



(21)申請案號：109124636

(22)申請日：中華民國 109 (2020) 年 07 月 21 日

(51)Int. Cl. : **C07K14/005 (2006.01)**

C12N7/00 (2006.01)

C12N15/86 (2006.01)

A61K48/00 (2006.01)

(30)優先權：2019/07/23 美國

62/877,508

2019/07/23 美國

62/877,516

2019/07/23 美國

62/877,524

2019/07/23 美國

62/877,532

2019/07/23 美國

62/877,540

2019/07/23 美國

62/877,551

2019/07/23 美國

62/877,561

2019/07/23 美國

62/877,577

(71)申請人：美商健生生物科技公司 (美國) JANSSEN BIOTECH, INC. (US)

美國

(72)發明人：佩利 威廉 勞埃德 三世 PERRY, WILLIAM LLOYD, III (US) ; 湯寇維茲 布

萊恩 TOMKOWICZ, BRIAN (US)

(74)代理人：陳長文

申請實體審查：無 申請專利範圍項數：106 項 圖式數：9 共 215 頁

(54)名稱

用於生物製造之合成遺傳因子

(57)摘要

描述用於腺相關病毒(Adeno-Associated Viruse, AAV)之經改良生產的重組建構體、細胞、及手段。亦描述使用該建構體及細胞以產生重組 AAV 的方法。

Recombinant constructs, cells and means for improved production of Adeno-Associated Viruses (AAVs) are described. Also described are methods of using the constructs and cells to produce recombinant AAVs.

指定代表圖：

得分	預期	方法	識別	陽性	間隙
697位元(1798)	0.0	組成基質調整	341/531(64%)	415/531(78%)	2/531(0%)

Query	1	MELKNIVNSYNITNILGYLRRSRQDMEREKRTGEDTLTEQKELMNKILTAIEIPYELKME	60
		MELK+IVNSYN+T I+GYLRRSRQD-EREKRTGEDTLTEQKELMNKILT IEIPYE + E	
Sbjct	1	MELKHIVNSYNVTKIIGYLRRSRQDVEREKRTGEDTLTEQKELMNKILTGIEIPYETRTE	60
Query	61	IGSGESIDGRPVFKECLKDLEEGKYQAIHAVKEITRLSRGSYS DAGQIVNLLQSKRLIIIT	120
		IGSGESI+GRPVFK CL DL GK+QAIHAVKEITRLSRGSYS DAG+IVNLL KR+IIIT	
Sbjct	61	IGSGESIEGRPVFKSCLADLRSGKFQAIHAVKEITRLSRGSYS DAGEIVNLLNEKRIIIT	120
Query	121	PYKVYDPRNPVDMRQIRFELFMAREEFEMTRERMTGAKYTYAAQGKWISGLAPYGYQLNK	180
		PYK-YDPRNPVD RQIRFELFMAREEFEMTRERM GAK+TYAAQGKWISGLAP+GY+LNK	
Sbjct	121	PYK-YDPRNPVDARQIRFELFMAREEFEMTRERMNGAKFTYAAQGKWISGLAPFGYKLNK	180
Query	181	KTSKLDPVEDEAKVVQLIFNIFLNGLNGKDYSYTAIAASHLTNLQIPTPSGKKRWNQYTIK	240
		+TS-L+P +++ VV+LIF+IFLNGL+GKD SYTAIA+HL+ LQ TP G KRW++ T++	
Sbjct	181	RTSRLEPSDEDKVVVKLIFDIFLNGLDGKDLSYTAIATHLSKLQFTTPRGGKRWSKDTVR	240
Query	241	AILQNEVYIGTVKYKVREKTKDGKRTIRPEKEQIVVQDAHAPIIDKEQFQQSQVKIANKV	300
		ILQNE Y+G V+YK RE TKDGK+ RPE E IVV DAH PII+KE F+ Q KI NKV	
Sbjct	241	KILQNEAYMGRVRYKARETTKDGKKVFRPESEHIVVDDAHEPIINKEDFEAVQEIKKNKV	300
Query	301	PLLPNKDEFELSELAGVCTCSKCGEPLSKYESKRIRKNKDGTSVYHVKSLTCKKNKCTY	360
		PLLP +E +ELAG+C CS CG+ L K+ES+ RKNKDG T S +HVK L CK NKCT	
Sbjct	301	PLLPVVTSYEPNELAGICVCSVCGKSLQKFESSEYNRKNKDGTS SYFHVKLLICKINKCTS	360
Query	361	VRYNDVENAILDYLSSLNDLNDSTLTKHINSMLSKYEDDNSNMKTKQMSEHLSQKEKEL	420
		VRY VE AIL+YL L L ++ L I + E +NS KT +QM +QK+KEL	
Sbjct	361	VRYEYVEEAILEYLEQLIALENNKLKATIEKSMEAAETNNSE-KTSEQMLVQANQKQKEL	419
Query	421	KNKENFIFDKYESGIYSDELFLKRKAALDEEFKELQNAKNELNGLQDTQSEIDSNTVRNN	480
		+NK FIF+K+ESGIY+DE+FL+RKAA+++E +++ K EL+ + + E D N R N	
Sbjct	420	ENKLT FIFEFK FESC IYTDEMFLQRKAAIEKEVADIKKLKQELSMTFEVK EKD VNEFRVN	478
Query	481	INKIIDQYHIESSSEKKNELLRMVLKDVIVNMTQKRKGPIPAQFEITPILR	531
		I+ -- Y S KNE LR + +++ MT+KR+GPIPA+F I P+LR	
Sbjct	479	ISDVVKFYKESKSRGLKNEKLRSIFDFIVLEMT EKRRGPIPAKFNIYPVLR	529

【圖1】



202116794

【發明摘要】

【中文發明名稱】

用於生物製造之合成遺傳因子

【英文發明名稱】

SYNTHETIC GENETIC ELEMENTS FOR BIOMANUFACTURE

【中文】

描述用於腺相關病毒(Adeno-Associated Viruse, AAV)之經改良生產的重組建構體、細胞、及手段。亦描述使用該建構體及細胞以產生重組AAV的方法。

【英文】

Recombinant constructs, cells and means for improved production of Adeno-Associated Viruses (AAVs) are described. Also described are methods of using the constructs and cells to produce recombinant AAVs.

【指定代表圖】

圖1

【代表圖之符號簡單說明】

無

【發明說明書】

【中文發明名稱】

用於生物製造之合成遺傳因子

【英文發明名稱】

SYNTHETIC GENETIC ELEMENTS FOR BIOMANUFACTURE

【技術領域】

【先前技術】

【0001】 腺相關病毒(adeno-associated virus, AAV)具有線性單股DNA (ssDNA)基因體，其在末端處具有兩個反向末端重複(inverted terminal repeat, ITR)。ITR側接兩個病毒基因-*rep*（複製）及*cap*（殼體），其分別編碼非結構及結構蛋白。該*rep*基因透過使用兩種啟動子及交替剪接編碼四種調節蛋白Rep78、Rep68、Rep52、及Rep40。更具體而言，Rep78及Rep68係轉錄自P5啟動子且Rep 40及Rep52係轉錄自P19啟動子（其嵌入於Rep78及Rep68讀框內）。P5及P19啟動子係由腺病毒E1A基因活化且在細胞中係活性的，諸如使用該等腺病毒E1基因轉化的HEK293。這些Rep蛋白係關於AAV基因體複製。該*cap*基因（透過轉譯之交替剪接及起始）產生三種殼體蛋白VP1（病毒蛋白1）、VP2及VP3，其等會組裝成病毒的近球形蛋白殼。該AAV病毒不編碼聚合酶，因此仰賴用於基因體複製的細胞聚合酶。

【0002】 若該AAV *rep*及*cap*基因可穩定地整合或維持在細胞中，則哺乳動物細胞中之AAV的大規模生產可係可能的，且之後經誘導以在高密度培養物中生產AAV。然而，Rep蛋白的表現對於宿主細胞可係細胞毒性的或細胞生長抑制性的(cytostatic)，使其難以在表現*rep*基因的宿主

中培養穩定的細胞系，諸如表現該等腺病毒E1基因（諸如HEK293細胞）者。因為AAV編碼四種具有重疊讀框的Rep蛋白，這是起因於使用兩種啟動子及交替剪接，因此使用誘導型啟動子控制*rep*基因表現並不容易。

【0003】該等四種Rep蛋白之細胞毒性或細胞生長抑制性本質已防止穩定細胞系的發展，該等細胞系可使用原生*rep/cap*啟動子生產高效價的AAV (Clark *et al.* (1995) Hum. Gene Ther. 6:1329-1341; Chadeuf *et al.* (2000) J. Gene med. 2:260-268)。數個群組已試圖重組地調節Rep表現。Yang將P5啟動子用小鼠金屬硫蛋白啟動子置換。雖然在HEK293中的穩定植株展現出金屬誘導的*rep78*表現，但*rep50*及*rep42*表現（由內部P19啟動子驅動）僅在低位準下偵測到，並且細胞的生長速率實質上降低 (Yang *et al.* (1994) J. Virol 68: 4847-4856)。Ogasawara用含有位於loxP兩側之填充片段之普遍存在的啟動子置換P5啟動子，其可由Cre重組酶活化。*rep52*、*rep40*、或*cap*基因均未在被腺病毒-Cre感染的穩定植株中誘發，其表示組成型*rep52/rep40*表現亦對細胞有害 (Ogasawara *et al.* (1999) J. Gen. Virol. 80: 2477-2480)。

【0004】Xiao及同事們描述了另一種調控*rep*表現的方法 (Qiao *et al.* (2002) J. Virol. 76: 13015-13027; Yuan *et al.* (2011) Hum. Gene Ther. 22:613-624)。Xiao將人工內含子插入至所有四種Rep蛋白共享的編碼區之該*rep*基因中，並將含有單獨的poly(A)序列或與*puro*（嘌呤黴素抗性基因）組合的loxP側接止擋匣插入至該內含子中。抑制所有該等Rep蛋白的表現，使HEK293細胞中的穩定細胞系得以產生。透過重組loxP位點將Cre重組酶遞送（藉由腺病毒感染）至細胞中來切除該止擋匣，使得全長pre-mRNA得以轉錄。然後藉由RNA剪接來精確地移除剩餘的內含子序

列，恢復所有四種Rep蛋白之編碼序列，並因此起始來自經整合之ITR側接轉殖基因的AAV之生產。然而，由於Cre重組酶識別兩個相同的loxP位點，故該等loxP位點在重組後維持相同，由於Cre催化連接及切除反應兩者，因此可能進行其他重組。

【0005】 AAV *rep*基因僅在亦表現該腺病毒E1（早期區域1）基因的細胞中表現。已在不表現該等E1基因（包括HeLa）的宿主中建構數種穩定*rep/cap*細胞系(Clark *et al* (1995) Hum. Gen. Therap. 6: 1329-1341; Yang *et al.* (1994) J. Virol. 68: 4847-4856; Gao *et al.* (1998) Hum. Gen. Ther. 9: 2353-2362)、A549 (Gao *et al.* (2002) Mol Ther. 5:644-659)、及Vero (Beal *et al.* (2007)第10屆美國基因治療年會，Seattle, WA，2007年5月30日至6月3日)。這些細胞系的最大缺點在於AAV生產需要完整的E1（且通常具有複製能力）腺病毒，其可能會由於AAV病毒製劑的污染物而增加安全性風險。

【0006】 已描述了使用數種不同的病毒來提供輔助功能並將該重組轉殖基因及/或AAV基因傳遞至人類細胞之AAV生產系統，包括Herpes (Thomas *et al.* (2009) Hum Gene Ther. 20:861-70; Clement *et al.* (2009) Hum Gene Ther. 20:796-806)、牛痘病毒(Wang *et al.* (2017) Mol. Ther. Methods Clin Devel. 7: 146-155.)、及腺病毒(Fisher *et al.* (1996) Hum gene Ther. 7: 2079-2087; Gao *et al.* (1998) Hum Gene Ther. 2353-2362. Liu *et al.* (1999) Gene Ther 6: 293-299)。這些方法需要生產數種不同的病毒（且在一些情況下為重組宿主細胞系）。AAV亦已生產於使用桿狀病毒之昆蟲細胞中(Mietzsch *et al.* (2014) Hum Gene Ther. 25:212-22; Aslanidi *et al.* (2009) Proc Natl Acad Sci USA. 106:5059-5064;

