgcccccta ttgacgtcaa	660
tgacggtaaa tggcccgccf ggcattatgc ccagtacatg accttacggg actttcctac	720
ttggcagtaf atctacgtat tagtcatcgc tattaccatg gtcgaggtga gccccacgtt	780
ctgcttcact ctccccatct cccccccctc cccacccccca attttgtatt tatttatfff	840
ttaattatff tgtgcagcga tgggggcggg gggggggggg gggcgcgcg caggcggggc	900
ggggcggggc gaggggcggg gcggggcgag gcggagaggt gcggcggcag ccaatcagag	960
cggcgcgctc cgaaagtffc cttttatggc gaggcggcgg cggcggcggc cctataaaaa	1020
gcgaagcgcg cggcgggcgg gagtgcctgc gcgctgccff cgcctccgtgc cccgctccgc	1080
cgcgcctcg cgcgcctccg cccggctctg actgaccgcg ttactccac aggtgagcgg	1140
gcgggacggc ccttctctc cgggctgtaa ttagcgttg gtttaatgac ggcttgtffc	1200
ttttctgtgg ctgcgtgaaa gccttgaggg gctccgggag ggccctttgt gcggggggag	1260
cggctcgggg ggtgcgtgcg tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgt	1320
gccccggcggc tgtgagcgt gcgggcgcgg cgcggggctt tgtgcgtcc gcagtgtgcg	1380
cgaggggagc gcggccgggg gcggtgcccc gcggtgcggg gggggctgcg aggggaacaa	1440
aggctgcgtg cggggtgtgt gcgtgggggg gtgagcaggg ggtgtgggcg cgtcggtcgg	1500
gctgcaacct cccctgcacc cccctccccg agttgctgag cacggccccg cttcgggtgc	1560
ggggctccgt acggggcgtg gcgcggggct cgcctgccg ggccgggggt ggcggcaggt	1620
gggggtgccg ggccgggcgg ggccgcctcg ggccggggag ggctcggggg aggggcgcgg	1680
cggcccccg agcgccggcg gctgtcgagg cgcggcgagc cgcagccatt gccttttatg	1740
gtaatcgtgc gagagggcgc agggacttcc tttgtcccaa atctgtgcgg agccgaaatc	1800
tgggaggcgc cgcgcaccc cctctagcgg gcgcggggcg aagcggtgcg gcgccggcag	1860
gaaggaaatg ggcggggagg gccttcgtgc gtcgccgcg cgcctcccc ttctccctct	1920

ccagcctcgg ggctgtccgc ggggggacgg ctgcccttcgg gggggacggg gcagggcggg	1980
gttcggcttc tggcgtgtga ccggcggtc tagagcctct gctaaccatg ttcatgcctt	2040
cttccttttc ctacagctcc tgggcaacgt gctggttatt gtgctgtctc atcattttgg	2100
caaagaattc cgccaccatg tctatggggg ctccctcgctc cctgctgctg gcactggccg	2160
ccgggctggc tgtcgcaaga ccacctaata tcgtcctgat ttttgcacac gatctgggat	2220
acggcgacct gggatgctat ggccacccaa gctccaccac acccaacctg gaccagctgg	2280
cagcaggagg cctgcggttc accgacttct acgtgccagt gagcctgtgc accccctcca	2340
gagccgccct gctgacaggc aggctgccag tgcgcatggg catgtatcct ggcgtgctgg	2400
tgccatctag caggggcggc ctgccactgg aggaggtgac cgtggcagag gtgctggcag	2460
ccagaggcta cctgacagga atggccggca agtggcacct gggagtggga ccagagggag	2520
ccttcctgcc ccctcaccag ggcttcacc ggtttctggg cateccittat tctcacgacc	2580
agggcccatg ccagaacctg acctgttttc caccagcaac accatgcgac ggaggatgtg	2640
atcagggcct ggtgccaatc ccactgctgg caaatctgag cgtggaggca cagcctccat	2700
ggctgcctgg cctggaggca agatacatgg ccttcgcccc cgacctgatg gcagatgcac	2760
agcggcagga tagaccttc tttctgtact atgcctccca ccacaccac tatccacagt	2820
tcagcggcca gtcccttgcc gagaggtccg gaaggggacc attcggcgac tctctgatgg	2880
agctggatgc cgccgtgggc accctgatga cagcaatcgg cgacctgggc ctgctggagg	2940
agacactggt catcttcacc gccgataacg gccctgagac aatgcggatg tctagaggcg	3000
gatgcagcgg cctgctgaga tgtggcaagg gaaccacata cgagggaggc gtgcgcgagc	3060
ctgccctggc attttggcca ggacacatcg cacctggagt gaccacgag ctggcctcct	3120
ctctggacct gctgccaaca ctggccgccc tggcaggagc acctctgcca aatgtgacct	3180
tggacggctt cgatctgagc ccactgctgc tgggaaccgg caagtccct aggcagtctc	3240
tgttccttta cccctcctat cctgatgagg tgcggggcgt gtttgccgtg agaaccggca	3300
agtacaaggc ccacttcttt acacagggt ctgcccacag cgacaccaca gcagatccag	3360
catgccacgc cagctcctct ctgaccgcac acgagccacc tctgctgtac gacctgtcca	3420

aggatcccgg	cgagaactat	aatctgctgg	gaggagtggc	aggagcaacc	cctgaggtgc	3480
tgcaggccct	gaagcagctg	cagctgctga	aggcacagct	ggacgcagca	gtgacattcg	3540
gccaagcca	ggtggccaga	ggcgaggatc	ccgccctgca	gatctgttgc	caccccggt	3600
gcacccaag	acctgcctgt	tgccattgcc	ccgaccaca	cgcctaagat	tctagagtcg	3660
agccgcggac	tagtaacttg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	3720
tcacaaattt	cacaaataaa	gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	3780
tcatcaatgt	atcttaggtc	tagatacgta	gataagtagc	atggcgggtt	aatcattaac	3840
tacaaggaac	ccctagtgat	ggagttggcc	actccctctc	tgcgcgctcg	ctcgctcact	3900
gaggccgggc	gaccaaaggt	cgcccgacgc	ccgggctttg	cccgggcggc	ctcagtgagc	3960
gagcgagcgc	gcagagaggg	agtggccaaa	gatccccggg	taccgaggac	gaattctcta	4020
gatatcgctc	aatactgacc	atttaaatac	tacctgacct	ccatagcaga	aagtcaaaag	4080
cctccgaccg	gaggcttttg	acttgatcgg	cacgtaagag	gttccaactt	tcaccataat	4140
gaaataagat	cactaccggg	cgtatttttt	gagttatcga	gattttcagg	agctaaggaa	4200
gctaaaatga	gccatattca	acgggaaacg	tcttgctcga	ggccgcgatt	aaattccaac	4260
atggatgctg	atttatatgg	gtataaatgg	gctcgcgata	atgtcgggca	atcaggtgcg	4320
acaatctatc	gattgtatgg	gaagccccgat	gcgccagagt	tgtttctgaa	acatggcaaa	4380
ggtagcgttg	ccaatgatgt	tacagatgag	atggtcaggc	taaactggct	gacggaattt	4440
atgcctcttc	cgaccatcaa	gcattttatc	cgtactcctg	atgatgcatg	gttactcacc	4500
actgcgatcc	cagggaaaac	agcattccag	gtattagaag	aatatcctga	ttcaggtgaa	4560
aatattgttg	atgcgctggc	agtgttcctg	cgccggttgc	attcgattcc	tgtttgtaat	4620
tgtcctttta	acggcgatcg	cgtatttcgt	ctcgctcagg	cgcaatcacg	aatgaataac	4680
ggtttggttg	gtgcgagtga	ttttgatgac	gagcgtaatg	gctggcctgt	tgaacaagtc	4740
tggaaagaaa	tgcataagct	tttgccattc	tcaccggatt	cagtcgtcac	tcatggtgat	4800
ttctcacttg	ataaccttat	ttttgacgag	gggaaattaa	taggttgtat	tgatgttgga	4860
cgagtcggaa	tcgcagaccg	ataccaggat	cttgccatcc	tatggaactg	cctcggtgag	4920

ttttctcctt cattacagaa acggcttttt caaaaatatg gtattgataa tcctgatatg	4980
aataaattgc agtttcactt gatgctcgat gagtttttct gagggcccaa atgtaatcac	5040
ctggctcacc ttcgggtggg cctttctgcg ttgctggcgt ttttccatag gctccgcccc	5100
cctgacgagc atcacaaaaa tcgatgctca agtcagaggt ggcgaaaccc gacaggacta	5160
taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg	5220
ccgcttaccg gatacctgtc cgcttttctc ccttcgggaa gcgtggcgct ttctcatagc	5280
tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac	5340
gaaccccccg ttcagcccgga ccgctgcgcc ttatccggta actatcgtct tgagtccaac	5400
ccggtaaagac acgacttata gccactggca gcagccactg gtaacaggat tagcagagcg	5460
aggtatgtag gcggtgctac agagtcttg aagtggtagc ctaactacgg ctacactaga	5520
agaacagtat ttggtatctg cgctctgctg aagccagtta cctcggaaaa agagttggta	5580
gctcttgatc cggcaaacaa accaccgctg gtagcggtag tttttttgtt tgcaagcagc	5640
agattacgcg cagaaaaaaaa ggatectcaag aagatccttt gattttctac cgaagaaagg	5700
cccaccgctg aaggtgagcc agtgagttga ttgcagtcca gttacgctgg agtctgaggc	5760
tcgtcctgaa tgatatcaag cttgaattcg tt	5792

- <210> 52
- <211> 6342
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 52	
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgatgctcaa gtcagaggtg	60
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg	120
ctctcctggt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag	180
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc	240
caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa	300

ctatcgtctt	gagtccaacc	cggtaagaca	cgacttateg	ccactggcag	cagccactgg	360
taacaggatt	agcagagcga	ggtatgtagg	cggtgctaca	gagttcttga	agtgggtggc	420
taactacggc	tacactagaa	gaacagtatt	tggtatctgc	gctctgctga	agccagttac	480
ctcggaaaaa	gagttggtag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcggtggt	540
ttttttgttt	gcaagcagca	gattacgcgc	agaaaaaag	gatctcaaga	agatcctttg	600
attttctacc	gaagaaaggc	ccaccctga	aggtgagcca	gtgagttgat	tgcagtcacg	660
ttacgctgga	gtctgaggct	cgctctgaat	gatatcaagc	ttgaattcgt	gtcaggtggc	720
acttttcggg	gaaatgtggc	atgcctgcat	ttggccactc	cctctctgcg	cgtctgctcg	780
ctcactgagg	ccgggcgacc	aaaggtcgcc	cgacgcccgg	gctttgcccc	ggcggcctca	840
gtgagcgagc	gagcgcgcag	agagggagtg	gccaaactcca	tactagggg	ttcctggagg	900
ggtggagtcg	tgacgtgaat	tacgtcatag	ggttagggag	gtcctgcata	tgcggccgcg	960
atcttcaata	ttggccatta	gccatattat	tcattggtta	tatagcataa	atcaatattg	1020
gctattggcc	attgcatacg	ttgtatctat	atcataatat	gtacatttat	attggctcat	1080
gtccaatatg	accgccatgt	tggcattgat	tattgactag	ttattaatag	taatcaatta	1140
cggggtcatt	agttcatagc	ccatatatgg	agttccgcgt	tacataactt	acggtaaatz	1200
gccccctgg	ctgaccgccc	aacgaccccc	gccattgac	gtcaataatz	acgtatgttc	1260
ccatagtaac	gccaataggg	actttccatt	gacgtcaatz	ggtggagtat	ttacggtaaa	1320
ctgcccactt	ggcagtacat	caagtgtatc	atatgccaa	tcgccccct	attgacgtca	1380
atgacggtaa	atggcccgcc	tggcattatg	ccagttacat	gaccttacgg	gactttecta	1440
cttggcagta	catctacgta	ttagtcateg	ctattaccat	ggtcgaggtg	agccccacgt	1500
tctgcttcac	tctccccatc	tccccccct	ccccacccc	aattttgtat	ttatttattt	1560
tttaattatt	ttgtgcagcg	atggggggcgg	gggggggggg	ggggcgcgcg	ccaggcgggg	1620
cggggcgggg	cgagggggcgg	ggcggggcga	ggcggagagg	tgcggcggca	gccaatcaga	1680
gcggcgcgct	ccgaaagttt	ccttttatgg	cgaggcggcg	gcggcgggcg	ccctataaaa	1740
agcgaagcgc	gcggcgggcg	ggagtcgctg	cgcgctgcct	tcgccccgtg	ccccgctccg	1800

ccgccgcctc	gcgccgcccg	ccccggctct	gactgaccgc	gttactccca	caggtgagcg	1860
ggcgggacgg	cccttctcct	ccgggctgta	attagcgctt	ggtttaatga	cggcttgttt	1920
cttttctgtg	gctgcgtgaa	agccttgagg	ggctccggga	gggccctttg	tgcgggggga	1980
gcggctcggg	gggtgcgtgc	gtgtgtgtgt	gcgtggggag	cgccgcgtgc	ggctccgcgc	2040
tgcccggcgg	ctgtgagcgc	tgcgggcgcg	gcgcggggct	ttgtgcgctc	cgcagtgtgc	2100
gcgaggggag	cgcggccggg	ggcggtgccc	cgcggtgcgg	ggggggctgc	gaggggaaca	2160
aaggctgcgt	gcgggggtgtg	tgcgtggggg	ggtgagcagg	gggtgtgggc	gcgtcggtcg	2220
ggctgcaacc	ccccctgcac	ccccctcccc	gagttgctga	gcacggcccc	gcttcgggtg	2280
cggggctccg	tacggggcgt	ggcgcggggc	tcgccgtgcc	gggcgggggg	tggcggcagg	2340
tgggggtgcc	gggcggggcg	gggccgcctc	ggcccgggga	gggctcgggg	gaggggcgcg	2400
gcggcccccg	gagcgccggc	ggctgtcgag	gcgcggcgag	ccgcagccat	tgccttttat	2460
ggtaatcgtg	cgagagggcg	cagggacttc	ctttgtccca	aatctgtgcg	gagccgaaat	2520
ctgggaggcg	ccgccgcacc	ccctctagcg	ggcgcggggc	gaagcggtgc	ggcgcgggca	2580
ggaaggaaat	gggcggggag	ggccttcgtg	cgtcgccgcg	ccgccgtccc	cttctccctc	2640
tccagcctcg	gggctgtccg	cggggggacg	gctgccttcg	ggggggacgg	ggcagggcgg	2700
ggttcggctt	ctggcgtgtg	accggcggt	ctagagcctc	tgctaaccat	gttcatgcct	2760
tcttcttttt	cctacagctc	ctgggcaacg	tgctggttat	tgtgctgtct	catcattttg	2820
gcaaagaatt	ccgccaccat	gtctatgggg	gtcctcgcgt	ccctgctgct	ggcactggcc	2880
gccgggctgg	ctgtcgcaag	accacctaat	atcgtcctga	tttttgcaga	cgatctggga	2940
tacggcgacc	tgggatgcta	tggccaccca	agctccacca	cacccaacct	ggaccagctg	3000
gcagcaggag	gcctgcggtt	caccgacttc	tacgtgccag	tgagcctgtg	cacccctcc	3060
agagccgccc	tgctgacagg	caggctgcca	gtgcgcgatg	gcatgtatcc	tggcgtgctg	3120
gtgccatcta	gcaggggcgg	cctgccactg	gaggaggtga	ccgtggcaga	ggtgctggca	3180
gccagaggct	acctgacagg	aatggccggc	aagtggcacc	tgggagtggg	accagaggga	3240
gccttcctgc	cccccacca	gggtttccac	cggtttctgg	gcateccetta	ttctcacgac	3300

cagggcccat gccagaacct gacctgtttt ccaccagcaa caccatgcga cggaggatgt	3360
gatcagggcc tggtgccaat cccactgctg gcaaattctga gcgtggaggc acagcctcca	3420
tggctgcctg gcctggaggc aagatacatg gccttcgccc acgacctgat ggcagatgca	3480
cagcggcagg atagaccttt ctttctgtac tatgcctccc accacacca ctatccacag	3540
ttcagcggcc agtcctttgc cgagaggtec ggaaggggac cattcggcga ctctctgatg	3600
gagctggatg ccgccgtggg caccctgatg acagcaatcg gcgacctggg cctgctggag	3660
gagacactgg tcattcttcac cgccgataac ggccctgaga caatgcggat gtctagaggc	3720
ggatgcagcg gcctgctgag atgtggcaag ggaaccacat acgagggagg cgtgcgcgag	3780
cctgccctgg cattttggcc aggacacatc gcacctggag tgaccacga gctggcctcc	3840
tctctggacc tgctgccaac actggccgcc ctggcaggag cacctctgcc aaatgtgacc	3900
ctggacggct tcgatctgag cccactgctg ctgggaaccg gcaagtcccc taggcagtct	3960
ctgttctttt acccctccta tcctgatgag gtgcggggcg tgtttgccgt gagaaccggc	4020
aagtacaagg cccacttctt tacacagggc tctgcccaca gcgacaccac agcagatcca	4080
gcatgccacg ccagctcctc tctgaccgca caggagccac ctctgctgta cgacctgtcc	4140
aaggatcccg gcgagaacta taatctgctg ggaggagtgg caggagcaac ccctgaggtg	4200
ctgcaggccc tgaagcagct gcagctgctg aaggcacagc tggacgcagc agtgacattc	4260
ggcccaagcc aggtggccag aggcgaggat cccgccctgc agatctgttg ccaccccggc	4320
tgcaccccaa gacctgcctg ttgccattgc cccgaccac acgcctaaga ttctagagtc	4380
gagccgcgga ctagtaactt gtttattgca gcttataatg gttacaaata aagcaatagc	4440
atcaciaaatt tcaciaataa agcatttttt tctctgcatt ctagttgtgg tttgtccaaa	4500
ctcatcaatg tatcttacct gcaggctcac cagtgtttgt gactgggaac tctccctgcc	4560
aaatattggc ataatgctgt cctttaggtt gcagcttatt gcccagggg aacagtctgt	4620
tgtgcagtcc accccggcag gaatactccc attctgcctc tgttggtaac cttttcccag	4680
cccaggtgca gtatgccact gcatcattcc aggacacatg cagcacaggg tggtcaggcc	4740
tgtggaggat agttgagtct ggtccctctg ggtgtctcca attggctcct ttaacaggca	4800

gccaccaggg	ggctgcagcc	actgcctggt	ggatattggt	cttcacctgc	tcacttagca	4860
tgccctcaaa	aacaaaactg	tcaccaaatt	tctcagccctc	tgtaagggtat	ccagtgtgtgt	4920
ttacaaatft	ctcaaattct	gtgtttgaca	cttcataggc	atccatatag	aaggcatcaa	4980
ttgtgactct	cctagctggt	gcttcaccat	cctgcttgat	ctgagggtca	tcagttccca	5040
tagtaaaaac	tcctgcaggt	ctagatacgt	agataagtag	catggcggggt	taatcattaa	5100
ctacaaggaa	cccctagtga	tggagttggc	cactccctct	ctgcgcgctc	gctcgtcac	5160
tgaggccggg	cgaccaaagg	tcgcccgaag	cccgggcttt	gcccgggcgg	cctcagtgag	5220
cgagcgagcg	cgcagagagg	gagtggccaa	agatccccgg	gtaccgagga	cgaattctct	5280
agatatcgct	caatactgac	catttaaatt	atacctgacc	tccatagcag	aaagtcaaaa	5340
gcctccgacc	ggaggctttt	gacttgatcg	gcacgtaaga	ggttccaact	ttcaccataa	5400
tgaaataaga	tcactaccgg	gcgtatfttt	tgagttatcg	agatftttcag	gagctaagga	5460
agctaaaatg	agccatattc	aacgggaaac	gtcttgctcg	aggccgcgat	taaattccaa	5520
catggatgct	gatttatatg	ggtataaatg	ggctcgcgat	aatgtcgggc	aatcaggtgc	5580
gacaatctat	cgattgtatg	ggaagcccga	tgcgccagag	ttgtttctga	aacatggcaa	5640
aggtagcgft	gccaatgatg	ttacagatga	gatggtcagg	ctaaactggc	tgacggaatt	5700
tatgcctctt	ccgaccatca	agcattttat	ccgtactcct	gatgatgcat	ggttactcac	5760
cactgcgata	ccagggaaaa	cagcattcca	ggtattagaa	gaatatcctg	attcaggtga	5820
aaatattgft	gatgcgctgg	cagtgttccct	gcgccggttg	cattcgattc	ctgtttgtaa	5880
ttgtccfttt	aacggcgata	gcgtatfttcg	tctcgtctcag	gcgcaatcac	gaatgaataa	5940
cggftttggft	ggtgcgagtg	atfttgatga	cgagcgtaat	ggctggccctg	ttgaacaagt	6000
ctggaaagaa	atgcataagc	ftttgccatt	ctcaccggat	tcagtcgtca	ctcatgggtga	6060
ftttctactt	gataacctta	fttttgacga	ggggaaatta	ataggfttgta	ttgatgttgg	6120
acgagtcgga	atcgagacc	gataccagga	tcttgccatc	ctatgggaact	gcctcgggtga	6180
gtfttctcct	tcattacaga	aacggctfttt	tcaaaaatat	ggtattgata	atcctgatat	6240
gaataaatfg	cagftttcact	tgatgctcga	tgagftttttc	tgagggccca	aatgtaatca	6300

cctggctcac cttcgggtgg gcctttctgc gttgctggcg tt	6342
<210> 53	
<211> 6612	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成聚核苷酸	
<400> 53	
cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaa cttgcatgcc	60
tgcatttggc cactccctct ctgcgcgctc gctcgtcac tgaggccggg cgaccaaagg	120
tgcgccgacg cccgggcctt gcccgggcgg cctcagtgag cgagcgagcg cgcagagagg	180
gagtggccaa ctccatcact aggggttcct ggaggggtgg agtcgtgacg tgaattacgt	240
catagggtta gggaggctct gcagatcttc aatattggcc attagccata ttattcattg	300
gttatatagc ataaatcaat attggctatt ggccattgca tacgttgtat ctatatcata	360
atatgtacat ttatatggc tcatgtccaa tatgaccgcc atgttggcat tgattattga	420
ctagttatta atagtaatca attacggggt cattagtcca tagccatat atggagttec	480
gcgttacata acttacggta aatggcccg cctggctgacc gccaacgac ccccgcccat	540
tgacgtcaat aatgacgtat gtcccatag taacgccaat agggacttcc cattgacgtc	600
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	660
caagtccgcc ccctattgac gtcaatgacg gtaaattggc cgcctggcat tatgcccagt	720
acatgacctt acgggacttt cctacttggc agtacatcta cgtattagtc atcgtatta	780
ccatggtcga ggtgagcccc acgttctgct tcactctccc catctcccc ccctccccac	840
ccccaatttt gtatttatTT attttttaat tattttgtgc agcgatgggg gcggggggggg	900
gggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga	960
gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc	1020
ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgcgt	1080
gccttcgccc cgtgccccgc tccgcgcgc cctcgcgcgc cccgccccgg ctctgactga	1140

ccgcgttact	cccacaggtg	agcgggcggg	acggcccttc	tcctccgggc	tgtaattagc	1200
gcttggttta	atgacggctt	gtttcttttc	tgtggctgcg	tgaaagcctt	gaggggctcc	1260
gggagggccc	tttgtgcggg	gggagcggct	cgggggggtgc	gtgcgtgtgt	gtgtgcgtgg	1320
ggagcgcgc	gtgcggctcc	gcgctgcccc	gcggctgtga	gcgctgcggg	cgcggcgcgg	1380
ggcttttgtc	gctccgcagt	gtgcgcgagg	ggagcgcggc	cgggggcggt	gccccgcggt	1440
gcgggggggg	ctgcgagggg	aacaaaggct	gcgtgcgggg	tgtgtgcgtg	gggggggtgag	1500
cagggggtgt	gggcgcgtcg	gtcgggctgc	aacccccct	gcacccccct	ccccgagttg	1560
ctgagcacgg	cccggcttcg	ggtgcggggc	tccgtacggg	gcgtggcgcg	gggctcgccg	1620
tgccgggcgg	ggggtggcgg	caggtggggg	tgccgggcgg	ggcggggccg	cctcggggccg	1680
gggagggctc	gggggagggg	cgcggcggcc	cccggagcgc	cggcggctgt	cgaggcgcgg	1740
cgagccgcag	ccattgcctt	ttatggtaat	cgtgcgagag	ggcgcaggga	cttcctttgt	1800
cccaaatctg	tgcggagccg	aaatctggga	ggcgccgccg	cacccccctc	agcgggcgcg	1860
gggcgaagcg	gtgcggcgcc	ggcaggaagg	aaatgggcgg	ggagggcctt	cgtgcgtcgc	1920
cgcgccgccg	tccccctctc	cctctccagc	ctcggggctg	tccgcggggg	gacggctgcc	1980
ttcggggggg	acggggcagg	gcggggttcg	gcttctggcg	tgtgaccggc	ggctctagag	2040
cctctgctaa	ccatgttcat	gccttcttct	ttttcctaca	gctcctgggc	aacgtgctgg	2100
ttattgtgct	gtctcatcat	tttggcaaag	aattccgcca	ccatgtctat	gggggctcct	2160
cgtccctgc	tgctggcact	ggccgccggg	ctggctgtcg	caagaccacc	taatatcgtc	2220
ctgatttttg	cagacgatct	gggatacggc	gacctgggat	gctatggcca	cccaagctcc	2280
accacacca	acctggacca	gctggcagca	ggaggcctgc	ggttcaccga	cttctacgtg	2340
ccagtgagcc	tgtgcacccc	ctccagagcc	gccctgctga	caggcaggct	gccagtgcgc	2400
atgggcatgt	atcctggcgt	gctggtgcca	tctagcaggg	gcggcctgcc	actggaggag	2460
gtgaccgtgg	cagaggtgct	ggcagccaga	ggctacctga	caggaatggc	cggcaagtgg	2520
cacctgggag	tgggaccaga	gggagccttc	ctgccccctc	accagggttt	ccaccggttt	2580
ctgggcatcc	cttattctca	cgaccagggc	ccatgccaga	acctgacctg	ttttccacca	2640

gcaacaccat	gcgacggagg	atgtgatcag	ggcctgggtgc	caatcccact	gctggcaaat	2700
ctgagcgtgg	aggcacagcc	tccatggctg	cctggcctgg	aggcaagata	catggccttc	2760
gcccacgacc	tgatggcaga	tgcacagcgg	caggatagac	ctttctttct	gtactatgcc	2820
tcccaccaca	cccactatcc	acagttcagc	ggccagtcct	ttgccgagag	gtccggaagg	2880
ggaccattcg	gcgactctct	gatggagctg	gatgccgccg	tgggcaccct	gatgacagca	2940
atcggcgacc	tgggcctgct	ggaggagaca	ctggtcatct	tcaccgccga	taacggccct	3000
gagacaatgc	ggatgtctag	aggcggatgc	agcggcctgc	tgagatgtgg	caagggaacc	3060
acatacgagg	gaggcgtgcg	cgagcctgcc	ctggcatttt	ggccaggaca	catcgcacct	3120
ggagtgaccc	acgagctggc	ctcctctctg	gacctgctgc	caacactggc	cgccctggca	3180
ggagcacctc	tgccaaatgt	gaccctggac	ggcttcgatc	tgagcccact	gctgctggga	3240
accggcaagt	cccctaggea	gtctctgttc	ttttacccct	cctatcctga	tgaggtgcgg	3300
ggcgtgtttg	ccgtgagaac	cggcaagtac	aaggcccact	tctttacaca	gggctctgcc	3360
cacagcgaca	ccacagcaga	tccagcatgc	cacgccagct	cctctctgac	cgcacacgag	3420
ccacctctgc	tgtacgacct	gtccaaggat	cccggcgaga	actataatct	gctgggagga	3480
gtggcaggag	caaccctga	ggtgctgcag	gccctgaagc	agctgcagct	gctgaaggca	3540
cagctggacg	cagcagtgac	attcggccca	agccagggtg	ccagaggcga	ggatcccgcc	3600
ctgcagatct	gttgccaccc	cggctgcacc	ccaagacctg	cctgttgcca	ttgccccgac	3660
ccacacgcct	aagattctag	agtcgagccg	cggactagta	acttgtttat	tgcagcttat	3720
aatggttaca	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	tttttcaactg	3780
cattctagtt	gtggtttgtc	caaactcadc	aatgtatctt	aggtctagat	acgtagataa	3840
gtagcatggc	gggttaatca	ttaactacaa	ggaaccacct	gtgatggagt	tggccactcc	3900
ctctctgcgc	gctcgctcgc	tcactgaggc	cgggcgcacca	aaggtcgccc	gacgcccggg	3960
ctttgccccg	gcggcctcag	tgagcgagcg	agcgcgcaga	gagggagtgg	ccaaagatcc	4020
ccgggtaccg	agctcgaatt	cgtaatcatg	tcatagctgt	ttcctgtgtg	aaattgttat	4080
ccgctcacaa	ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	ctgggggtgcc	4140

taatgagtga gctaactcac attaatfgcg ttgcgctcac tgcccgcctt ccagtcggga	4200
aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt	4260
attggcgaac ttttgctgag ttgaaggatc agatcacgca tcttcccgac aacgcagacc	4320
gttccgtggc aaagcaaaag ttcaaaatca gtaaccgtca gtgccgataa gttcaaagtt	4380
aaacctggtg ttgataccaa cattgaaacg ctgatcgaaa acgcgctgaa aaacgctgct	4440
gaatgtgcga gcttcttccg ctctctcgct cactgactcg ctgcgctcgg tcgttcggct	4500
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga	4560
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc	4620
cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg	4680
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg	4740
aagctccctc gtgcgctctc ctgttccgac cctgccgcct accggatacc tgtccgcctt	4800
tctcccttcg ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcgg	4860
gtaggtcggt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg	4920
cgccttatcc ggtaactatc gtcttgagtc caacccggt aacacacgact tatcgccact	4980
ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcggtg ctacagagtt	5040
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggt aatgcgctct	5100
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac	5160
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc	5220
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg atccgtegag	5280
aggtctgcct cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc	5340
cagccagaaa gtgagggagc cacggttgat gagagctttg ttgtaggtgg accagttggt	5400
gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg	5460
atccttcaac tcagcaaaag ttcgatttat tcaacaaagc cacgttgtgt ctcaaaatct	5520
ctgatgttac attgcacaag ataaaaatat atcatcatga acaataaaac tgtctgctta	5580
cataaacagt aatacaaggg gtgttatgag ccatattcaa cgggaaacgt cttgctcgaa	5640

gccgcgatta aattccaaca tggatgctga tttatatggg tataaatggg ctcgcgataa	5700
tgtcggggcaa tcaggtgcga caatctatcg attgtatggg aagcccgatg cgccagagtt	5760
gtttctgaaa catggcaaag gtagcgttgc caatgatggt acagatgaga tggtcagact	5820
aaactggctg acggaattta tgcctcttcc gaccatcaag cattttatcc gtactcctga	5880
tgatgcatgg ttactcacca ctgcgatecc cgggaaaaca gcattccagg tattagaaga	5940
atatcctgat tcaggtgaaa atattgttga tgcgctggca gtgttcctgc gccggttgca	6000
ttcgattcct gtttgtaatt gtccttttaa cagcgatcgc gtatttcgtc tcgctcaggc	6060
gcaatcacga atgaataacg gtttggttga tgcgagtgat tttgatgacg agcgtaatgg	6120
ctggcctggt gaacaagtct ggaaagaaat gcataagctt ttgccattct caccggattc	6180
agtcgtcact catggtgatt tctcacttga taaccttatt tttgacgagg ggaaattaat	6240
aggttgtatt gatgttggac gagtcggaat cgcagaccga taccaggatc ttgccatcct	6300
atggaactgc ctcggtgagt tttctccttc attacagaaa cggccttttc aaaaatatgg	6360
tattgataat cctgatatga ataaattgca gtttcatttg atgctcgatg agtttttcta	6420
atcagaattg gttaattggg tgtaacactg gcagagcatt acgctgactt gacgggacgg	6480
cggccttggt gaataaatcg cattcgccat tcaggctgcg caactgttgg gaagggcgat	6540
cgggtgcgggc ctcttcgcta ttacgccage tggcgaaagg gggatgtgct gcaaggcgat	6600
taagttgggt aa	6612

<210> 54
<211> 918
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 54	
ggcatcctaa aaaatattca gtggaaacgt aaaaacatta aagactgatt aaacatcgca	60
gcatgacaca gatitagcaa ctgagcataa ataatttgac tcggatactg ctccaaaatc	120
cgaagaggac caatttcttc caggaggaca actacctcgt cctctgcaga cccctctcct	180

cggcagctga aggagtgtgg ccaatctgcc tccacctccc cgcggacccc ctactctcag	240
gacctcctgc agcacccecaa actggaagtg gccgctgcag acccaaggac gaggggcacg	300
cgggagccgg cagccctagt ggagcggttg gagatgttga ggtgggaggg tcacccaggt	360
ggggtgaggc tggggtaggt agcggagtga acggcttccg aagctctggg ccgccccag	420
gttggactaa gcaggcgctc tgtcttcgcc cccgcccagg gtgggcgtct cctgaggact	480
ccccgccaca cctgaccoga gaccgcgcgc ccagcctaga acgtttcccc gacccagcgt	540
agggccgccg cgactggcgg gcgagggtcg gcgggaggcc tggcgaaccc gggggcggga	600
ccaggcgggc aaggcccggc tgccgcagcg ccgctctgcg cgaggcggct ccgccgcggc	660
ggagggatac ggcgcaccat atatatatcg cggggcgag actcgcgctc cggcagtgg	720
gctgggagtg tcgtggacgc cgtgccgtta ctcgtagtca ggcggcggcg caggcggcgg	780
cggcggcata gcgcacagcg cgccttagca gcagcagcag cagcagcggc atcggaggta	840
ccccgccgt cgcagcccc gcgctggtgc agccaccctc gctccctctg ctcttcctcc	900
cttcgctcgc accaagag	918

- <210> 55
- <211> 953
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 55	
aattcggtac cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata	60
tatggagttc cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga	120
ccccgceca ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt	180
ccattgacgt caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt	240
gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca	300
ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt	360
catcgctatt accatggtcg aggtgagccc cacgttctgc ttcactctcc ccattctccc	420

cccccccca cccccaattt tgtatttatt tatttttttaa ttattttgtg cagcgatggg	480
ggcggggggg gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg	540
ggcgaggcgg agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt	600
tatggcgagg cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt	660
cgtgcgacg ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cggccgcccc	720
ggctctgact gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctcgg	780
gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc	840
ttgaggggct ccgggagcta gaggctctgc taacctgtt catgccttct tcttttctct	900
acagctcctg ggcaacgtgc tggttattgt gctgtctcat cattttggca aag	953

<210> 56
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 56	
gtagataagt agcatggcgg gttaatcatt aactaca	37

<210> 57
<211> 180
<212> DNA
<213> 人工序列

<220>
<223> 3' ITR

<400> 57	
gtagataagt agcatggcgg gttaatcatt aactacaagg aaccctagt gatggagttg	60
gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaaa ggtcgcgccga	120
cggccgggct ttgcccgggc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc	180

<210> 58
<211> 380

<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 58
ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60

catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgccca 120

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 180

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240

aagtgtatca tatgccaaagt ccgcccccta ttgacgtcaa tgacggtaaa tggcccgcc 300

ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat 360

tagtcatcgc tattaccatg 380

<210> 59
<211> 1246
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 59
tcgaggtgag ccccacgttc tgcttcactc tccccatctc cccccctcc ccacccccaa 60

ttttgtattt atttatTTTT taattatTTT gtgcagcgat gggggcgggg gggggggggg 120

ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg 180

cggcggcagc caatcagagc ggcgcgctcc gaaagtTTTc ttttatggcg aggcggcggc 240

ggcggcgggc ctataaaaag cgaagcgcgc ggcgggcggg agtcgctgcg cgctgccttc 300

gccccgtgcc ccgtccgcc gccgcctcgc gccgcccgcc ccggctctga ctgaccgcgt 360

tactcccaca ggtgagcggg cgggacggcc cttctcctcc gggctgtaat tagcgcttgg 420

tttaatgacg gcttgTTTTt tttctgtggc tgcgtgaaag ccttgagggg ctccgggagg 480

gccctttgtg cgggggggagc ggctcggggg gtgcgtgcgt gtgtgtgtgc gtggggagcg 540

ccgcgtgcgg ctccgcgctg cccggcggt gtgagcgtg cgggcgcggc gcggggcttt 600

gtgcgctccg cagtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg	660
ggggctgcga ggggaacaaa ggctgcgtgc ggggtgtgtg cgtggggggg tgagcagggg	720
gtgtgggcgc gtcggtcggg ctgcaacccc cctgcaccc cctcccccga gttgctgagc	780
acggccccggc ttcgggtgcg gggctccgta cggggcgctg gcgggggctc gccgtgccgg	840
gcgggggggtg gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg	900
gctcggggga ggggcgcggc ggcccccgga gcgccggcgg ctgtcgaggc gcggcgagcc	960
gcagccattg ccttttatgg taatcgtgcg agagggcgca gggacttcct ttgtcccaa	1020
tctgtgcgga gccgaaatct gggaggcgcc gccgcacccc ctctagcggg cgcggggcga	1080
agcggtgccg cggcggcagg aaggaaatgg gcggggaggg ctttcgtgcg tcgccgcgcc	1140
gccgtcccct tctccctctc cagcctcggg gctgtccgcg gggggacggc tgccttcggg	1200
ggggacgggg cagggcgggg ttcggcttct ggcgtgtgac cggcgg	1246

<210> 60
<211> 95
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 60	
cctctgctaa ccatgttcat gccttcttct ttttctaca gctcctgggc aacgtgctgg	60
ttattgtgct gtctcatcat tttggcaaag aattc	95

<210> 61
<211> 1061
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 61	
tagggaggtc ctgcacgtta cataacttac ggtaaattgc ccgcctggct gaccgccc	60
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac	120

tttccattga cgtcaatggg tggagtatit acggtaaact gcccacttgg cagtacatca	180
agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaata ggcccgcctg	240
gcattatgcc cagtacatga ccttatggga ctttccctact tggcagtaca tctacgtatt	300
agtcatecgt attaccatgg tcgaggtgag ccccacgttc tgcttcactc tccccatctc	360
ccccccctcc ccacccccaa ttttgtatit atttatitit taattatitit gtgcagcgat	420
gggggcgggg ggggggggggg gcgcgcgcca ggcggggcgg ggcgggggcga ggggcggggc	480
ggggcgaggc ggagaggtgc ggcggcagcc aatcagagcg gcgcgctccg aaagtittcct	540
tttatggcga ggcgggcgcg gcggcgggccc tataaaaagc gaagcgcgcg gcgggcgggga	600
gtcgtcgcg gctgccttcg ccccgctgcc cgctccgccc ccgcctcgcg ccgcccggcc	660
cggcctcgac tgaccgcgtt actaaaacag gtaagtcgg cctccgcgcc gggttttggc	720
gcctcccgcg ggcgcccccc tcttcacggc gacgcctgcc acgtcagacg aagggcgcag	780
cgagcgtcct gatccttcg cccggacgct caggacagcg gcccgctgct cataagactc	840
ggccttagaa cccagtatc agcagaagga cattttagga cgggacttgg gtgactctag	900
ggcactgggt ttctttccag agagcggaac aggcgaggaa aagtagtccc ttctcggcga	960
ttctgcggag ggatctccgt ggggcggtga acgccgatga tgcctctact aacctgttc	1020
atgttttctt tttttttcta caggtcctgg gtgacgaaca g	1061

- <210> 62
- <211> 1527
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 62	
atgagcatgg gcgccccag aagcctgtta cttgcttttag ctgctggcct tgcagtggca	60
aggcccccta acatcgtgct gatctttgca gatgacttgg gatatgggga tcttggttgt	120
tatggccacc catcaagcac aactcccaat ctggatcagt tggtgcagg aggtctgagg	180
tttacagact tttatgttcc agtctccctg tgcactcctt ctgggctgc cctgcttact	240

gggaggctcc ctgtgagaat gggatatgtac cctggagtgt tggteccatc cagcagggga	300
gggctgcccc tggaagaggt gacagtggca gaggtgctgg cagcacgagg ctatctgact	360
ggcatggcag gcaagtggca cctgggtgta gggccagagg gtgctttcct gcctcccat	420
cagggctttc ataggtttct gggaatecca tactctcatg accaaggacc ctgccagaac	480
ctcacctgtt tccccctgc aacaccatgt gatgggggct gtgatcaagg tctggttcct	540
ataccactgc ttgctaatct ttcagtggaa gctcaaccac cctggctgcc tggcttggag	600
gctagataca tggccttcgc acatgatctg atggcagatg cccagagaca agataggcct	660
ttcttcctct actatgcate tcaccacacc cactatectc agttctcagg ccaatcattt	720
gctgagcgta gtggcagggg cccatttggg gacagtttga tggaactgga tgccgcagtt	780
ggtaccctca tgacagcaat aggggactta ggtttgctgg aggaaacatt ggtaattttc	840
acagctgata atggccctga gacaatgaga atgtctaggg gaggctgctc tggctctctg	900
aggtgtggta aagggactac atatgaggga ggagtgaggg aaccagctct tgccttttgg	960
ccaggtcaca tagccccctgg agttacacat gaactagctt ctccctgga cttgcttcct	1020
acactggcag ccctggcagg tgccctctc cctaatagtaa ctttagatgg atttgacctc	1080
tctccactac ttttagggac agggaaaagt ccaaggcagt cttattctt ctatccttc	1140
taccagatg aggtgagggg tgtttttgcc gtgaggactg ggaaatacaa agctcatttt	1200
tttaccagg gatcagctca ttcagacacc acagctgac ctgcctgtca tgccagcagt	1260
agcttgacag cacatgagcc tcccttactg tatgacctga gcaaggaccc aggggagAAC	1320
tataacctgc ttgggggggt tgctggggcc acccagaag tgcttcaggc actaaagcag	1380
ctgcaactgc ttaaagcaca gttggatgct gcagtgacct ttggcccttc ccaggtggcc	1440
agaggcgagg atcccgccct gcagatctgc tgccaccag gctgcacacc cagacctgcc	1500
tgctgtcact gccccgaccc acacgcc	1527

- <210> 63
- <211> 57
- <212> DNA
- <213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 63
gctactaact tcagcctgct gaagcaggct ggagacgtgg aggagaacct tggacct 57

<210> 64
<211> 1122
<212> DNA
<213> 智人

<400> 64
atggctgccc cagccctggg gctggtgtgt ggcagatgcc ctgagctggg cctggtgctg 60

cttctcctgc tgctgagcct cctgtgtggt gctgctggct ctcaggaagc agggacagga 120

gcaggagcag gttctctggc tggctcatgc ggttgtggga cccccagag gccaggggct 180

catgggtcct ctgcagctgc ccacaggtac tcaagggaag caaatgcccc tggccccgta 240

cctggggaaa ggcaacttgc tactccaag atggttccta tccctgcagg agtttttact 300

atgggaactg atgacctca gatcaagcag gatggtgaag caccagctag gagagtcaca 360

attgatgcct tctatatgga tgcctatgaa gtgtcaaaca cagaatttga gaaatttgta 420

aacagcactg gataccttac agaggctgag aaatttggtg acagttttgt ttttgaaggc 480

atgctaagtg agcaggtgaa gaccaatatc caacaggcag tggctgcagc cccctggtgg 540

ctgcctgtta aaggagccaa ttggagacac ccagagggac cagactcaac tatectccac 600

aggcctgacc accctgtgct gcatgtgtcc tggaatgatg cagtggcata ctgcacctgg 660

gctgggaaaa ggttaccaac agaggcagaa tgggagtatt cctgccgggg tggactgcac 720

aacagactgt tcccctgggg caataagctg caacctaaag gacagcatta tgccaatatt 780

tggcagggag agtteccagt cacaacact ggtgaggatg gcttccaggg aactgcccct 840

gtggatgctt tcccacccaa tggctatggg ttgtacaata tagttgggaa tgcctgggag 900

tggacttctg actggtggac ggtccatcac agtgtggaag agacactgaa cccaaagggg 960

ccccctcag gcaaggacag agtcaagaaa ggtggctctt atatgtgtca cagaagctat 1020

tgctacagat ataggtgtgc tgcaagaagt cagaacaccc ctgacagctc agctagcaat 1080

ctgggattta gatgtgcagc agatagactc cccaccatgg ac	1122
<210> 65	
<211> 3739	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成聚核苷酸	
<400> 65	
ggcatcctaa aaaatattca gtggaaacgt aaaaacatta aagactgatt aaacatcgca	60
gcatgacaca gatttagcaa ctgagcataa ataatttgac tcggatactg ctccaaaatc	120
cgaagaggac caattttcttc caggaggaca actacctcgt cctctgcaga cccctctcct	180
cggcagctga aggagtgtgg ccaatctgcc tccacctccc cgcggacccc ctactctcag	240
gacctcctgc agcaccccaa actggaagtg gccgctgcag acccaaggac gaggggcacg	300
cgggagccgg cagccctagt ggagcggttg gagatgttga ggtgggaggg tcaccaggt	360
ggggtgaggc tggggtaggt agcggagtga acggcttccg aagctctggg ccgccccag	420
gttggactaa gcaggcgctc tgtcttcgcc cccgcccagg gtgggcgtct cctgaggact	480
ccccgccaca cctgaccoga gaccgcgcgc ccagcctaga acgcttcccc gaccacagct	540
agggccgccg cgactggcgg gcgagggtcg gcgggaggcc tggcgaacct gggggcgggga	600
ccaggcgggc aaggcccggc tgccgcagcg ccgctctgcg cgaggcggct ccgccgcggc	660
ggagggatac ggcgcaccat atatatatcg cggggcgcag actcgcgctc cggcagtgg	720
gctgggagtg tcgtggacgc cgtgccgtta ctcgtagtca ggcggcggcg caggcggcgg	780
cggcggcata gcgcacagcg cgccttagca gcagcagcag cagcagcggc atcggaggt	840
ccccgccgt cgcagcccc gcgctggtgc agccaccctc gctccctctg ctcttctctc	900
cttcgctcgc accaagaggt aagggtttaa gggatggttg gttggtgggg tattaatgtt	960
taattacctg gagcacctgc ctgaaateac tttttttcag gttgggccac ccgccgccac	1020
catgagcatg ggcgccccca gaagcctgtt acttgcttta gctgctggcc ttgcagtggc	1080
aaggccccct aacatcgtgc tgatctttgc agatgacttg ggatatgggg atcttggttg	1140

ttatggccac ccatcaagca caactcccaa tctggatcag ttggctgcag gaggtctgag	1200
gtttacagac ttttatgttc cagtctccct gtgcactcct tctcgggctg ccctgcttac	1260
tgggaggctc cctgtgagaa tgggtatgta ccctggagtg ttgggtcccat ccagcagggg	1320
agggctgccc ctggaagagg tgacagtggc agaggtgctg gcagcacgag gctatctgac	1380
tggcatggca ggcaagtggc acctgggtgt agggccagag ggtgctttcc tgccctccca	1440
tcagggtttt cataggtttc tgggaatccc atactctcat gaccaaggac cctgccagaa	1500
cctcacctgt tccccccctg caacaccatg tgatgggggc tgtgatcaag gtctggttcc	1560
tataccactg cttgctaate tttcagtgga agctcaacca ccctggctgc ctggcttgga	1620
ggctagatac atggccttcg cacatgatct gatggcagat gccagagac aagataggcc	1680
tttcttcttc tactatgcat ctaccacac ccactatcct cagttctcag gccaatcatt	1740
tgctgagcgt agtggcaggg gcccatttgg ggacagtttg atggaactgg atgccgcagt	1800
tggtagcctc atgacagcaa taggggactt aggtttgctg gaggaacat tggtaatttt	1860
cacagctgat aatggccctg agacaatgag aatgtctagg ggaggctgct ctggtcttct	1920
gaggtgtggt aaagggacta catatgaggg aggagtgagg gaaccagctc ttgccttttg	1980
gccaggtcac atagccctg gagttacaca tgaactagct tcttccctgg acttgcttcc	2040
tacactggca gccctggcag gtgccccctc ccctaatagta actttagatg gatttgacct	2100
ctctccacta cttttaggga cagggaaaag tccaaggcag tccttattct tctatccttc	2160
ctaccagat gaggtgaggg gtgtttttgc cgtgaggact gggaaataca aagctcattt	2220
ttttaccag ggatcagctc attcagacac cacagctgat cctgcctgtc atgccagcag	2280
tagcttgaca gcacatgagc ctcccttact gtatgacctg agcaaggacc caggggagaa	2340
ctataacctg cttggggggg ttgctggggc caccacagaa gtgcttcagg cactaaagca	2400
gctgcaactg cttaaagcac agttggatgc tgcagtgacc tttggccctt ccaggtggc	2460
cagaggcgag gatcccgccc tgcagatctg ctgccacca ggctgcacac ccagacctgc	2520
ctgctgtcac tgccccgacc cacacgccgg cageggagct actaacttca gcctgctgaa	2580
gcaggctgga gacgtggagg agaaccctgg acctatggct gcccagccc tggggctggt	2640

gtgtggcaga	tgccctgagc	tgggcctggt	gctgcttctc	ctgctgctga	gcctcctgtg	2700
tggctgctgct	ggctctcagg	aagcagggac	aggagcagga	gcaggttctc	tggctggctc	2760
atgcggttgt	gggaccccc	agaggccagg	ggctcatggg	tcctctgcag	ctgcccacag	2820
gtactcaagg	gaagcaaatg	cccctggccc	cgtacctggg	gaaaggcaac	ttgctcactc	2880
caagatgggt	cctatccctg	caggagtitt	tactatggga	actgatgacc	ctcagatcaa	2940
gcaggatggt	gaagcaccag	ctaggagagt	cacaattgat	gccttctata	tggatgccta	3000
tgaagtgtca	aacacagaat	ttgagaaatt	tgtaaacagc	actggatacc	ttacagaggc	3060
tgagaaattt	ggtgacagtt	ttgtttttga	aggcatgcta	agtgagcagg	tgaagaccaa	3120
tatccaacag	gcagtggctg	cagccccctg	gtggctgcct	gttaaaggag	ccaattggag	3180
acacccagag	ggaccagact	caactatcct	ccacaggcct	gaccaccctg	tgctgcatgt	3240
gtcctggaat	gatgcagtgg	catactgcac	ctgggctggg	aaaaggttac	caacagaggc	3300
agaatgggag	tattcctgcc	ggggtggact	gcacaacaga	ctgttcccct	ggggcaataa	3360
gctgcaacct	aaaggacagc	attatgccaa	tatttggcag	ggagagttec	cagtcacaaa	3420
cactggtgag	gatggcttcc	agggaactgc	ccctgtggat	gctttcccac	ccaatggcta	3480
tgggttgtac	aatatagtgt	ggaatgcctg	ggagtggact	tctgactggg	ggacgggtcca	3540
tcacagtgtg	gaagagacac	tgaacccaaa	ggggcccccc	tcaggcaagg	acagagtcaa	3600
gaaaggtggc	tcttatatgt	gtcacagaag	ctattgctac	agatataggt	gtgctgcaag	3660
aagtcagaac	acccttgaca	gtcagctag	caatctggga	tttagatgtg	cagcagatag	3720
actccccacc	atggactga					3739

- <210> 66
- <211> 54
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400>	66																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			</
-------	----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

<210> 67
<211> 3686
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 67	
aattcgggtac cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata	60
tatggagttc cgcgttacat aacttacggg aaatggcccc cctggctgac cgcccaacga	120
ccccgcccc ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt	180
ccattgacgt caatgggtgg actattttacg gtaaactgcc cacttggcag tacatcaagt	240
gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc cgccttggca	300
ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt	360
catcgctatt accatggteg aggtgagccc cacgtttctgc ttcactctcc ccctctcccc	420
ccccccccca cccccaattt tgtattttatt ttttttttaa ttattttgtg cagcgatggg	480
ggcgggggggg ggggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg	540
ggcgaggcgg agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttctttt	600
tatggcgagg cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt	660
cgtgcgacg ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc	720
ggctctgact gaccgcgtta ctcccacagg tgagcgggcg ggaaggccct tctcctcgg	780
gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc	840
ttgaggggct ccgggagcta gagcctctgc taaccatgtt catgccttct tctttttcct	900
acagctcctg ggcaacgtgc tggttattgt gctgtctcat cattttggca aaggctagcg	960
ccgccaccat gagcatgggc gccccagaa gcctgttact tgcttttagct gctggccttg	1020
cagtggcaag gccccctaac atcgtgctga tctttgcaga tgacttggga tatggggatc	1080
ttggttgfta tggccacca tcaagcacia ctcccactct ggatcagttg gctgcaggag	1140
gtctgaggtt tacagacttt tatgttccag tctccctgtg cactccttct cgggctgccc	1200

tgccttactgg	gaggctccct	gtgagaatgg	gatatgtaccc	tggagtgttg	gtcccatcca	1260
gcaggggagg	gctgcccctg	gaagaggtga	cagtggcaga	ggtgctggca	gcacgaggct	1320
atctgactgg	catggcaggc	aagtggcacc	tgggtgtagg	gccagagggt	gctttcctgc	1380
ctcccatca	gggctttcat	aggtttcttg	gaatcccata	ctctcatgac	caaggaccct	1440
gccagaacct	cacctgtttc	ccccctgcaa	caccatgtga	tgggggctgt	gatcaaggtc	1500
tggttcctat	accactgctt	gctaatacttt	cagtggaagc	tcaaccaccc	tggctgcctg	1560
gcttggaggc	tagatacatg	gccttcgcac	atgatctgat	ggcagatgcc	cagagacaag	1620
ataggccttt	cttcctctac	tatgcatctc	accacaccca	ctatcctcag	ttctcaggcc	1680
aatcatttgc	tgagcgtagt	ggcaggggcc	catttgggga	cagtttgatg	gaactggatg	1740
cgcagtttg	tacctcatg	acagcaatag	gggacttagg	tttgctggag	gaaacattgg	1800
taattttcac	agctgataat	ggccctgaga	caatgagaat	gtctagggga	ggctgctctg	1860
gtcttctgag	gtgtggtaaa	gggactacat	atgagggagg	agtgagggaa	ccagctcttg	1920
ccttttggcc	aggtcacata	gcccctggag	ttacacatga	actagcttct	tccctggact	1980
tgcttcctac	actggcagcc	ctggcaggtg	ccccctctcc	taatgtaact	ttagatggat	2040
ttgacctctc	tccactactt	ttagggacag	ggaaaagtcc	aaggcagtcc	ttattcttct	2100
atccttccta	cccagatgag	gtgaggggtg	tttttgccgt	gaggactggg	aaatacaaag	2160
ctcatttttt	taccagggga	tcagctcatt	cagacaccac	agctgatcct	gcctgtcatg	2220
ccagcagtag	cttgacagca	catgagcctc	ccttactgta	tgacctgagc	aaggaccag	2280
gggagaacta	taacctgctt	gggggggttg	ctggggccac	cccagaagtg	cttcaggcac	2340
taaagcagct	gcaactgctt	aaagcacagt	tggatgctgc	agtgaccttt	ggcccttccc	2400
aggtggccag	aggcgaggat	cccgccctgc	agatctgctg	ccaccagggc	tgcacacca	2460
gacctgcctg	ctgtcactgc	cccgaccac	acgccggcag	cggagctact	aacttcagcc	2520
tgctgaagca	ggctggagac	gtggaggaga	accctggacc	tatggctgcc	ccagccctgg	2580
ggctggtgtg	tggcagatgc	cctgagctgg	gcctggtgct	gcttctcctg	ctgctgagcc	2640
tctgtgtgg	tgctgctggc	tctcaggaag	cagggacagg	agcaggagca	ggttctctgg	2700

ctggctcatg cggttggtggg accccccaga ggccaggggc tcatgggtcc tctgcagctg	2760
cccacaggta ctcaaggga gcaaatgccc ctggccccgt acctggggaa aggcaacttg	2820
ctcactccaa gatggttctt atccctgcag gagtttttac tatgggaact gatgaccctc	2880
agatcaagca ggatggtgaa gcaccagcta ggagagtcac aattgatgcc ttctatatgg	2940
atgcctatga agtgtcaaac acagaatttg agaaatttgt aaacagcact ggatacctta	3000
cagaggctga gaaatttggt gacagttttg tttttgaagg catgctaagt gagcaggatga	3060
agaccaatat ccaacaggca gtggctgcag cccctgggtg gctgcctgtt aaaggagcca	3120
attggagaca cccagaggga ccagactcaa ctatcctcca caggcctgac caccctgtgc	3180
tgcattgtgc ctggaatgat gcagtggcat actgcacctg ggctgggaaa aggttaccaa	3240
cagaggcaga atgggagtat tctgcccggg gtggactgca caacagactg ttcccctggg	3300
gcaataagct gcaacctaaa ggacagcatt atgccaatat ttggcaggga gattcccag	3360
tcacaaacac tggtgaggat ggcttccagg gaactgcccc tgtggatgct ttcccacca	3420
atggctatgg gttgtacaat atagttagga atgcctggga gtggacttct gactggtgga	3480
cggtccatca cagtgtggaa gagacactga acccaaaggg gccccctca ggcaaggaca	3540
gagtcaagaa aggtggctct tatatgtgtc acagaagcta ttgctacaga tataggtgtg	3600
ctgcaagaag tcagaacacc cctgacagct cagctagcaa tctgggattt agatgtgcag	3660
cagatagact cccaccatg gactga	3686

- <210> 68
- <211> 4346
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 68	
ttggcactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc	60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaaactca tcactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag	180

ggttagggag gtcctgcata tgcggccgcg gcatacctaaa aaatatcag tggaaacgta	240
aaaacattaa agactgatta aacatcgcag catgacacag atttagcaac tgagcataaa	300
taatttgact cggatactgc tccaaaatcc gaagaggacc aatttcttcc aggaggacaa	360
ctacctcgtc ctctgcagac ccctctcctc ggcagctgaa ggagtgtggc caatctgcct	420
ccacctcccc gcggaccccc tactctcagg acctcctgca gcacccccaaa ctggaagtgg	480
ccgctgcaga cccaaggacg aggggcacgc gggagccggc agccctagtg gagcggttgg	540
agatgttgag gtgggagggg caccaggtg gggtgaggct ggggtaggta gcggagtgaa	600
cggcttccga agctctgggc cgccccagg ttggactaag caggcgtct gtcttcgccc	660
ccgcccaggg tgggcgtctc ctgaggactc cccgccacac ctgacctgag accgcgcgcc	720
cagcctagaa cgcttccccg acccagcgtg gggccgccgc gactggcggg cgagggtcgg	780
cgggaggcct ggccaacccg ggggcgggac caggcgggca aggcccggt gccgcagcgc	840
cgctctgcgc gaggcggctc cgccgcggcg gagggatacg gcgcaccata tatatatcgc	900
ggggcgcaga ctgcgcctcc ggcagtgggt ctgggagtgt cgtggacgcc gtgccgttac	960
tcgtagtcag gcggcggcgc aggcggcggc ggcggcatag cgcacagcgc gccttagcag	1020
cagcagcagc agcagcggca tcggaggtac ccccgccgtc gcagcccccg cgctggtgca	1080
gccacctcg ctccctctgc tcttcctccc ttcgctcgca ccaagaggta agggtttaag	1140
ggatggttgg ttggtggggt attaatgttt aattacctgg agcacctgcc tgaaatcact	1200
ttttttcagg ttgggccacc cgccgccacc atgagcatgg gcgccccag aagcctgtta	1260
cttgctttag ctgctggcct tgcagtggca agggcccccta acatcgtgct gatctttgca	1320
gatgacttgg gatatgggga tcttggttgt tatggccacc catcaagcac aactcccaat	1380
ctggatcagt tggtgcagg aggtctgagg ttacagact tttatgttcc agtctccctg	1440
tgcactcctt ctcgggctgc cctgcttact gggaggctcc ctgtgagaat gggtatgtac	1500
cctggagtgt tggtcccatc cagcagggga gggctgcccc tggaagaggt gacagtggca	1560
gaggtgctgg cagcacgagg ctatctgact ggcatggcag gcaagtggca cctgggtgta	1620
gggccagagg gtgctttcct gcctccccat cagggccttc ataggtttct gggaatccca	1680

tactctcatg accaaggacc ctgccagaac ctcacctgtt tccccctgc aacaccatgt	1740
gatgggggct gtgatcaagg tctggttcct ataccactgc ttgctaatct ttcagtggaa	1800
gctcaaccac cctggctgcc tggcttggag gctagataca tggccttcgc acatgatctg	1860
atggcagatg cccagagaca agataggcct ttcttctct actatgcac tcaccacacc	1920
cactatcctc agttctcagg ccaatcattt gctgagcgta gtggcagggg cccatttggg	1980
gacagtttga tggaactgga tgccgcagtt ggtaccctca tgacagcaat aggggactta	2040
ggtttgctgg aggaaacatt ggtaattttc acagctgata atggccctga gacaatgaga	2100
atgtctaggg gaggctgctc tggctctctg aggtgtggta aagggactac atatgaggga	2160
ggagtgaggg aaccagctct tgccttttgg ccaggtcaca tagcccttg agttacacat	2220
gaactagctt ctccctgga ctgcttctc acactggcag ccctggcagg tgccctctc	2280
cctaataata ctttagatgg atttgacctc tctccactac ttttagggac agggaaaagt	2340
ccaaggcagt ccttattctt ctatccttcc taccagatg aggtgagggg tgtttttgcc	2400
gtgaggactg ggaaatacaa agctcatttt tttaaccagg gatcagctca ttcagacacc	2460
acagctgac ctcctgtca tgccagcagt agcttgacag cacatgagcc tcccttactg	2520
tatgacctga gcaaggaccc aggggagaaac tataacctgc ttgggggggt tgctggggcc	2580
acccagaag tgcttcaggc actaaagcag ctgcaactgc ttaaagcaca gttggatgct	2640
gcagtgacct ttggcccttc ccaggtggcc agaggcgagg atcccgccct gcagatctgc	2700
tgccaccag gctgcacacc cagacctgcc tgctgtcact gccccgacc acacgccggc	2760
agcggagcta ctaacttcag cctgctgaag caggctggag acgtggagga gaaccctgga	2820
cctatggctg cccagccct ggggctggtg tgtggcagat gccctgagct gggcctggtg	2880
ctgcttctcc tgctgctgag cctcctgtgt ggtgctgctg gctctcagga agcagggaca	2940
ggagcaggag caggttctct ggctggctca tgcggttggt ggacccccca gaggccaggg	3000
gctcatgggt cctctgcagc tgcccacagg tactcaaggg aagcaaatgc ccctggcccc	3060
gtacctgggg aaaggcaact tgctcactcc aagatgggtc ctatccctgc aggagttttt	3120
actatgggaa ctgatgaccc tcagatcaag caggatggtg aagcaccagc taggagagtc	3180

acaattgatg ccttctatat ggatgcctat gaagtgtcaa acacagaatt tgagaaattt	3240
gtaaacagca ctggatacct tacagaggct gagaaatttg gtgacagttt tgtttttgaa	3300
ggcatgctaa gtgagcaggt gaagaccaat atccaacagg cagtggctgc agccccctgg	3360
tggctgcctg ttaaaggagc caattggaga caccagagg gaccagactc aactatcctc	3420
cacaggcctg accacctgt gctgcatgtg tcttggaatg atgcagtggc atactgcacc	3480
tgggctggga aaaggttacc aacagaggca gaatgggagt attcctgccg gggtggactg	3540
cacaacagac tgttcccctg gggcaataag ctgcaaccta aaggacagca ttatgccaat	3600
atttggcagg gagagtccc agtcacaaac actggtgagg atggcttcca gggaactgcc	3660
cctgtggatg ctttcccacc caatggctat gggttgtaca atatagttgg gaatgcctgg	3720
gagtggactt ctgactggtg gacggtccat cacagtgtgg aagagacact gaacccaaag	3780
gggccccct caggcaagga cagagtcaag aaaggtggct cttatatgtg tcacagaagc	3840
tattgctaca gatataggtg tgctgcaaga agtcagaaca cccctgacag ctcagctagc	3900
aatctgggat ttagatgtgc agcagataga ctccccacca tggactgaga tccagacatg	3960
ataagataca ttgatgagtt tggacaaacc acaactagaa tgcagtgaaa aaaatgcttt	4020
atttgtgaaa tttgtgatgc tattgcttta tttgtaacca ttataagctg caataaaca	4080
gttaacaaca acaattgcat tcattttatg tttcaggttc agggggaggt gtgggaggtt	4140
ttttaaacct gcaggtctag atacgtagat aagtagcatg gcgggttaat cattaactac	4200
aaggaacccc tagtgatgga gttggccact cctctctctgc gcgctcgtc gctcactgag	4260
gccgggcgac caaaggtcgc ccgacgcccg ggctttgccc gggcggcctc agtgagcgag	4320
cgagcgcgca gagagggagt ggccaa	4346