Cecchini *et al.* (2011) Hum Gene Ther. 22:1021-1030)。昆蟲細胞與人細胞中產生的AAV在功能上是否等效仍是一個懸而未決的問題。

【0007】 對於改良具有重組建構體及細胞的AAV之生產仍有需求。

【發明內容】

【0008】 在一個態樣中，本文提供一種非天然存在的核酸分子，其包含經修飾腺相關病毒(AAV) *rep*基因，該基因具有編碼四種Rep蛋白Rep78、Rep68、Rep52、及Rep40之AAV *rep*基因及插入至由該等四種Rep蛋白共享之該*rep*基因的編碼序列中的人工內含子，其中該人工內含子包含止擋匣(stop cassette)，該止擋匣係插入在該人工內含子之5'剪接位點下游及分支位點上游，且該止擋匣以5'至3'的順序包含：(a)具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點，較佳的是具有SEQ ID NO:7之核苷酸序列的*attP*位點；(b)剪接受體；(c)終止子；及(d)具有與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attB*位點，較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點。

【0009】 在一個實施例中，該剪接受體包含SEQ ID NO:17之核苷酸序列。

【0010】 在一個實施例中，該終止子包含多腺苷酸化信號。在一個實施例中，該終止子進一步包含SEQ ID NO: 19之核苷酸序列。

【0011】 在一個實施例中，該止擋匣包含編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶(neomycin phosphotransferase)表現匣。

【0012】 在一個實施例中，該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:15之核苷酸序列。

【0013】 在一個實施例中，該AAV *rep*基因包含AAV1至AAV8中之一者之*rep*基因、或其混合。在一個實施例中，該AAV *rep*基因包含人類AAV2之該*rep*基因，該基因具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至2202。在一個實施例中，該人工內含子係插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號996至1905之間。在一個實施例中，該人工內含子係直接插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號1052、1061、1712、1906、1022、1112、1475、1514、1700、1742、1784或1340、較佳的是核苷酸編號1052的下游。

【0014】 在一個態樣中，本文提供一種非天然存在的核酸分子，其包含經修飾AAV *rep*基因，該基因以5'至3'的順序包含：(a)具有SEQ ID NO:55之核苷酸序列的AAV *rep*基因之5'部分；(b)人工內含子，其以5'至3'的順序包含：(i)具有SEQ ID NO:14之核苷酸序列的5'內含子片段；(ii)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)具有SEQ ID NO:17之核苷酸序列的剪接受體；(3)具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)具有SEQ ID NO:19之核苷酸序列的終止子；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(iii)具有SEQ ID NO:15之核苷酸序列的3'內含子片段；及(c)具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。

【0015】 在一個態樣中，本文提供一種非天然存在的核酸分子，其

包含經修飾AAV *rep*基因，該基因以5'至3'的順序包含：(a)具有SEQ ID NO:73之核苷酸序列的AAV *rep*基因之5'部分；(b)人工內含子，其以5'至3'的順序包含：(i)具有SEQ ID NO:14之核苷酸序列的5'內含子片段；(ii)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)具有SEQ ID NO:17之核苷酸序列的剪接受體；(3)具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)具有SEQ ID NO:19之核苷酸序列的終止子；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(iii)具有SEQ ID NO:66之核苷酸序列的3'內含子片段；及(c)具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。在一個實施例中，該止擋匣包含SEQ ID NO: 16之核苷酸序列。

【0016】 在一個實施例中，該非天然存在的核酸分子進一步包括編碼三種殼體蛋白VP1、VP2及VP3的AAV *cap*基因。在一個實施例中，該AAV *cap*基因包含AAV1至AAV9及AAVDJ中之一者、或其混合之*cap*基因。在一個實施例中，該AAV *cap*基因包含具有GenBank存取號AY530579.1之核苷酸序列的人類AAV9之*cap*基因。在一個實施例中，該AAV *cap*基因進一步包含多腺苷酸化信號（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號4411至4466之AAV2的多腺苷酸化信號）及強化子（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至313之AV2 *rep* P5啟動子），其中該多腺苷酸化信號及該強化子兩者皆係該*cap*基因之該編碼序列的下游。在一個實施例中，該非天然存在的核酸分子進一步包括由該AAV *cap*基因下游的一對AAV反向末端重複(ITR)側接的轉殖基因。

【0017】 在一個實施例中，該非天然存在的核酸分子又進一步包括

該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：(a)具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；(b)具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；(c)具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；(d)具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因子06；(e)具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07；(f)具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12；(g)具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13；(h)具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35；(i)具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36；(j)具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因子52；(k)具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53；及(l)來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子，更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及SEQ ID NO:25之核苷酸序列。在一個實施例中，該非天然存在的核酸分子包含該經修飾AAV *rep*基因上游之該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第一間隔序列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自由下列所組成之群組：(a) SEQ ID NO:67之核苷酸序列；及(b) SEQ ID NO:68之核苷酸序列。在一個實施例中，該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；較佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多

腺苷酸化信號具有SEQ ID NO:23之核苷酸序列。

【0018】 在一個態樣中，本文提供一種非天然存在的核酸分子，其以5'至3'的順序包含：(A)第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；(B)經修飾AAV *rep*基因，其以5'至3'的順序包含：(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:55之核苷酸序列；(ii)人工內含子，其以5'至3'的順序包含：(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；(b)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；(3)編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:15之核苷酸序列；(iii)該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；(C) AAV *cap*基因，較佳地該AAV *cap*基因包含SEQ ID NO:57之核苷酸序列；(D)由一對AAV ITR側接之轉殖基因，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及(E)第二絕緣子，較佳地該第二絕緣子具有SEQ ID NO:25之核苷酸序列。

【0019】 在一個態樣中，本文提供一種非天然存在的核酸分子，其

以5'至3'的順序包含：(A)第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；(B)經修飾AAV *rep*基因，其以5'至3'的順序包含：

(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:73之核苷酸序列；(ii)人工內含子，其以5'至3'的順序包含：(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；(b)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；(3)編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:66之核苷酸序列；(iii)該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；(C) AAV *cap*基因；(D)由下列側接之轉殖基因：(1)一對AAV ITR，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及(2)一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列。

【0020】 在一個態樣中，本文提供一種載體，其包含以上所述之非天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:12之核苷酸序列。

【0021】 在一個態樣中，本文提供一種載體，其包含以上所述之非

天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:70之核苷酸序列。

【0022】 在一個態樣中，本文提供一種製造以上所述之非天然存在的核酸分子之方法。在具體實施例中，本文提供一種製造載體之方法，該載體包含以上所述之非天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:12之核苷酸序列。在另一實施例中，本文提供一種製造載體之方法，該載體包含以上所述之非天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:70之核苷酸序列。

【0023】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾腺相關病毒(AAV) *rep*基因，該基因具有編碼四種Rep蛋白Rep78、Rep68、Rep52、及Rep40之AAV *rep*基因及插入至由該等四種Rep蛋白共享之該*rep*基因的編碼序列中的人工內含子，其中該人工內含子包含止擋匣，該止擋匣係插入在該人工內含子之5'剪接位點下游及分支位點上游，且該止擋匣以5'至3'的順序包含：(a)具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點，較佳的是具有SEQ ID NO:7之核苷酸序列的*attP*位點；(b)剪接受體；(c)終止子；及(d)具有與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attB*位點，較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點。

【0024】 在一個實施例中，該剪接受體包含SEQ ID NO:17之核苷

酸序列。

【0025】 在一個實施例中，該終止子包含多腺苷酸化信號。在一個實施例中，該終止子進一步包含SEQ ID NO: 19之核苷酸序列。

【0026】 在一個實施例中，該止擋匣包含編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶(neomycin phosphotransferase)表現匣。

【0027】 在一個實施例中，該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:15之核苷酸序列。在另一實施例中，該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:66之核苷酸序列。

【0028】 在一個實施例中，該AAV *rep*基因包含AAV1至AAV8中之一者之*rep*基因、或其混合。在一個實施例中，該AAV *rep*基因包含人類AAV2之該*rep*基因，該基因具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至2202。在一個實施例中，該人工內含子係插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號996至1905之間。在一個實施例中，該人工內含子係直接插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號1052、1061、1712、1906、1022、1112、1475、1514、1700、1742、1784或1340、較佳的是核苷酸編號1052的下游。

【0029】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾AAV *rep*基因，該基因以5'至3'的順序包含：(a)具有SEQ ID NO:55之核苷酸序列的AAV *rep*基因之5'部分；(b)人工內含子，其以5'至3'的順序包含：(i)具有SEQ ID NO:14之核苷酸序

列的5'內含子片段；(ii)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)具有SEQ ID NO:17之核苷酸序列的剪接受體；(3)具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)具有SEQ ID NO:19之核苷酸序列的終止子；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(iii)具有SEQ ID NO:15之核苷酸序列的3'內含子片段；及(c)具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。

【0030】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾AAV *rep*基因，該基因以5'至3'的順序包含：(a)具有SEQ ID NO:73之核苷酸序列的AAV *rep*基因之5'部分；(b)人工內含子，其以5'至3'的順序包含：(i)具有SEQ ID NO:14之核苷酸序列的5'內含子片段；(ii)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)具有SEQ ID NO:17之核苷酸序列的剪接受體；(3)具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)具有SEQ ID NO:19之核苷酸序列的終止子；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(iii)具有SEQ ID NO:66之核苷酸序列的3'內含子片段；及(c)具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。

【0031】 在一個實施例中，該止擋匣包含SEQ ID NO: 16之核苷酸序列。

【0032】 在一個實施例中，以上所述之細胞進一步包括編碼三種殼體蛋白VP1、VP2及VP3的AAV *cap*基因。在一個實施例中，該AAV *cap*基因包含AAV1至AAV9及AAVDJ中之一者、或其混合之*cap*基因。在一

個實施例中，該AAV *cap*基因包含具有GenBank存取號AY530579.1之核苷酸序列的人類AAV9之*cap*基因。在一個實施例中，AAV *cap*基因包含AAV9的混合之*cap*基因。

【0033】 在一個實施例中，該AAV *cap*基因進一步包含多腺苷酸化信號（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號4411至4466之AAV2的多腺苷酸化信號）及強化子（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至313之AAV2 *rep* P5啟動子），其中該多腺苷酸化信號及該強化子兩者皆係該*cap*基因之該編碼序列的下游。

【0034】 在一個實施例中，包含*cap*基因之細胞進一步包括轉殖基因，該轉殖基因由該AAV *cap*基因下游的一對AAV反向末端重複(ITR)側接。在一個實施例中，該細胞進一步包括該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：(a)具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；(b)具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；(c)具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；(d)具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因子06；(e)具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07；(f)具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12；(g)具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13；(h)具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35；(i)具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36；(j)具有GenBank存取號

AY190757.1之核苷酸序列的抗抑制因子52；(k)具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53；及(l)來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子，更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及SEQ ID NO:25之核苷酸序列。在一個實施例中，該細胞包含該經修飾AAV *rep*基因上游之該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第一間隔序列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自由下列所組成之群組：(a) SEQ ID NO:67之核苷酸序列；及(b) SEQ ID NO:68之核苷酸序列。

【0035】 在一個實施例中，該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；較佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列。

【0036】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子以5'至3'的順序包含：(A)第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；(B)經修飾AAV *rep*基因，其以5'至3'的順序包含：(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:55之核苷酸序列；(ii)人工內含子，其以5'至3'的順序包含：(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；(b)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；(3)編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)終止子，較

佳地該終止子具有SEQ ID NO:19之核苷酸序列；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:15之核苷酸序列；(iii)該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；(C) AAV *cap*基因，較佳地該AAV *cap*基因包含SEQ ID NO:57之核苷酸序列；(D)由一對AAV ITR側接之轉殖基因，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及(E)第二絕緣子，較佳地該第二絕緣子具有SEQ ID NO:25之核苷酸序列。