<210> 69
<211> 4492
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 69

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc	60
cgacgccccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tcactagggg ttcttgagg ggtggagtcg tgacgtgaat tacgtcatag	180
ggttagggag gtctgcata tgcggccgca cctaggtcat tctggcctcc cctccctca	240
aggccagtea ttctggcctg tcttccccg aaggccagtc attctggcct cccctcccc	300
caaggccagt cattctggcc tccccctcc ttaaggccag agtactatcg attcacacaa	360
aaaaccaaca cactattgca atgaaaataa atttcttta ttaagcttaa ttcggtacct	420
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg	480
cgttacataa cttacggtaa atggcccgc tggtgaccg cccaacgacc ccgcccatt	540
gacgtcaata atgacgtatg tcccatagt aacgccaata gggactttcc attgacgtca	600
atgggtggac tatttacggt aaactgcccc cttggcagta catcaagtgt atcatatgcc	660
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	720
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	780
catggtcgag gtgagcccca cgttctgctt cactctcccc atctcccccc cctccccacc	840
cccaattttg tatttattta ttttttaatt attttgtgca gcgatggggg cggggggggg	900
ggggggggcgc gcgccaggcg gggcggggcg gggcgagggg cggggcgggg cgaggcggag	960
aggtgcggcg gcagccaatc agagcggcgc gctccgaaag ttctctttta tggcgaggcg	1020
gcggcggcgg cggccctata aaaagcgaag cgcgcggcgg gcgggagtcg ctgcgacgt	1080
gccttcgccc cgtgccccgc tccgccgcc cctcgcgcc cccgccccgg ctctgactga	1140
ccgcgttact cccacaggtg agcgggcggg acggcccttc tctccgggc tgtaattagc	1200
gcttggttta atgacggctt gtttcttttc tgttgctgcg tgaaagcctt gaggggctcc	1260
gggagctaga gcctctgcta accatgttca tgctttcttc ttttccctac agctcctggg	1320
caacgtgctg gttattgtgc tgtctcatca ttttgcaaaa ggctagcgcc gccacatga	1380
gcatgggcgc cccagaagc ctgttacttg ctttagctgc tggccttgca gtggcaaggc	1440
cccctaacat cgtgctgac tttgcagatg acttgggata tggggatctt ggttgttatg	1500

gccacccatc aagcacaact cccaatctgg atcagttggc tgcaggaggt ctgaggttta	1560
cagactttta tgttccagtc tccctgtgca ctccctctcg ggctgccctg cttactggga	1620
ggctccctgt gagaatgggt atgtaccctg gagtgttggc cccatccagc aggggagggc	1680
tgccccctgga agaggtgaca gtggcagagg tgctggcagc acgaggctat ctgactggca	1740
tggcaggcaa gtggcacctg ggtgtagggc cagagggcgc tttcctgcct ccccatcagg	1800
gctttcatag gtttctggga atcccatact ctcatgacca aggaccctgc cagaacctca	1860
cctgtttccc ccctgcaaca ccatgtgatg ggggctgtga tcaaggctctg gttcctatac	1920
cactgcttgc taatctttca gtggaagctc aaccaccctg gctgcctggc ttggaggcta	1980
gatacatggc cttcgcacat gatctgatgg cagatgccc gagacaagat aggcctttct	2040
tcctctacta tgcctctcac cacaccact atcctcagtt ctcaggccaa tcatttgctg	2100
agcgtagtgg caggggcccc tttggggaca gtttgatgga actggatgcc gcagttggta	2160
ccctcatgac agcaataggg gacttaggtt tgctggagga aacattggta attttcacag	2220
ctgataatgg ccctgagaca atgagaatgt ctaggggagg ctgctctggt cttctgaggt	2280
gtggtaaagg gactacatat gagggaggag tgagggaacc agctcttgcc ttttggccag	2340
gtcacatagc ccctggagtt acacatgaac tagctttctc cctggacttg cttcctacac	2400
tggcagccct ggcaggtgcc cctctcccta atgtaacttt agatggattt gacctctctc	2460
cactactttt agggacaggg aaaagtccaa ggcagtcctt attcttctat ccttcctacc	2520
cagatgaggt gaggggtgtt tttgccgtga ggactgggaa atacaaagct cattttttta	2580
cccagggatc agctcattca gacaccacag ctgatcctgc ctgtcatgcc agcagtagct	2640
tgacagcaca tgagcctccc ttactgtatg acctgagcaa ggaccaggg gagaactata	2700
acctgcttgg ggggggttgct gggggccacc cagaagtgct tcaggcacta aagcagctgc	2760
aactgcttaa agcacagttg gatgctgcag tgacctttgg cccttcccag gtggccagag	2820
gcgaggatcc cgccctgcag atctgctgcc acccaggctg cacaccaga cctgcctgct	2880
gtcactgccc cgaccacac gccggcagcg gagctactaa cttcagcctg ctgaagcagg	2940
ctggagacgt ggaggagaac cctggaccta tggctgcccc agccctgggg ctggtgtgtg	3000

gcagatgccc	tgagctgggc	ctggtgctgc	ttctcctgct	gctgagcctc	ctgtgtggtg	3060
ctgctggctc	tcaggaagca	gggacaggag	caggagcagg	ttctctggct	ggctcatgcg	3120
gttgtgggac	ccccagagg	ccaggggctc	atgggtcctc	tgcagctgcc	cacaggtact	3180
caagggaagc	aaatgcccc	ggccccgtac	ctggggaaag	gcaacttgct	cactccaaga	3240
tggttcctat	ccctgcagga	gtttttacta	tgggaactga	tgaccctcag	atcaagcagg	3300
atggtgaagc	accagctagg	agagtcacaa	ttgatgccct	ctatatggat	gcctatgaag	3360
tgtcaaacac	agaatttgag	aaatttgtaa	acagcactgg	ataccttaca	gaggctgaga	3420
aatttggtga	cagttttgtt	tttgaaggca	tgctaagtga	gcaggtgaag	accaatatcc	3480
aacaggcagt	ggctgcagcc	ccctggtggc	tgccctgtta	aggagccaat	tggagacacc	3540
cagagggacc	agactcaact	atcctccaca	ggcctgacca	ccctgtgctg	catgtgtcct	3600
ggaatgatgc	agtggcatac	tgcacctggg	ctgggaaaag	gttaccaaca	gaggcagaat	3660
gggagtattc	ctgccggggt	ggactgcaca	acagactgtt	cccctggggc	aataagctgc	3720
aacctaaagg	acagcattat	gccaatattt	ggcagggaga	gttcccagtc	acaaacactg	3780
gtgaggatgg	cttccaggga	actgccccct	tggatgcttt	cccacccaat	ggctatgggt	3840
tgtacaatat	agttgggaat	gcctgggagt	ggactttctga	ctggtggacg	gtccatcaca	3900
gtgtggaaga	gacactgaac	ccaaaggggc	ccccctcagg	caaggacaga	gtcaagaaag	3960
gtggctctta	tatgtgtcac	agaagctatt	gctacagata	taggtgtgct	gcaagaagtc	4020
agaacacccc	tgacagctca	gctagcaatc	tgggatttag	atgtgcagca	gatagactcc	4080
ccaccatgga	ctgagatcca	gacatgataa	gatacattga	tgagtttgga	caaaccacaa	4140
ctagaatgca	gtgaaaaaaaa	tgctttattt	gtgaaatttg	tgatgctatt	gctttatttg	4200
taaccattat	aagctgcaat	aaacaagtta	acaacaacaa	ttgcattcat	tttatgtttc	4260
aggttcaggg	ggaggtgtgg	gaggtttttt	aaacctgcag	gtctagatac	gtagataagt	4320
agcatggcgg	gttaatcatt	aactacaagg	aacccttagt	gatggagttg	gccactccct	4380
ctctgcgcgc	tcgtctgctc	actgaggccg	ggcgaccaaa	ggtcgcccga	cgcgccgggt	4440
ttgcccgggc	ggcctcagtg	agcgagcgag	cgcgcagaga	gggagtggcc	aa	4492

<210> 70
<211> 7537
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 70	
aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcactcatta ggcaccccag	60
gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt	120
cacacaggaa acagctatga ccatgattac gccaaacctta gatccccggg taccgagctc	180
gaattcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact	240
taatcgcttt gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgcac	300
cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa tggcgcctga tgcggtattt	360
tctccttacg catctgtgcg gtatttcaca ccgcatatgg tgcactctca gtacaatctg	420
ctctgatgcc gcatagttaa gccagccccg acacccgcca acaccgctg acgcgcctg	480
acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct ccgggagctg	540
catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg gcctcgtgat	600
acgcctattt ttataggtta atgtcatgat aataatggtt tcttagacgt caggtggcac	660
ttttcgggga aatgtggcat gcctgcattt ggccactccc tctctgcgcg ctgctcgt	720
cactgaggcc gggcgaccaa aggtcgcccc acgcccgggc tttgcccggg cggcctcagt	780
gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt cctggagggg	840
tggagtcgtg acgtgaatta cgtcataggg ttagggaggt cctgcatatg cggccgcggc	900
atcctaataa atattcagtg gaaacgtaaa aacattaaag actgattaaa catcgcagca	960
tgacacagat ttagcaactg agcataaata atttgactcg gatactgctc caaaatccga	1020
agaggaccaa tttcttccag gaggacaact acctcgctct ctgcagacce ctctcctcgg	1080
cagctgaagg agtgtggcca atctgcctcc acctccccgc ggaccccccta ctctcaggac	1140
ctcctgcagc accccaaact ggaagtggcc gctgcagacc caaggacgag gggcacgcgg	1200

gagccggcag ccctagtgga gcggttggag atgttgaggt gggaggggtca cccaggtggg	1260
gtgaggctgg ggtaggtagc ggagtgaacg gcttccgaag ctctgggccg cccccaggtt	1320
ggactaagca ggcgctctgt cttegcccc gcccagggtg ggcgtctcct gaggactccc	1380
cgccacacct gacccgagac cgcgcgcccc gcctagaacg ctteccccgac ccagcgtagg	1440
gccgccgcga ctggcgggcg agggtcggcg ggaggcctgg cgaacccggg ggcgggacca	1500
ggcgggcaag gcccggctgc cgcagcgccg ctctgcgcga ggcggctccg ccgcggcgga	1560
gggatacggc gcacatatata tatatcgcg ggcgcagact cgcgctccgg cagtggtgct	1620
gggagtgtcg tggacgccgt gccgttactc gtagtcaggc ggcggcgag gcggcgggcg	1680
cggcatagcg cacagcgcg cttagcagca gcagcagcag cagcggcatc ggaggtaccc	1740
cgcgcgtcgc agcccccgcg ctgggtgcagc caccctcgct ccctctgctc ttcctccctt	1800
cgctcgcacc aagaggtaag ggtttaaggg atggttggtt ggtggggtat taatgtttaa	1860
ttacctggag cacctgcctg aaatcacttt ttttcaggtt gggccacccg ccgccaccat	1920
gagcatgggc gccccagaa gcctgttact tgcttttagct gctggccttg cagtggcaag	1980
gccccctaac atcgtgctga tctttgcaga tgacttggga tatggggatc ttggttgtta	2040
tggccacca tcaagcaca ctcccaatct ggatcagttg gctgcaggag gtctgaggtt	2100
tacagacttt tatgttccag tctccctgtg cactccttct cgggctgccc tgcttactgg	2160
gaggctccct gtgagaatgg gtatgtaccc tggagtgttg gtcccatcca gcaggggagg	2220
gctgccccctg gaagaggatga cagtggcaga ggtgctggca gcacgaggct atctgactgg	2280
catggcaggc aagtggcacc tgggtgtagg gccagagggt gctttcctgc ctcccatca	2340
gggctttcat aggtttcttg gaatcccata ctctcatgac caaggacct gccagaacct	2400
cacctgtttc cccctgcaa caccatgtga tgggggctgt gatcaaggtc tggttectat	2460
accactgctt gctaattctt cagtggaagc tcaaccaccc tggctgcctg gcttggaggc	2520
tagatacatg gccttcgcac atgatctgat ggcagatgcc cagagacaag ataggccttt	2580
cttcctctac tatgcatctc accacacca ctatcctcag ttctcaggcc aatcatttgc	2640
tgagcgtagt ggcaggggcc catttgggga cagtttgatg gaactggatg ccgcagttgg	2700

taccctcatg acagcaatag gggacttagg tttgctggag gaaacattgg taattttcac	2760
agctgataat ggccctgaga caatgagaat gtctagggga ggctgctctg gtcttctgag	2820
gtgtggtaaa gggactacat atgagggagg agtgagggaa ccagctcttg ccttttggcc	2880
aggtcacata gcccctggag ttacacatga actagcttct tccctggact tgcttccctac	2940
actggcagcc ctggcaggtg cccctctccc taatgtaaact ttagatggat ttgacctctc	3000
tccactactt ttagggacag ggaaaagtcc aaggcagtcc ttattcttct atccttccta	3060
cccagatgag gtgaggggtg tttttgccgt gaggactggg aaatacaaag ctcatttttt	3120
taccagggga tcagctcatt cagacaccac agctgatact gcctgtcatg ccagcagtag	3180
cttgacagca catgagcctc ccttactgta tgacctgagc aaggaccagc gggagaacta	3240
taacctgctt ggggggggtg ctggggccac ccagaagtg cttcaggcac taaagcagct	3300
gcaactgctt aaagcacagt tggatgctgc agtgaccttt ggcccttccc aggtggccag	3360
aggcgaggat cccgccctgc agatctgctg ccaccaggc tgcacacca gacctgcctg	3420
ctgtcactgc cccgaccac acgccggcag cggagctact aaattcagcc tgctgaagca	3480
ggctggagac gtggaggaga accctggacc tatggctgcc ccagccctgg ggctgggtgtg	3540
tggcagatgc cctgagctgg gcctgggtgt gcttctcctg ctgctgagcc tctgtgttg	3600
tgctgctggc tctcaggaag cagggacagg agcaggagca ggttctctgg ctggctcatg	3660
cggttgtggg accccccaga ggccaggggc tcatgggtcc tctgcagctg cccacaggt	3720
ctcaaggga gcaaattgcc ctggccccgt acctggggaa aggcaacttg ctcactccaa	3780
gatggttctt atccctgcag gaggttttac tatgggaact gatgaccctc agatcaagca	3840
ggatggtgaa gcaccagcta ggagagtcac aattgatgcc ttctatatgg atgcctatga	3900
agtgcaaac acagaatttg agaaatttgt aaacagcact ggatacetta cagaggctga	3960
gaaatttggt gacagttttg tttttgaagg catgctaagt gagcaggatga agaccaatat	4020
ccaacaggca gtggctgcag ccccctggtg gctgcctgtt aaaggagcca attggagaca	4080
cccagaggga ccagactcaa ctatcctcca caggcctgac caccctgtgc tgcattgtgc	4140
ctggaatgat gcagtggcat actgcacctg ggctgggaaa aggttaccaa cagaggcaga	4200

atgggagtat	tcctgccggg	gtggactgca	caacagactg	ttccccctggg	gcaataagct	4260
gcaacctaaa	ggacagcatt	atgccaatat	ttggcaggga	gagttcccag	tcacaaacac	4320
tggtgaggat	ggcttccagg	gaactgcccc	tgtggatgct	ttcccacca	atggctatgg	4380
gttgtacaat	atagttggga	atgcctggga	gtggacttct	gactggtgga	cgggccatca	4440
cagtgtggaa	gagacactga	acccaaaggg	gccccctca	ggcaaggaca	gagtcaagaa	4500
aggtggctct	tatatgtgtc	acagaagcta	ttgctacaga	tatagggtgtg	ctgcaagaag	4560
tcagaacacc	cctgacagct	cagctagcaa	tctgggattt	agatgtgcag	cagatagact	4620
ccccaccatg	gactgagatc	cagacatgat	aagatacatt	gatgagtttg	gacaaaccac	4680
aactagaatg	cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	ttgctttatt	4740
tgtaaccaat	ataagctgca	ataaacaagt	taacaacaac	aattgcattc	attttatgtt	4800
tcaggttcag	ggggaggtgt	gggaggtttt	ttaaacctgc	aggtctagat	acgtagataa	4860
gtagcatggc	gggttaatca	ttaactacaa	ggaaccctca	gtgatggagt	tggccactcc	4920
ctctctgcgc	gctcgctcgc	tcactgaggc	cgggcgacca	aaggtcgccc	gacgcccggg	4980
ctttgcccgg	gcggcctcag	tgagcgagcg	agcgcgcaga	gagggagtgg	ccaaagatcc	5040
ccgggtaccg	agctcgaatt	cactggccgt	cgttttacaa	cgtcgtgact	gggaaaaccc	5100
tggcgttacc	caacttaatc	gccttgcagc	acatccccct	ttcgccagct	ggcgtaatag	5160
cgaagaggcc	cgcaccgac	gcccttccca	acagttgcgc	agcctgaatg	gcgaatggcg	5220
cctgatgcgg	tattttctcc	ttacgcatct	gtgcggtatt	tcacaccgca	tatggtgcac	5280
tctcagtaca	atctgctctg	atgccgcata	gttaagccag	ccccgacacc	cgccaacacc	5340
cgtgacgcg	ccctgacggg	cttgtctgct	cccggcatcc	gcttacagac	aagctgtgac	5400
cgtctccggg	agctgcatgt	gtcagagggt	ttcaccgtea	tcaccgaaac	gcgcgagacg	5460
aaagggcctc	gtgatacgcc	tatttttata	ggttaatgtc	atgataataa	tggtttctta	5520
gacgtcaggt	ggcacttttc	ggggaaatgt	gcgcggaacc	cctatttggt	tatttttcta	5580
aatacatcca	aatatgtatc	cgtcatgag	acaataaccc	tgataaatgc	ttcaataata	5640
ttgaaaaagg	aagagtatga	gtattcaaca	tttccgtgtc	gcccttattc	ccttttttgc	5700

ggcattttgc	cttcctgttt	ttgctcacc	agaaacgctg	gtgaaagtaa	aagatgctga	5760
agatcagttg	ggtgcacgag	tgggttacat	cgaactggat	ctcaacagcg	gtaagatcct	5820
tgagagtttt	cgccccgaag	aacgttttcc	aatgatgagc	acttttaaag	ttctgctatg	5880
tggcgcggta	ttatcccgt	ttgacgccgg	gcaagagcaa	ctcggtcgcc	gcatacacta	5940
ttctcagaat	gacttggttg	agtactcacc	agtcacagaa	aagcatctta	cggatggcat	6000
gacagtaaga	gaattatgca	gtgctgccat	aaccatgagt	gataaacactg	cggccaactt	6060
acttctgaca	acgatcggag	gaccgaagga	gctaaccgct	tttttgcaca	acatggggga	6120
tcatgtaact	cgccttgatc	gttgggaacc	ggagctgaat	gaagccatac	caaacgacga	6180
gcgtgacacc	acgatgcctg	tagcaatggc	aacaacgttg	cgcaaactat	taactggcga	6240
actacttact	ctagcttccc	ggcaacaatt	aatagactgg	atggaggcgg	ataaagttag	6300
aggaccactt	ctgcgctcgg	cccttccggc	tggctggttt	attgctgata	aatctggagc	6360
cggtgagcgt	gggtctcgcg	gtatcattgc	agcactgggg	ccagatggta	agccctcccg	6420
tatcgtagtt	atctacacga	cggggagtca	ggcaactatg	gatgaacgaa	atagacagat	6480
cgctgagata	ggtgcctcac	tgattaagca	ttggtaaactg	tcagaccaag	tttactcata	6540
tatactttag	attgatttaa	aacttcattt	ttaatttaaa	aggatctagg	tgaagatcct	6600
ttttgataat	ctcatgacca	aaatccctta	acgtgagttt	tcgttccact	gagcgtcaga	6660
ccccgtagaa	aagatcaaag	gatcttcttg	agatcctttt	tttctgcgcg	taatctgctg	6720
cttgcaaaca	aaaaaaccac	cgtaccagc	ggtggtttgt	ttgccggatc	aagagctacc	6780
aactcttttt	ccgaaggtaa	ctggcttcag	cagagcgcag	ataccaaata	ctgttcttct	6840
agtgtagccg	tagttaggcc	accacttcaa	gaactctgta	gcaccgccta	catacctcgc	6900
tctgctaate	ctgttaccag	tggctgctgc	cagtggcgat	aagtcgtgtc	ttaccggggt	6960
ggactcaaga	cgatagttac	cggataaggc	gcagcggtcg	ggctgaacgg	ggggttcgtg	7020
cacacagccc	agcttgagac	gaacgacct	caccgaactg	agatacctac	agcgtgagct	7080
atgagaaagc	gccacgcttc	ccgaagggag	aaaggcggac	aggtatccgg	taagcggcag	7140
ggtcggaaca	ggagagcgca	cgagggagct	tccaggggga	aacgcctgg	atctttatag	7200

tectgtcggg	tttcgccacc	tctgacttga	gcgtcgat	ttgtgatgct	cgtcaggggg	7260
gcggagccta	tggaaaaacg	ccagcaacgc	ggccttttta	cggttcctgg	ccttttgctg	7320
gccttttgct	cacatgttct	ttcctgcgtt	atccccgat	tctgtggata	accgtattac	7380
cgcctttgag	tgagctgata	ccgctcgccg	cagccgaacg	accgagcgca	gcgagtcagt	7440
gagcgaggaa	gcggaagagc	gccaataacg	caaaccgcct	ctccccgcgc	gttggccgat	7500
tcattaatgc	agctggcacg	acaggtttcc	cgactgg			7537

<210> 71
<211> 6335
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 71	
gtcaggtggc	acttttcggg gaaatgtggc atgcctgcat ttggccactc cctctctgcg 60
cgctcgctcg	ctcactgagg ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg 120
ggcggcctca	gtgagcgagc gagcgcgcag agagggagtg gccaaactcca tcaactagggg 180
ttcctggagg	ggtggagtcg tgacgtgaat tacgtcatag ggttagggag gtcctgcata 240
tgcggccgca	cctaggteat tctggcctcc cctccctca aggccagtca ttctggcctg 300
tccttccccg	aaggccagtc attctggcct cccctcccc caaggccagt cattctggcc 360
ttccccctcc	ttaaggccag agtactatcg attcacacaa aaaaccaaca cactattgca 420
atgaaaataa	atttccttta ttaagcttaa ttcggtaccc tagttattaa tagtaatcaa 480
ttacggggtc	attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 540
atggcccgcc	tggetgaccg cccaacgacc ccgcccatt gacgtcaata atgacgtatg 600
ttcccatagt	aacgccaata gggactttcc attgacgtca atgggtggac tatttacggt 660
aaactgccca	cttggcagta catcaagtgt atcatatgcc aagtacgcc cctattgacg 720
tcaatgacgg	taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 780
ctacttggca	gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca 840

cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg tatttattta	900
ttttttaatt attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg	960
gggcggggcg gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc	1020
agagcggcgc gctccgaaag tttcctttta tggcgaggcg gcggcggcgg cggccctata	1080
aaaagcgaag cgcgcggcgg gcgggagtcg ctgcgacgt gccttcgccc cgtgccccgc	1140
tccgccgccg cctcgcgccg cccgccccgg ctctgactga ccgcgttact cccacaggtg	1200
agcgggcggg acggcccttc tectccgggc tgtaattagc gcttggttta atgacggctt	1260
gtttcttttc tgtggctgcg tgaaagcctt gaggggctcc gggagctaga gcctctgcta	1320
accatgttca tgccttcttc ttttctctac agctcctggg caacgtgctg gttattgtgc	1380
tgtctcatca ttttggcaaa ggctagcgcc gccaccatga gcatgggcgc cccagaagc	1440
ctgttacttg ctttagctgc tggccttgca gtggcaaggc cccctaacat cgtgctgac	1500
tttgcagatg acttgggata tggggatctt ggttgttatg gccaccctc aagcacaact	1560
ccaatctgg atcagttggc tgcaggaggc ctgaggttta cagactttta tgttccagtc	1620
tccctgtgca ctctttctcg ggctgccttg ctactggga ggctccctgt gagaatgggt	1680
atgtaccctg gagtgttggt cccatccagc aggggagggc tgccctgga agaggtgaca	1740
gtggcagagg tgctggcagc acgaggctat ctgactggca tggcaggcaa gtggcacctg	1800
ggtgtagggc cagagggtgc tttcctgcct ccccatcagg gctttcatag gtttctggga	1860
atcccatact ctcatgacca aggaccctgc cagaacctca cctgtttccc ccctgcaaca	1920
ccatgtgatg ggggctgtga tcaaggctcg gttcctatac cactgcttgc taatctttca	1980
gtggaagctc aaccaccctg gctgcctggc ttggaggcta gatacatggc cttcgcacat	2040
gatctgatgg cagatgceca gagacaagat aggcctttct tctctacta tgcatctcac	2100
cacaccact atcctcagtt ctcaggccaa tcatttgctg agcgtagtgg caggggcca	2160
tttggggaca gtttgatgga actggatgcc gcagttggta ccctcatgac agcaataggg	2220
gacttaggtt tgctggagga aacattggta attttcacag ctgataatgg ccctgagaca	2280
atgagaatgt ctaggggagg ctgctctggt cttctgaggt gtggtaaagg gactacatat	2340

gagggaggag	tgagggaaacc	agctcttgcc	ttttggccag	gtcacatagc	ccctggagtt	2400
acacatgaac	tagcttcttc	cctggacttg	cttcctacac	tggcagccct	ggcaggtgcc	2460
cctctcccta	atgtaacttt	agatggattt	gacctctctc	cactactttt	agggacaggg	2520
aaaagtccaa	ggcagtcctt	attctttctat	ccttcctacc	cagatgaggt	gaggggtggt	2580
tttgccgtga	ggactgggaa	atacaaagct	cattttttta	cccagggatc	agctcattca	2640
gacaccacag	ctgatactgc	ctgtcatgcc	agcagtagct	tgacagcaca	tgagcctccc	2700
ttactgtatg	acctgagcaa	ggaccacagg	gagaactata	acctgcttgg	gggggttgct	2760
ggggccaccc	cagaagtgct	tcaggcacta	aagcagctgc	aactgcttaa	agcacagttg	2820
gatgctgcag	tgacctttgg	cccttcccag	gtggccagag	gcgaggatcc	cgccctgcag	2880
atctgctgcc	accagggctg	cacaccacaga	cctgcctgct	gtcactgccc	cgaccacac	2940
gccggcagcg	gagctactaa	cttcagcctg	ctgaagcagg	ctggagacgt	ggaggagAAC	3000
cctggaccta	tggctgcccc	agccctgggg	ctggtgtgtg	gcagatgccc	tgagctgggc	3060
ctggtgctgc	ttctcctgct	gttgagcctc	ctgtgtggtg	ctgctggctc	tcaggaagca	3120
gggacaggag	caggagcagg	ttctctggct	ggctcatgcg	gttgtgggac	ccccagagg	3180
ccaggggctc	atgggtcctc	tgcagctgcc	cacaggtact	caagggaagc	aaatgccct	3240
ggccccgtac	ctggggaaag	gcaacttgct	cactccaaga	tggttcctat	ccctgcagga	3300
gtttttacta	tgggaactga	tgacctcag	atcaagcagg	atggtgaagc	accagctagg	3360
agagtcacaa	ttgatgcctt	ctatatggat	gcctatgaag	tgtcaaacac	agaatttgag	3420
aaatttgtaa	acagcactgg	ataccttaca	gaggctgaga	aatttggtga	cagttttgtt	3480
tttgaaggca	tgctaagtga	gcaggtgaag	accaatatcc	aacaggcagt	ggctgcagcc	3540
ccctggtggc	tgctgtttaa	aggagccaat	tggagacacc	cagagggacc	agactcaact	3600
atcctccaca	ggcctgacca	ccctgtgctg	catgtgtcct	ggaatgatgc	agtggcatac	3660
tgcacctggg	ctgggaaaag	gttaccaaca	gaggcagaat	gggagtattc	ctgccggggt	3720
ggactgcaca	acagactggt	ccccTggggc	aataagctgc	aacctaaagg	acagcattat	3780
gccaatatTT	ggcagggaga	gttcccagtc	acaAACactg	gtgaggatgg	cttccaggga	3840

actgcccctg tggatgcttt cccacceaat ggctatgggt tgtacaatat agttgggaat	3900
gcctgggagt ggacttctga ctgggtggacg gtccatcaca gtgtggaaga gacactgaac	3960
ccaaaggggc cccctcagg caaggacaga gtcaagaaag gtggctctta tatgtgtcac	4020
agaagctatt gctacagata taggtgtgct gcaagaagtc agaacacccc tgacagctca	4080
gctagcaatc tgggatttag atgtgcagca gatagactcc ccaccatgga ctgagatcca	4140
gacatgataa gatacattga tgagtttggg caaaccacaa ctagaatgca gtgaaaaaaaa	4200
tgctttattt gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat	4260
aaacaagtta acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg	4320
gaggtttttt aaacctgcag gtctagatac gtagataagt agcatggcgg gttaatcatt	4380
aactacaagg aacccttagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc	4440
actgaggccg ggcgaccaaa ggtcgcccga cggccgggct ttgcccgggc ggccctcagt	4500
agcgagcgag cgcgcagaga gggagtggcc aaagatcccc gggtaccgag gacgaattct	4560
ctagatatcg ctcaatactg accatttaaa tcatacctga cctccatagc agaaagtcaa	4620
aagcctccga ccggaggctt ttgacttgat cggcacgtaa gaggttccaa ctttcaccat	4680
aatgaaataa gatcactacc gggcgtattt tttgagttat cgagattttc aggagctaag	4740
gaagctaaaa tgagccatat tcaacgggaa acgtcttgct cgaggccgcg attaaattcc	4800
aacatggatg ctgatttata tgggtataaa tgggctcgcg ataatgtcgg gcaatcaggt	4860
gcgacaatct atcgattgta tgggaagccc gatgcgccag agttgtttct gaaacatggc	4920
aaaggtagcg ttgccaatga tgttacagat gagatggcca ggctaaactg gctgacggaa	4980
tttatgcctc ttccgaccat caagcatttt atccgtactc ctgatgatgc atggttactc	5040
accactgcga tcccagggaa aacagcattc caggtattag aagaatatcc tgattcaggt	5100
gaaaatattg ttgatgcgtc ggcagtgttc ctgcgccggt tgcattcgat tcctgtttgt	5160
aattgtcctt ttaacggcga tcgcgtattt cgtctcgctc aggcgcgaatc acgaatgaat	5220
aacggtttgg ttggtgcgag tgattttgat gacgagcgta atggctggcc tgttgaacaa	5280
gtctggaaag aaatgcataa gcttttgcca ttctcaccgg attcagtcgt cactcatggt	5340

gatttctcac ttgataacct tatttttgac gaggggaaat taataggttg tattgatgtt	5400
ggacgagtcg gaatcgcaga ccgataccag gatcttgcca tcctatggaa ctgcctcgg	5460
gagttttctc cttcattaca gaaacggctt tttcaaaaat atgggtattga taatcctgat	5520
atgaataaat tgcagtttca cttgatgctc gatgagtttt tctgagggcc caaatgtaat	5580
cacctggctc accttcgggt gggcctttct gcgttgctgg cgtttttcca taggctcgc	5640
ccccctgacg agcatcaca aaatcgatgc tcaagtcaga ggtggcgaaa cccgacagga	5700
ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc	5760
ctgccgctta ccgataacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat	5820
agctcacgct gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg	5880
cacgaacccc ccgttcagcc cgaccgtgc gccttatccg gtaactatcg tcttgagtcc	5940
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga	6000
gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact	6060
agaagaacag tatttggtat ctgcgctctg ctgaagccag ttacctcgga aaaagagttg	6120
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc	6180
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgattttc taccgaagaa	6240
aggcccaccc gtgaaggtga gccagtgagt tgattgcagt ccagttacgc tggagtctga	6300
ggctcgtcct gaatgatatc aagcttgaat tcgtt	6335

<210> 72
<211> 1527
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 72	
atgtctatgg gggtcctcgc ctccctgctg ctggcactgg ccgccgggct ggctgtcgca	60
agaccaccta atatcgtcct gattttttgca gacgatctgg gatacggcga cctgggatgc	120
tatggccacc caagctccac cacacccaac ctggaccagc tggcagcagg aggcctgcgg	180

ttcaccgact tctacgtgcc agtgagcctg tgcacccccct ccagagccgc cctgctgaca	240
ggcaggctgc cagtgcgcat gggcatgtat cctggcgtgc tggtgccatc tagcaggggc	300
ggcctgccac tggaggaggt gaccgtggca gaggtgctgg cagccagagg ctacctgaca	360
ggaatggccg gcaagtggca cctgggagtg ggaccagagg gagecttcct gccccctcac	420
cagggettcc accggttttct gggcatecct tattctcacg accagggccc atgccagaac	480
ctgacctgtt ttccaccagc aacaccatgc gacggaggat gtgatcaggg cctggtgcca	540
atcccactgc tggcaaatct gagcgtggag gcacagcctc catggctgcc tggcctggag	600
gcaagataca tggccttcgc ccacgacctg atggcagatg cacagcggca ggatagacct	660
ttctttctgt actatgcctc ccaccacacc cactatccac agttcagcgg ccagtccttt	720
gccgagaggt ccggaagggg accattcggc gactctctga tggagctgga tgccgcctg	780
ggcacctga tgacagcaat cggcgacctg ggctgctgg aggagacact ggtcatcttc	840
accgccgata acggccctga gacaatgcgg atgtctagag gcggatgcag cggcctgctg	900
agatgtggca agggaaccac atacgaggga ggcgtgcgcg agcctgccct ggcattttgg	960
ccaggacaca tcgcacctgg agtgaccac gagctggcct cctctctgga cctgctgcca	1020
acactggccg ccctggcagg agcacctctg ccaaattgtga ccctggacgg cttcgatctg	1080
agcccactgc tgctgggaac cggcaagtcc cctaggcagt ctctgttctt ttacccctcc	1140
tatcctgatg aggtgcgggg cgtgttttgc gtgagaaccg gcaagtacaa ggcccacttc	1200
tttacacagg gctctgccc cagcgacacc acagcagatc cagcatgcca cgccagctcc	1260
tctctgaccg cacacgagcc acctctgctg tacgacctgt ccaaggatcc cggcgagAAC	1320
tataatctgc tgggaggagt ggcaggagca acccctgagg tgctgcaggc cctgaagcag	1380
ctgcagctgc tgaaggcaca gctggacgca gcagtgacat tcggcccaag ccaggtggcc	1440
agaggcgagg atcccgccct gcagatctgc tgccaccag gctgcacacc cagacctgcc	1500
tgetgtcact gccccgaccc acacgcc	1527

<210> 73
<211> 237
<212> DNA

<213> 智人

<400> 73

ggggacgttt gccaggactg cattcagatg gtgactgaca tccagactgc tgtacggacc	60
aactccacct ttgtccaggc ctgggtggaa catgtcaagg aggagtgtga ccgcctgggc	120
cctggcatgg ccgacatatg caagaactat atcagccagt attctgaaat tgctatccag	180
atgatgatgc acatgcaacc caaggagatc tgtgcgctgg ttgggttctg tgatgag	237

<210> 74

<211> 1833

<212> DNA

<213> 人工序列

<220>

<223> 合成聚核苷酸

<400> 74

atgtctatgg gggtctctcg ctccctgctg ctggcactgg ccgccgggct ggctgtcgca	60
agaccaccta atatcgtcct gatTTTTgca gacgatctgg gatacggcga cctgggatgc	120
tatggccacc caagctccac cacacccaac ctggaccagc tggcagcagg aggcctgcgg	180
ttcacgact tctacgtgcc agtgagcctg tgcacccccct ccagagccgc cctgctgaca	240
ggcaggctgc cagtgcgcat gggcatgtat cctggcgtgc tggtgccatc tagcaggggc	300
ggcctgccac tggaggaggt gaccgtggca gaggtgctgg cagccagagg ctacctgaca	360
ggaatggccg gcaagtggca cctgggagtg ggaccagagg gagccttcct gccccctcac	420
cagggcttcc accggtttct gggcatecct tattctcagc accagggccc atgccagaac	480
ctgacctgtt ttccaccagc aacaccatgc gacggaggat gtgatcaggg cctggtgcca	540
atcccactgc tggcaaatct gagcgtggag gcacagcctc catggctgcc tggcctggag	600
gcaagataca tggccttcgc ccacgacctg atggcagatg cacagcggca ggatagacct	660
ttctttctgt actatgcctc ccaccacacc cactatccac agttcagegg ccagtccttt	720
gccgagaggt ccggaagggg accattcggc gactctctga tggagctgga tgccgccgtg	780
ggcacctga tgacagcaat cggcgacctg ggctgtctgg aggagacact ggtcatcttc	840
accgccgata acggccctga gacaatgcgg atgtctagag gcggatgcag cggcctgctg	900

agatgtggca agggaaccac atacgaggga ggcgtgcgcg agcctgccct ggcattttgg	960
ccaggacaca tcgcacctgg agtgacccac gagctggcct cctctctgga cctgctgcca	1020
acactggccg ccctggcagg agcacctctg ccaaattgtga ccctggacgg cttcgatctg	1080
agcccactgc tgctgggaac cggcaagtcc cctaggcagt ctctgttctt ttaccctcc	1140
tatcctgatg aggtgcgggg cgtgtttgcc gtgagaaccg gcaagtacaa ggcccacttc	1200
tttacacagg gctctgcccc cagcgacacc acagcagatc cagcatgcca cgccagctcc	1260
tctctgaccg cacacgagcc acctctgctg tacgacctgt ccaaggatcc cggcgagAAC	1320
tataatctgc tgggaggagt ggcaggagca acccctgagg tgctgcaggc cctgaagcag	1380
ctgcagctgc tgaaggcaca gctggacgca gcagtgacat tcggcccaag ccaggtggcc	1440
agaggcgagg atcccgcctt gcagatctgc tgccaccag gctgcacacc cagacctgcc	1500
tgctgtcact gccccgacct acacgccggc agcggagcta ctaacttcag cctgctgaag	1560
caggctggag acgtggagga gaacctgga cctggggacg tttgccagga ctgcattcag	1620
atggtgactg acatccagac tgctgtacgg accaactcca cttttgtcca ggccttggtg	1680
gaacatgtca aggaggagtg tgaccgcctg ggccctggca tggccgacat atgcaagaac	1740
tatatcagcc agtattctga aattgctatc cagatgatga tgcacatgca acccaaggag	1800
atctgtgcgc tggttgggtt ctgtgatgag tga	1833

- <210> 75
- <211> 3698
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成聚核苷酸

<400> 75	
ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgcca	120
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240

aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggcccgct	300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360
tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcaact ctccccatct	420
cccccccctc ccacccccca attttgtatt tatttatttt ttaattattt tgtgcagcga	480
tgggggcggg gggggggggg gggcgcgcg caggcggggc ggggcggggc gaggggcggg	540
gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg	660
gagtcgctgc gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc	720
cccggctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ctttctcctc	780
cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa	840
gccttgaggg gctccgggag ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg	900
tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgct gcccggcggc tgtgagcgt	960
gcgggcgcgg cgcggggctt tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg	1020
gcggtgcccc gcggtgcggg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt	1080
gcgtgggggg gtgagcaggg ggtgtgggcg cgtcggtcgg gctgcaacce cccctgcacc	1140
cccctccccg agttgctgag cacggcccgg cttcgggtgc ggggctccgt acggggcgtg	1200
gcgcgggggt cgccgtgccg ggcggggggt ggcggcaggt gggggtgccg ggcggggcgg	1260
ggccgcctcg ggccggggag ggctcggggg aggggcgcgg cggcccccg agcgccggcg	1320
gctgtcgagg cgcggcgagc cgcagccatt gccttttatg gtaatcgtgc gagagggcgc	1380
agggacttcc tttgtcccaa atctgtgcgg agccgaaatc tgggaggcgc cgccgcacce	1440
cctctagcgg gcgcggggcg aagcgggtgcg gcgccggcag gaaggaaatg ggcggggagg	1500
gccttcgtgc gtcgccgcgc cgccgtcccc ttctccctct ccagcctcgg ggtgtccgc	1560
ggggggacgg ctgccttcgg gggggacggg gcagggcggg gttcggcttc tggcgtgtga	1620
ccggcggtc tagagcctct gtaacctg ttcatgccct cttctttttc ctacagctcc	1680
tgggcaacgt gctggttatt gtgctgtctc atcattttgg caaagaattc cgccaccatg	1740

tctatggggg	ctcctcgctc	cctgctgctg	gcactggccg	ccgggctggc	tgtcgcaaga	1800
ccacctaata	tcgtcctgat	ttttgcagac	gatctgggat	acggcgacct	gggatgctat	1860
ggccacccaa	gtccaccac	acccaacctg	gaccagctgg	cagcaggagg	cctgcggttc	1920
accgacttct	acgtgccagt	gagcctgtgc	accccccca	gagccgccct	gctgacaggc	1980
aggctgccag	tgcgcatggg	catgtatcct	ggcgtgctgg	tgccatctag	caggggcggc	2040
ctgccactgg	aggaggtgac	cgtggcagag	gtgctggcag	ccagaggcta	cctgacagga	2100
atggccggca	agtggcacct	gggagtggga	ccagagggag	ccttcctgcc	ccctcaccag	2160
ggcttccacc	ggtttctggg	catcccttat	tctcacgacc	agggcccatg	ccagaacctg	2220
acctgttttc	caccagcaac	accatgcgac	ggaggatgtg	atcagggcct	ggtgccaatc	2280
ccactgctgg	caaatctgag	cgtggaggca	cagcctccat	ggctgcctgg	cctggaggca	2340
agatacatgg	ccttcgcccc	cgacctgatg	gcagatgcac	agcggcagga	tagacctttc	2400
tttctgtact	atgcctcccc	ccacaccac	tatccacagt	tcagcggcca	gtcctttgcc	2460
gagaggtccg	gaaggggacc	attcggcgac	tctctgatgg	agctggatgc	cgccgtgggc	2520
accctgatga	cagcaatcgg	cgacctgggc	ctgctggagg	agacactggt	catcttcacc	2580
gccgataacg	gccctgagac	aatgcggatg	tctagaggcg	gatgcagcgg	cctgctgaga	2640
tgtggcaagg	gaaccacata	cgagggaggc	gtgcgcgagc	ctgccctggc	attttggcca	2700
ggacacatcg	cacctggagt	gacccacgag	ctggcctcct	ctctggacct	gctgccaaca	2760
ctggccgccc	tggcaggagc	acctctgcca	aatgtgacce	tggacggctt	cgatctgagc	2820
ccactgctgc	tgggaaccgg	caagtccctt	aggcagtcct	tgttctttta	cccctcctat	2880
cctgatgagg	tgcggggcgt	gtttgccgtg	agaaccggca	agtacaaggc	ccacttcttt	2940
acacagggct	ctgcccacag	cgacaccaca	gcagatccag	catgccacgc	cagctcctct	3000
ctgaccgcac	acgagccacc	tctgctgtac	gacctgtcca	aggatcccgg	cgagaactat	3060
aatctgctgg	gaggagtggc	aggagcaacc	cctgaggtgc	tgcaggccct	gaagcagctg	3120
cagctgctga	aggcacagct	ggacgcagca	gtgacattcg	gcccgaagcca	ggtggccaga	3180
ggcgaggatc	ccgccctgca	gatctgctgc	caccaggtct	gcacaccag	acctgcctgc	3240

tgtcactgcc	ccgacccaca	cgccggcagc	ggagctacta	acttcagcct	gctgaagcag	3300
gctggagacg	tggaggagaa	ccctggacct	ggggacgttt	gccaggactg	cattcagatg	3360
gtgactgaca	tccagactgc	tgtacggacc	aactccacct	ttgtccaggc	cttggtggaa	3420
catgtcaagg	aggagtgtga	ccgcctgggc	cctggcatgg	ccgacatatg	caagaactat	3480
atcagccagt	attctgaaat	tgctatccag	atgatgatgc	acatgcaacc	caaggagatc	3540
tgtgcgctgg	ttgggttctg	tgatgagtga	actagtaact	tgtttattgc	agcttataat	3600
ggttacaat	aaagcaatag	catcacaaat	ttcacaaata	aagcattttt	ttcactgcat	3660
tctagttgtg	gtttgtccaa	actcatcaat	gtatctta			3698

<210> 76
<211> 4231
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

ttggccactc	cctctctgcg	cgtctgctcg	ctcactgagg	ccgggcgacc	aaaggtcgcc	60
cgacgccccg	gccttgcccc	ggcggcctca	gtgagcgagc	gagcgcgcag	agagggagtg	120
gccaactcca	tcactagggg	ttcctggagg	ggtggagtcg	tgacgtgaat	tacgtcatag	180
ggttagggag	gtcctgcaga	tcttcaatat	tggccattag	ccatattatt	cattggttat	240
atagcataaa	tcaatattgg	ctattggcca	ttgcatacgt	tgtatctata	tcataatatg	300
tacatttata	ttggctcatg	tccaatatga	ccgccatgtt	ggcattgatt	attgactagt	360
tattaatagt	aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	420
acataactta	cggtaaattg	ccgccttggc	tgaccgcccc	acgacccccg	cccattgacg	480
tcaataatga	cgtatgttcc	catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	540
gtggagtatt	tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	600
ccgcccccta	ttgacgtcaa	tgacggtaaa	tggccccgct	ggcatttatgc	ccagtacatg	660
accttacggg	actttcctac	ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	720

gtcgaggtga gccccacgtt ctgcttcaact ctccccatct cccccccctc cccacccccca	780
atttttgtatt tatttatttt ttaattattt tgtgcagcga tgggggcggg gggggggggg	840
gggcgcgcgc caggcggggc ggggcggggc gaggggcggg gcggggcgag gcggagaggt	900
gcggcggcag ccaatcagag cggcgcgctc cgaaagtffc cttttatggc gaggcggcgg	960
cggcggcggc cctataaaaa gcgaagcgcg cggcggggcg gagtcgctgc gcgctgcctt	1020
cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc cccggctctg actgaccgcg	1080
ttactcccac aggtgagcgg gcgggacggc ctttctcctc cgggctgtaa ttagcgcttg	1140
gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa gccttgaggg gctccgggag	1200
ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg tgtgtgtgtg cgtggggagc	1260
gccgcgtgcg gctccgcgt gcccggcggc tgtgagcgt gcgggcgcgg cgcggggctt	1320
tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg gcggtgcccc gcggtgcggg	1380
gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt gcgtgggggg gtgagcaggg	1440
ggtgtgggcg cgtcggtcgg gctgcaacc cccctgcacc cccctcccc agttgctgag	1500
cacggccccg cttcgggtgc ggggctccgt acggggcgtg gcgcggggct cgcctgccc	1560
ggcgggggggt ggcggcaggt gggggtgccg ggcggggcgg ggccgcctcg ggccggggag	1620
ggctcggggg aggggcgcgg cggcccccg agcgccggcg gctgtcgagg cgcggcgagc	1680
cgcagccatt gccttttatg gtaatcgtgc gagagggcgc agggacttcc tttgtcccaa	1740
atctgtgcgg agccgaaatc tgggaggcgc cgccgcaccc cctctagcgg gcgcggggcg	1800
aagcgggtgc gcgccggcag gaaggaaatg ggcggggagg gccttcgtgc gtcgccgcgc	1860
cgcctgcccc ttctccctct ccagcctcgg ggctgtccgc ggggggacgg ctgccttcgg	1920
gggggacggg gcagggcggg gttcggcttc tggcgtgtga ccggcggtc tagagcctct	1980
gctaaccatg ttcatgcctt cttctttttc ctacagctcc tgggcaacgt gctggttatt	2040
gtgctgtctc atcattttgg caaagaattc cgccaccatg tctatggggg ctccctcctc	2100
cctgctgctg gcactggccg ccgggctggc tgtcgcaaga ccacctaaata tegtcctgat	2160
ttttgcagac gatctgggat acggcgacct gggatgctat ggccacccaa gctccaccac	2220

acccaacctg gaccagctgg cagcaggagg cctgcggttc accgacttct acgtgccagt	2280
gagcctgtgc accccctcca gagccgccct gctgacaggc aggctgccag tgcgcatggg	2340
catgtatcct ggctgtctgg tgccatctag caggggcggc ctgccactgg aggaggtgac	2400
cgtggcagag gtgctggcag ccagaggcta cctgacagga atggccggca agtggcacct	2460
gggagtggga ccagagggag ccttcctgcc ccctcaccag ggcttcacc ggtttctggg	2520
catcccttat tctcacgacc agggcccatg ccagaacctg acctgttttc caccagcaac	2580
accatgcgac ggaggatgtg atcagggcct ggtgccaatc cactgctgg caaatctgag	2640
cgtggaggca cagcctccat ggctgcctgg cctggaggca agatacatgg ccttcgcca	2700
cgacctgatg gcagatgcac agcggcagga tagaccttc tttctgtact atgcctccca	2760
ccacaccac tatccacagt tcagcggcca gtcctttgcc gagagggccg gaaggggacc	2820
attcggcgac tctctgatgg agctggatgc cgccgtgggc accctgatga cagcaatcgg	2880
cgacctgggc ctgctggagg agacactggt catcttcacc gccgataacg gccctgagac	2940
aatgcggatg tctagaggcg gatgcagcgg cctgctgaga tgtggcaagg gaaccacata	3000
cgagggaggc gtgcgcgagc ctgccctggc attttggcc ggacacatcg cacctggagt	3060
gaccacagag ctggcctcct ctctggacct gctgccaaca ctggccgccc tggcaggagc	3120
acctctgcca aatgtgacct tggacggett cgatctgagc cactgctgc tgggaaccgg	3180
caagtccctt aggcagtctc tgttctttta cccctcctat cctgatgagg tgcggggcgt	3240
gtttgccgtg agaaccggca agtacaaggc ccacttcttt acacagggct ctgcccacag	3300
cgacaccaca gcagatccag catgccacgc cagctcctct ctgaccgcac acgagccacc	3360
tctgctgtac gacctgtcca aggatcccgg cgagaactat aatctgctgg gaggagtggc	3420
aggagcaacc cctgaggtgc tgcaggccct gaagcagctg cagctgctga aggcacagct	3480
ggacgcagca gtgacattcg gccaagcca ggtggccaga ggcgaggatc ccgccctgca	3540
gatctgctgc caccaggtt gcacaccag acctgcctgc tgtcactgcc ccgaccaca	3600
cgccggcagc ggagctacta acttcagcct gctgaagcag gctggagacg tggaggagaa	3660
ccctggacct ggggacgttt gccaggactg cattcagatg gtgactgaca tccagactgc	3720

tgtacggacc aactccacct ttgtccaggc cttggtggaa catgtcaagg aggagtgtga	3780
ccgcctgggc cctggcatgg ccgacatatg caagaactat atcagccagt attctgaaat	3840
tgctatccag atgatgatgc acatgcaacc caaggagatc tgtgcgctgg ttgggttctg	3900
tgatgagtga actagtaact tgtttattgc agcttataat ggttacaaat aaagcaatag	3960
catcacaaat ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa	4020
actcatcaat gtatcttagg tctagatacg tagataagta gcatggcggg ttaatcatta	4080
actacaagga acccctagtg atggagttgg ccactccctc tctgcgcgct cgctcgctca	4140
ctgaggccgg gcgaccaaag gtcgcccgcac gcccgggctt tgcccgggcg gcctcagtga	4200
gcgagcgagc gcgcagagag ggagtggcca a	4231

<210> 77
<211> 6073
<212> DNA
<213> 人工序列

<220>
<223> 合成聚核苷酸

<400> 77 tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgatgctcaa gtcagaggtg	60
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg	120
ctctcctggt ccgacctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag	180
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc	240
caagctgggc tgtgtgcacg aacccccgt tcagcccgcac cgctgcgcct tatccggtaa	300
ctatcgtctt gagtccaacc cggtaaagaca cgacttatcg ccactggcag cagccactgg	360
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagtctctga agtggtggcc	420
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac	480
ctcggaaaaa gaggttgtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt	540
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg	600
attttctacc gaagaaaggc ccaccgctga aggtgagcca gtgagttgat tgcagtccag	660

ttacgctgga gtctgaggct cgtcctgaat gatatcaagc ttgaattcgt gtcaggtggc	720
acttttcggg gaaatgtggc atgcctgcat ttggccactc cctctctgcg cgctcgctcg	780
ctcactgagg ccggggcgacc aaaggctcgcc cgacgcccgg gctttgcccc ggcggcctca	840
gtgagcgagc gagcgcgag agagggagtg gccaaactcca tcactagggg ttcctggagg	900
ggtggagtcg tgacgtgaat tacgtcatag ggttaggggag gtccctgcaga tcttcaatat	960
tggccattag ccatattatt cattggttat atagcataaa tcaatattgg ctattggcca	1020
ttgcatacgt tgtatctata tcataatatg tacatttata ttggctcatg tccaatatga	1080
ccgccatgtt ggcatlgatt attgactagt tattaatagt aatcaattac ggggtcatta	1140
gttcatagcc catatatgga gtccgcgtt acataactta cggtaaattg cccgcctggc	1200
tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg	1260
ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg	1320
gcagtacatc aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa	1380
tggcccgctt ggcatltgc ccagtacatg accttacggg actttcctac ttggcagtac	1440
atctacgtat tagtcategc tattaccatg gtcgaggtga gccccacgtt ctgcttcact	1500
ctccccatct cccccccctc cccaccccc aTTTTgtatt tatttatTTT ttaattattt	1560
tgtgcagcga tggggggcggg gggggggggg gggcgcgcg caggcggggc ggggcggggc	1620
gaggggcggg gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc	1680
cgaagtttc cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg	1740
cggcgggcgg gagtgcctgc gcgtgcctt cgcctcgtgc cccgtccgc cgcgcctcg	1800
cgcgccccgc cccggctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc	1860
ccttctcctc cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg	1920
ctgcgtgaaa gccttgaggg gctccgggag ggccctttgt gcggggggag cggctcgggg	1980
ggtgcgtgcg tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgt gcccggcggc	2040
tgtgagcgt gcgggcgcgg cgcggggctt tgtgcgtcc gcagtgtgcg cgaggggagc	2100
gcggccgggg gcggtgcccc gcggtgcgg gggggctgcg aggggaacaa aggctgcgtg	2160

cggggtgtgt	gcgtgggggg	gtgagcaggg	ggtgtgggcg	cgtcggtcgg	gctgcaaccc	2220
ccctgcacc	ccctccccc	agttgctgag	cacggcccgg	cttcgggtgc	ggggctccgt	2280
acggggcgtg	gcgcggggct	cgccgtgccg	ggcgggggggt	ggcggcaggt	gggggtgccg	2340
ggcggggcgg	ggccgcctcg	ggccggggag	ggctcggggg	aggggcgcgg	cggcccccg	2400
agcgccggcg	gctgtcgagg	cgcggcgagc	cgcagccatt	gccttttatg	gtaatcgtgc	2460
gagagggcgc	agggacttcc	tttgtcccaa	atctgtgcgg	agccgaaatc	tgggaggcgc	2520
cggcgcaccc	cctctagcgg	gcgcggggcg	aagcggtgcg	gcgccggcag	gaaggaaatg	2580
ggcggggagg	gccttcgtgc	gtcgccgcgc	cgccgtcccc	ttctccctct	ccagcctcgg	2640
ggctgtccgc	ggggggacgg	ctgcccttcg	gggggacggg	gcagggcggg	gttcggcttc	2700
tggcgtgtga	ccggcggtc	tagagcctct	gctaaccatg	tteatgcctt	cttccttttc	2760
ctacagctcc	tgggcaacgt	gctggttatt	gtgctgtctc	atcattttgg	caaagaattc	2820
cgccaccatg	tctatggggg	ctcctcgctc	cctgctgctg	gcactggccg	ccgggctggc	2880
tgtcgcaaga	ccaccttaata	tcgtcctgat	ttttgcagac	gatctgggat	acggcgacct	2940
gggatgctat	ggccacccaa	gtccaccac	acccaacctg	gaccagctgg	cagcaggagg	3000
cctgcggttc	accgacttct	acgtgccagt	gagcctgtgc	accccccca	gagccgccct	3060
gctgacaggc	aggetgccag	tgcgcatggg	catgtatcct	ggcgtgctgg	tgccatctag	3120
caggggcggc	ctgccactgg	aggaggtgac	cgtggcagag	gtgctggcag	ccagaggcta	3180
cctgacagga	atggccggca	agtggcacct	gggagtggga	ccagagggag	ccttcctgcc	3240
ccctcaccag	ggcttcaccc	ggtttctggg	catcccttat	tctcacgacc	agggcccatg	3300
ccagaacctg	acctgttttc	caccagcaac	accatgcgac	ggaggatgtg	atcagggcct	3360
ggtgccaatc	ccactgctgg	caaatctgag	cgtggaggca	cagcctccat	ggctgcctgg	3420
cctggaggca	agatacatgg	ccttcgccca	cgacctgatg	gcagatgcac	agcggcagga	3480
tagacctttc	tttctgtact	atgcctccca	ccacacccac	tatccacagt	tcagcggcca	3540
gtcctttgcc	gagaggtccg	gaaggggacc	attcggcgac	tctctgatgg	agctggatgc	3600
cgccgtgggc	accctgatga	cagcaatcgg	cgacctgggc	ctgctggagg	agacactggt	3660

catcttcacc	gccgataacg	gccctgagac	aatgcggatg	tctagaggcg	gatgcagcgg	3720
cctgctgaga	tgtggcaagg	gaaccacata	cgagggaggc	gtgcgcgagc	ctgccctggc	3780
atTTtggcca	ggacacatcg	cacctggagt	gacccacgag	ctggcctcct	ctctggacct	3840
gctgccaaca	ctggccgccc	tggcaggagc	acctctgcca	aatgtgacct	tggacggctt	3900
cgatctgagc	ccactgctgc	tgggaaccgg	caagtccccct	aggcagtctc	tgttctttta	3960
ccccctctat	cctgatgagg	tgcggggcgt	gtttgccgtg	agaaccggca	agtacaaggc	4020
ccacttcttt	acacagggct	ctgcccacag	cgacaccaca	gcagatccag	catgccacgc	4080
cagctcctct	ctgaccgcac	acgagccacc	tctgctgtac	gacctgtcca	aggatcccgg	4140
cgagaactat	aatctgctgg	gaggagtggc	aggagcaacc	cctgaggtgc	tgcaggccct	4200
gaagcagctg	cagctgctga	aggcacagct	ggacgcagca	gtgacattcg	gcccagcca	4260
ggtggccaga	ggcgaggatc	ccgccctgca	gatctgctgc	caccaggt	gcacaccag	4320
acctgcctgc	tgtcactgcc	ccgaccacaca	cgccggcagc	ggagctacta	acttcagcct	4380
gctgaagcag	gctggagacg	tggaggagaa	ccctggacct	ggggacgttt	gccaggactg	4440
cattcagatg	gtgactgaca	tccagactgc	tgtacggacc	aactccacct	ttgtccaggc	4500
cttgggtggaa	catgtcaagg	aggagtgtga	ccgcctgggc	cctggcatgg	ccgacatatg	4560
caagaactat	atcagccagt	attctgaaat	tgctatccag	atgatgatgc	acatgcaacc	4620
caaggagatc	tgtgcgctgg	ttgggttctg	tgatgagtga	actagtaact	tgtttattgc	4680
agcttataat	ggttacaaat	aaagcaatag	catcacaaat	ttcacaaata	aagcattttt	4740
ttcactgcat	tctagtttgtg	gtttgtccaa	actcatcaat	gtatcttagg	tctagatacg	4800
tagataagta	gcatggcggg	ttaatcatta	actacaagga	acccttagtg	atggagttag	4860
ccactccctc	tctgcgcgct	cgctcgctca	ctgaggccgg	gcgaccaaag	gtcgcccgac	4920
gcccggggctt	tgccccggcg	gcctcagtga	gcgagcgagc	gcgcagagag	ggagtggcca	4980
aagatccccg	ggtaccgagg	acgaattctc	tagatatcgc	tcaatactga	ccatttaaat	5040
catacctgac	ctccatagca	gaaagtcaaa	agcctccgac	cggaggtttt	tgacttgatc	5100
ggcacgtaag	aggttccaac	tttcaccata	atgaaataag	atcactaccg	ggcgtatttt	5160

ttgagttatc gagattttca ggagctaagg aagctaaaat gagccatatt caacgggaaa	5220
cgtcttgctc gaggccgcga ttaaattcca acatggatgc tgatttatat gggtataaat	5280
gggctcgcga taatgtcggg caatcaggtg cgacaatcta tcgattgtat gggaagcccg	5340
atgcgccaga gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg	5400
agatggtcag gctaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta	5460
tccgtactcc tgatgatgca tggttactca ccactgcgat cccagggaaa acagcattcc	5520
aggtattaga agaatatcct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc	5580
tgcgccggtt gcattcgatt cctgtttgta attgtccttt taacggcgat cgcgtatttc	5640
gtctcgtca ggcgcaatca cgaatgaata acggtttggt tggatgcgagt gattttgatg	5700
acgagcgtaa tggctggcct gttgaacaag tctggaaaga aatgcataag cttttgccat	5760
tctcaccgga ttcagtcgtc actcatgggtg atttctcact tgataacctt atttttgacg	5820
aggggaaatt aataggttgt attgatgttg gacgagtcgg aatcgcagac cgataccagg	5880
atcttgccat cctatggaac tgcctcgggtg agttttctcc ttcattacag aaacggcttt	5940
ttcaaaaata tggatttgat aatcctgata tgaataaatt gcagtttcac ttgatgctcg	6000
atgagttttt ctgagggccccc aaatgtaatc acctggctca ccttcgggtg ggcctttctg	6060
cgttgctggc gtt	6073

- <210> 78
- <211> 42
- <212> DNA
- <213> 人工序列

- <220>
- <223> 合成核酸序列

<400> 78	
ggaaaaccaa taccaaaccc tctattagga ttggactcaa ca	42

- <210> 79
- <211> 3458
- <212> DNA
- <213> 人工序列

<220>

<223> 合成核酸序列

<400> 79

ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgcca	120
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	240
aagtgtatca tatgccaagt ccgcccccta ttgacgtcaa tgacggtaaa tggcccgct	300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360
tagtcatcgc tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct	420
cccccccctc cccacccccca attttgtatt tatttatttt ttaattattt tgtgcagcga	480
tggggggcggg ggggggggggg gggcgcgcg caggcggggc ggggcggggc gaggggcggg	540
gcggggcgag gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc	600
cttttatggc gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg	660
gagtcgctgc gcgtgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgccgc	720
cccggctctg actgaccgcg ttactccac aggtgagcgg gcgggacggc ctttctctc	780
cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa	840
gccttgaggg gctccgggag ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg	900
tgtgtgtgtg cgtggggagc gccgcgtgcg gctccgcgct gcccggcggc tgtgagcgt	960
gcggggcgcg cgcggggctt tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg	1020
gcggtgcccc gcggtgcggg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt	1080
gcgtgggggg gtgagcaggg ggtgtgggcg cgtcggtcgg gctgcaacc cccctgcacc	1140
ccccccccg agttgctgag cacggccccg cttcgggtgc ggggctccgt acggggcgtg	1200
gcgcggggct cgccgtgccg ggcggggggt ggcggcaggt gggggtgccg ggcggggcgg	1260
ggccgcctcg ggccggggag ggctcggggg aggggcgcgg cgccccccgg agcgccggcg	1320
gctgtcgagg cgcggcgagc cgcagccatt gccttttatg gtaatcgtgc gagagggcgc	1380

agggacttcc	tttgtcccaa	atctgtgcgg	agccgaaatc	tgggaggcgc	cgccgcaccc	1440
cctctagcgg	gcgcggggcg	aagcggtgcg	gcgccggcag	gaaggaaatg	ggcggggagg	1500
gccttcgtgc	gtcgccgcgc	cgccgtcccc	ttctccctct	ccagccctcgg	ggctgtccgc	1560
ggggggacgg	ctgccttcgg	gggggacggg	gcagggcggg	gttcggcttc	tggcgtgtga	1620
ccggcggctc	tagagcctct	gctaaccatg	ttcatgccct	cttctttttc	ctacagctcc	1680
tgggcaacgt	gctggttatt	gtgctgtctc	atcattttgg	caaagaattc	cgccaccatg	1740
tctatggggg	ctcctcgctc	cctgctgctg	gcactggccg	ccgggctggc	tgtcgcaaga	1800
ccaccttaata	tcgtcctgat	ttttgcagac	gatctgggat	acggcgacct	gggatgctat	1860
ggccacccaa	gtcccaccac	acccaacctg	gaccagctgg	cagcaggagg	cctgcggttc	1920
accgacttct	acgtgccagt	gagcctgtgc	accccccca	gagccgccct	gctgacaggc	1980
aggctgccag	tgcgcatggg	catgtatcct	ggcgtgctgg	tgccatctag	caggggcggc	2040
ctgccactgg	aggaggtgac	cgtggcagag	gtgctggcag	ccagaggcta	cctgacagga	2100
atggccggca	agtggcacct	gggagtggga	ccagagggag	ccttcctgcc	ccctcaccag	2160
ggcttcaccc	ggtttctggg	catcccttat	tctcacgacc	agggcccatg	ccagaacctg	2220
acctgttttc	caccagcaac	accatgcgac	ggaggatgtg	atcagggcct	ggtgccaatc	2280
ccactgctgg	caaatctgag	cgtggaggca	cagcctccat	ggctgcctgg	cctggaggca	2340
agatacatgg	ccttcgcccc	cgacctgatg	gcagatgcac	agcggcagga	tagacctttc	2400
tttctgtact	atgcctccca	ccacacccac	tatccacagt	tcagcggcca	gtcctttgcc	2460
gagaggtccg	gaaggggacc	attcggcgac	tctctgatgg	agctggatgc	cgccgtgggc	2520
accctgatga	cagcaatcgg	cgacctgggc	ctgctggagg	agacactggt	catcttcacc	2580
gccgataacg	gccctgagac	aatgcggatg	tctagaggcg	gatgcagcgg	cctgctgaga	2640
tgtggcaagg	gaaccacata	cgagggaggc	gtgcgcgagc	ctgccctggc	attttggcca	2700
ggacacatcg	cacctggagt	gaccacgag	ctggcctcct	ctctggacct	gctgccaaca	2760
ctggccgccc	tggcaggagc	acctctgcca	aatgtgacce	tggacggctt	cgatctgagc	2820
ccactgctgc	tgggaaccgg	caagtccctt	aggcagtctc	tgtttcttta	ccccctctat	2880

cctgatgagg tgcggggcgt gtttgccgtg agaaccggca agtacaaggc ccacttcctt	2940
acacagggct ctgcccacag cgacaccaca gcagatccag catgccacgc cagctcctct	3000
ctgaccgcac acgagccacc tctgctgtac gacctgtcca aggatcccgg cgagaactat	3060
aatctgctgg gaggagtggc aggagcaacc cctgaggtgc tgcaggccct gaagcagctg	3120
cagctgctga aggcacagct ggacgcagca gtgacattcg gccaagcca ggtggccaga	3180
ggcgaggatc ccgccctgca gatctgttgc caccocggct gcacccaag acctgcctgt	3240
tgccattgcc ccgaccaca cgccggaaaa ccaataccaa accctctatt aggattggac	3300
tcaacataag attctagagt cgagccgcgg actagtaact tgttttattgc agcttataat	3360
ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttcactgcat	3420
tctagttgtg gtttgtccaa actcatcaat gtatctta	3458

<210> 80
<211> 3991
<212> DNA
<213> 人工序列

<220>
<223> 合成核酸序列

ttggccactc cctctctgcg cgctcgtctg ctcaactgagg ccgggcgacc aaaggtcgcc	60
cgacgccccg gctttgcccg ggccggcctca gtgagcgagc gagcgcgcag agagggagtg	120
gccaactcca tcaactagggg ttccctggagg ggtggagtcg tgacgtgaat tacgtcatag	180
ggttagggag gtccctgcaga tcttcaatat tggccattag ccatattatt cattggttat	240
atagcataaa tcaatattgg ctattggcca ttgcatacgt tgtatctata tcataatatg	300
tacatttata ttggctcatg tccaatatga ccgccatgtt ggcattgatt attgactagt	360
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt	420
acataactta cggtaaattg cccgcctggc tgaccgcca acgacccccg cccattgacg	480
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	540
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt	600

ccgcccccta ttgacgtcaa tgacggtaaa tggcccgctt ggcattatgc ccagtacatg	660
accttacggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	720
gtcgaggtga gccccacgtt ctgcttcaact ctccccatct cccccccctc cccacccccca	780
attttgtatt tatttatttt ttaattattt tgtgcagcga tgggggcggg gggggggggg	840
gggcgcgcgc caggcggggc ggggcggggc gaggggcggg gcggggcgag gcggagaggt	900
gcggcggcag ccaatcagag cggcgcgcctc cgaaagtttc cttttatggc gaggcggcgg	960
cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg gagtcgctgc gcgctgcctt	1020
cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc cccggctctg actgaccgcg	1080
ttactcccac aggtgagcgg gcgggacggc cttctcctc cgggctgtaa ttagcgcttg	1140
gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa gccttgaggg gctccgggag	1200
ggccctttgt gcggggggag cggctcgggg ggtgcgtgcg tgtgtgtgtg cgtggggagc	1260
gccgcgtgcg gctccgcgct gcccggcggc tgtgagcgtt gcgggcgcgg cgcggggctt	1320
tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg gcggtgcccc gcggtgcggg	1380
gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt gcgtgggggg gtgagcaggg	1440
ggtgtgggcg cgtcggtcgg gctgcaacc cccctgcacc cccctcccc agttgctgag	1500
cacggccccg cttcgggtgc ggggctccgt acggggcgtg gcgcgggggt cgccgtgccg	1560
ggcgggggggt ggcggcaggt gggggtgccg ggcggggcgg ggccgcctcg ggccggggag	1620
ggctcggggg aggggcgcgg cgccccccgg agcgccggcg gctgtcgagg cgcggcgagc	1680
cgcagccatt gccttttatg gtaatcgtgc gagagggcgc agggacttcc tttgtcccaa	1740
atctgtgcgg agccgaaatc tgggaggcgc cgccgcaccc cctctagcgg gcgcggggcg	1800
aagcggtgcg gcgccggcag gaaggaaatg ggcggggagg gccttcgtgc gtcgccgcgc	1860
cgccgtcccc ttctccctct ccagcctcgg ggctgtccgc ggggggacgg ctgccttcgg	1920
gggggacggg gcagggcggg gttcggcttc tggcgtgtga ccggcggctc tagagcctct	1980
gctaaccatg ttcatgcctt cttctttttc ctacagctcc tgggcaacgt gctggttatt	2040
gtgctgtctc atcattttgg caaagaattc cgccaccatg tctatggggg ctccctcgtc	2100

cctgctgctg gcactggccg ccgggctggc tgtcgcaaga ccacctata tcgtcctgat	2160
ttttgcagac gatctgggat acggcgacct gggatgctat ggccacccaa gctccaccac	2220
acccaacctg gaccagctgg cagcaggagg cctgcgggtc accgacttct acgtgccagt	2280
gagcctgtgc accccctcca gagccgccct gctgacaggc aggctgccag tgcgcatggg	2340
catgtatcct ggctgtctgg tgccatctag caggggctgc ctgccactgg aggaggtgac	2400
cgtggcagag gtgctggcag ccagaggcta cctgacagga atggccggca agtggcacct	2460
gggagtggga ccagagggag ccttcctgcc cctcaccag ggcttcacc ggtttctggg	2520
catcccttat tctcacgacc agggcccatg ccagaacctg acctgttttc caccagcaac	2580
accatgcgac ggaggatgtg atcagggcct ggtgccaatc cactgctgg caaatctgag	2640
cgtggaggca cagcctccat ggctgcctgg cctggaggca agatacatgg ccttcgcca	2700
cgacctgatg gcagatgcac agcggcagga tagaccttc tttctgtact atgcctccca	2760
ccacaccac tatccacagt tcagcggcca gtcccttgcc gagagggtcc gaaggggacc	2820
attcggcgac tctctgatgg agctggatgc cgccgtgggc acctgatga cagcaatcgg	2880
cgacctgggc ctgctggagg agacactggt catcttcacc gccgataacg gccctgagac	2940
aatgcggatg tctagaggcg gatgcagcgg cctgctgaga tgtggcaagg gaaccacata	3000
cgagggaggc gtgcgcgagc ctgccctggc attttggcc ggacacatcg cacctggagt	3060
gaccacgag ctggcctcct ctctggacct gctgccaaaca ctggccgcc tggcaggagc	3120
acctctgcca aatgtgacct tggacggctt cgatctgagc cactgctgc tgggaaccgg	3180
caagtcacct aggcagtctc tgttctttta cccctcctat cctgatgagg tgcggggcgt	3240
gtttgccgtg agaaccggca agtacaaggc ccacttcttt acacagggt ctgcccacag	3300
cgacaccaca gcagatccag catgccacgc cagctcctct ctgaccgcac acgagccacc	3360
tctgctgtac gacctgtcca aggatcccg cgagaactat aatctgctgg gaggagtggc	3420
aggagcaacc cctgaggtgc tgcaggccct gaagcagctg cagctgctga aggcacagct	3480
ggacgcagca gtgacattcg gccaagcca ggtggccaga ggcgaggatc ccgccctgca	3540
gatctgttgc caccceggct gcacccaag acctgcctgt tgccattgcc ccgaccaca	3600

cgccggaaaa ccaataccaa accctctatt aggattggac tcaacataag attctagagt	3660
cgagccgcgg actagtaact tgtttattgc agcttataat ggttacaaat aaagcaatag	3720
catcaciaat ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa	3780
actcatcaat gtatcttagg tctagatacg tagataagta gcatggcggg ttaatcatta	3840
actacaagga acccctagtg atggagttgg ccactccctc tctgcgcgct cgctcgcctca	3900
ctgaggccgg gcgaccaaag gtcgcccgcg gcccgggctt tgcccgggcg gcctcagtga	3960
gcgagcgagc gcgcagagag ggagtggcca a	3991