【0037】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子以5'至3'的順序包含：(A)第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；(B)經修飾AAV *rep*基因，其以5'至3'的順序包含：(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:73之核苷酸序列；(ii)人工內含子，其以5'至3'的順序包含：(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；(b)止擋匣，其以5'至3'的順序包含：(1)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(2)剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；(3)編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(4)終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及(5)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(c) 3'內含子片段，較佳地

該3'內含子片段具有SEQ ID NO:66之核苷酸序列；(iii)該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；(C) AAV *cap*基因；及(D)由下列側接之轉殖基因：(i)一對AAV ITR，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及(ii)一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列。

【0038】 在一個實施例中，該非天然存在的核酸分子係游離型，其具有SEQ ID NO:12之核苷酸序列。在另一實施例中，該非天然存在的核酸分子係游離型，其具有SEQ ID NO:70之核苷酸序列。

【0039】 在一個實施例中，該細胞進一步包括編碼重組酶之核酸分子，該重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的胺基酸序列；較佳地，該核酸包含與SEQ ID NO: 3之核苷酸序列至少85%、至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%同一的核苷酸序列；更佳地，該細胞包含重組 Δ E1/ Δ E3腺病毒血清型5 (Ad5) 病毒，該病毒編碼具有SEQ ID NO:2之胺基酸序列的重組酶。

【0040】 在一個實施例中，該細胞進一步包括腺病毒E1A及E1B基因，較佳地該細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6

細胞。

【0041】 在一個態樣中，本文提供一種生產包含轉殖基因之重組 AAV 的方法，其包含：(A)獲得第一宿主細胞，其包含：(i)經修飾 AAV *rep* 基因，該基因以 5' 至 3' 的順序包含：(a) AAV *rep* 基因之 5' 部分，較佳地該 AAV *rep* 基因具有 SEQ ID NO:55 之核苷酸序列；(b)人工內含子，其以 5' 至 3' 的順序包含：(1) 5' 內含子片段，較佳地該 5' 內含子片段具有 SEQ ID NO:14 之核苷酸序列；(2)止擋匣，其以 5' 至 3' 的順序包含：(aa)具有 SEQ ID NO:7 之核苷酸序列的 *attP* 位點；(bb)剪接受體，較佳地該剪接受體具有 SEQ ID NO:17 之核苷酸序列；(cc)編碼可選擇標記之基因，較佳的是具有 SEQ ID NO:18 之核苷酸序列的新黴素磷酸轉移酶表現匣；(dd)終止子，較佳地該終止子具有 SEQ ID NO:19 之核苷酸序列；及 (ee)具有 SEQ ID NO:8 或 SEQ ID NO:9 之核苷酸序列的 *attB* 位點；及 (3) 3' 內含子片段，較佳地該 3' 內含子片段具有 SEQ ID NO:15 之核苷酸序列；(c)該 AAV *rep* 基因之 3' 部分，較佳地該 AAV *rep* 基因之該 3' 部分具有 SEQ ID NO:56 之核苷酸序列；(ii) AAV *cap* 基因，較佳地該 AAV *cap* 基因包含 SEQ ID NO:57 之核苷酸序列；及 (iii)由一對 AAV ITR 側接之轉殖基因，較佳地，該 ITR 具有 SEQ ID NO:20 之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有 SEQ ID NO:21 之核苷酸序列且該多腺苷酸化信號具有 SEQ ID NO:23 之核苷酸序列；(B)將該第一宿主細胞用重組腺病毒感染，該重組腺病毒包含編碼具有與 SEQ ID NO:2 之胺基酸序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或 100% 序列同一性的重組酶之重組酶基

因，以獲得進一步含有該重組酶基因的第二宿主細胞；(C)在生產包含該轉殖基因之該重組AAV的條件下生長該第二宿主細胞；及(D)可選地收集該重組AAV。

【0042】 在一個態樣中，本文提供一種生產包含轉殖基因之重組AAV的方法，其包含：(A)獲得第一宿主細胞，其包含：(i)經修飾AAV *rep*基因，該基因以5'至3'的順序包含：(a) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因具有SEQ ID NO:73之核苷酸序列；(b)人工內含子，其以5'至3'的順序包含：(1) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；(2)止擋匣，其以5'至3'的順序包含：(aa)具有SEQ ID NO:7之核苷酸序列的*attP*位點；(bb)剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；(cc)編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；(dd)終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及(ee)具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及(3) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:66之核苷酸序列；(c)該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:66之核苷酸序列；(ii) AAV *cap*基因；及(iii)由下列側接之轉殖基因：(a)一對AAV ITR，較佳地，該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及(b)一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列；(B)將該第一宿主細胞用重組腺病毒感染，該重組

腺病毒包含編碼具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%序列同一性的重組酶之重組酶基因，以獲得進一步含有該重組酶基因的第二宿主細胞；(C)在生產包含該轉殖基因之該重組AAV的條件下生長該第二宿主細胞；及(D)可選地收集該重組AAV。

【0043】 在一個實施例中，該第一宿主細胞進一步包含該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：(a)具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；(b)具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；(c)具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；(d)具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因子06；(e)具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07；(f)具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12；(g)具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13；(h)具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35；(i)具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36；(j)具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因子52；(k)具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53；及(l)來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子，更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及SEQ ID NO:25之核苷酸序列。

【0044】 在一個實施例中，該第一宿主細胞包含該經修飾AAV *rep*

基因上游之該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第一間隔序列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自由下列所組成之群組：(a) SEQ ID NO:67之核苷酸序列；及(b) SEQ ID NO:68之核苷酸序列。

【0045】 在一個實施例中，該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、及該第二絕緣子。在一個實施例中，該第一宿主細胞係藉由將核酸分子引入至該細胞中而獲得，該核酸分子以5'至3'的順序包含該第一絕緣子、該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、及該第二絕緣子、較佳地包含SEQ ID NO:12之核苷酸序列的質體。

【0046】 在一個實施例中，該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、該第一間隔序列、及該第二間隔序列。在一個實施例中，該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、該第一間隔序列、及該第二間隔序列、較佳地包含SEQ ID NO:70之核苷酸序列的質體。

【0047】 在一個實施例中，該重組腺病毒係重組 Δ E1/ Δ E3腺病毒血清型5 (Ad5)病毒，該病毒包含SEQ ID NO:3之核苷酸序列。

【0048】 在一個實施例中，該宿主細胞包含腺病毒E1A及E1B基

因，較佳地該細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6細胞。

【0049】 在一個實施例中，用於生長該第二宿主細胞之條件包含用2-胺基嘌呤培養該第二細胞。在一個實施例中，該2-胺基嘌呤濃度係小於約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約1 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約10 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約100 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約1.25 mM。

【0050】 在一個實施例中，用2-胺基嘌呤培養該第二細胞係在用重組腺病毒感染該第一宿主細胞後約24小時起始。。

【0051】 在一個態樣中，本文提供一種組成物，其包含如上所述之包含編碼重組酶之核酸分子的細胞、及2-胺基嘌呤。在一個實施例中，該2-胺基嘌呤濃度係小於約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約1 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約10 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約100 μ M至約1.25 mM。在一個實施例中，該2-胺基嘌呤濃度係約1.25 mM。

【0052】 在一個態樣中，本文提供一種非天然存在的核酸分子，其包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。在一個實施例中，該非天然存在的核酸分子包括與SEQ ID NO:3之核苷酸序列具有至少85%、86%、87%、88%、89%、90%、91%、

92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核苷酸序列。

【0053】 在一個態樣中，本文提供一種載體，其包含該非天然存在的核酸分子，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。

【0054】 在一個態樣中，本文提供一種載體，其包含該非天然存在的核酸分子，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:3之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。

【0055】 在一個實施例中，該載體進一步包括啟動子，較佳的是可操作地連接至編碼該絲胺酸重組酶之該核苷酸序列的巨細胞病毒(CMV)啟動子。

【0056】 在一個實施例中，該載體進一步包括多腺苷酸化信號，諸如猴病毒40 (SV40)多腺苷酸化信號，其可操作地連接至編碼該絲胺酸重組酶的該核苷酸序列。

【0057】 在一個實施例中，該載體係DNA質體。在一些實施例中，該載體係重組腺病毒載體。

【0058】 在一個實施例中，該載體係一種重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列在CMV啟動子的控制下編碼具有SEQ ID NO:2之胺基酸序列的絲胺酸重組酶，其中該核苷

酸序列係進一步可操作地連接至SV40多腺苷酸化信號(NC_001669.1, nt 2550至2774)。

【0059】 在一個態樣中，本文提供一種細胞，其包含非天然存在的核酸分子，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。在一個實施例中，該細胞包括與SEQ ID NO: 3之核苷酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%同一性的核苷酸序列。

【0060】 在一個態樣中，本文提供一種細胞，其包括包含該非天然存在的核酸分子之載體，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。在另一態樣中，本文提供一種細胞，其包括包含該非天然存在的核酸分子之載體，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:3之胺基酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。

【0061】 在一個實施例中，該細胞進一步包括啟動子，較佳的是可操作地連接至編碼該絲胺酸重組酶之該核苷酸序列的巨細胞病毒(CMV)啟動子。

【0062】 在一個實施例中，該細胞進一步包括多腺苷酸化信號，諸如猴病毒40 (SV40)多腺苷酸化信號，其可操作地連接至編碼該絲胺酸重組酶的該核苷酸序列。

【0063】 在一個實施例中，該載體係DNA質體。在一些實施例中，該載體係重組腺病毒載體。

【0064】 在一個實施例中，該重組腺病毒載體包括一種重組 Δ E1/ Δ E3腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列在CMV啟動子的控制下編碼具有SEQ ID NO:2之胺基酸序列的絲胺酸重組酶，其中該核苷酸序列係進一步可操作地連接至SV40多腺苷酸化信號(NC_001669.1, nt 2550至2774)。

【0065】 在一個實施例中，該細胞包括腺病毒E1A及E1B基因，較佳地該細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6細胞。

【0066】 在一個態樣中，本文提供一種在細胞中進行位點特異性重組的方法，該方法包含：(a)獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attP$ 位點（較佳的是具有SEQ ID NO:7之核苷酸序列的 $attP$ 位點）及與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attB$ 位點（較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 $attB$ 位點）；(b)將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%、

86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；及(c)在使該絲胺酸重組酶催化在該等*attP*與*attB*位點之間該位點特異性重組的條件下生長該細胞。

【0067】 在一個態樣中，本文提供一種藉由在細胞中進行位點特異性重組之程序所產生的產物，該程序包含：(a)獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點（較佳的是具有SEQ ID NO:7之核苷酸序列的*attP*位點）及與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attB*位點（較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點）；(b)將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；及(c)在使該絲胺酸重組酶催化在該等*attP*與*attB*位點之間該位點特異性重組的條件下生長該細胞。

【0068】 在一個態樣中，本文提供一種用於獲得來自細胞之產物的程序，該程序包含：(a)獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點（較佳的是具有SEQ ID NO:7之核苷酸序列的*attP*位點）及與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%

同一的核苷酸序列的*attB*位點（較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點）；(b)將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；(c)在使該絲胺酸重組酶催化在該等*attP*與*attB*位點之間該位點特異性重組的條件下生長該細胞；及(d)從該細胞中生產並回收產物。

【0069】 在一個態樣中，本文提供一種非天然存在的系統，其包含：用於經AAV媒介之重組之手段，其中該手段可選地包含轉殖基因因子。在一個態樣中，本文提供一種用於轉移該非天然存在的系統之手段，其包含：用於經AAV媒介之重組之手段，其中該手段可選地包含轉殖基因因子。

【0070】 在一個態樣中，本文提供一種非天然存在的系統，其包含：用於重組包含下列之系統的重組手段：用於經AAV媒介之重組之手段，其中該手段可選地包含轉殖基因因子，其中該重組手段包括在催化期間使用至少一種絲胺酸殘基。在一個態樣中，本文提供一種用於轉移該非天然存在的系統之手段，其包含：用於重組包含下列之系統的重組手段：用於經AAV媒介之重組之手段，其中該手段可選地包含轉殖基因因子，其中該重組手段包括在催化期間使用至少一種絲胺酸殘基。