<210> 81
<211> 6654
<212> DNA
<213> 人工序列

<220>
<223> 合成核酸序列

<400> 81 cgccagggtt ttcccagtea cgacgttgta aaacgacggc cagtgccaaag cttgcatgcc	60
tgcatTTggc cactccctct ctgcgcgctc gctcgcctac tgaggccggg cgaccaaagg	120
tcgcccgcg cccgggcttt gcccgggcgg cctcagtgag cgagcgagcg cgcagagagg	180
gagtggccaa ctccatcact aggggttctt ggaggggttg agtcgtgacg tgaattacgt	240
catagggtta gggaggctct gcagatcttc aatattggcc attagccata ttattcattg	300
gttatatagc ataaatcaat attggctatt ggccattgca tacgttgtat ctatatcata	360
atatgtacat ttatatTggc tcatgtccaa tatgaccgcc atgttggcat tgattattga	420
ctagttatta atagtaatca attacggggT cattagttca tagcccatat atggagtTcc	480
gcgttacata acttacggta aatggccccg ctggctgacc gccaacgcac ccccgcccat	540
tgacgtcaat aatgacgtat gttcccatag taacgccaat agggaCTTc cattgacgtc	600
aatgggtTga gtattTacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	660
caagtcgcgc ccctattgac gtcaatgacg gtaaattggc cgcttggeat tatgcccagt	720
acatgacctt acgggacttt cctactTggc agtacatcta cgtattagtc atcgctatta	780

ccatggtcga ggtgagcccc acgttctgtct tcactctccc catctccccc cctcccccac	840
ccccaatttt gtatttatatt attttttaat tatttttgtgc agcgatgggg gcgggggggg	900
ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga	960
gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttcctttt atggcgaggc	1020
ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgcgct	1080
gccttcgccc cgtgccccgc tccgcgcgcg cctcgcgcgc cccgccccgg ctctgactga	1140
ccgcgttact cccacagggtg agcgggcggg acggcccttc tcctccgggc tgtaattagc	1200
gcttggttta atgacggctt gtttcttttc tgttgctgcg tgaaagcctt gaggggctcc	1260
gggagggccc tttgtgcggg gggagcggct cgggggggtgc gtgcgttgt gtgtgcgtgg	1320
ggagcgcgcg gtgcggctcc gcgctgcccg gcggctgtga gcgctgcggg cgcggcgcgg	1380
ggcttttgtc gctccgcagt gtgcgcgagg ggagcgcggc cgggggcggg gccccgcgg	1440
gcgggggggg ctgcgagggg aacaaaggct gcgtgcgggg tgtgtgcgtg ggggggtgag	1500
cagggggtgt gggcgcgtcg gtcgggctgc aacccccct gcacccccct ccccgagttg	1560
ctgagcacgg cccggcttcg ggtgcggggc tccgtacggg gcgtggcgcg gggctcgcgc	1620
tgccgggcgg ggggtggcgg caggtggggg tgccgggcgg ggcggggccg cctcgggccg	1680
gggagggctc gggggagggg cgcggcggcc cccggagcgc cggcggctgt cgaggcgcgg	1740
cgagccgcag ccattgcctt ttatggtaat cgtgcgagag ggcgcaggga cttcctttgt	1800
cccaaatctg tgcggagccg aaatctggga ggcgccgcg cccccctct agcgggcgcg	1860
gggcgaagcg gtgcggcgcc ggcaggaagg aaatgggcgg ggagggcctt cgtgcgtcgc	1920
cgcgccgcg tcccccttc cctctccagc ctcggggctg tccgcggggg gacggctgcc	1980
ttcggggggg acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag	2040
cctctgctaa ccatgttcat gccttcttct ttttctaca gctcctgggc aacgtgctgg	2100
ttatttgtct gtctcatcat tttggcaaag aattccgcca ccatgtctat gggggctcct	2160
cgtccctgc tgttggaact ggccgccggg ctggctgtcg caagaccacc taatatcgtc	2220
ctgatttttg cagacgatct gggatacggc gacctgggat gctatggcca cccaagctcc	2280

accacacca	acctggacca	gctggcagca	ggaggcctgc	ggttcaccga	cttctacgtg	2340
ccagtgagcc	tgtgcacccc	ctccagagcc	gccctgctga	caggcaggct	gccagtgcgc	2400
atgggcatgt	atcctggcgt	gctggtgcc	tctagcaggg	gcggcctgcc	actggaggag	2460
gtgaccgtgg	cagaggtgct	ggcagccaga	ggctacctga	caggaatggc	cggcaagtgg	2520
cacctgggag	tgggaccaga	gggagccttc	ctgccccctc	accagggctt	ccaccggttt	2580
ctgggcatcc	cttattctca	cgaccagggc	ccatgccaga	acctgacctg	ttttccacca	2640
gcaacaccat	gcgacggagg	atgtgatcag	ggcctgggtgc	caatcccact	gctggcaaat	2700
ctgagcgtgg	aggcacagcc	tccatggctg	cctggcctgg	aggcaagata	catggccttc	2760
gcccacgacc	tgatggcaga	tgcacagcgg	caggatagac	ctttctttct	gtactatgcc	2820
tcccaccaca	cccactatcc	acagttcagc	ggccagtcct	ttgccgagag	gtccggaagg	2880
ggaccattcg	gcgactctct	gatggagctg	gatgccgccg	tgggcaccct	gatgacagca	2940
atcggcgacc	tgggcctgct	ggaggagaca	ctggtcatct	tcaccgccga	taacggccct	3000
gagacaatgc	ggatgtctag	aggcggatgc	agcggcctgc	tgagatgtgg	caagggaacc	3060
acatacgagg	gaggcgtgcg	cgagcctgcc	ctggcatttt	ggccaggaca	catcgcacct	3120
ggagtgacct	acgagctggc	ctcctctctg	gacctgctgc	caacactggc	cgccctggca	3180
ggagcacctc	tgccaaatgt	gaccctggac	ggcttcgata	tgagcccact	gctgctggga	3240
accggcaagt	cccctaggca	gtctctgttc	ttttaccctt	cctatcctga	tgaggtgcgg	3300
ggcgtgtttg	ccgtgagaac	cggcaagtac	aaggcccact	tctttacaca	gggctctgcc	3360
cacagegaca	ccacagcaga	tccagcatgc	cacgccagct	cctctctgac	cgcacacgag	3420
ccacctctgc	tgtacgacct	gtccaaggat	cccggcgaga	actataatct	gctgggagga	3480
gtggcaggag	caaccctga	ggtgctgcag	gccctgaagc	agctgcagct	gctgaaggca	3540
cagctggacg	cagcagtgac	attcggccca	agccaggtag	ccagaggcga	ggatcccgcc	3600
ctgcagatct	gttgccacct	cggctgcacc	ccaagacctg	cctgttgcca	ttgccccgac	3660
ccacacgccg	gaaaaccaat	accaaaccct	ctattaggat	tggactcaac	ataagattct	3720
agagtcgagc	cgcggactag	taacttgttt	attgcagctt	ataatggtta	caaataaagc	3780

aatagcatca	caaatttcac	aaataaagca	tttttttcac	tgcattctag	ttgtggtttg	3840
tccaaactca	tcaatgtatc	ttaggtctag	atacgtagat	aagtagcatg	gcgggttaat	3900
cattaactac	aaggaacccc	tagtgatgga	gttggccact	ccctctctgc	gcgctcgctc	3960
gctcactgag	gccgggcgac	caaaggtcgc	ccgacgcccc	ggctttgccc	gggcggcctc	4020
agtgagcgag	cgagcgcgca	gagagggagt	ggccaaagat	ccccgggtac	cgagctcgaa	4080
ttcgtaatca	tgtcatagct	gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	4140
aacatacgag	ccggaagcat	aaagtgtaaa	gcctggggtg	cctaatagag	gagctaactc	4200
acattaattg	cgttgcgctc	actgccccgt	ttccagtcgg	gaaacctgtc	gtgccagctg	4260
cattaatgaa	tcggccaacg	cgcggggaga	ggcggtttgc	gtattggcga	acttttgctg	4320
agttgaagga	tcagatcacg	catcttcccc	acaacgcaga	ccgttccgtg	gcaaagcaaa	4380
agttcaaaat	cagtaaccgt	cagtgccgat	aagttcaaag	ttaaacctgg	tgttgatacc	4440
aacattgaaa	cgtgatcga	aaacgcgctg	aaaaacgctg	ctgaatgtgc	gagcttcttc	4500
cgttccctcg	ctcactgact	cgtgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	4560
tcactcaaag	gcggtaatac	ggttatecac	agaatcaggg	gataacgcag	gaaagaacat	4620
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	4680
ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgtcaagtc	agaggtggcg	4740
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	4800
tcctgttccg	acctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	4860
ggcgttttct	caatgctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgctccaa	4920
gctgggctgt	gtgcacgaac	ccccgttca	gcccgaccgc	tgcgccttat	ccggtacta	4980
tcgtcttgag	tccaacccg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	5040
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	5100
ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	5160
cggaaaaaga	gttggtagct	cttgatccgg	caaacaaacc	accgctggta	gcggtggttt	5220
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	5280

cttttctacg gggctctgacg ctcagtggaa cgatccgtcg agaggtctgc ctcgtgaaga	5340
aggtgttgct gactcatacc aggcctgaat cgccccatca tccagccaga aagtgaggga	5400
gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga acttttgctt	5460
tgccacggaa cggctctgctg tgtcgggaag atgcgtgac tgatccttca actcagcaaa	5520
agttcgatth attcaacaaa gccacgttgt gtctcaaaat ctctgatgtt acattgcaca	5580
agataaaaa atatcatcat gaacaataaa actgtctgct tacataaaca gtaatacaag	5640
gggtgttatg agccatattc aacgggaaac gtcttgctcg aagccgcgat taaattccaa	5700
catggatgct gatttatatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc	5760
gacaatctat cgattgtatg ggaagcccga tgcgccagag ttgtttctga aacatggcaa	5820
aggtagcgtt gccaatgatg ttacagatga gatggtcaga ctaaactggc tgacggaatt	5880
tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgeat ggttactcac	5940
cactgcgac cccgggaaaa cagcattcca ggtattagaa gaatatcctg attcaggtga	6000
aaatattgtt gatgcgctgg cagtgttcc tgcgccgttg cattcgattc ctgtttgtaa	6060
ttgtcccttt aacagcgac gcgtatttcg tctcgctcag gcgcaatcac gaatgaataa	6120
cggtttggtt gatgcgagtg attttgatga cgagcgtaat ggctggcctg ttgaacaagt	6180
ctggaaagaa atgcataagc ttttgccatt ctcaccggat tcagtcgtca ctcatgggtga	6240
tttctcactt gataacctta tttttgacga ggggaaatta ataggttgta ttgatgttgg	6300
acgagtcgga atcgcagacc gataccagga tcttgccatc ctatggaaact gcctcgggtga	6360
gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat	6420
gaataaattg cagtttcatt tgatgctcga tgagtthttc taatcagaat tggttaattg	6480
gttgtaacac tggcagagca ttacgctgac ttgacgggac ggcggctttg ttgaataaat	6540
cgcattcgcc attcaggctg cgcaactgtt gggaagggcg atcggtgccg gcctcttcgc	6600
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaa	6654

【發明申請專利範圍】

【請求項1】 一種用於在細胞中表現芳基硫酸酯酶 A（ARSA）多肽之方法，該方法包含用重組腺相關病毒（rAAV）轉導該細胞，該重組腺相關病毒包含：

（a）AAV 衣殼，其包含 AAV 衣殼蛋白質；及

（b）轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序列之轉錄調節元件。

【請求項2】 如請求項 1 之方法，其中該細胞為神經元及/或神經膠細胞，視情況其中該細胞為中樞神經系統及/或周邊神經系統之神經元及/或神經膠細胞。

【請求項3】 如請求項 1 之方法，其中該細胞為選自由以下組成之群組的中樞神經系統區域之細胞：脊髓、運動皮質、感覺皮質、海馬區、殼核、小腦，視情況，小腦核，及其任何組合。

【請求項4】 如請求項 1 之方法，其中該細胞為選自由以下組成之群組之細胞：運動神經元、星形膠質細胞、寡樹突神經膠細胞、中樞神經系統中之大腦皮質之細胞、周邊神經系統之感覺神經元、許旺細胞（Schwann cell）及其任何組合。

【請求項5】 如請求項 1 至 4 中任一項之方法，其中該細胞在哺乳動物個體中且以可有效轉導該個體中之該細胞之量向該個體投與該 AAV。

【請求項6】 一種用於治療患有異染性白質失養症（MLD）之個體之方法，該方法包含向該個體投與有效量之 rAAV，該 rAAV 包含：

（a）AAV 衣殼，其包含衣殼蛋白質；及

（b）轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序

列之轉錄調節元件。

【請求項7】 如前述請求項中任一項之方法，其中該以緘默方式改變之 ARSA 編碼序列編碼 SEQ ID NO: 23 中所示之胺基酸序列。

【請求項8】 如請求項 7 之方法，其中該以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO: 14、62 或 72 中所示之核苷酸序列。

【請求項9】 如前述請求項中任一項之方法，其中該轉錄調節元件包含選自由以下組成之群組的元件中之一或多者：細胞巨大病毒（CMV）強化子元件、雞 β -肌動蛋白（CBA）啟動子、小型雞 β -肌動蛋白（SmCBA）啟動子、鈣調蛋白 1（CALM1）啟動子、蛋白脂質蛋白質 1（PLP1）啟動子、膠質原纖維酸性蛋白（GFAP）啟動子、突觸蛋白 2（SYN2）啟動子、金屬硫蛋白 3（MT3）啟動子及其任何組合。

【請求項10】 如前述請求項中任一項之方法，其中該轉錄調節元件包含與選自由以下組成之群組的序列至少 90%一致之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。

【請求項11】 如前述請求項中任一項之方法，其中該轉錄調節元件包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。

【請求項12】 如前述請求項中任一項之方法，其中該轉錄調節元件自 5'至 3'包含 SEQ ID NO: 58、25 及 32 中所示之核苷酸序列。

【請求項13】 如前述請求項中任一項之方法，其中該轉錄調節元件包含 SEQ ID NO: 36 中所示之核苷酸序列。

【請求項14】 如前述請求項中任一項之方法，其中該轉移基因組進一步包含位於該以緘默方式改變之 ARSA 編碼序列 3'-聚腺苷酸化序列。

【請求項15】 如請求項 14 之方法，其中該聚腺苷酸化序列為外源聚腺苷酸化序列。

【請求項16】 如請求項 15 之方法，其中該外源聚腺苷酸化序列為 SV40 聚腺苷酸化序列。

【請求項17】 如請求項 16 之方法，其中該 SV40 聚腺苷酸化序列包含 SEQ ID NO: 42 中所示之核苷酸序列。

【請求項18】 如前述請求項中任一項之方法，其中該轉移基因組進一步包含填充序列。

【請求項19】 如前述請求項中任一項之方法，其中該轉移基因組進一步包含位於該以緘默方式改變之 ARSA 編碼序列 3'-之填充序列。

【請求項20】 如請求項 18 或 19 之方法，其中該填充序列位於該聚腺苷酸化序列的 3'。

【請求項21】 如前述請求項中任一項之方法，其中該轉移基因組包含選自由以下組成之群組之序列：SEQ ID NO: 41、44、46、65、67、75 及 79。

【請求項22】 如前述請求項中任一項之方法，其中該轉移基因組進一步包含該基因組的 5'-5'反向末端重複（5' ITR）核苷酸序列，及該基因組的 3'-3'反向末端重複（3' ITR）核苷酸序列。

【請求項23】 如請求項 22 之方法，其中該 5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95%序列一致性，且該 3' ITR 核苷酸序列與 SEQ ID NO: 19 具有至少 95%序列一致性。

【請求項24】 如請求項 22 之方法，其中該 5' ITR 核苷酸序列與 SEQ ID NO: 26 具有至少 95%序列一致性，且該 3' ITR 核苷酸序列與 SEQ ID NO: 27 具有至少 95%序列一致性。

【請求項25】 如請求項 22 之方法，其中該 5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95%序列一致性，且該 3' ITR 核苷酸序列與 SEQ ID NO: 57 具有至少 95%序列一致性。

【請求項26】 如前述請求項中任一項之方法，其中該轉移基因組包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 47、48、49、68、69、76 及 80。

【請求項27】 如請求項 5 至 25 中任一項之方法，其中異染性白質失養症與芳基硫酸酯酶 A（ARSA）基因突變相關。

【請求項28】 如請求項 6 至 27 中任一項之方法，其中該個體為人類個體。

【請求項29】 如前述請求項中任一項之方法，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項30】 如請求項 29 之方法，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項31】 如請求項 30 之方法，其中：

（a）該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，

且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；

(b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；

(c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；

(d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或

(e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項32】 如請求項 30 之方法，其中該衣殼蛋白質包含 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列。

【請求項33】 如前述請求項中任一項之方法，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項34】 如請求項 33 之方法，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO:

16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項35】 如請求項 34 之方法，其中：

- (a) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；
- (b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；
- (c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；
- (d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或
- (e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項36】 如請求項 34 之方法，其中該衣殼蛋白質包含 SEQ ID NO: 2、

3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列。

【請求項37】 如前述請求項中任一項之方法，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項38】 如請求項 37 之方法，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 68 之胺基酸為 V；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ

ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項39】 如請求項 38 之方法，其中：

(a) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；

(b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 Y；

(c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；

(d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；

(e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；

(f) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；

(g) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；

(h) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或

(i) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項40】 如請求項 38 之方法，其中該衣殼蛋白質包含 SEQ ID NO: 2、

3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列。

【請求項41】 一種 rAAV，其包含：

(a) AAV 衣殼，其包含 AAV 衣殼蛋白質；及

(b) 轉移基因組，其包含可操作地連接至以緘默方式改變之 ARSA 編碼序列之轉錄調節元件。

【請求項42】 如請求項 41 之 rAAV，其中該以緘默方式改變之 ARSA 編碼序列編碼 SEQ ID NO:23 中所示之胺基酸序列。

【請求項43】 如請求項 42 之 rAAV，其中該以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO:14 中所示之核苷酸序列。

【請求項44】 如請求項 42 之 rAAV，其中該以緘默方式改變之 ARSA 編碼序列包含 SEQ ID NO:62 或 72 中所示之核苷酸序列。

【請求項45】 如請求項 41 至 44 中任一項之 rAAV，其中該轉錄調節元件包含選自由以下組成之群組的元件中之一或多者：細胞巨大病毒 (CMV) 強化子元件、雞 β -肌動蛋白 (CBA) 啟動子、小型雞 β -肌動蛋白 (SmCBA) 啟動子、鈣調蛋白 1 (CALM1) 啟動子、蛋白脂質蛋白質 1 (PLP1) 啟動子、膠質原纖維酸性蛋白 (GFAP) 啟動子、突觸蛋白 2 (SYN2) 啟動子、金屬硫蛋白 3 (MT3) 啟動子及其任何組合。

【請求項46】 如請求項 45 之 rAAV，其中該轉錄調節元件包含與選自由以下組成之群組的序列至少 90%一致之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。

【請求項47】 如請求項 45 之 rAAV，其中該轉錄調節元件包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 25、32、36、54、55 及 58。

【請求項48】 如請求項 45 之 rAAV，其中該轉錄調節元件自 5'至 3'包含

SEQ ID NO: 58、25 及 32 中所示之核苷酸序列。

【請求項49】 如請求項 45 之 rAAV，其中該轉錄調節元件包含 SEQ ID NO: 36 中所示之核苷酸序列。

【請求項50】 如請求項 41 至 49 中任一項之 rAAV，其中該轉移基因組進一步包含位於該以緘默方式改變之 ARSA 編碼序列 3'-聚腺苷酸化序列。

【請求項51】 如請求項 50 之 rAAV，其中該聚腺苷酸化序列為外源聚腺苷酸化序列。

【請求項52】 如請求項 51 之 rAAV，其中該外源聚腺苷酸化序列為 SV40 聚腺苷酸化序列。

【請求項53】 如請求項 52 之 rAAV，其中該 SV40 聚腺苷酸化序列包含 SEQ ID NO: 42 中所示之核苷酸序列。

【請求項54】 如請求項 41 至 53 中任一項之 rAAV，其中該轉移基因組包含選自由以下組成之群組之序列：SEQ ID NO: 41、44、46、65、67、75 及 79。

【請求項55】 如請求項 41 至 54 中任一項之 rAAV，其中該轉移基因組進一步包含該基因組的 5'-5'反向末端重複（5' ITR）核苷酸序列，及該基因組的 3'-3'反向末端重複（3' ITR）核苷酸序列。

【請求項56】 如請求項 55 之 rAAV，其中該 5' ITR 核苷酸序列與 SEQ ID NO: 18 具有至少 95%序列一致性，且該 3' ITR 核苷酸序列與 SEQ ID NO: 19 具有至少 95%序列一致性。

【請求項57】 如請求項 41 至 55 中任一項之 rAAV，其中該轉移基因組包含選自由以下組成之群組之核苷酸序列：SEQ ID NO: 47、48、49、68、69、76 及 80。

【請求項58】 如請求項 41 至 55 中任一項之 rAAV，其中該轉移基因組之核苷酸序列由選自由以下組成之群組之核苷酸序列組成：SEQ ID NO: 47、48、

49、68、69、76 及 80。

【請求項59】 如請求項 41 至 58 中任一項之 rAAV，其中該轉移基因組之核苷酸序列由 SEQ ID NO: 48 中所闡述之核苷酸序列組成。

【請求項60】 如請求項 41 至 59 中任一項之 rAAV，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項61】 如請求項 60 之 rAAV，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項62】 如請求項 61 之 rAAV，其中：

(a) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；

(b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，

該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；

(c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；

(d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或

(e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項63】 如請求項 61 之 rAAV，其中該衣殼蛋白質包含 SEQ ID NO: 2、3、4、6、7、10、11、12、13、15、16 或 17 之胺基酸 203-736 之胺基酸序列。

【請求項64】 如請求項 41 至 63 中任一項之 rAAV，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項65】 如請求項 64 之 rAAV，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸

501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項66】 如請求項 65 之 rAAV，其中：

- (a) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；
- (b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；
- (c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；
- (d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或
- (e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項67】 如請求項 65 之 rAAV，其中該衣殼蛋白質包含 SEQ ID NO: 2、3、4、5、6、7、9、10、11、12、13、15、16 或 17 之胺基酸 138-736 之胺

基酸序列。

【請求項68】 如請求項 41 至 67 中任一項之 rAAV，其中該衣殼蛋白質包含與 SEQ ID NO: 2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺基酸序列具有至少 95%序列一致性之胺基酸序列。

【請求項69】 如請求項 68 之 rAAV，其中：該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 68 之胺基酸為 V；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 151 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 160 之胺基酸為 D；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 206 之胺基酸為 C；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 590 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G 或 Y；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C；或該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G。

【請求項70】 如請求項 69 之 rAAV，其中：

- (a) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 2 之胺基酸為 T，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 312 之胺基酸為 Q；
- (b) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 65 之胺基酸為 I，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 Y；
- (c) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 77 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 690 之胺基酸為 K；
- (d) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 119 之胺基酸為 L，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 468 之胺基酸為 S；
- (e) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 626 之胺基酸為 G，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 718 之胺基酸為 G；
- (f) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 296 之胺基酸為 H，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 464 之胺基酸為 N，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 681 之胺基酸為 M；
- (g) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 687 之胺基酸為 R；
- (h) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 346 之胺基酸為 A，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R；或
- (i) 該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 501 之胺基酸為 I，該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 505 之胺基酸為 R，且該衣殼蛋白質中對應於 SEQ ID NO: 16 之胺基酸 706 之胺基酸為 C。

【請求項71】 如請求項 69 之 rAAV，其中該衣殼蛋白質包含 SEQ ID NO:

2、3、4、5、6、7、8、9、10、11、12、13、15、16 或 17 之胺基酸 1-736 之胺

基酸序列。

【請求項72】 一種醫藥組成物，其包含如請求項 41 至 71 中任一項之 rAAV。

【請求項73】 一種聚核苷酸，其包含 SEQ ID NO: 14、62 及 72 中所示之核
酸序列。

【請求項74】 一種用於製備 rAAV 之封裝系統，其中該封裝系統包含

(a) 第一核苷酸序列，其編碼一或多種 AAV Rep 蛋白質；

(b) 第二核苷酸序列，其編碼如請求項 41 至 71 中任一項之 AAV 之衣殼蛋
白質；及

(c) 第三核苷酸序列，其包含如請求項 41 至 71 中任一項之 AAV 之 rAAV
基因組序列。

【請求項75】 如請求項 74 之封裝系統，其中該封裝系統包含有包含該第
一核苷酸序列及該第二核苷酸序列之第一載體，及包含該第三核苷酸序列之第
二載體。

【請求項76】 如請求項 74 或 75 之封裝系統，其進一步包含第四核苷酸序
列，該第四核苷酸序列包含一或多種輔助病毒基因。

【請求項77】 如請求項 76 之封裝系統，其中該第四核苷酸序列包含於第
三載體內。

【請求項78】 如請求項 74 至 77 中任一項之封裝系統，其中該第四核苷酸
序列包含一或多種來自選自由以下組成之群組的病毒之基因：腺病毒、疱疹病
毒、痘瘡病毒及細胞巨大病毒（CMV）。

【請求項79】 如請求項 74 至 78 中任一項之封裝系統，其中該第一載體、
第二載體及/或該第三載體為質體。

【請求項80】 一種用於 rAAV 之重組製備之方法，該方法包含在可產生該

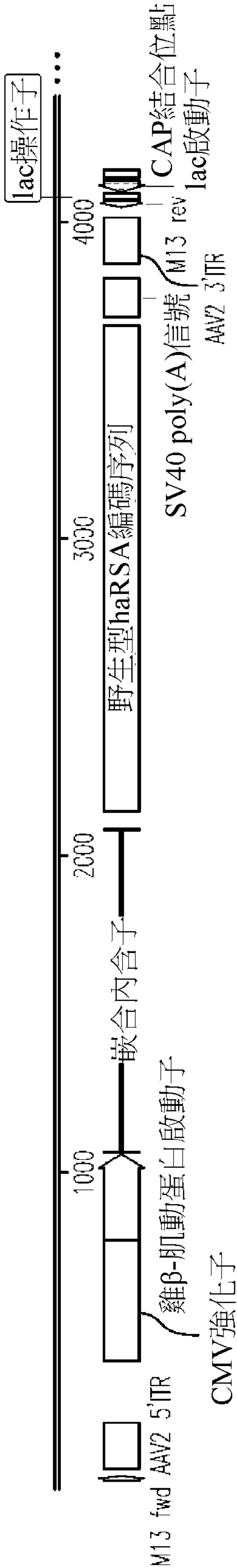
rAAV 之條件下將如請求項 74 至 79 中任一項之封裝系統引入細胞中。

【請求項81】 如請求項 41 至 71 中任一項之 rAAV、如請求項 72 之醫藥組成物或如請求項 73 之聚核苷酸，其適用作藥劑。

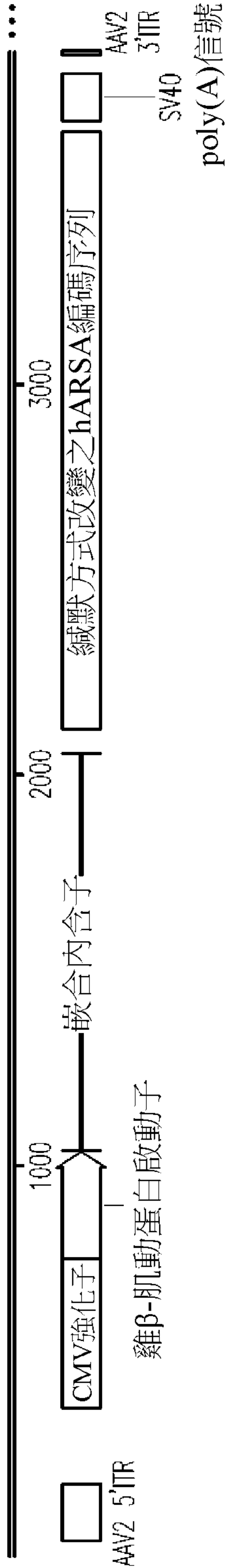
【請求項82】 如請求項 41 至 71 中任一項之 rAAV、如請求項 72 之醫藥組成物或如請求項 73 之聚核苷酸，其係用於治療 MLD。

【請求項83】 如請求項 41 至 71 中任一項之 rAAV、如請求項 72 之醫藥組成物或如請求項 73 之聚核苷酸，其係用於用以治療患有 MLD 之個體之方法中，該方法包含向該個體投與有效量之該 rAAV、該醫藥組成物或該聚核苷酸。