【0071】 在一個態樣中，本文提供一種用於製造分子的手段，其中該用於製造分子的手段包含上述手段中之任一者並且能夠被複製。

【0072】 在一個態樣中，本文提供一種用於經AAV媒介之位點特異性重組之程序，其包含：(a)用於執行獲得細胞之功能的步驟，其包含用

於經AAV媒介之重組之手段，其中該手段可選地包含轉殖基因因子；(b) 在催化期間使用至少一種絲胺酸殘基以使位點特異性重組的條件下，用於執行生長該細胞的功能之步驟。在一個實施例中，該程序包括獲得產物，其中可選地該產物係治療性產物。

【0073】發明之其他態樣、特徵、及優點，自下文揭露（包括實施方式與其較佳實施例，以及附加之申請專利範圍）觀之，係顯而易見者。

【圖式簡單說明】

【0074】前述發明內容以及下文中本申請案之較佳實施例的實施方式在結合附圖閱讀時可更有利理解。然而應理解的是，本申請案並不限於圖式中所示確切實施例。

〔圖1〕顯示SPBetac2整合酶蛋白(SEQ ID NO:1, query)與在沙福芽孢桿菌(*Bacillus safensis*)菌株CCMA-560 (SEQ ID NO:2, Sbjct)之基因體中經識別的推定絲胺酸重組酶的比對統計及序列比對，序列ID: WP_029708089.1，長度為535個胺基酸。該二種蛋白質在來自胺基酸1-529之蛋白質位準下具有64%序列同一性。此推定絲胺酸重組酶在本文中係命名為SR21（絲胺酸重組酶21）。

〔圖2〕顯示代表插入前基因座之菌株的識別：CCMA-560 DNA序列(query)與沙福芽孢桿菌菌株Fairview contig56_1 (Sbjct)的全基因體散彈槍序列之核苷酸464352至464839的比對統計及序列比對，序列ID: NZ_JFBY01000018.1，長度為568093個核苷酸。

〔圖3〕顯示SR21重組酶attP及attB位點。該等attP及attB位點係由圍繞中央二核苷酸重組交換位點（劃底線）的雙對稱(dyad symmetry)組成。半部位點經編號。在attB序列中引入空格以顯示由先前研究推斷的鋅

帶結構域(ZD)及重組酶結構域(RD)結合的序列比對(Rutherford *et al.* (2013) *Nucleic Acids Res.* 41:8341-8356)。與ZD或RD結構域之三或四個同一的殘基以粗體表示。所示者為該 *attP* (SEQ ID NO:7)與兩個交替的 *attB*序列(SEQ ID NO:8)及(SEQ ID NO:9)的比對。

〔圖4〕繪示報導基因的重組酶活化。質體P41編碼兩個報導子基因轉錄本。由EF1 α 啟動子驅動的第一者具有組成型活性，並在綠色螢光蛋白(GFP)與藉由自切割F2A胜肽連接子連接的海腎螢光素酶(*renilla luciferase*)之間編碼融合蛋白。由CMV所驅動的第三轉錄物包括SR21重組酶*attB*位點(SEQ ID NO:9)，接著是由該P2A自切割胜肽連接子及SR21 *attP*位點連接的編碼mCherry及螢火蟲螢光素酶之反向融合蛋白編碼區。螢光素酶及mCherry皆不表現，因為其等相對於該啟動子處於相反定向。當表現SR21重組酶時，該等*attB*及*attP*序列重組，其導致該等報導基因倒置且螢火蟲螢光素酶及mCherry表現。

〔圖5〕顯示根據本申請案之實施例所產生之經純化的重組AAV樣本中之AAV殼體蛋白。從用質體P439穩定轉染的細胞中純化樣本，使其在Hyperflask容器中以20 MOI (A)及40 MOI (B)生長並感染，並且經受PAGE及銀染色。

〔圖6〕繪示根據本申請案之實施例具有人工內含子之*rep/cap*表現匣，該人工內含子具有插入其中的止擋匣。

〔圖7〕繪示根據本申請案之實施例的載體（質體P439）。

〔圖8〕繪示在P439中藉由RT-PCR所識別的RNA剪接位點之位置及序列。頂部圖式代表止擋匣切除之後的REP基因結構。REP之5'及3'半部係由該 β -肌動蛋白內含子之上游半部(SEQ IDNO:14)、該SR21 AttL因子

(SEQ ID 35)、及 β -肌動蛋白內含子之下游半部(SEQ ID NO:15)隔開。在(2)該 β -肌動蛋白剪接供體(SEQ ID NO:71)與(3) β -肌動蛋白剪接受體(SEQ ID NO:72)之間的剪接係由實線表示。在(1) 5'REP序列中之上游剪接供體(SEQ ID No:64)與(3) 3' β -肌動蛋白受體(SEQ ID NO:72)之間的剪接係用虛線顯示。剪接供體及受體之序列顯示如下。小寫字母序列表示該內含子序列。

〔圖9〕繪示根據本申請案之實施例之載體（質體P600）。

【實施方式】

【0075】 本申請案主張於2019年7月23日申請之美國臨時申請案第62/877,508號；2019年7月23日申請之美國臨時申請案第62/877,516號；2019年7月23日申請之美國臨時申請案第62/877,524號；2019年7月23日申請之美國臨時申請案第62/877,532號；2019年7月23日申請之美國臨時申請案第62/877,540號；2019年7月23日申請之美國臨時申請案第62/877,551號；2019年7月23日申請之美國臨時申請案第62/877,561號；及2019年7月23日申請之美國臨時申請案第62/877,577號等之優先權，其等全文以引用方式併入本文中。

電子提交序列表之參照

【0076】 本申請案含有序列表，該序列表已經由EFS-Web以ASCII格式序列表電子提交，檔案名稱為「14620-192-185_SEQ_LISTING」，創建日期為2020年7月16日，且檔案大小為152,403位元組。經EFS-Web提交之序列表係本說明書之一部分，其全文以引用方式併入本文中。

【0077】 各篇公開案、論文、及專利已於先前技術及整份說明書引用或描述；此等參考文獻之各者全文係以引用方式併入本文中。在本說明

書中所包括之對於文件、行動、材料、裝置、物品、或其類似者的論述，目的在於提供關於本發明的脈絡。此等論述並非承認，任一或所有此等情事形成了關於任何所揭示或請求之發明的先前技術部分。

【0078】 除非另有定義，否則本文中所使用之所有技術及科學用語，均與本發明有關技術領域中具有通常知識者所通常了解之意義相同。在其他方面，在本文中所使用的某些用語具有如本說明書所闡述之意義。在本文中所引用的所有專利、已公開專利申請案及公開案係以引用方式併入，猶如全文說明於本文中。

【0079】 必須注意的是，本文及附加之申請專利範圍中所使用之單數形式「一(a/an)」及「該(the)」皆包括複數指稱，除非上下文另有明確說明。

【0080】 除非以其他方式指示，在一系列因子之前的用語「至少(at least)」應理解為係指該序列中的每一因子。所屬技術領域中具有通常知識者將認可或僅使用例行實驗即可確定本文所述之本發明的特定實施例的許多等效物。該等等效物意欲涵蓋於本發明中。

【0081】 在整份說明書及申請專利範圍中，除非上下文另有要求，否則詞語「包含(comprise)」及諸如「包含(comprises)」及「包含(comprising)」的變化形式，將理解為暗示包括所述整數或步驟或整數或步驟之群組，但不排除任何其他整數或步驟或整數或步驟之群組。當用於本文時，用語「包含(comprising)」可經用語「含有(containing)」或「包括(including)」取代，或有時候在本文中使用用語「具有(having)」。

【0082】 當用於本文時，「由...組成(consisting of)」排除所請之因子中未指定之任何因子、步驟、或成分。當用於本文時，「基本上由...

組成(consisting essentially of)」不排除不會實質影響申請專利範圍之基本及新穎特徵的材料或步驟。在本文中，當任何前述用語「包含(comprising)」、「含有(containing)」、「包括(including)」、及「具有(having)」用於本申請案之態樣或實施例之上下文時，可由用語「由...組成(consisting of)」或「基本上由...組成(consisting essentially of)」置換。

【0083】 如本文中所使用，多個所述因子之間的聯合用語「及/或(and/or)」係理解為涵蓋個別及組合選項兩者。例如，其中兩個因子係藉由「及/或」結合，第一選項係指第一因子在沒有第二因子的情況下之適用性。第二選項係指第二因子在沒有第一因子的情況下之適用性。第三選項係指第一及第二因子一起的適用性。這些選項之任一者應理解為落入該含義內，並因此滿足如本文中所使用之用語「及/或」之要求。該等選項之多於一者的並行適用性亦應理解為落入該含義內，並因此滿足用語「及/或」之要求。

【0084】 除非另有說明，在本文中描述之任何數值（諸如濃度或濃度範圍）應理解為在所有情況下皆受到用語「約(about)」之修飾。因此，數值一般包括記載值之 $\pm 10\%$ 。例如，濃度1 mg/mL包括0.9 mg/mL至1.1 mg/mL。同樣地，1 mg/mL至10 mg/mL之濃度範圍包括0.9 mg/mL至11 mg/mL。如本文中所使用，明示使用之數值範圍包括所有可能的子範圍、在該範圍內之所有個別數值，包括在該等範圍內之整數及數值之分數，除非上下文以其他方式清楚指示。

【0085】 當參考胺基酸序列使用時，用語「百分比(%)序列同一性」或「%同一性(% identity)」或「與...%同一(% identical to)」描述相

較於構成胺基酸序列總長度的胺基酸殘基編號，二或更多個比對的胺基酸序列之相同胺基酸的匹配數（「命中」）。換句話說，使用比對，當比較序列並對其進行比對以獲得最大對應性時（如使用所屬技術領域中已知的序列比較演算法測量），或當手動比對及目視檢查時，可判定二或更多個序列胺基酸殘基相同之百分比（例如，在該等胺基酸序列全長上90%、91%、92%、93%、94%、95%、97%、98%、99%或100%同一性）。判定序列同一性所比較的序列可能因胺基酸的（多個）取代、（多個）添加、或（多個）缺失而不同。用於比對蛋白質序列之合適程式係所屬技術領域中具有通常知識者所習知。例如，蛋白質序列的百分比序列同一性可用諸如CLUSTALW、Clustal Omega、FASTA、或BLAST之程式，例如使用NCBI BLAST演算法來判定(Altschul SF, *et al* (1997), *Nucleic Acids Res.* 25:3389-3402)。

【0086】如本文中所使用，「非天然存在的(non-naturally occurring)」核酸或多肽係指在自然界中不存在之核酸或多肽。可在實驗室及/或製造設定中合成、處理、製造、及/或以其他方式操縱「非天然存在的」核酸或多肽。在一些情況下，非天然存在的核酸或多肽可包含經處理、加工、或操縱的天然存在的核酸或多肽，以展現該天然存在的核酸或多肽在處理前不存在的性質。如本文中所使用，「非天然存在的」核酸或多肽可係與所發現之天然來源單離或分離的核酸或多肽，且其缺乏與其在天然來源相關聯之序列的共價鍵。可重組或經由其他方法（諸如化學合成）來製成「非天然存在的」核酸或多肽。

【0087】如本文中所使用，用語「混合(hybrid)」當用於表示AAV *cap*基因時，旨在意指包括與不同血清型殼體的部分結合之一種血清型殼

體的部分之 cap 基因。該用語亦包括AAV cap 基因變異體，其中該天然存在的AAV血清型序列含有一或多個非天然存在之突變。

【0088】如本文中所使用，用語「間隔序列(spacer sequence)」旨在意指非編碼核苷酸的區域，除了分離其他遺傳因子外，該核苷酸不具有表觀功能。

【0089】如本文中所使用，用語「可操作地連接的(optionally linked)」係指鍵聯或併接(juxtaposition)，其中所述之組分係處於允許其等以其預期方式起作用的關係。例如，如果啟動子影響編碼序列的轉錄，則其可操作地連接至該編碼序列，或者可操作地連接至所關注之胺基酸序列的信號序列能夠在膜上分泌或易位所關注之胺基酸序列。

【0090】為了幫助本申請案的讀者，已將說明書分成各種不同的段落或章節，或指向本申請案之各種實施例。這些分段不應被視為是將來段落或章節或實施例與另一段落或章節或實施例實質斷開。相反地，所屬技術領域中具有通常知識者將理解，本說明書具有廣泛的應用並且涵蓋可設想的各種章節、段落、及文句之所有組合。任何實施例之討論僅意指為例示性並且不意欲暗示本揭露之範疇（包括申請專利範圍）限於這些實施例。例如，儘管本文所述之本申請案的非天然存在的核酸或重組載體之實施例（例如，質體DNA或病毒載體）可含有特定組分，包括但不限於某些啟動子序列、強化子、或調節序列、內含子、AAV Rep及/或Cap之編碼序列、多腺苷酸化信號等以特定順序排列，所屬技術領域中具有通常知識者將理解，本文所揭示之概念可同等地應用於以其他順序排列的其他組分，該等其他組分可用於本申請案之核酸或載體中。本申請案設想以任何組合使用任何可應用組分，該組合具有可用於本申請案之核酸或載體中的任何

序列，無論是否明確描述特定組合。

【0091】 如本文中所使用，「載體(vector)」係用來載送遺傳物質至細胞中之核酸分子，在該細胞中可複製及/或表現該遺傳物質。可使用鑒於本揭露之所屬技術領域中具有通常知識者已知之任何載體。載體之實例包括但不限於質體、病毒載體（噬菌體、動物病毒、及植物病毒）、黏接質體、及人工染色體（例如，YAC）。較佳地，載體係DNA質體。所屬技術領域中具有通常知識者可透過鑒於本揭露之標準重組技術建構本申請案之載體。

【0092】 本申請案之載體可係表現載體。如本文中所使用，用語「表現載體(expression vector)」係指任何類型的遺傳建構體，該遺傳建構體包含編碼能夠被轉錄之RNA的核酸。表現載體包括（但不限於）用於重組蛋白表現之載體（諸如DNA質體或病毒載體）、及用於將核酸遞送進入對象體內以於該對象之組織中表現的載體（諸如DNA質體或病毒載體）。所屬技術領域中具有通常知識者將理解，表現載體之設計可取決於諸如待轉化的宿主細胞之選擇、所欲之蛋白質表現位準等因素。

【0093】 在本申請案之一些實施例中，載體係非病毒載體。非病毒載體之實例包括但不限於DNA質體、細菌人工染色體、酵母人工染色體、噬菌體等。較佳地，非病毒載體係DNA質體。「DNA質體(DNA plasmid)」，其與「DNA質體載體(DNA plasmid vector)」、「質體DNA(plasmid DNA)」或「質體DNA載體(plasmid DNA vector)」可互換使用，係指能夠在合適的宿主細胞中自主複製之雙股及大致上圓形的DNA序列。用於表現經編碼多核苷酸之DNA質體一般包含複製起點、多個選殖位點、及可選擇標記，其例如可係抗生素抗性基因。可使用之DNA質

體的實例包括但不限於用於眾所周知的表現系統中之市售表現載體（包括原核系統及真核系統兩者），諸如 pSE420 (Invitrogen, San Diego, Calif.)，其可用於大腸桿菌中蛋白質的生產及/或表現；pYES2 (Invitrogen, Thermo Fisher Scientific)，其可用於酵母的啤酒酵母菌 (*Saccharomyces cerevisiae*) 菌株中之生產及/或表現；MAXBAC[®] 完整桿狀病毒表現系統 (Thermo Fisher Scientific)，其可用於在昆蟲細胞中生產及/或表現；pcDNA[™] 或 pcDNA3[™] (Life Technologies, Thermo Fisher Scientific)，其可用於哺乳動物細胞中高位準組成型蛋白質表現；及 pVAX 或 pVAX-1 (Life Technologies, Thermo Fisher Scientific)，其可用於在大多數哺乳動物細胞中所關注之蛋白質的高位準瞬時表現。任何市售可得的 DNA 質體之主鏈可經修飾以最佳化該宿主細胞中的蛋白質表現，諸如反轉某些因子（例如，複製起點及/或抗生素抗性匣）之定向、置換質體之內源性啟動子（例如，抗生素抗性匣中的啟動子）、及/或藉由使用常規技術及輕易可得的起始材料來置換編碼轉錄蛋白之多核苷酸序列（例如，抗生素抗性基因之編碼序列）。（參見例如 Sambrook *et al.*, *Molecular Cloning a Laboratory Manual*, Second Ed. Cold Spring Harbor Press (1989)）。

【0094】較佳地，DNA 質體係適用於在哺乳動物宿主細胞中蛋白質表現的表現載體。適用於在哺乳動物宿主細胞中表現蛋白質的表現載體包括但不限於 pUC、pcDNA[™]、pcDNA3[™]、pVAX、pVAX-1、ADVAX、NTC8454 等。例如，載體可基於 pUC57，其含有 pUC 複製起點及安比西林抗性基因 (SEQ ID NO:30)。其可進一步包含哺乳動物嘌呤黴素抗性基因匣，其係建構自疱疹病毒胸苷激酶基因啟動子 (SEQ ID

NO:26)、嘌呤黴素N-乙醯基轉移酶編碼區(SEQ ID NO:27)、及來自牛生長激素基因之多腺苷酸化信號(SEQ ID NO:28)。該載體亦可包含艾司坦-巴爾病毒(EBV) OriP複製起點片段(SEQ ID NO:29)，其代表EBV之「雙對稱(Dyad Symmetry)」區域及「重複家族(Family of Repeat)」區域之組合。

【0095】 本申請案之載體亦可係病毒載體。通常，病毒載體係經基因工程改造的病毒，其攜帶經修飾病毒DNA或非傳染性但仍含有病毒啟動子及轉殖基因的RNA，因此允許轉殖基因轉譯通過病毒啟動子。因為病毒載體經常缺乏感染序列，所以其等需要輔助病毒或包裝線用於大規模轉染之包裝線。可使用的病毒載體之實例包括但不限於腺病毒載體、腺相關病毒載體、痘病毒載體、腸病毒載體、委內瑞拉馬腦炎(Venezuelan Equine Encephalitis)病毒載體、Semliki森林病毒載體、菸草鑲嵌病毒載體、慢病毒載體等。該載體亦可係非病毒載體。

【0096】 較佳地，病毒載體係一種腺病毒載體，例如重組腺病毒載體。如本文中所使用，用語「重組腺病毒載體(recombinant adenovirus vector)」及「重組腺病毒載體(recombinant adenoviral vector)」可互換使用，且係指經基因工程改造的腺病毒，其經設計以將所關注之多核苷酸插入真核細胞，使得該多核苷酸隨後表現。可用作本發明之病毒載體的腺病毒之實例包括彼等具有或衍生自下列血清型者：Ad2、Ad5、Ad11、Ad12、Ad24、Ad26、Ad34、Ad35、Ad40、Ad48、Ad49、Ad50、Ad52（例如，RhAd52）、及Pan9（亦稱為AdC68）；這些載體可衍生自例如人類、黑猩猩(chimpanzee)（例如ChAd1、ChAd3、ChAd7、ChAd8、ChAd21、ChAd22、ChAd23、ChAd24、ChAd25、ChAd26、

ChAd27.1 、 ChAd28.1 、 ChAd29 、 ChAd30 、 ChAd31.1 、 ChAd32 、 ChAd33 、 ChAd34 、 ChAd35.1 、 ChAd36 、 ChAd37.2 、 ChAd39 、 ChAd40.1 、 ChAd41.1 、 ChAd42.1 、 ChAd43 、 ChAd44 、 ChAd45 、 ChAd46 、 ChAd48 、 ChAd49 、 ChAd49 、 ChAd50 、 ChAd67 、 或 SA7P) 、 或 恆 河 猴 (rhesus) 腺 病 毒 (例 如 rhAd51 、 rhAd52 、 或 rhAd53) 。 重 組 腺 病 毒 載 體 可 例 如 衍 生 自 人 類 腺 病 毒 (HAdV 、 或 AdHu) 、 或 猿 猴 (simian) 腺 病 毒 (諸 如 黑 猩 猩 或 大 猩 猩 (gorilla) 腺 病 毒 (ChAd 、 AdCh 、 或 SAdV) 、 或 恆 河 猴 腺 病 毒 (rhAd) 。

【0097】 較佳地，腺病毒載體係重組人類腺病毒載體，例如重組人類腺病毒血清型5，或重組人類腺病毒血清型26、4、35、7、48等中之任一者。可用於本申請案之重組病毒載體可使用鑒於本揭露之所屬技術領域中已知的方法來製備。例如，鑒於基因密碼之簡併性，數個核酸序列可經設計以編碼相同多肽。編碼所關注蛋白質之多核苷酸可任選地經密碼子最佳化，以確保在該宿主細胞中之適當表現（例如，細菌細胞或哺乳動物細胞）。密碼子最佳化(codon-optimization)係所屬技術領域中廣泛應用的技術，且用於獲得密碼子最佳化之多核苷酸的方法係鑒於本揭露所屬技術領域中具有通常知識者所熟知。

【0098】 非天然存在的核酸分子或載體可包含一或多個表現匣。「表現匣(expression cassette)」係引導細胞機器以製造RNA及蛋白質的核酸分子或載體的一部分。表現匣可包含啟動子序列、開讀框、3'-未轉譯區(UTR)，其可選地包含多腺苷酸化信號。開讀框(open reading frame, ORF)係含有從開始密碼子至止擋密碼子之所關注蛋白質（例如，Rep、Cap、所關注重組酶或重組蛋白質）之編碼序列的讀框。將表現匣之調節

因子可操作地連接至編碼所關注蛋白質之多核苷酸序列。

【0099】 本申請案之非天然存在的核酸分子或載體可含有各種調節序列。如本文中所使用，用語「調節序列(regulating sequence)」係指任何允許、促成或調節核酸分子的功能性調節之序列，其包括複製、重複、轉錄、剪接、轉譯、穩定性及/或傳輸核酸或其衍生物之一者（亦即 mRNA）進入宿主細胞或生物體。調節因子包括但不限於啟動子、強化子、多腺苷酸化信號、轉譯終止密碼子、核糖體結合因子、轉錄終止子、選擇標記、複製起點等。

【0100】 非天然存在的核酸分子或載體可包含啟動子序列，較佳的是在表現匣內，以控制所關注蛋白質的表現。用語「啟動子(promoter)」係以其習知意義使用，且係指起始可操作地連接的核苷酸序列的轉錄之核苷酸序列。啟動子位於其轉錄之核苷酸序列附近的同一股上。啟動子可係組成型、誘導型、或阻遏型。啟動子可係天然存在或合成的。啟動子可衍生自包括病毒、細菌、真菌、植物、昆蟲、及動物之來源。啟動子可係同源啟動子（亦即，衍生自與載體相同的遺傳來源）或異源啟動子（亦即，衍生自不同載體或遺傳來源）。例如，若待採用的載體係DNA質體，則該啟動子對於該質體可係內源性的（同源）或衍生自其他來源（異源）。較佳地，該啟動子係位於該多核苷酸上游，該多核苷酸在表現匣內編碼所關注蛋白質。

【0101】 可使用之啟動子實例包括但不限於來自下列者之啟動子：猴病毒40 (SV40)、小鼠乳房腫瘤病毒(MMTV)啟動子、人類免疫缺乏病毒(HIV)啟動子（諸如牛免疫缺乏病毒(BIV)長末端重複(LTR)啟動子）、莫洛尼(Moloney)病毒啟動子、鳥白血病毒(ALV)啟動子、巨細胞病毒

(CMV)啟動子（諸如該CMV即時早期啟動子(CMV-IE)）、艾司坦-巴爾病毒(EBV)啟動子、或勞斯肉瘤病毒(RSV)啟動子。啟動子亦可係來自人類基因（諸如人類肌動蛋白、人類肌凝蛋白、人類血紅素、人類肌肉肌酸、或人類金屬硫蛋白）的啟動子。啟動子亦可係組織特異性啟動子，諸如天然或合成的肌肉或皮膚特異性啟動子。較佳地，啟動子係強的真核啟動子，諸如巨細胞病毒(CMV)啟動子(nt -672至+15)、EF1- α 啟動子、疱疹病毒胸苷激酶基因啟動子(SEQ ID NO:26)等。