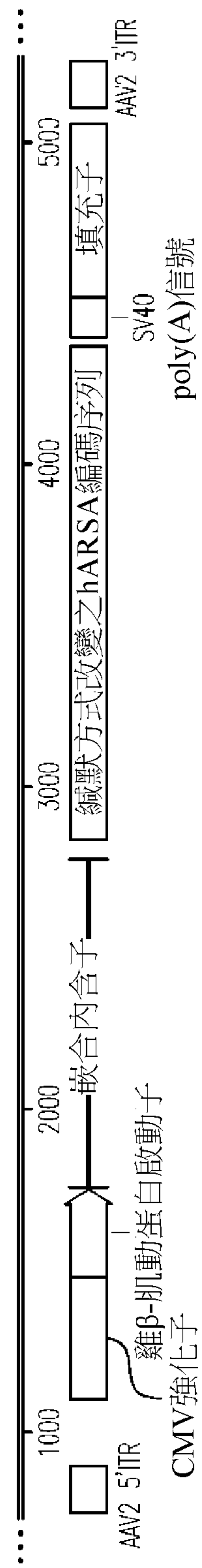
【發明圖式】



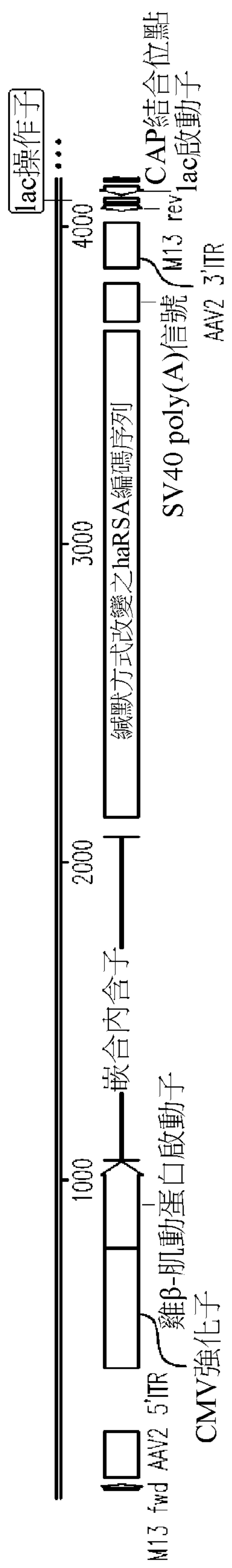
【圖1A】



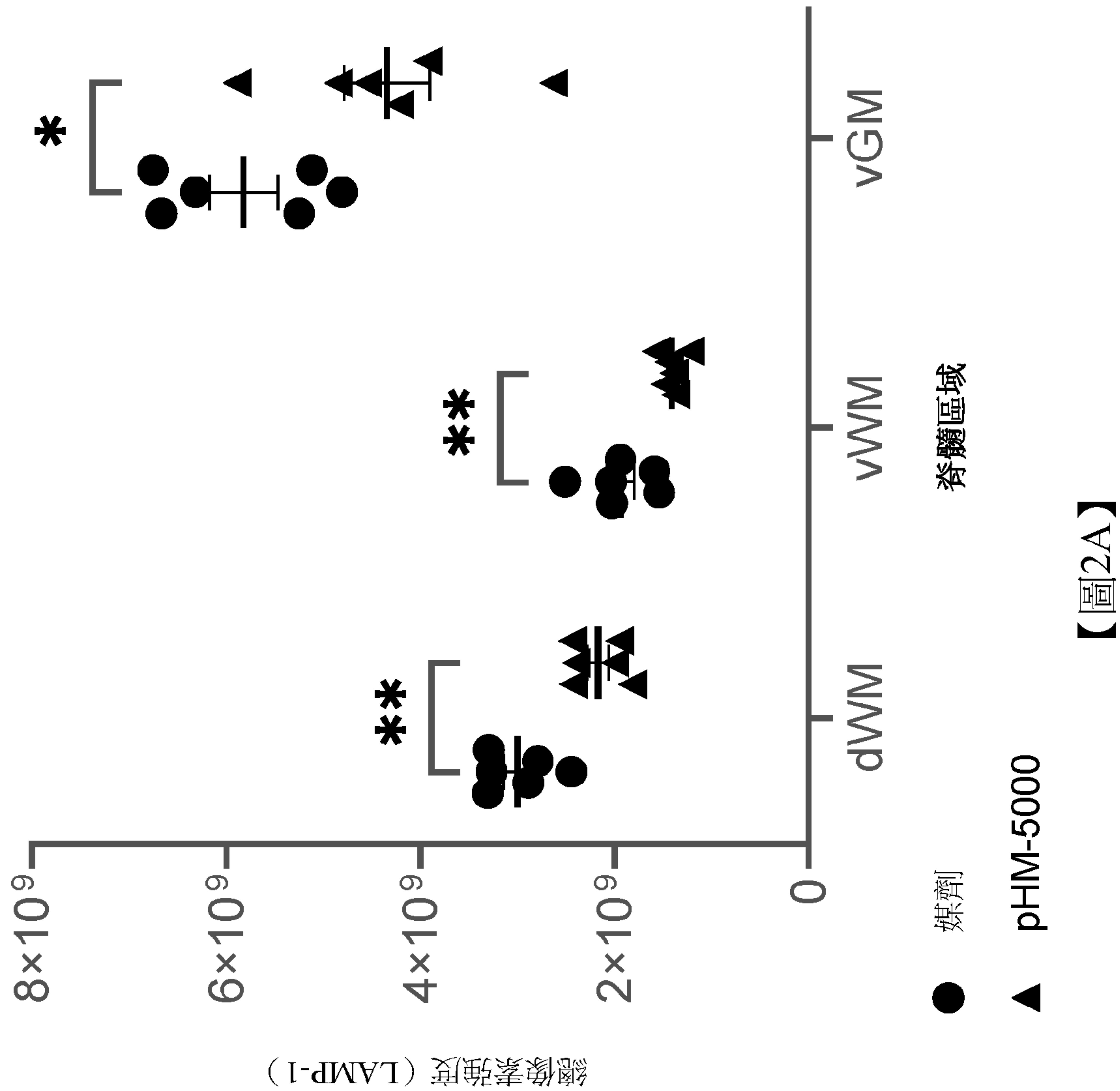
【圖1B】

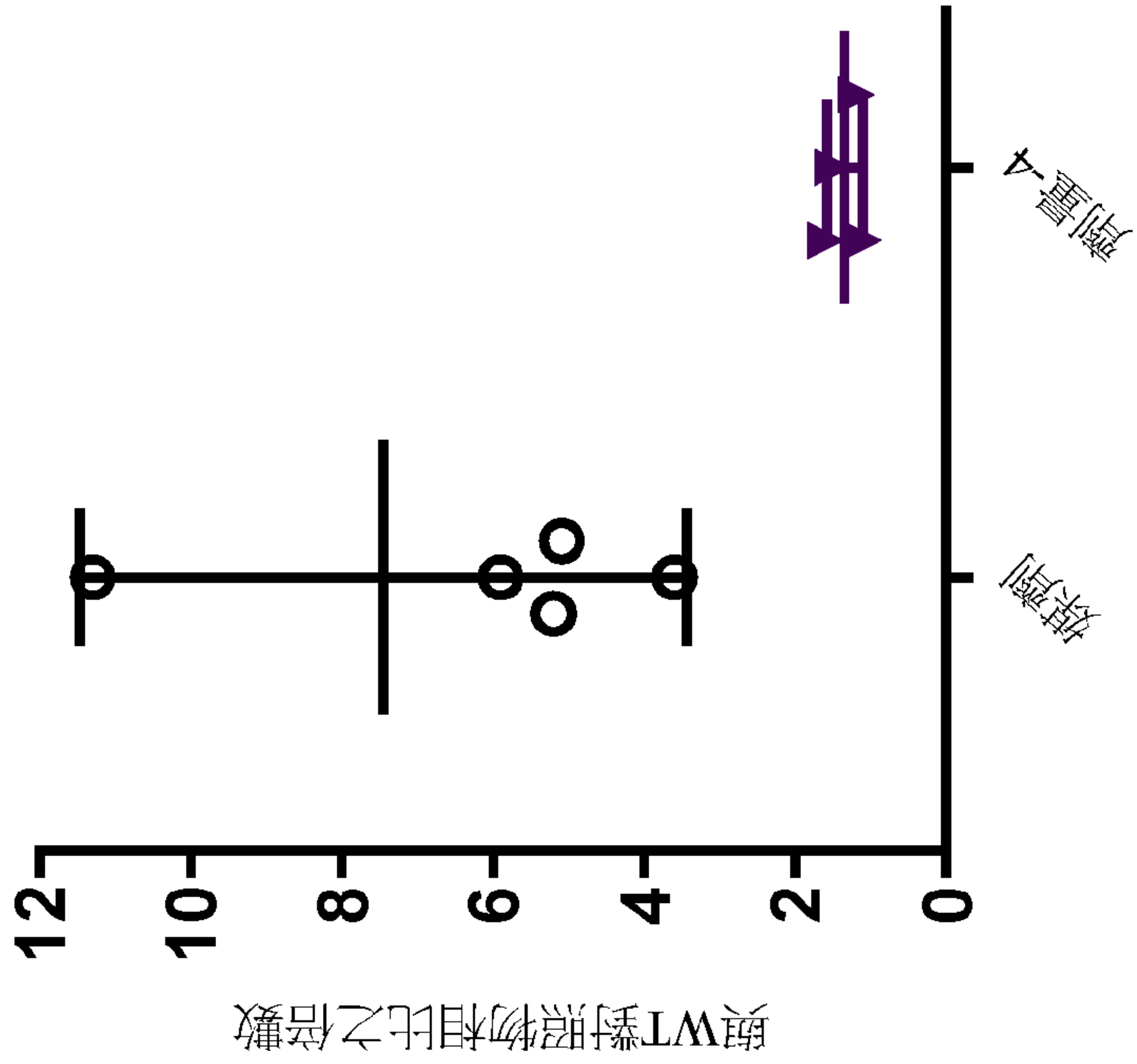


【圖1C】

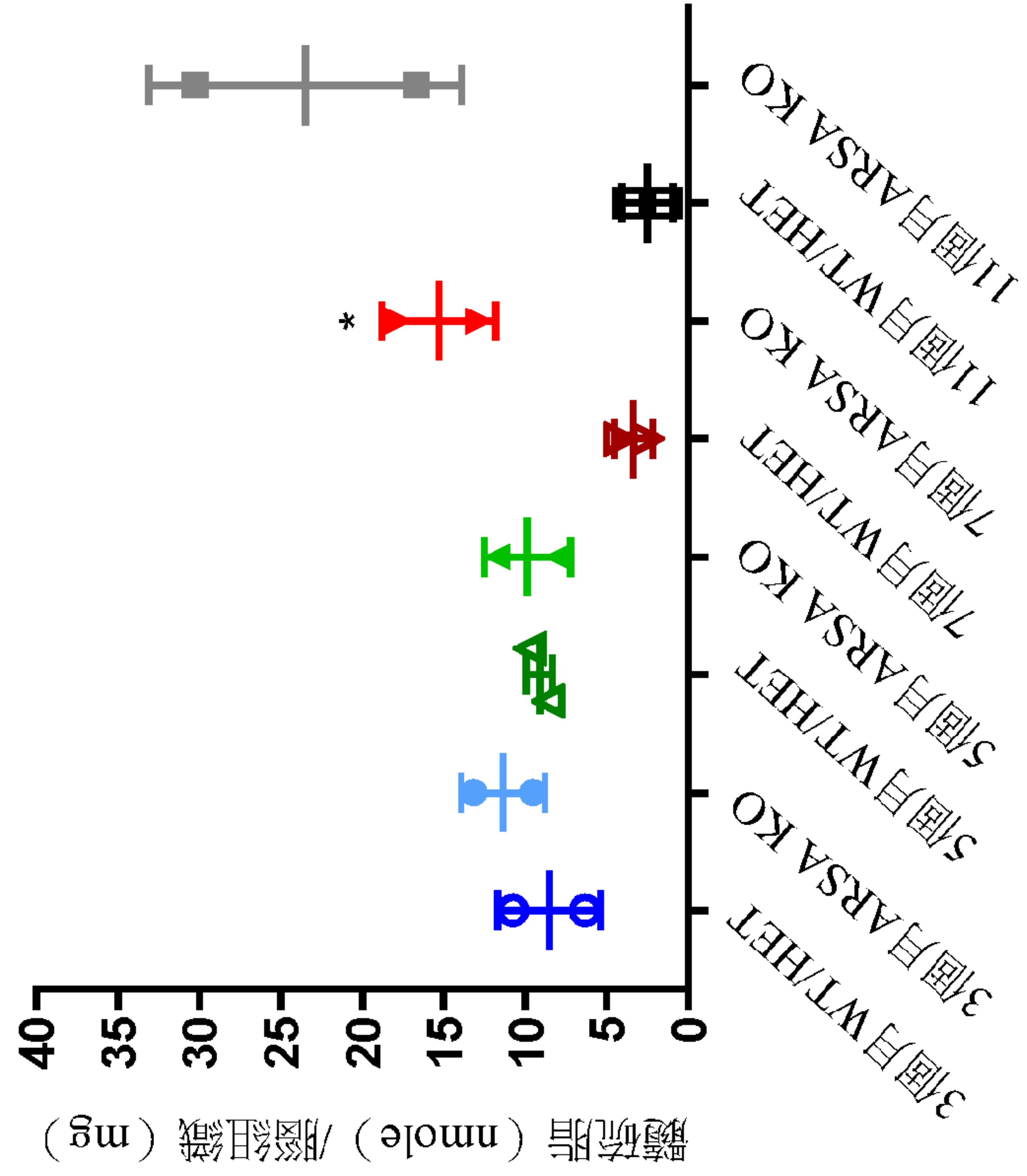


【圖1D】

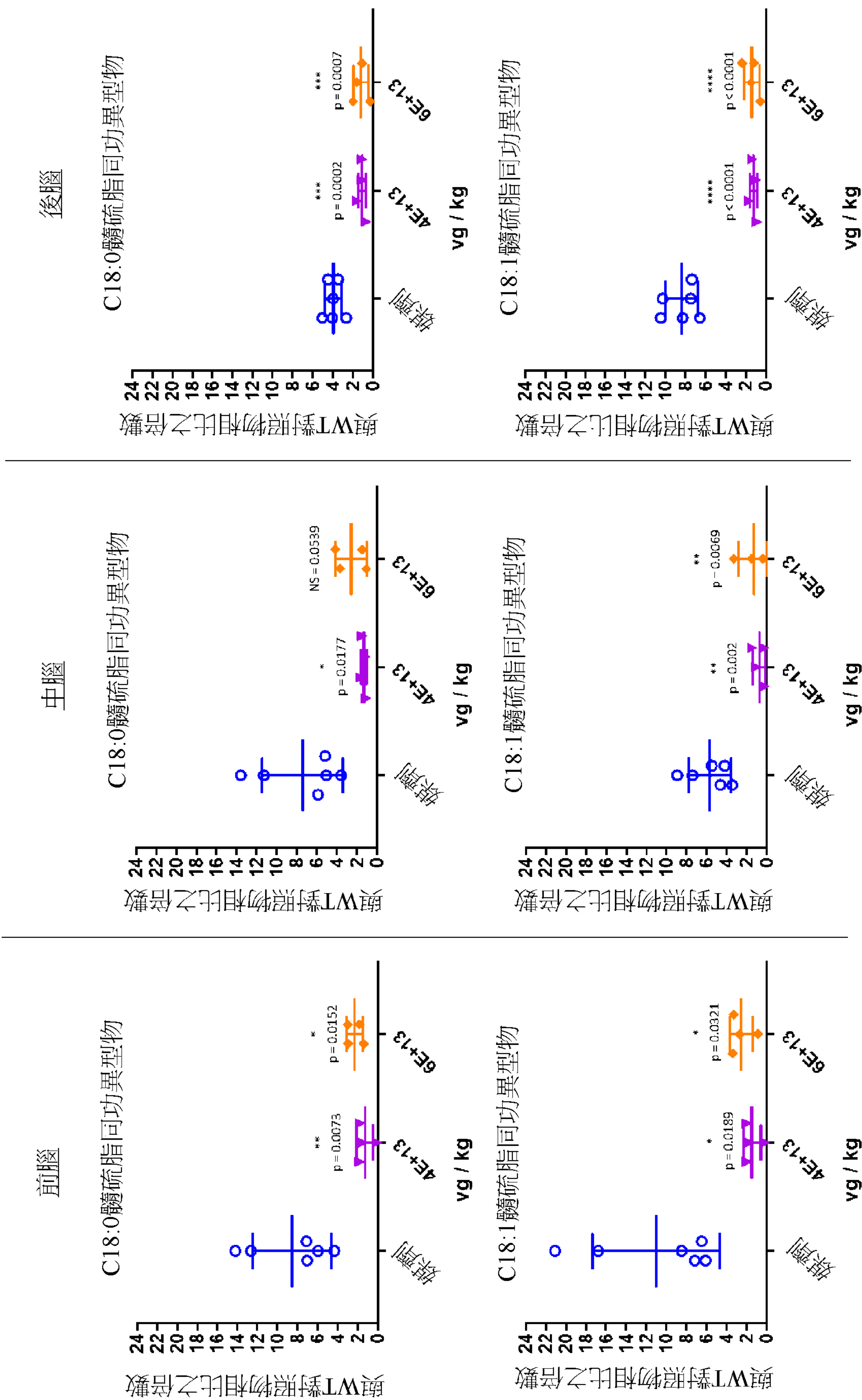




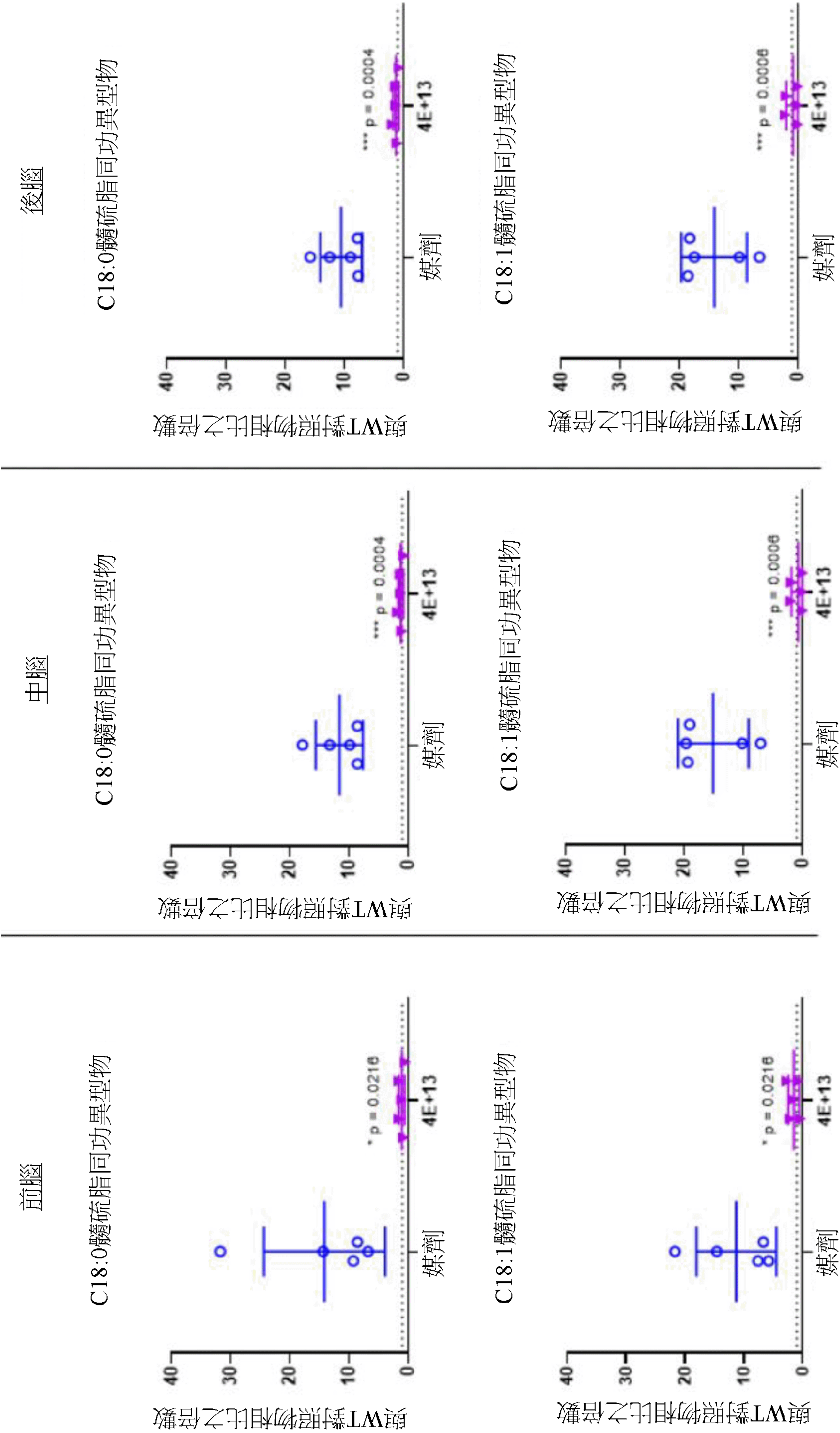
【圖2C】



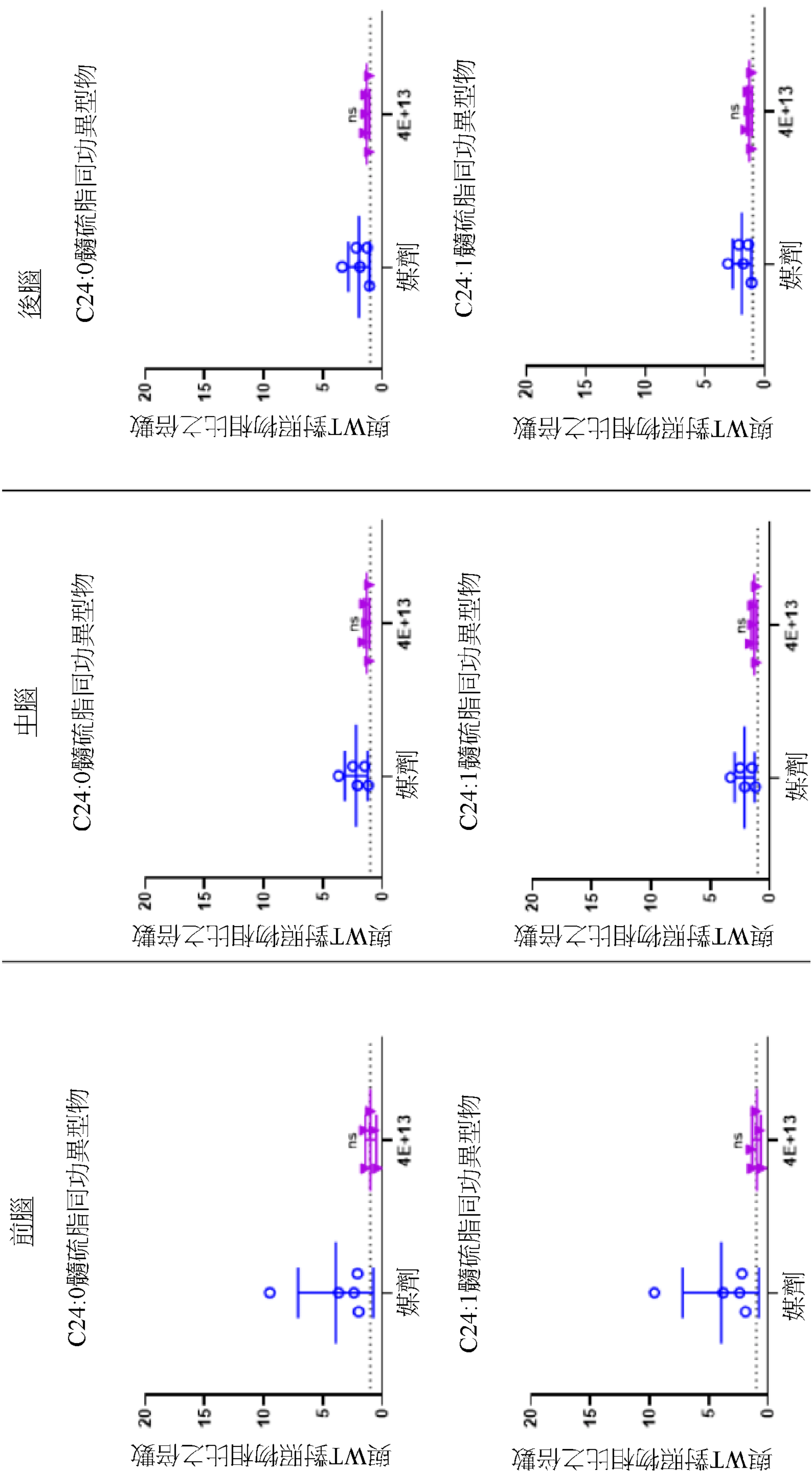
【圖2B】



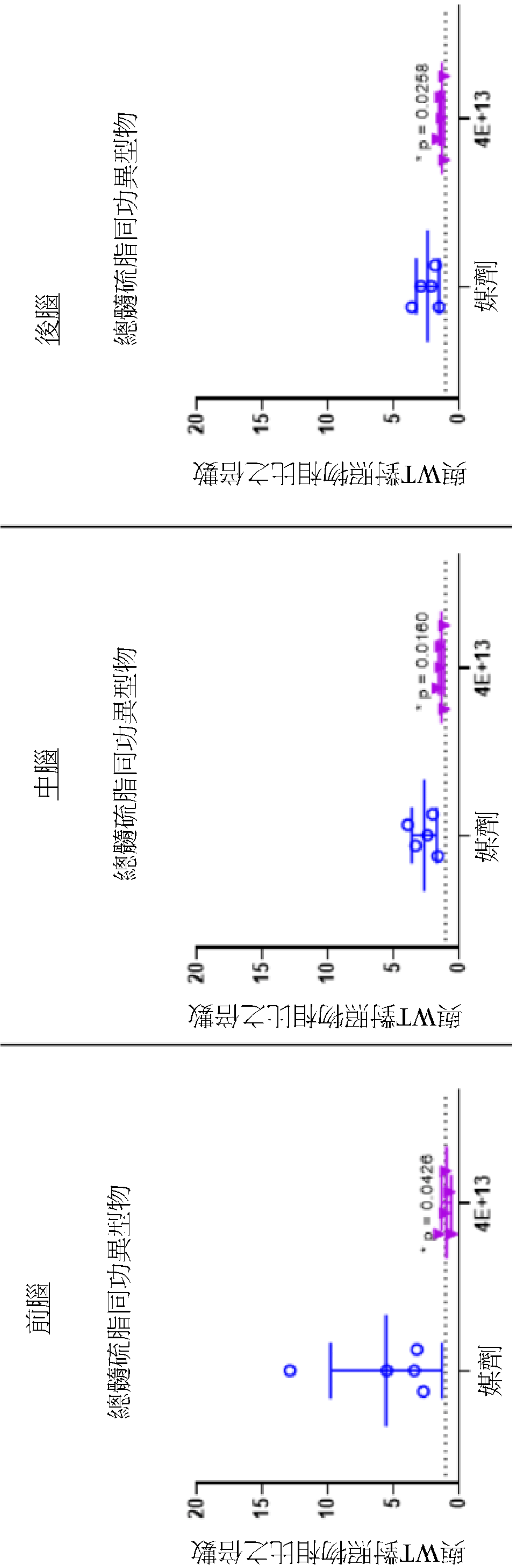
【圖2D】



【圖2E】

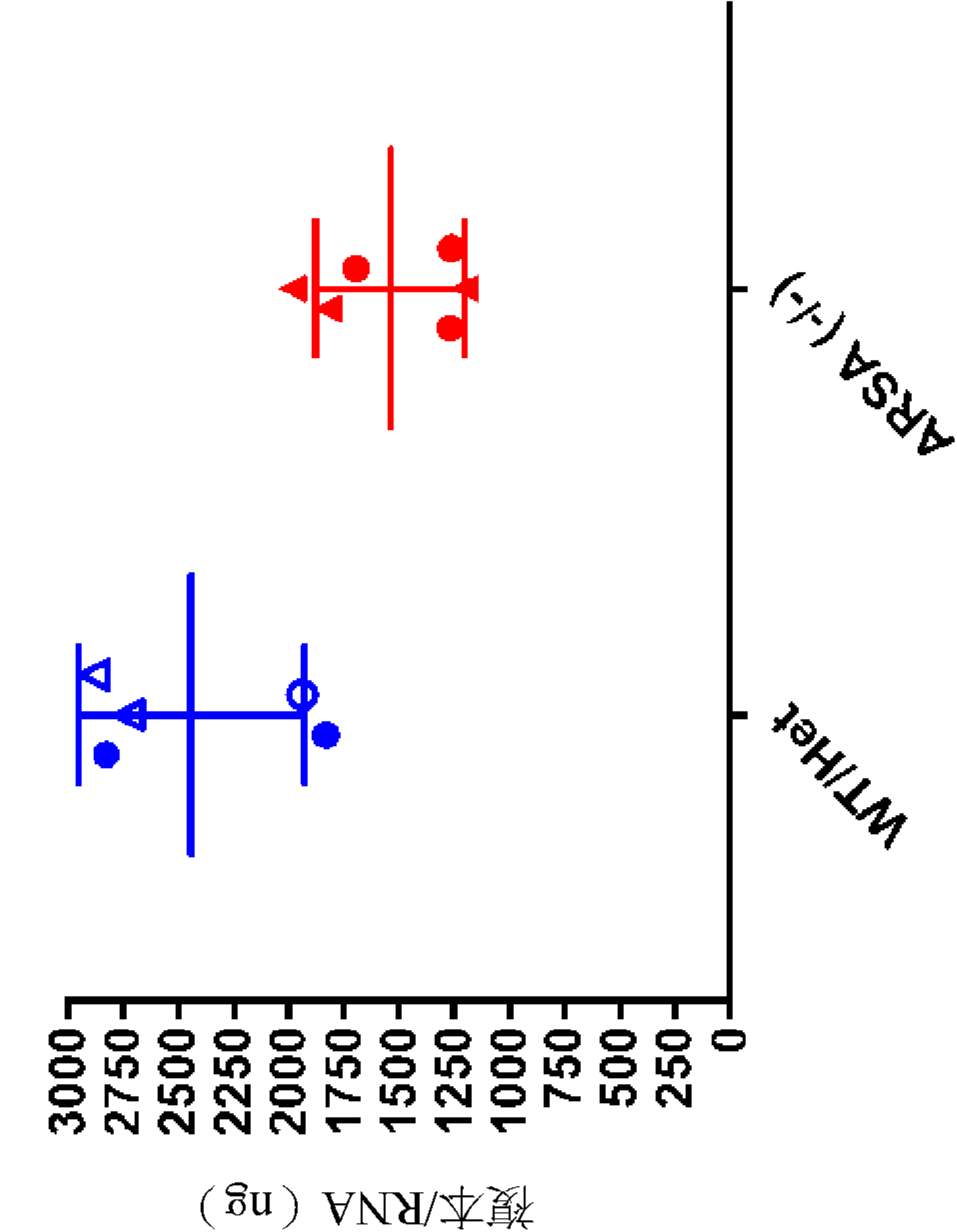


【圖2F】

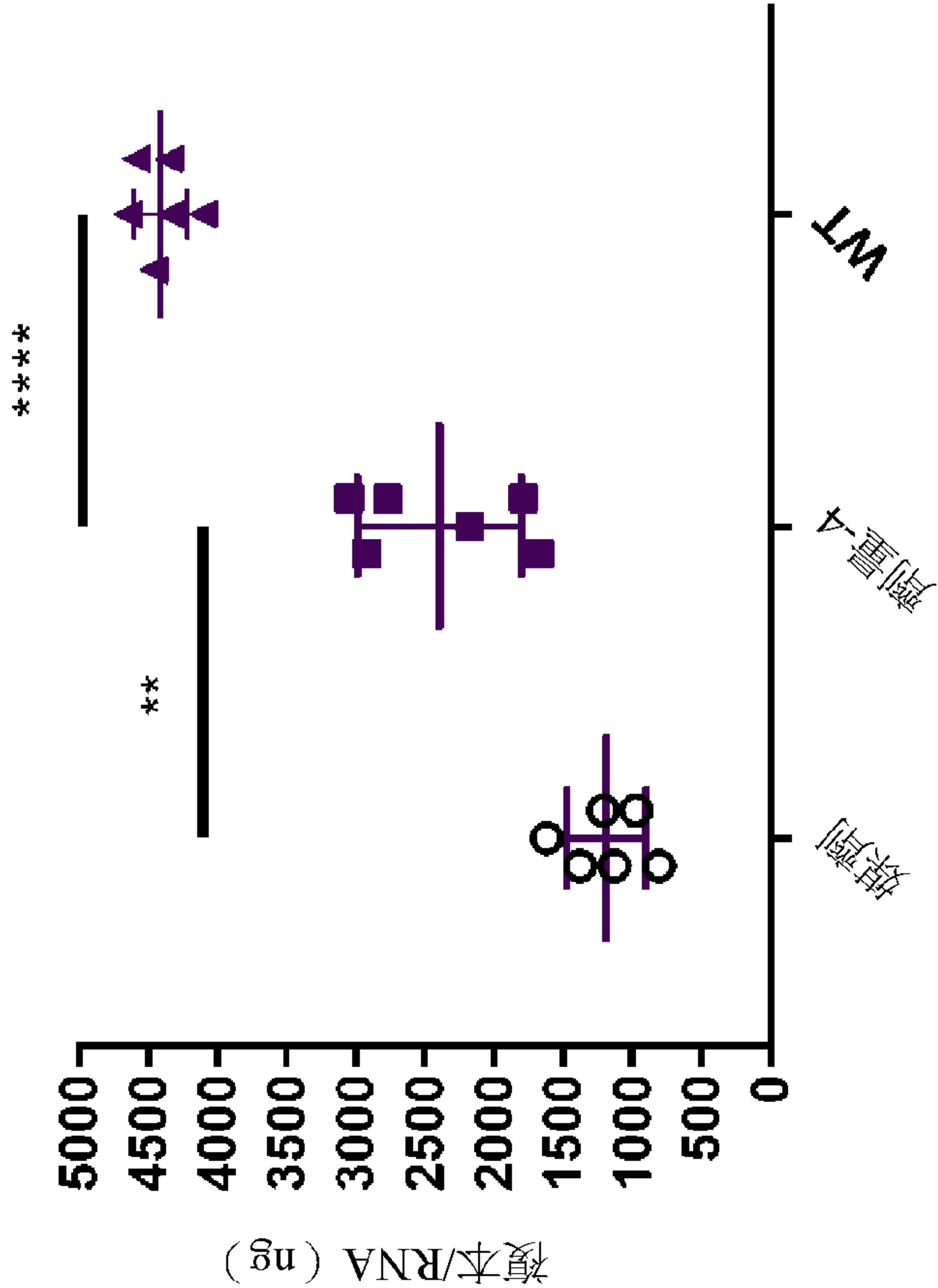


【圖2G】

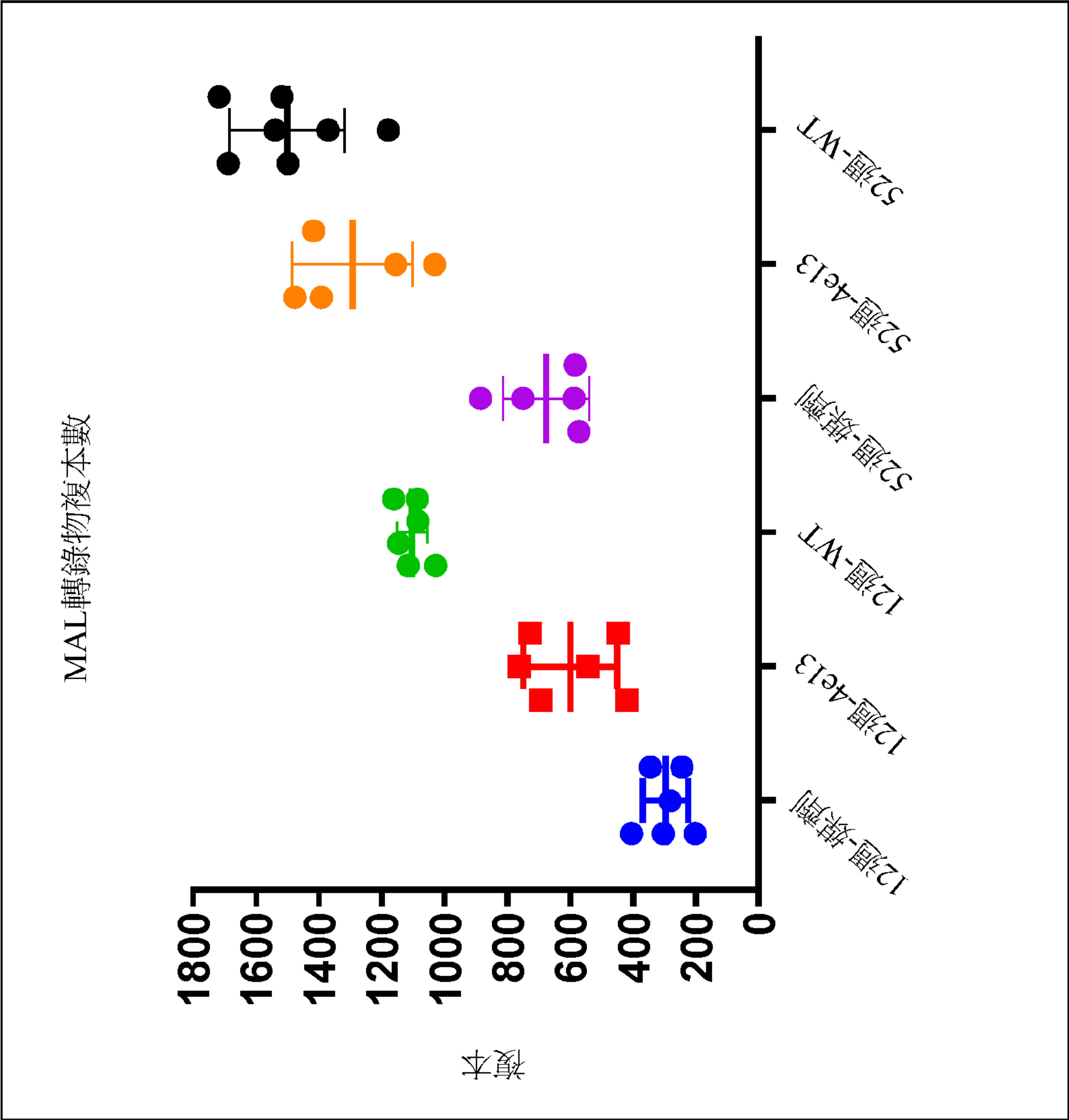