【0102】 非天然存在的核酸分子或載體可包含穩定經表現轉錄本、增強RNA轉錄本的核輸出、及/或改善轉錄轉譯偶合的額外多核苷酸序列。此類序列之實例包括多腺苷酸化信號及強化子序列。多腺苷酸化信號一般位於編碼序列下游，以用於該載體之表現匣內的所關注蛋白質（例如，Rep、Cap、重組酶）。當由轉錄因子結合時，該等強化子序列係調節DNA序列，其強化相關基因的轉錄。強化子序列較佳地在啟動子序列下游且可在載體之表現匣內的編碼序列下游或上游。

【0103】 可使用鑒於本揭露之所屬技術領域中具有通常知識者已知的任何多腺苷酸化信號。例如，該多腺苷酸化信號可係SV40多腺苷酸化信號（例如，SEQ ID NO:60）、AAV2多腺苷酸化信號(bp 4411-4466, NC_001401.2)、來自單純疱疹病毒胸苷激酶基因(SEQ ID NO:23)、LTR多腺苷酸化信號、牛生長激素(bGH)多腺苷酸化信號、人類生長激素(hGH)多腺苷酸化信號、或人類 β -球蛋白多腺苷酸化信號。較佳地，多腺苷酸化信號係牛生長激素(bGH)多腺苷酸化信號(SEQ ID NO:28)、具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號4411至4466之AAV2的多腺苷酸化信號、或SV40多腺苷酸化信號(SEQ ID NO:60)。

【0104】 可使用鑒於本揭露之所屬技術領域中具有通常知識者已知的任何一個強化子序列。例如，強化子序列可係人類肌動蛋白、人類肌凝蛋白、人類血紅素、人類肌肉肌酸、或病毒強化子，諸如來自CMV、HA、RSV、或EBV之一者。特定增強子之實例包括但不限於：土撥鼠HBV轉錄後調節因子(WPRE)、衍生自人載脂蛋白A1前驅(ApoAI)之內含子/外顯子序列、人類T細胞白血病毒第1型(HTLV-1)長末端重複(LTR)之未轉譯R-U5結構域、剪接增強子、合成兔 β -球蛋白內含子、或其任何組合。

【0105】 較佳地，強化子序列包含AAV之P5啟動子。該P5啟動子係該*rep*基因之編碼序列內部的順式作用Rep依賴性因子(CARE)的一部分。當在順式中存在時，CARE顯示出增加複製及殼體化作用。在一些AAV生產細胞系中，CARE對染色體整合的*rep*基因之擴增（如果不存在AAV ITR）亦係重要的。儘管不希望受到理論的束縛，但據信置於*cap*編碼序列下游的P5啟動子可能充當增加Cap表現的強化子，從而使AAV產生，並且其亦提供增強子活性，用於擴增整合至染色體中的基因。

【0106】 非天然存在的核酸分子或載體（諸如DNA質體）亦可包括細菌複製起點及抗生素抗性表現匣，用於在細菌細胞（例如大腸桿菌）中選擇及維持該質體。複製起點(origin of replication, ORI)係起始複製的序列，使得質體能夠再生並在細胞內存活。適用於本申請案之ORI之實例包括但不限於ColE1、pMB1、pUC、pSC101、R6K、及15A，較佳的是pUC。

【0107】 用於在細菌細胞中選擇及維持之載體一般包括可操作地連接至抗生素抗性基因的啟動子序列。較佳地，可操作地連接至抗生素抗性

基因的啟動子序列與可操作地連接至編碼所關注蛋白質之多核苷酸序列的啟動子序列不同。該抗生素抗性基因可經密碼子最佳化，且該抗生素抗性基因之序列組成通常經調節成細菌（例如，大腸桿菌）密碼子使用。可使用鑒於本揭露之所屬技術領域中具有通常知識者已知的任何抗生素抗性基因，包括但不限於康黴素抗性基因(Kan^r)、安比西林抗性基因(Amp^r)、及四環黴素抗性基因(Tet^r)，以及賦予對氯黴素、博萊黴素、大觀黴素(spectinomycin)、卡本西林(carbenicillin)等抗性的基因。

【0108】 用於在哺乳動物細胞中選擇及維持之載體一般包括可操作地連接至編碼賦予可選擇標記之蛋白質的基因。較佳地，該基因進一步包含多腺苷酸化信號。例如，哺乳動物嘌呤黴素抗性基因匣可包含疱疹病毒胸苷激酶基因啟動子(SEQ ID NO:26)、嘌呤黴素N-乙醯基轉移酶編碼區(SEQ ID NO:27)、及來自牛生長激素基因之多腺苷酸化信號(SEQ ID NO:28)。

【0109】 在人類細胞中重組AAV的製造需要AAV複製(*rep*)及殼體(*cap*)基因、腺病毒基因、及AAV-可包裝轉殖基因的表現，該轉殖基因由AAV反向末端重複(ITR)側接的表現匣組成。所有三種組分都可仰賴單獨的質體遞送至細胞以產生AAV，但是現有的轉染方法難以擴展成大規模培養。將一些該等因子摻入宿主細胞系中可使AAV生產更有效率，然而一些AAV及腺病毒基因係細胞生長抑制性或細胞毒性，因而限制了此方法。

【0110】 本申請案描述非天然存在的核酸分子、載體、細胞、及使該AAV *rep*基因可逆失活的方法，使得AAV *rep*基因、AAV *cap*基因及可包裝的轉殖基因可被維持及/或整合至合適的宿主細胞中並擴增。藉由表

現重組酶之重組腺病毒感染這些細胞會重新活化該等*rep*基因並誘導AAV複製及包裝。不同於由Xiao與同事所述的方法(Qiao *et al.* (2002) J. Virol. 76: 13015-13027; Yuan *et al.* (2011) Hum. Gene Ther. 22:613-624)，其使用Cre（一種識別兩個相同loxP位點並催化連接及切除反應兩者的酪胺酸重組酶），本發明使用絲胺酸重組酶21 (SR21)（一種由本申請案之發明人新表徵之絲胺酸重組酶）。與Cre不同的是，SR21識別該等*attP*及*attB*位點，其具有不同序列。在藉由SR21催化的連接反應之後，該等*attP*及*attB*位點重組及破壞，使得無額外重組的可能。因此，本申請案之方法可比由Cre催化更有效。本申請案之某些實施例包括額外的特徵，諸如在不同的人工內含子、強化子、絕緣子等中插入不同的止擋匣，其等進一步改善了先前技術之方法。本申請案之可逆失活/再活化系統允許在包裝細胞生長期間得以嚴密控制該AAV *rep*基因，從而避免Rep蛋白對該宿主細胞的細胞生長抑制性/細胞毒性作用。在載體生產期間，其亦提供對該AAV *rep*基因的強誘導作用及AAV載體的高產率。

絲胺酸重組酶

【0111】 由大絲胺酸重組酶家族之成員（諸如SR21）所催化的位點特異性重組不需要細胞機制即可進行同源重組。一般而言，其需要一種識別該等位點、斷裂、及連接該DNA的特異化重組酶。基於胺基酸序列同源性及機制相關性，大多數位點特異性重組酶分為兩個家族之一者：酪胺酸重組酶家族或絲胺酸重組酶家族。該等名稱源於保留的親核性胺基酸殘基，其用於在股交換期間攻擊DNA並與其變成共價連接。

【0112】 絲胺酸重組酶結合並重組單獨的重組識別位點，稱為「附接位點」：*attP*，「附接噬菌體」及*attB*，「附接細菌」染色體。該等

*attP*及*attB*位點係由圍繞中央二核苷酸重組交換位點的雙對稱組成。*attP*及*attB*位點的左右兩半部藉由鋅帶(ZD)及重組酶(RD)結構域與重組酶單體結合(Rutherford *et al.* (2013) *Nucleic Acids Res.* 41:8341-8356)。

【0113】 如以下實例中更詳細描述者，在本發明中在沙福芽孢桿菌菌株CCMA-560之基因體中新識別之絲胺酸重組酶，在本文中稱為「絲胺酸重組酶21」或「SR21」。在本發明中亦表徵由SR21識別的該等*attP*及*attB*位點。

【0114】 在一個大致態樣中，本申請案係關於一種非天然存在的核酸分子，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。較佳地，該非天然存在的核酸分子包含編碼絲胺酸重組酶的核苷酸序列，該絲胺酸重組酶具有SEQ ID NO:2之胺基酸序列。在一個實施例中，該非天然存在的核酸分子包含與SEQ ID NO:3之核苷酸序列具有至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核苷酸序列。

【0115】 在某些實施例中，本申請案係關於包含該非天然存在的核酸之載體。該載體可係表現載體，其在所關注之細胞（例如，細菌細胞或哺乳動物細胞）中表現該絲胺酸重組酶。在一個實施例中，該載體在CMV啟動子或本文所述或所屬技術領域中已知之任何其他合適的啟動子的控制下在哺乳動物細胞中表現該絲胺酸重組酶。在某些實施例中，該載體可進一步包括多腺苷酸化信號，諸如猴病毒40 (SV40)多腺苷酸化信號

或本文所述或所屬技術領域中已知的任何其他合適的多腺苷酸化信號。

【0116】 在一個實施例中，該載體係DNA質體，諸如具有SEQ ID NO:10之核苷酸序列的質體P175。

【0117】 在另一實施例中，該載體係病毒載體，諸如重組腺病毒載體。

【0118】 在一個實施例中，該載體係重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列編碼具有與SEQ ID NO:2之胺基酸序列至少85%同一性、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶，且該編碼序列係在哺乳動物細胞中有作用的啟動子的控制下。較佳地，該啟動子係CMV啟動子。更佳地，該重組Ad5載體以5'至3'的順序包含CMV啟動子，該CMV啟動子可操作地連接至編碼SEQ ID NO:2之胺基酸序列的核苷酸序列，該核苷酸序列係可操作地連接至SV40多腺苷酸化信號(NC_001669.1, nt 2550至2774)。在一個實施例中，該編碼SEQ ID NO:2之胺基酸序列的核苷酸序列係與SEQ ID NO:3相同，除了該細菌轉譯起始密碼子「TTG」係由「ATG」置換，並且引入三個點突變以破壞SEQ ID NO:3內的限制性核酸內切酶識別位點。這些限制性核酸內切酶識別位點係Xba I位點(TCTAGA)；Sac I位點(GAGCTC)；EcoRI位點(GAATTC)。

【0119】 可使用鑒於本揭露之所屬技術領域中已知的任何方法製成編碼本申請案之絲胺酸酶的載體。

【0120】 如以下實例中所更詳細描述者，本發明中識別用於本申請案之絲胺酸重組酶的*attP*及*attB*位點。在某些實施例中，本申請案之絲胺

酸重組酶識別包含SEQ ID NO:7之核苷酸序列或其變體的*attP*位點。在某些實施例中，本申請案之絲胺酸重組酶識別包含與SEQ ID NO:7至少90%、諸如至少91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點。

【0121】 在某些實施例中，本申請案之絲胺酸重組酶識別包含SEQ ID NO:8或SEQ ID NO:9之核苷酸序列或其變體的*attB*位點。在某些實施例中，本申請案之絲胺酸重組酶識別包含與SEQ ID NO:8或SEQ ID NO:9至少90%、諸如至少91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attB*位點。

【0122】 在一個實施例中，本申請案係關於一種在細胞中進行位點特異性重組的方法。該方法包含：

- 1) 獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%同一的核苷酸序列的*attP*位點及與SEQ ID NO:8或SEQ ID NO:9至少90%同一的核苷酸序列的*attB*位點；
- 2) 將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%同一性、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；及
- 3) 在使該絲胺酸重組酶催化在該等*attP*與*attB*位點之間該位點特異性重組的條件下生長該細胞。