4週時之MAL轉錄物



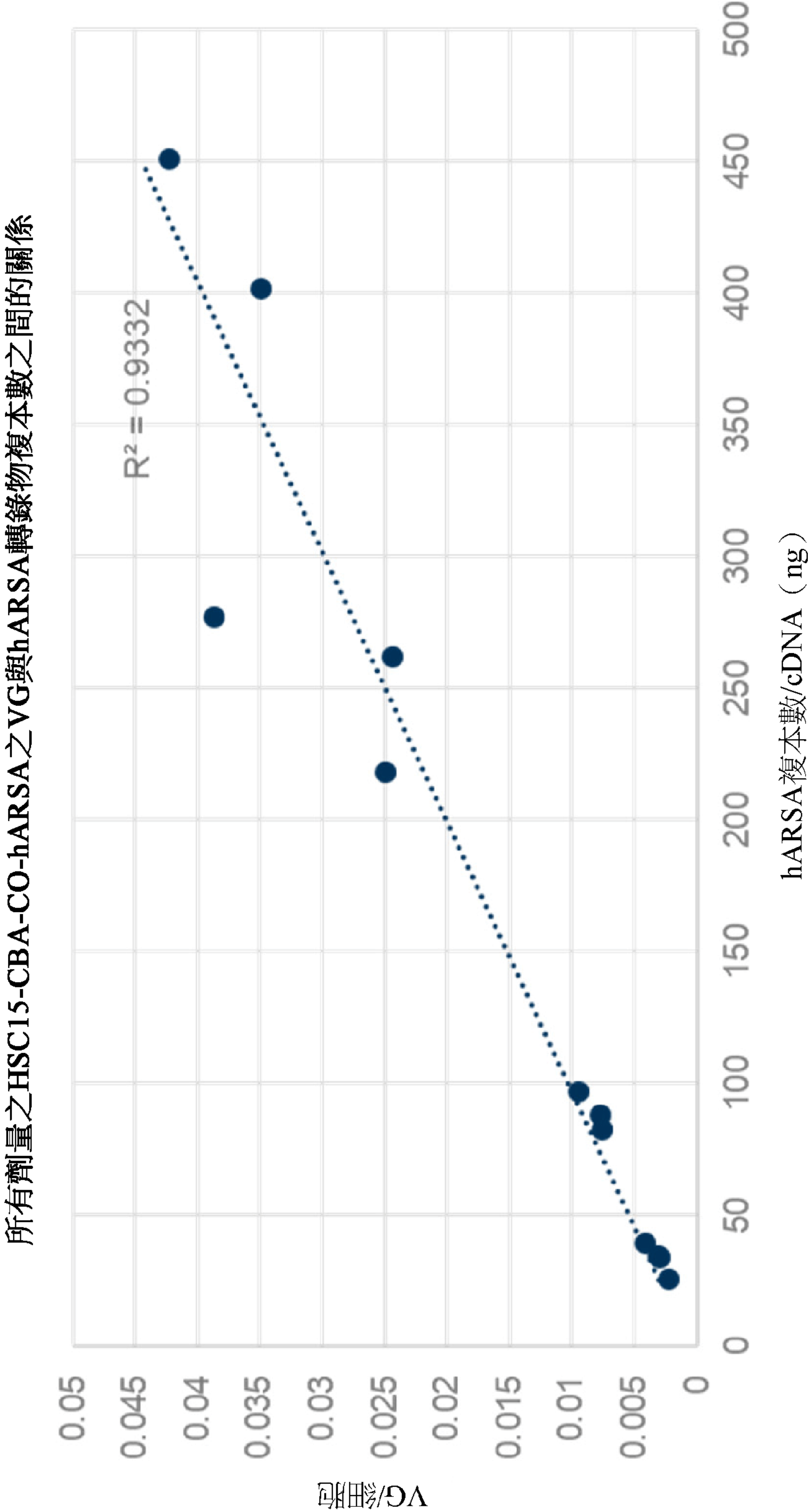
【圖3A】



【圖3B】



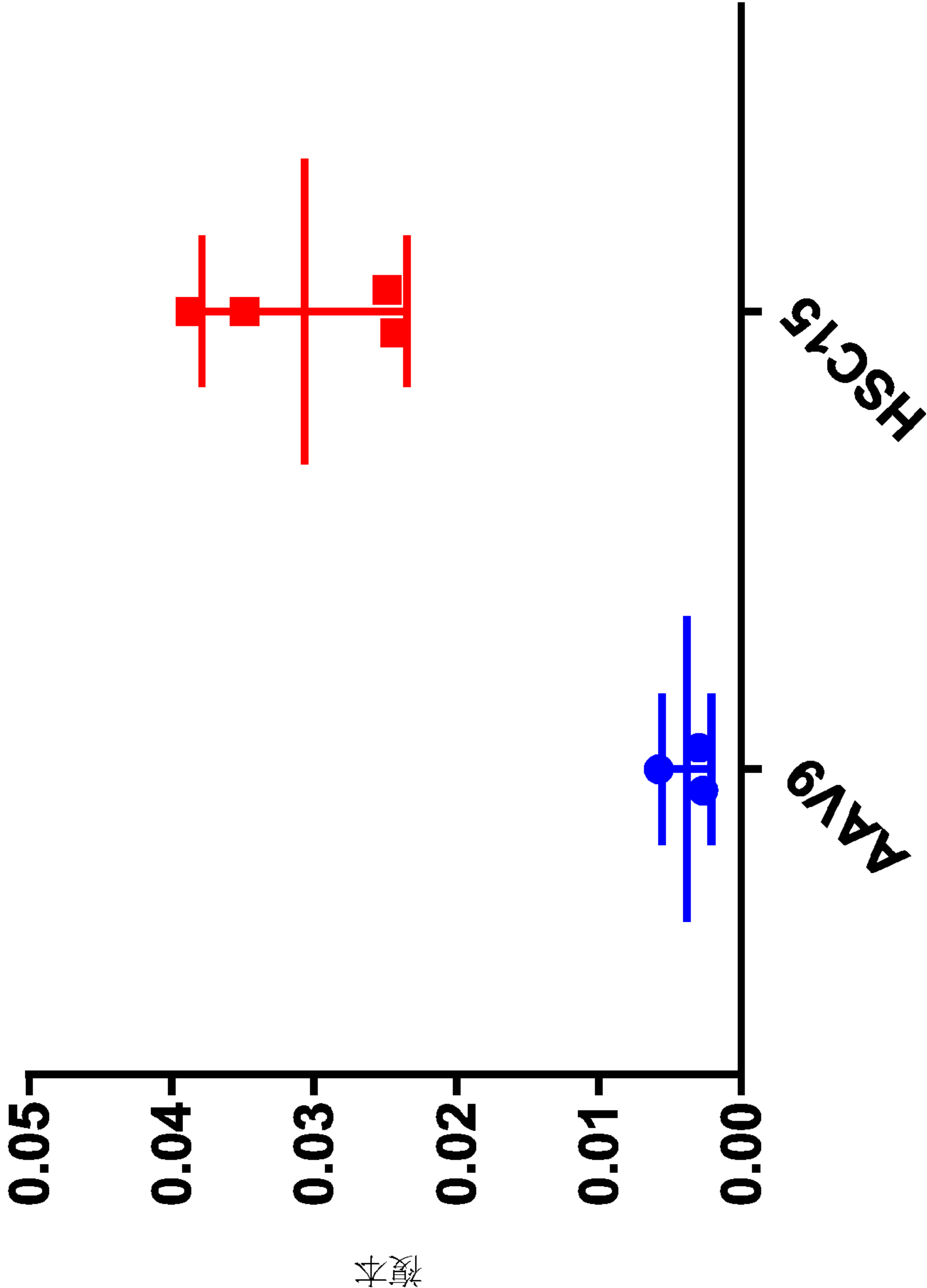
【圖3C】



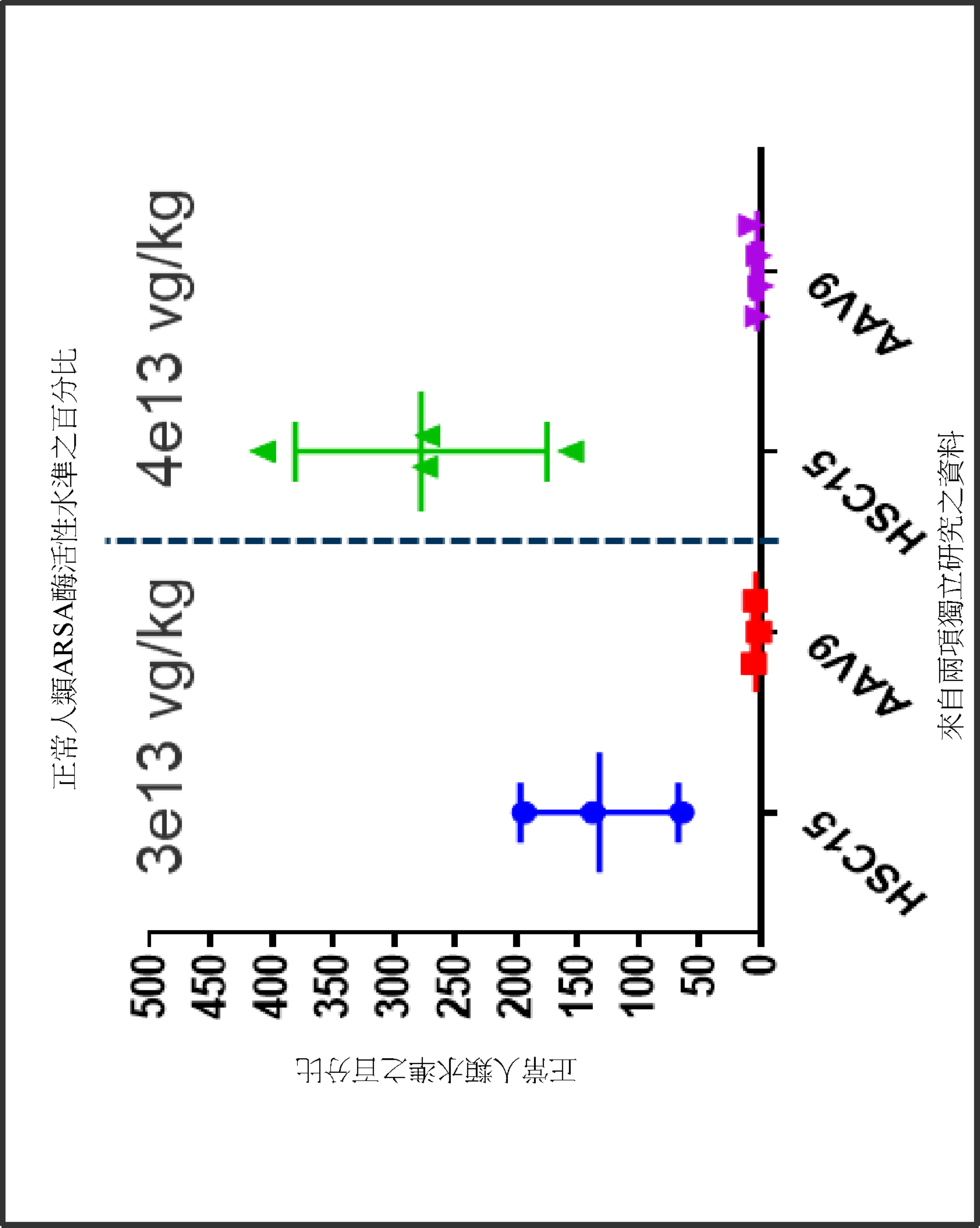
【圖4】

在2e13 vg/kg劑量下之以每個細胞計的VG

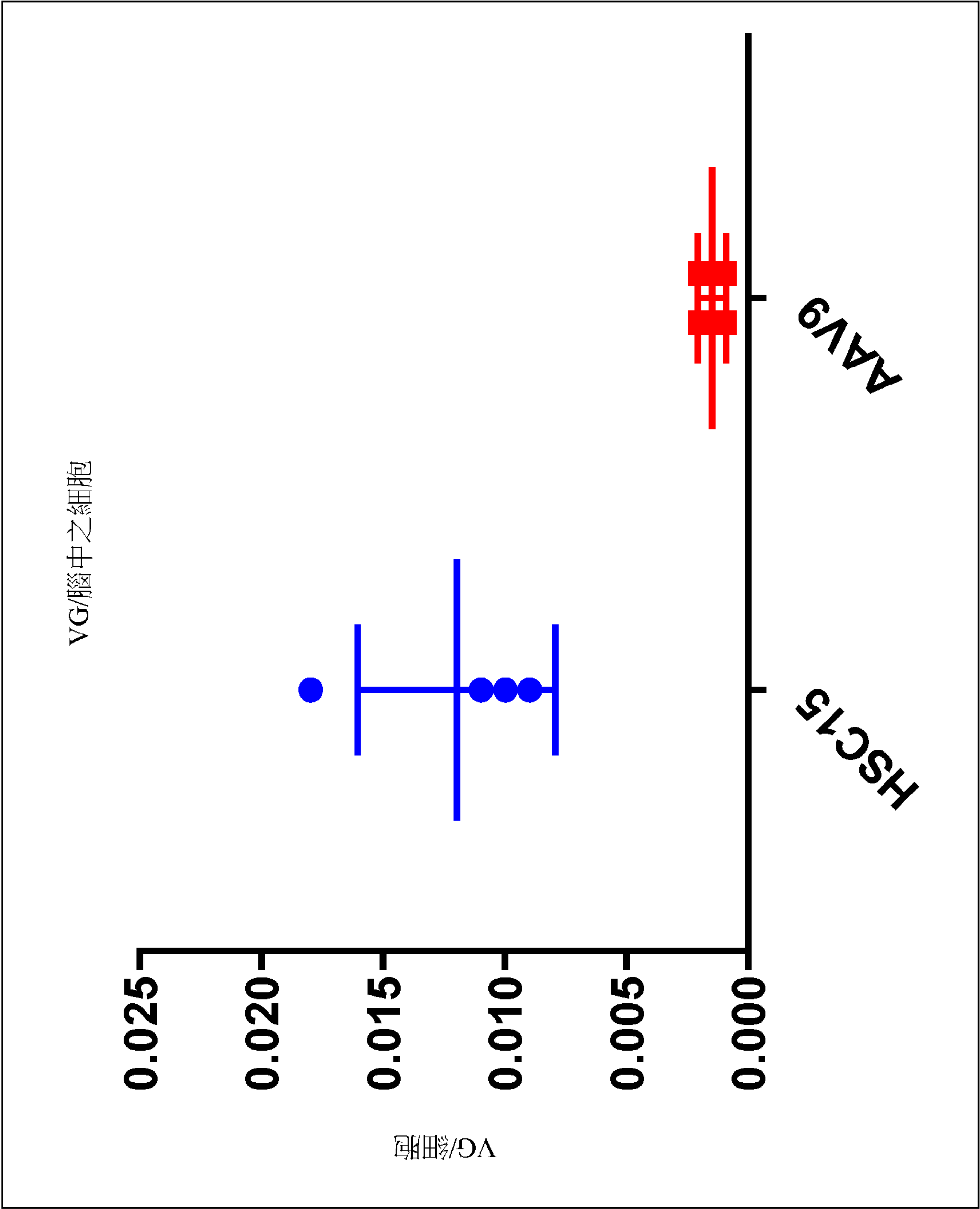
HSC15-CBA-CO-hARSA



【圖5】



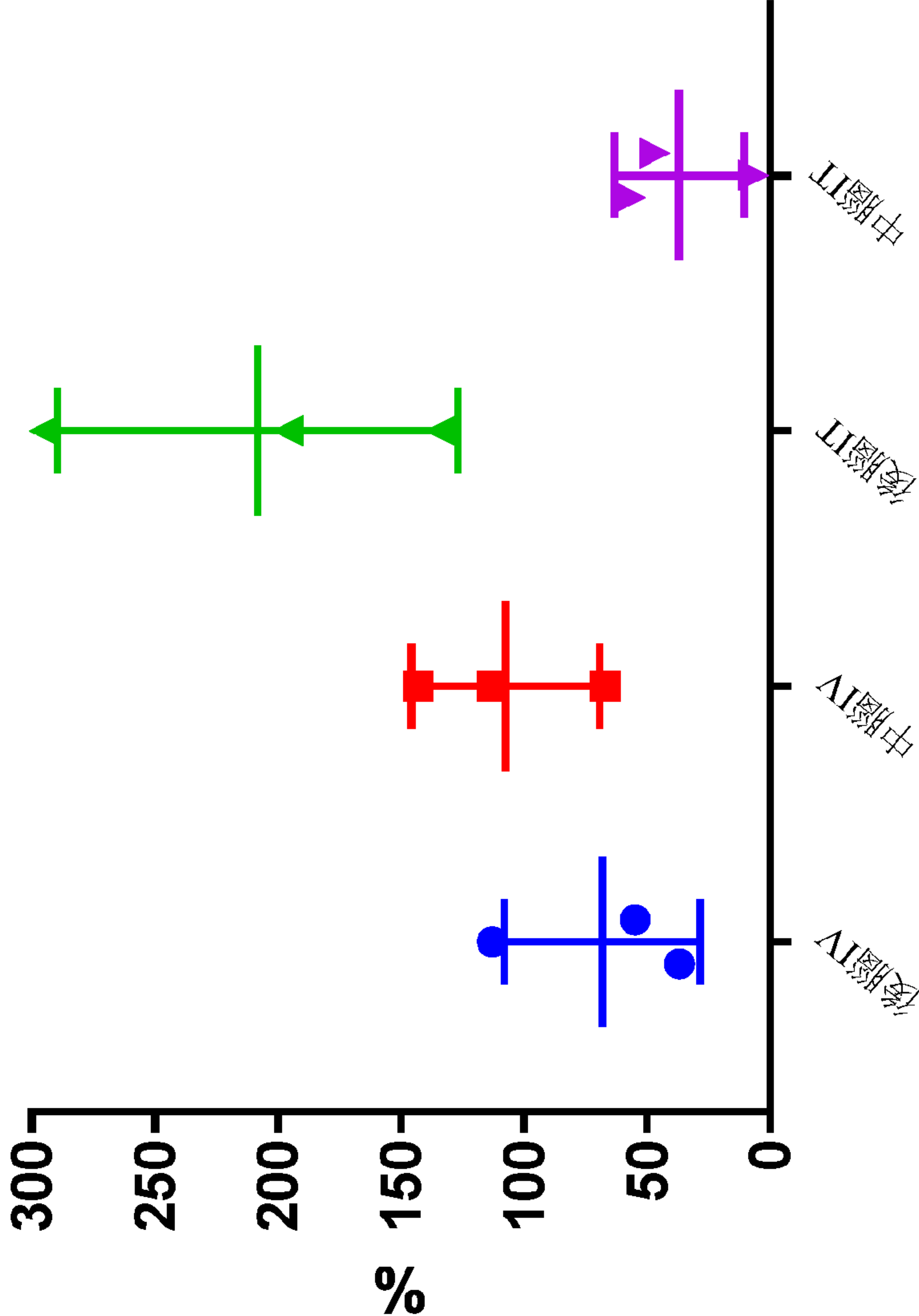
【圖6】



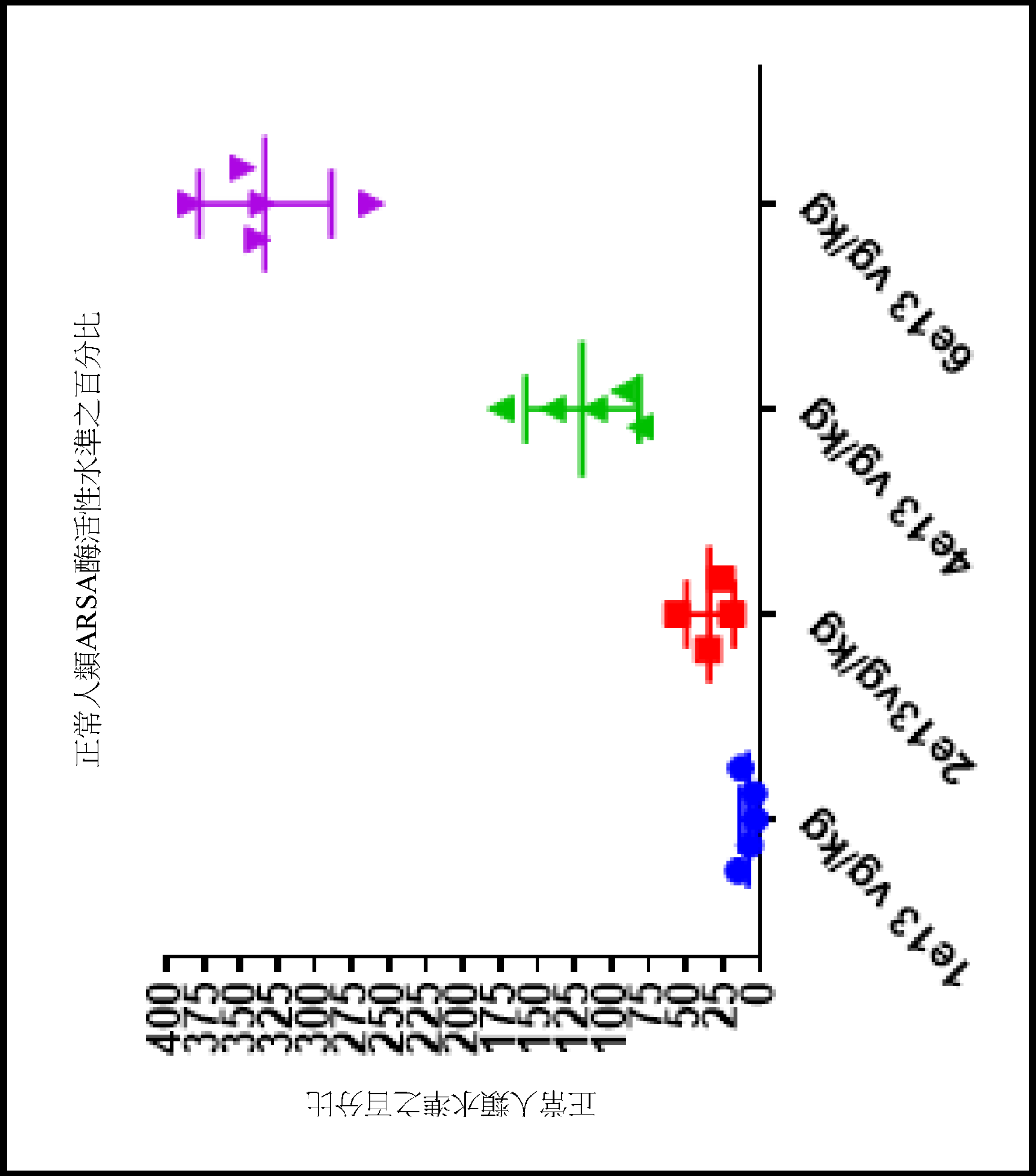
【圖7】

正常人類ARSA酶活性之百分比

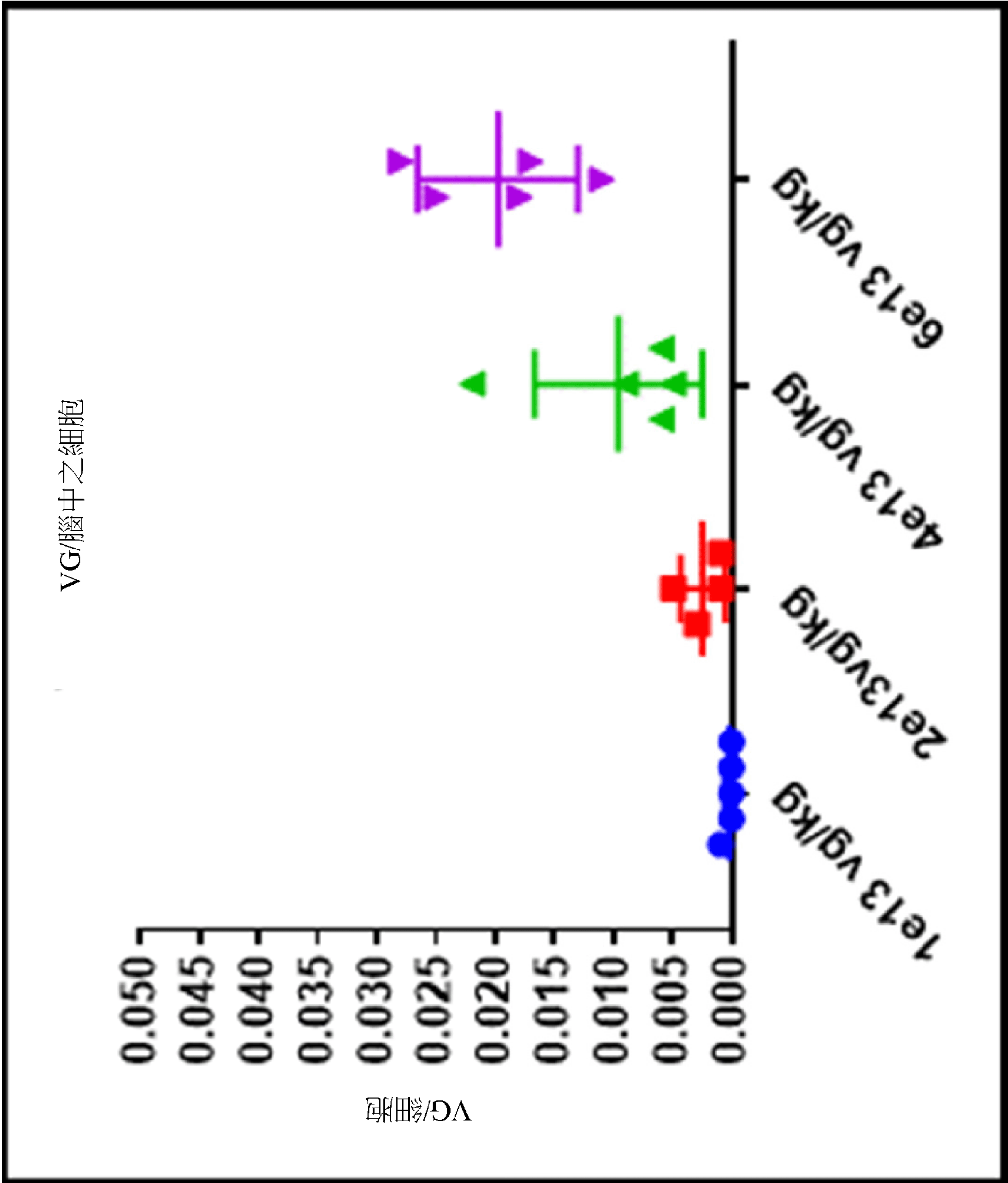
HSC15-CO-hARSA



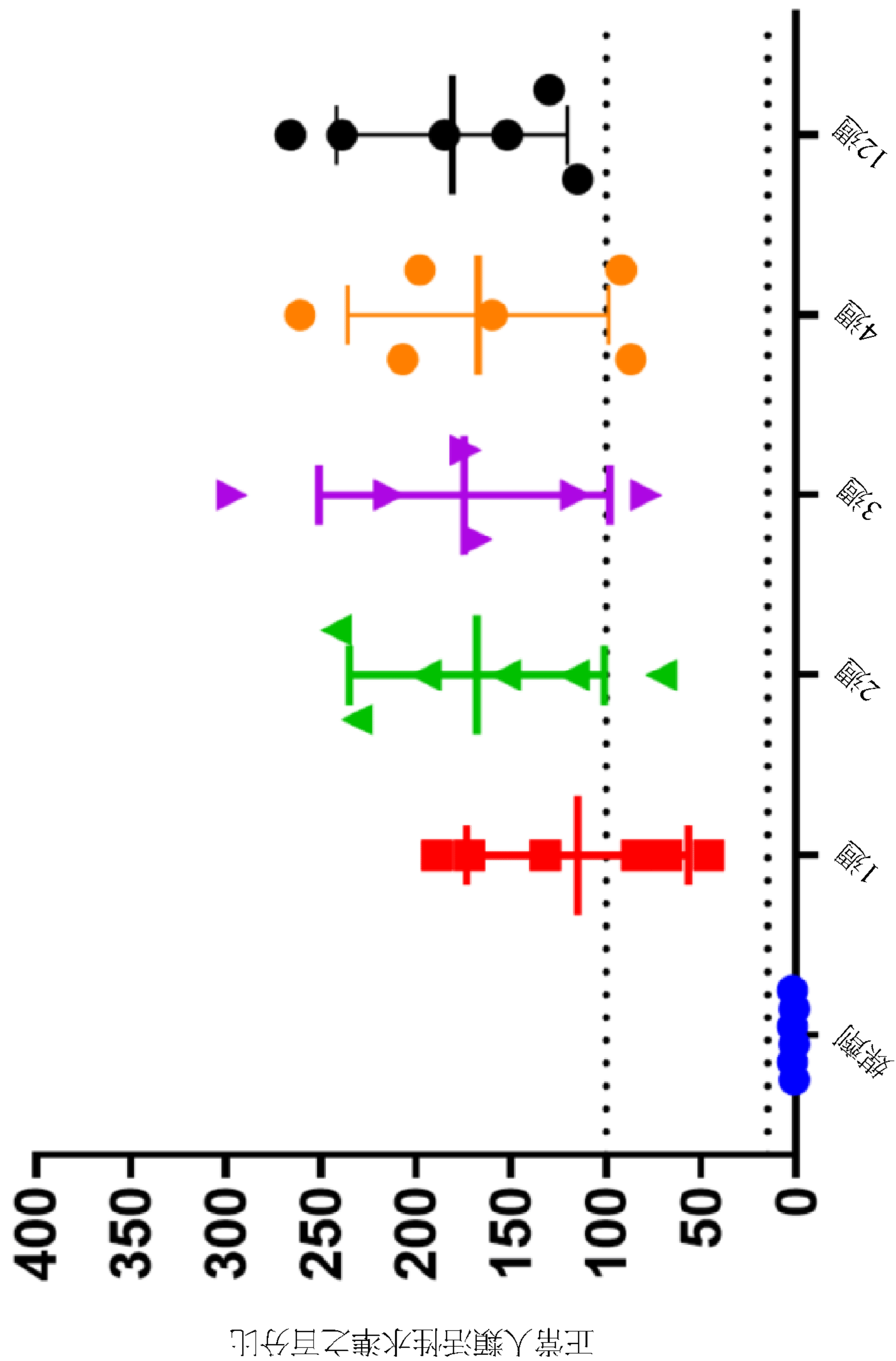
【圖8】



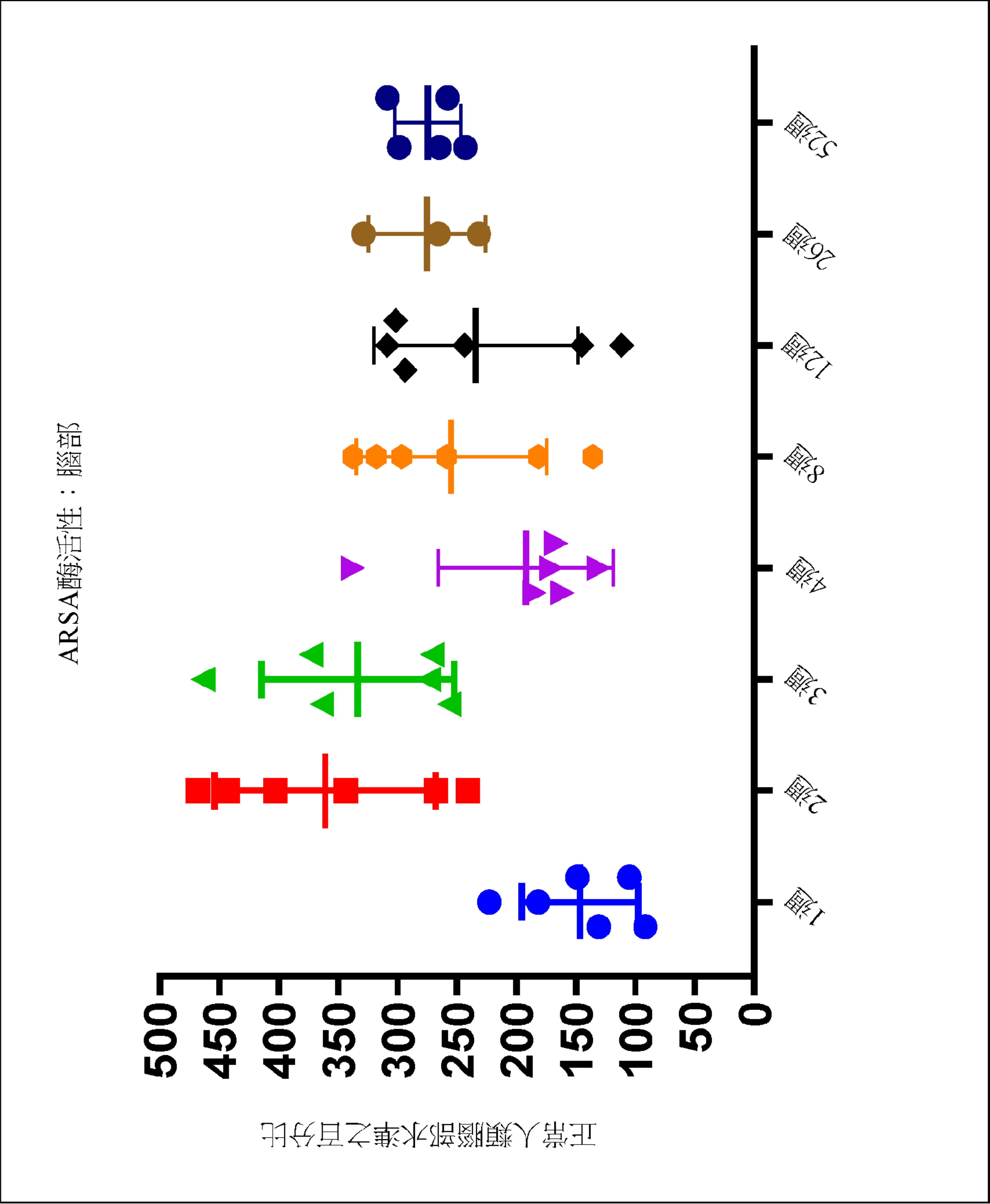
【圖9A】



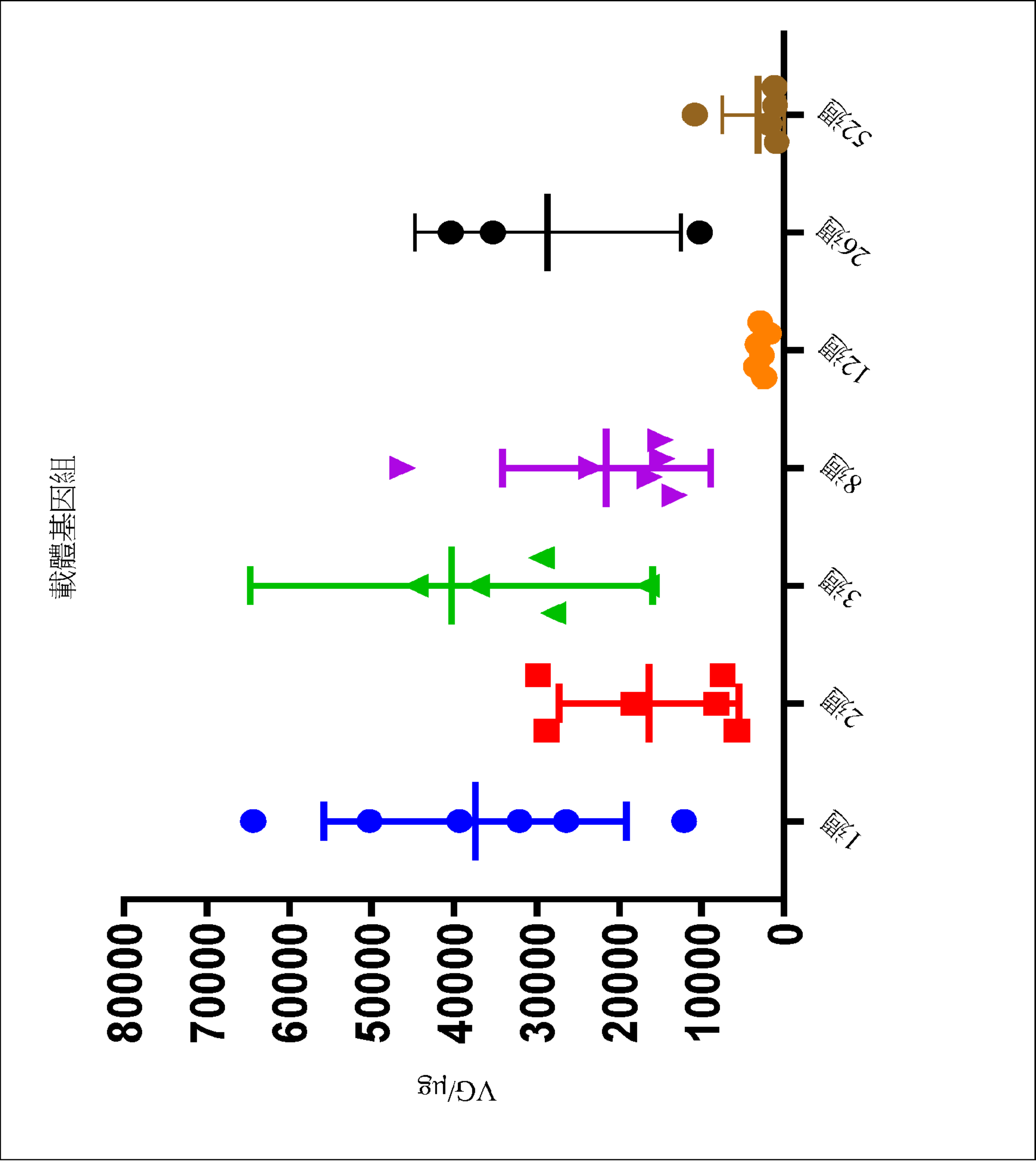
【圖9B】