【0123】 在一較佳實施例中，本申請案係關於一種在細胞中進行位點特異性重組的方法。該方法包含：

- 1) 獲得包含核酸分子的細胞，該核酸分子具有SEQ ID NO:7之

核苷酸序列的 $attP$ 位點及SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 $attB$ 位點；

2) 將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有SEQ ID NO:2之胺基酸的絲胺酸重組酶；及

3) 在使該絲胺酸重組酶催化在該等 $attP$ 與 $attB$ 位點之間該位點特異性重組的條件下生長該細胞。

用於生產重組AAV之建構體、細胞、及方法

【0124】 如以下實例中所說明，本申請案之新識別的絲胺酸重組酶可用來改善重組AAV的生產。

經修飾AAV *rep*基因建構體

【0125】 在一個大致態樣中，本申請案係關於一種非天然存在的核酸分子，該核酸分子包含經修飾腺相關病毒(AAV) *rep*基因，該基因具有編碼四種Rep蛋白Rep78、Rep68、Rep52、及Rep40之AAV *rep*基因及插入至由該等四種Rep蛋白共享之該*rep*基因的編碼序列中的人工內含子。該人工內含子包含止擋匣，該止擋匣係插入在該人工內含子之5'剪接位點下游及分支位點上游，且該止擋匣以5'至3'的順序包含：(i)具有與SEQ ID No:7至少90%、諸如至少91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attP$ 位點，(ii)剪接受體；(iii)終止子；及(iv)具有與SEQ ID NO:8或SEQ ID NO:9至少90%、諸如至少91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attB$ 位點。較佳地，該 $attP$ 位點具有SEQ ID NO:7之核苷酸序列且該 $attB$ 位點具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列。

【0126】 如本文中所使用，「內含子(intron)」係經廣泛地定義為可

藉由RNA剪切而移除的核苷酸序列。「RNA剪接(RNA splicing)」意指從前驅mRNA切除內含子以形成成熟mRNA。本文中所使用之「人工內含子(artificial intron)」係指不是基因之天然發生的內含子但仍可藉由RNA剪接而移除的核苷酸序列。例如，「人工內含子(artificial intron)」可係具有經插入止擋匣之天然存在的內含子。

【0127】內含子（包括人工內含子）含有5'剪接位點或接合點(junction)、剪接受體或分支點、及3'剪接位點或剪接接合點(junction)。用語「5'剪接位點」或「5'剪接接合點」意指外顯子-內含子接合點的位置，其中該接合點係在基因或核酸片段之5'片段的3'端與該內含子之5'端之間，並且包括在RNA剪接所需的內含子之5'端處的共通序列。用語「剪接受體」或「分支點」係指核苷酸，通常係腺苷酸，位於距3'剪接位點大約20至50 bp處，其在RNA剪接期間的第一次酯交換反應期間幫助形成套索結構(lariat structure)。用語「3'剪接位點」或「3'剪接接合點」意指外顯子-內含子接合點的位置，其中該接合點係在基因或核酸片段之3'片段的5'端與該內含子之3'端之間，並且亦包括在RNA剪接所需的內含子之3'端處的共通序列。用語「共通序列(consensus sequence)」意指在RNA剪接所需的5'或3'剪接接合點中或與之相鄰的核苷酸；這些序列通常係不變的或高度保留的。

【0128】大量的mRNA的分析顯示某些核苷酸在典型的內含子及剪接接合點中係保留的。例如，內含子之幾乎不變的鹼基係5'-GU及3'-AG。這些5'及3'保留區兩側的某些鹼基通常以異常（非隨機）頻率出現。同樣保留的係分支點腺苷酸，通常距3'剪接位點20至50個鹼基。請參見例如Gao *et al* (2008) *Nucleic Acids Research* 36: 2257-2267的圖4，其顯示

在外顯子的情況下對於內含子的一般共識，Gao *et al* (2008)之全部內容以引用方式併入本文中。然而，內含子的中央區域的長度可在40至50,000個鹼基之間，因此通常不需要進行剪接。內含子藉由剪接體作為套索狀結構從RNA或前驅mRNA移除。經由二個依序的酯交換反應來進行外顯子一起的剪接。

【0129】可藉由所屬技術領域中已知的任何方法來完成內含子插入經表現序列中。側接的外顯子環境以及實際使用的內含子序列在是否可有效剪接出新的「內含子」方面發揮作用。可藉由電腦模擬(*in silico*)製造複合序列並使用線上剪接預測程式來測試適用於本發明之內含子，以尋找對於有效的RNA剪接給出足夠高的得分之*rep*基因序列與內含子序列的組合。鑒於本揭露，可測試及最佳化基因體或合成序列中的任一個內含子以用於本發明之建構體。

【0130】為了破壞Rep78、Rep68、Rep52、及Rep40之所有四種*rep*開讀框的表現，較佳地將人工內含子插入至由該四種Rep蛋白共享之*rep*基因的編碼序列。因此，在某些實施例中，為了破壞所有四種ORF，將人工內含子插入至AAV2 (NC_001401.2)的核苷酸996及至多1905之後或另一個AAV *rep*基因中的對應位置。但是，為了使止擋匣在插入人工內含子中時起作用，較佳係將內含子插入該*rep*基因中，並儘可能地在上游。

【0131】另外，該內含子插入位點上游及下游的外顯子環境對於定義什麼將作為可能的插入位點很重要，例如，在上述外顯子的情況下對於內含子的一般共識。在一個實施例中，該共通序列CAG[^]G（其中[^]標記該插入所在的位置）如下發生在AAV2中的*rep*基因的相關區域中，其中數

字表示插入之前AAV的最後一個核苷酸：1052、1061、1712、及1906。在另一實施例中，該共通序列AAG^G發生在AAV2的位置1022（如Qiao所使用的）、1112、1475、1514、1700、1742、及1784中。其他共通序列（諸如AAG^A）發生在例如AAV2的核苷酸1340處。鑒於本揭露，亦可在其他AAV的*rep*基因中識別較佳的插入位點。

【0132】 可用於本發明的人工內含子可衍生自任何來源，諸如來自基因體庫。如下所述，可藉由聚合酶連鎖反應(PCR)使用引子從人類DNA獲得內含子。能夠在細胞中進行RNA剪接的任何內含子皆可用於本發明之方法。在以下實例中，該內含子係人類β-肌動蛋白基因的內含子。

【0133】 根據本申請案之實施例，除了經由人工內含子進行RNA剪接外，亦可經由插入該人工內含子的止擋匣進行DNA剪接來調節Rep蛋白的表現。該止擋卡匣包含由該等*attP*及*attB*位點側接之轉錄終止子，該等位點由絲胺酸重組酶特異性識別，諸如本發明中所特徵化者。在一個實施例中，該終止子包含一或多個多腺苷酸化信號。在另一實施例中，該終止子包含另一個用於有效轉錄終止的序列，諸如來自該多腺苷酸化信號下游人類β-球蛋白基因的序列，其編碼自切割RNA模體，較佳地具有SEQ ID NO:19之核苷酸序列。其他終止子亦可用於本發明，諸如可裂解自身RNA的錘頭核酶(ribozyme)。請參見West (2008) *Molecular Cell* 29:600-610之使用其他核酶置換該β球蛋白因子，及Kharma (2016) *Nucleic Acids Res.* 44:e39之設計核酶的描述，兩者之內容以全文引用方式併入本文中。

【0134】 在一個實施例中，該止擋匣進一步包含編碼可選擇標記之基因。在一個實施例中，該可選擇標記基因包含新黴素磷酸轉移酶表現匣

(*neo*) (SEQ ID NO:18)，其係由哺乳動物啟動子（例如，小鼠磷酸甘油酯激酶1）及細菌（例如，*Lac zya*）啟動子並接著多腺苷酸化信號（諸如來自SV40）所驅動。該基因分別在哺乳動物及細菌細胞中賦予對新黴素及康黴素的抗性。儘管不希望受到理論的束縛，但據信除了提供用於細胞系發育的可選擇標記外，可選擇標記基因可進一步阻斷該*rep*基因的轉錄，從而增加含有該經修飾*rep*基因之宿主細胞的穩定性。可用於本發明之其他可選擇標記基因包括但不限於抗生素選擇基因（嘌呤黴素、潮黴素、博萊黴素）、代謝基因（例如，麩醯胺酸合成酶或次黃嘌呤鳥嘌呤之磷酸核糖基轉移酶(HPRT)）、病毒標記（諸如mCherry）、酶（諸如 β -半乳糖酶）、分泌的鹼性磷酸酶、或任何其他合適的標記基因。

【0135】 在另一實施例中，該止擋匣包含剪接受體以預防該止擋匣從初級mRNA轉錄本中剪接出來。可使用任何天然存在的剪接受體位點或合成序列，前提是該剪接受體未被跳過。根據本申請案之實施例，該剪接受體含有符合該共通序列(yTnAynn)的分支點序列，其中y係C或T，且n係任何核苷酸、多嘧啶通道(4-24 nt)、「AG」二核苷酸、及20至80 bp之真核基因外顯子序列（或當置於內含子序列旁時充當外顯子之合成序列）。NetGene2剪接預測軟體應將該序列識別為剪接受體位點(www.cbs.dtu.dk/services/NetGene2/; Brunak, S., Engelbrecht, J., and Knudsen, S.: Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence, Journal of Molecular Biology, 1991, 220, 49-65)置信得分為0.4或更好（或用類似的剪接預測軟體）；得分接近1.0更佳。在一個實施例中，該剪接受體包含來自小鼠HPRT基因之SEQ ID NO:17之核苷酸序列（NC_000086.7，核苷酸53001998至53002138），

加上來自人類刺豚鼠(agouti)傳訊蛋白之29 nt區 (NC_000020.11 , 核苷酸34262765至34262793) 。

【0136】 根據本申請案之實施例，該止擋匣係插入在該人工內含子之5'剪接供體位點下游及「分支位點」上游。可在兩個位點之間的任何位置插入該止擋匣，前提是該插入不會損壞該等位點的功能。在一個實施例中，該止擋匣係插入在該兩個位點中間。在以下實例中所述之一個例示性實施例中，該止擋匣係插入至人類 β -肌動蛋白基因之內含子中，使得該5'內含子片段具有SEQ ID NO:14之核苷酸序列且該3'內含子片段具有SEQ ID NO:15之核苷酸序列。

【0137】 如本文所提供，在一些實施例中，該3'內含子片段可包括間隔序列而使得該REP/CAP基因太大以致於無法包裝在AAV中。例如，該AAV包裝限制係大約5.0 kb。因此，可根據本文提供之揭露內容產生使該REP/CAP基因大於大約5.0 kb的間隔序列。在一些實施例中，該間隔序列係插入在該3'內含子片段中的2 kb隨機間隔子。因此，在以下實例中所述之例示性實施例中，該止擋匣係插入至人類 β -肌動蛋白基因之內含子中，使得該5'內含子片段具有SEQ ID NO:14之核苷酸序列且該3'內含子片段具有SEQ ID NO:66之核苷酸序列。然而，應理解的是該間隔序列不需要是2 kb，且可係導致REP/CAP基因大於大約5.0 kb的任何長度。

【0138】 任何AAV *rep*基因可包括在本發明之經修飾*rep*基因中。例如，該AAV *rep*基因可包含AAV1至AAV8中之一者之*rep*基因、或其混合。該AAV *rep*基因之序列可得自例如GenBank，對於各種AAV基因體具有下列之GenBank存取號：AAV1，GenBank存取號NC_002077.1；AAV2，GenBank存取號NC_001401.2；AAV3，GenBank存取號

NC_001729.1 ； AAV4 ， GenBank 存取號 NC_001829.1 ； AAV5 ， GenBank 存取號 NC_006152.1 ； AAV6 ， GenBank 存取號 AF028704.1 ； AAV7 ， GenBank 存取號 NC_006260.1 ； 及 AAV8 ， GenBank 存取號 NC_006261.1 。

【0139】 在以下實例中，使用具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至2202的人類AAV2之*rep*基因。

【0140】 在一些實施例中，可對該*rep*基因中之隱蔽性剪接位點進行修飾以移除此位點處之剪接。例如，可進行對DNA序列的同義突變，其中該DNA序列係經突變的，但該突變不改變經編碼的胺基酸。