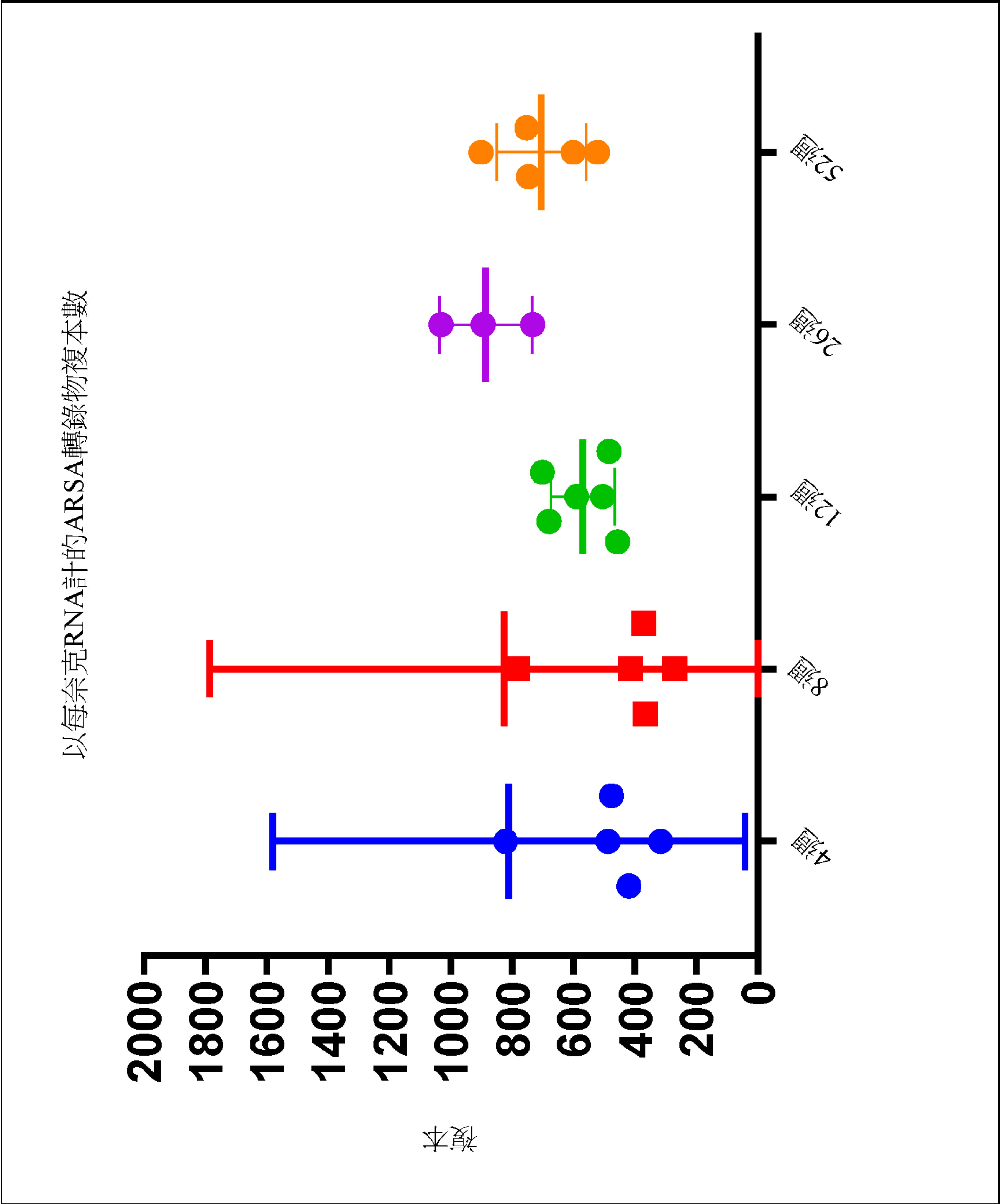
96 回



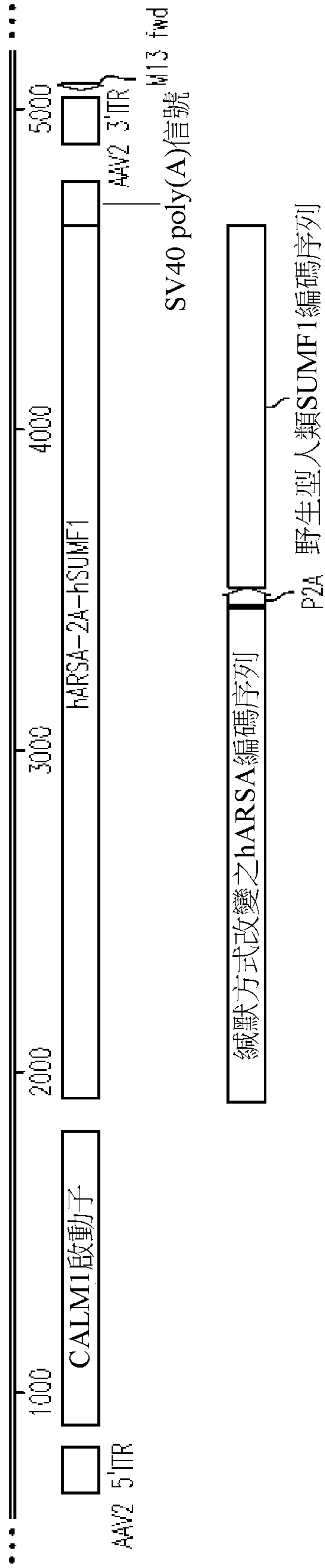
【圖9D】



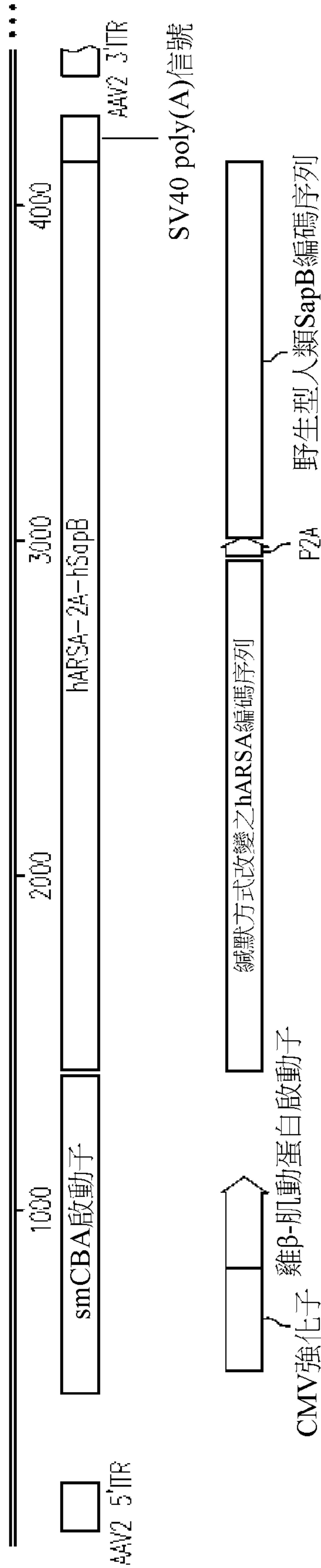
【圖9E】



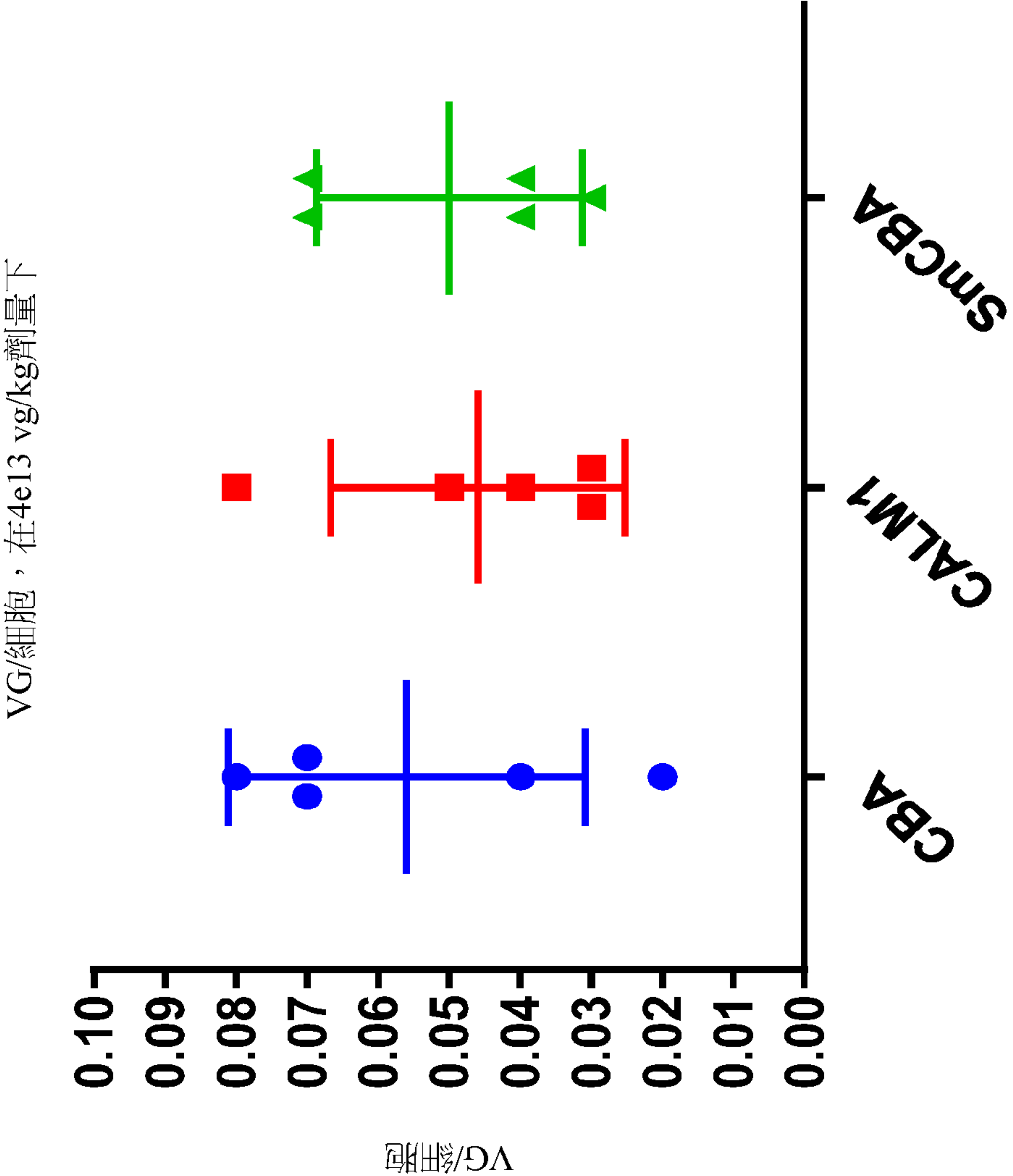
【圖9F】



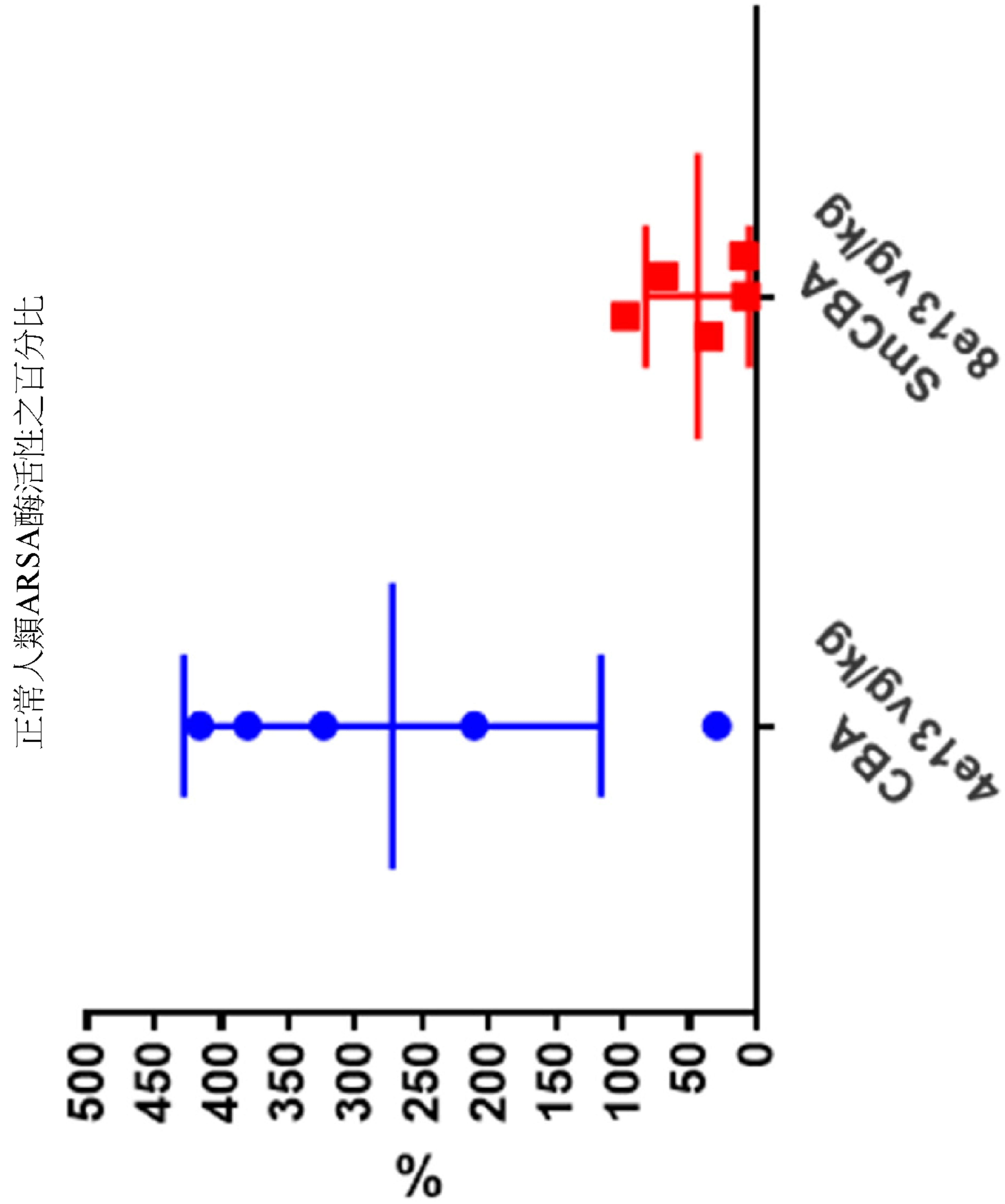
【圖10A】



【圖10B】

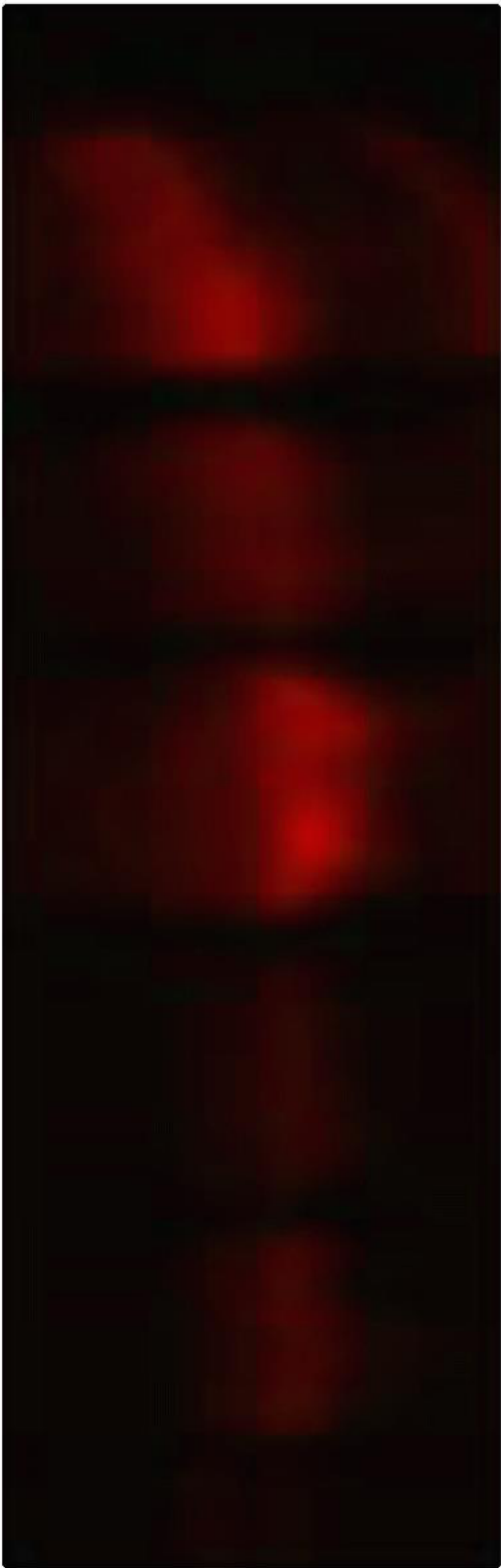


【圖11】



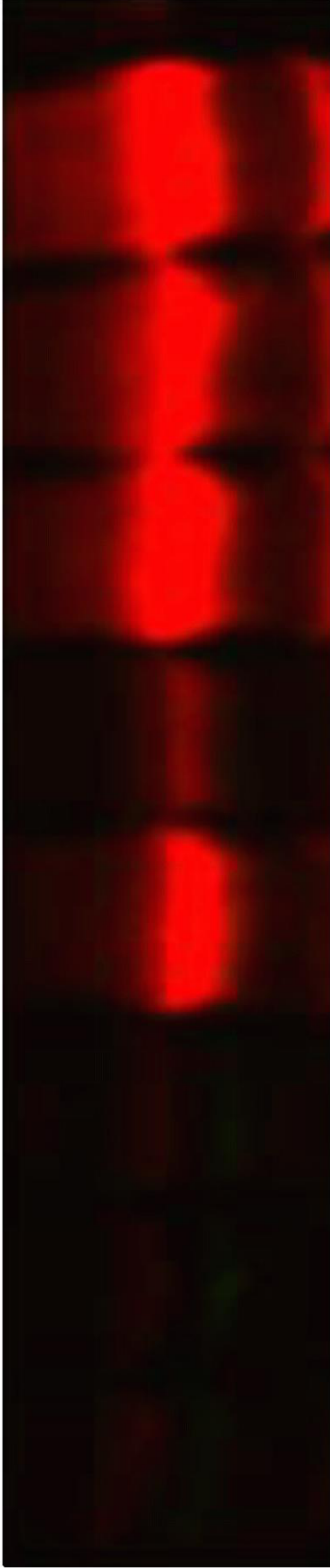
【圖12】

SmCBA - 8e13vg/kg



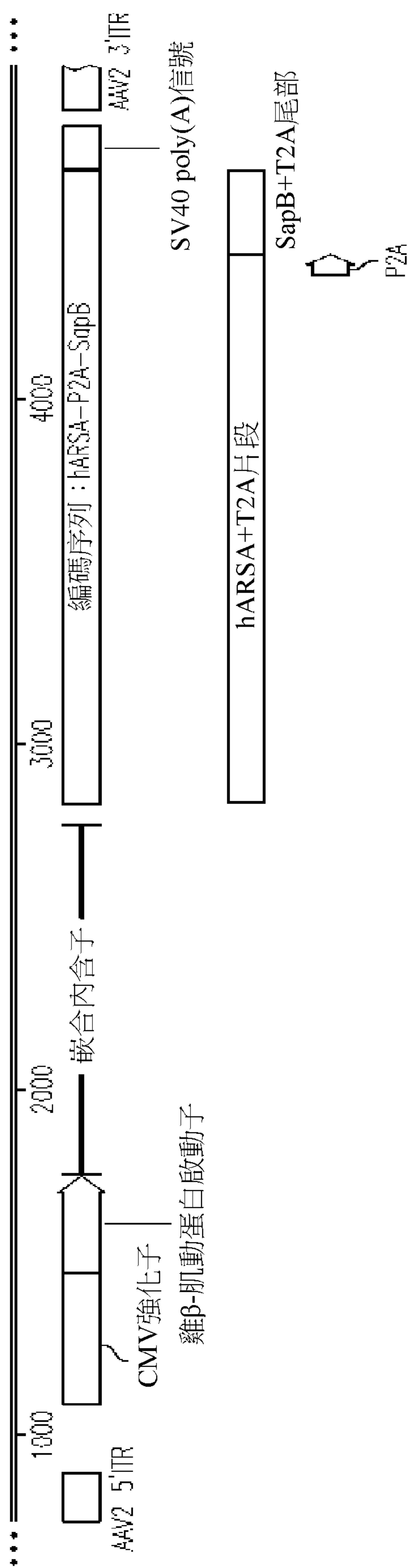
抗hARSA

CBA - 4e13 vg/kg

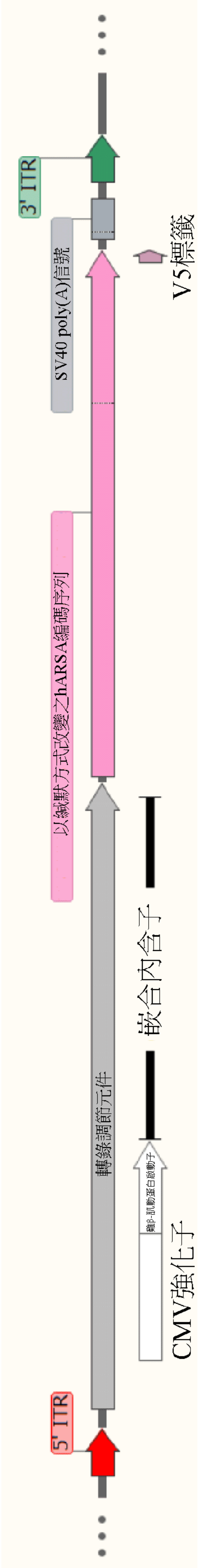


媒劑

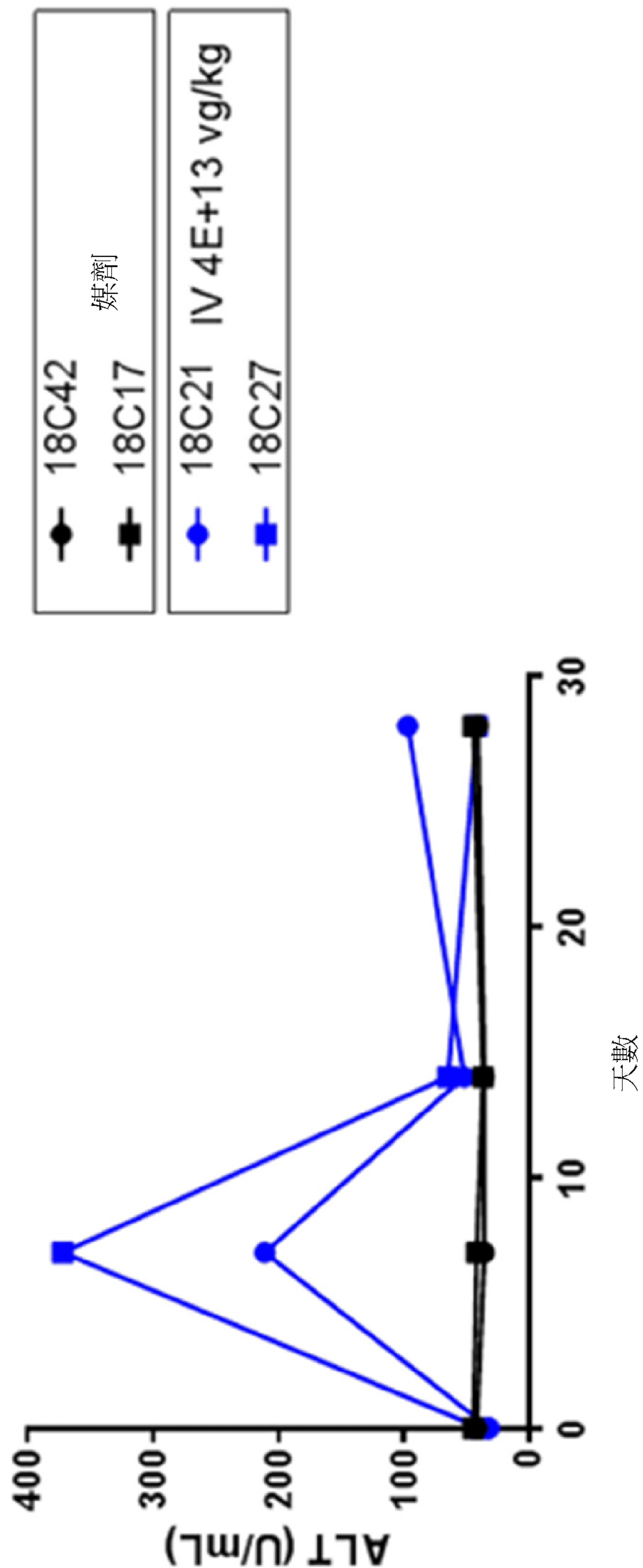
【圖13】



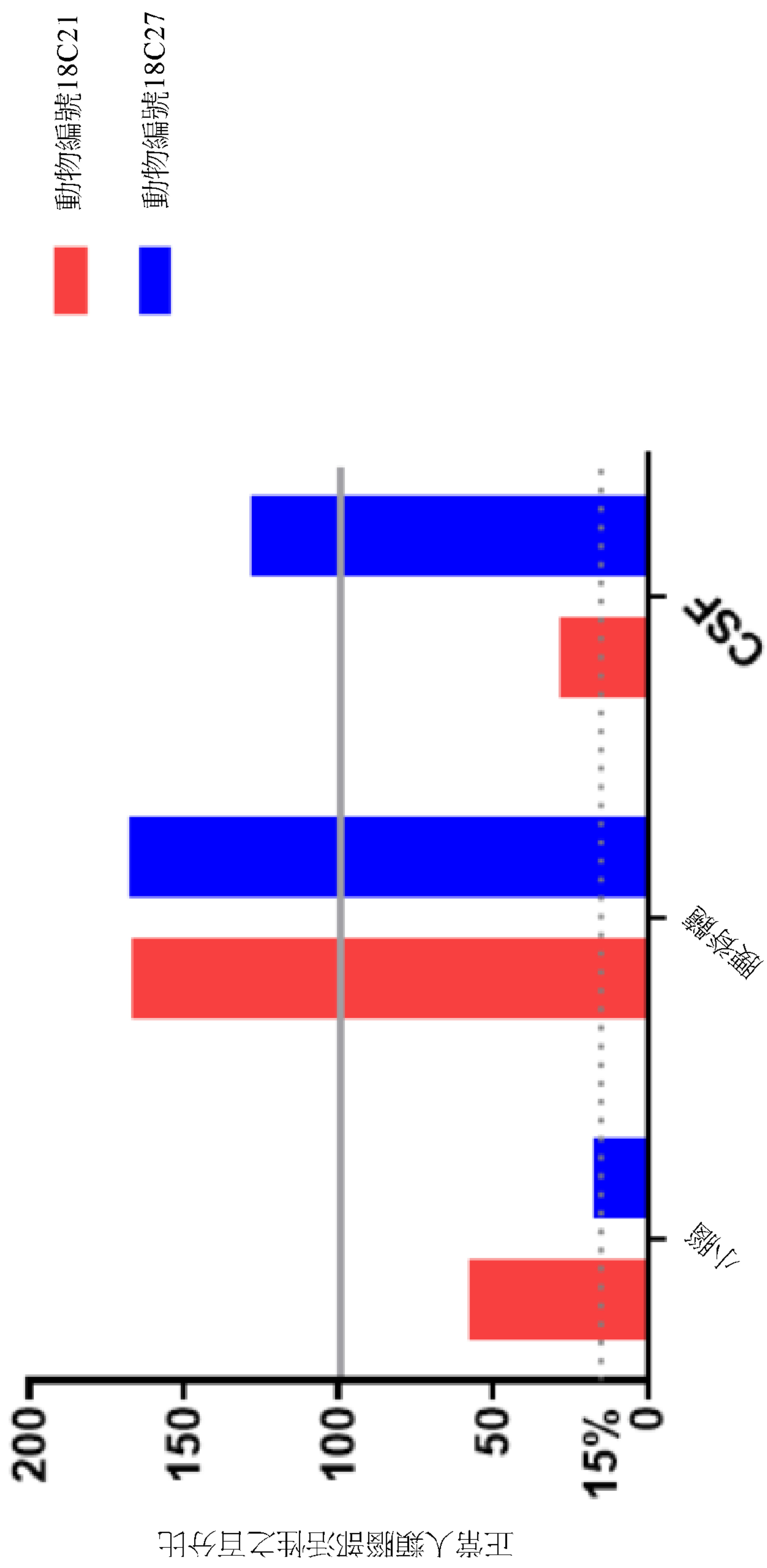
【圖14】



【圖15】



【圖16】



【圖17】