具有經修飾AAV *rep*基因及AAV *cap*基因之建構體

【0141】 在另一大致態樣中，本申請案係關於非天然存在的核酸分子，該核酸分子包含本申請案之經修飾AAV *rep*基因及AAV *cap*基因、或其混合。較佳地，該AAV *cap*基因係該經修飾AAV *rep*基因的下游。

【0142】 在一個實施例中，該AAV *cap*基因進一步包含可操作地連接至該基因之編碼序列的多腺苷酸化信號。在以下實例中所述之例示性實施例中，在該AAV9 *cap*編碼序列下游包括AAV2多腺苷酸化信號(bp 4411-4466, NC_001401.2)。

【0143】 在另一實施例中，該AAV *cap*基因進一步包含強化子。在以下實例中，該AAV2多腺苷酸化信號下游包括AAV2 *rep* P5啟動子(bp 190-313, NC_001401.2)。

【0144】 在某些實施例中，該AAV *cap*基因編碼該等殼體蛋白VP1、VP2、及VP3之全部三者。

【0145】 在其他實施例中，該AAV *cap*基因編碼該等殼體蛋白之小

於三者。例如，據報導AAV血清型1至5可成功包裝、複製、及轉導沒有VP2的細胞(Grieger *et al.*, J Virol. 2005年八月；79(15): 9933-9944)。因此，在一個實施例中，AAV *cap*基因編碼AAV1至AAV5中任一者之VP1及VP3（但不編碼VP2）、或其混合。

【0146】 任何AAV *cap*基因皆可用於本發明中。例如，AAV *cap*基因可係AAV1至AAV8、AAV9、AAVDJ中之一者之*cap*基因、或其混合。在一個實施例中，該*cap*基因係AAV9變體。該AAV *cap*基因的序列可得自例如，GenBank。請參見上述對於AAV1至AAV8基因體之GenBank存取號。該AAV9基因體具有GenBank存取號No. AY530579.1，且該AAVDJ具有GenBank蛋白質存取號3J1Q_A。

【0147】 在一個實施例中，於以下實例中所述，使用具有GenBank存取號AY530579.1之核苷酸序列的人類AAV9之*cap*開讀框。
具有經修飾AAV *rep*基因、AAV *cap*基因、及轉殖基因之建構體

【0148】 在另一大致態樣中，本申請案係關於非天然存在的核酸分子，該核酸分子包含本申請案之經修飾AAV *rep*基因、AAV *cap*基因、及由AAV反向末端重複(ITR)側接之轉殖基因。

【0149】 ITR係AAV之生物學中重要的順式-活性序列。ITR的關鍵角色係在AAV DNA複製。除了在AAV複製中的作用以外，ITR對於AAV基因體包裝、轉錄、非許可條件下之負調節、及位點特異性整合亦係重要的。

【0150】 在一個實施例中，130 bp ITR包含SEQ ID NO:20之核苷酸序列，其衍生自3' AAV2 ITR（核苷酸4535-4664，NC_001401.2）用來側接該轉殖基因。在另一實施例中，使用較短的經突變ITR。例如，對

於較短的基因，ITR突變為較短的基因，並且該基因可折疊成雙股形式以增加表現並在感染後加速表現。請參見McCarty 2008 Mol Ther. 2008; 16(10):1648-56。

【0151】在另一實施例中，該轉殖基因包含促進劑，較佳的是哺乳動物細胞中之啟動子功能。在以下所述之實例中，人類EF1- α 啟動子（包括外顯子1、內含子1、及內含子2之部分）(SEQ ID NO:21)係包括在該轉殖基因中。

【0152】在另一實施例中，該轉殖基因包含多腺苷酸化信號。在以下所述之實例中，來自單純疱疹病毒胸苷激酶基因(SEQ ID NO:23)之多腺苷酸化信號係包括在該轉殖基因中。

【0153】在又一實施例中，非天然存在的核酸分子包含一對側接經修飾AAV *rep*基因、AAV *cap*基因、及經ITR側接的轉殖基因之絕緣子。在另一實施例中，非天然存在的核酸分子包含經修飾AAV *rep*基因上游之單一絕緣子、AAV *cap*基因、及ITR側接的轉殖基因。在一個實施例中，該絕緣子包含阻斷基因表現之染色質相關阻遏的基因體因子(Kwaks *et al.* (2003) Nature Biotechnology 21: 554-558; Kwaks *et al.* (2003) Nature Biotechnology 21: 822)。

【0154】可在本發明中使用任何合適的絕緣子，諸如本文中所述者。在一個實施例中，該絕緣子係具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40。在另一實施例中，該絕緣子係具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40。在另一實施例中，在另一實施例中，該絕緣子係具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04。在另一實施例中，該絕緣子係具有GenBank存取號AY190750.1之核苷酸

序列的抗抑制因子06。在另一實施例中，該絕緣子係具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07。在另一實施例中，該絕緣子係具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12。在另一實施例中，該絕緣子係具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13。在另一實施例中，該絕緣子係具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35。在另一實施例中，該絕緣子係具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36。在另一實施例中，該絕緣子係具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因子52。在另一實施例中，該絕緣子係具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53。在另一實施例中，該絕緣子係來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子。

【0155】 包含一對絕緣子之非天然存在的核酸分子可具有與所關注基因片段相同或不同的一對側接絕緣子。較佳地，不同的絕緣子係用來作為一對側接於所關注的基因片段。在以下實例所述之一個例示性實施例中，人類抗抑制因子40 (AY190756.1, SEQ ID NO:24)及小鼠抗抑制因子40 (SEQ ID NO:25)係用作為該等絕緣子。在以下實例所述之另一例示性實施例中，人類抗抑制因子40 (AY190756.1, SEQ ID NO:24)係用作該絕緣體。

【0156】 如本文所提供，本揭露之建構體亦可包括在該AAV轉殖基因兩側上的間隔序列，以降低錯誤包裝其他載體組分的風險。在一個實施例中，該非天然存在的核酸分子分別包含該轉殖基因上游及下游的第一及第二間隔序列。在某些實施例中，該等間隔序列係2 kb間隔序列。在一具

體實施例中，該非天然存在的核酸分子包含該經修飾AAV *rep*基因上游之該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第一及第二間隔序列，其中該第一絕緣子及該第二間隔序列係獨立地選自由下列所組成之群組：(a) SEQ ID NO:67之核苷酸序列；及(b) SEQ ID NO:68之核苷酸序列。

用於生產重組AAV的細胞及方法

【0157】 來自本申請案之經修飾AAV *rep*基因的該Rep蛋白的表現受到DNA剪接及RNA剪接機制兩者的嚴格控制，並因此使含有該經修飾*rep*基因的穩定宿主細胞得以在生物反應器中產生並生長至高編號。對於AAV生產，首先將含有經修飾AAV *rep*基因、AAV *cap*基因、及由ITR側接之轉殖基因的穩定宿主細胞生長至高編號，接著用在經修飾AAV *rep*基因中識別該等*attP*及*attB*位點的絲胺酸重組酶之複製缺陷型腺病毒轉染。藉由該絲胺酸重組酶所催化之在該等*attP*與*attB*位點之間的位點特異性重組將該止擋匣剪接出來，導致產生包含由功能性內含子分開之5'及3' *rep*編碼序列兩者的前驅mRNA。然後藉由普遍存在的細胞機器（剪接體）切除內含子，產生編碼四種Rep蛋白的mRNA，從而以高滴度生產AVV。

【0158】 可藉由將細胞以編碼該等基因之一或多個核酸分子的細胞轉導來獲得含有經修飾AAV *rep*基因、AAV *cap*基因、及由ITR側接之轉殖基因的穩定宿主細胞。在一個實施例中，藉由將細胞以編碼經修飾AAV *rep*基因及AAV *cap*基因之第一核酸分子轉導以獲得包含該經修飾AAV *rep*基因及AAV *cap*基因的第一宿主細胞，並且進一步將該第一宿主細胞以編碼由ITR側接之轉殖基因的第二核酸分子轉導來獲得穩定宿主細胞。在一個實施例中，該經修飾AAV *rep*基因及該AAV *cap*基因係穩定地

整合至該第一宿主細胞之染色體中。在另一實施例中，該經修飾AAV *rep*基因及該AAV *cap*基因在該第一宿主細胞中維持游離型。由ITR側接之轉殖基因亦可穩定地整合至宿主細胞中或維持游離型(episomal)。

【0159】 在另一實施例中，藉由將細胞以編碼經修飾AAV *rep*基因、AAV *cap*基因、及由ITR側接之轉殖基因的核酸分子轉導來獲得穩定宿主細胞。該經修飾AAV *rep*基因、該AAV *cap*基因及由ITR側接之轉殖基因可穩定地整合至宿主細胞中或維持游離型。

【0160】 穩定宿主細胞可在表現絲胺酸酶之腺病毒被感染之前，使該穩定宿主細胞生長至高細胞密度。

【0161】 使用鑒於本揭露之所屬技術領域中已知的任何方法，將表現本申請案之絲胺酸重組酶的複製缺陷型腺病毒引入至穩定的宿主細胞中。在一個實施例中，該複製缺陷型腺病毒係包含重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列編碼與SEQ ID NO:2之胺基酸序列至少85%同一、較佳的是與SEQ ID NO:2 100%同一的胺基酸序列。例如，該腺病毒可包含與SEQ ID NO:3至少85%同一的核苷酸序列，較佳的是與SEQ ID NO:3至少95%同一。

【0162】 如本文所揭示，本揭露亦包括用於增加AAV生產的方法及組成物，其藉由使本文所述之細胞與2-胺基嘌呤(2-AP)接觸。在該腺病毒生命週期的晚期階段，該病毒會抑制宿主蛋白質合成。這部分係由於晚期腺病毒100千道耳頓(kDa)蛋白的作用所致，該蛋白將Mnk1激酶從該cap-起始錯合物eIF4F中置換出來，導致eIF4E的去磷酸化並抑制cap-依賴性mRNA的轉譯（參見例如，Cuesta (2004), J. Virology 78: 7707-7716）。腺病毒晚期基因轉錄物在其5'末端包含一個三方前導序列，該序

列藉由所謂的核糖體分流(shunting)機制啟動轉譯（參見例如，Yueh (2000) *Genes Dev* 14: 414-421）。在AAV生產細胞系的環境下，預期cap依賴性轉譯的抑制會阻斷AAV REP及CAP基因以及AAV複製及包裝所需的早期腺病毒蛋白的表達。因此，在一些實施例中，使用腺病毒誘導劑，用阻斷宿主蛋白質轉譯關閉的化學物質培養細胞，以提高AAV生產細胞系的效率。

【0163】 在某些實施例中，該阻斷宿主蛋白質轉譯關閉的化學品係2-胺基嘌呤(2-AP)。已顯示2-AP阻斷由腺病毒誘導之宿主蛋白質合成的關閉（參見例如，Zhang and Schneider (1994) *J. Virology* 68: 2544-2555; Huang and Schneider (1990) *PNAS* 87: 7115-7119）。用2-AP處理生產AAV的細胞能夠減少感染的細胞病變作用，包括恢復通常被晚期感染所降解的細胞角蛋白網絡(Zhang and Schneider (1994) *J. Virology* 68: 2544-2555)。2-AP在體外抑制許多激酶，包括RNA依賴性蛋白激酶PKR（亦稱為真核轉譯起始因子2 α 激酶2，EIF2AK2）(DeBenedetti (1983) *J Biol Chem*, 258: 14556-14562)，但無法阻斷腺病毒感染後發生的細胞中之PKR活化及eIF-2 α 的磷酸化(Huang and Schneider (1990) *PNAS* 87: 7115-7119)。2-AP確實將早期腺病毒DNA-結合蛋白(DBP)位準提高了10至20倍，而mRNA位準卻沒有提高(Huang and Schneider (1990) *PNAS* 87: 7115-7119)，與對cap依賴性轉譯的影響一致。

【0164】 因此，在一些實施例中，生產包含轉殖基因之重組AAV的方法包括用2-胺基嘌呤培養本揭露之細胞。在一些實施例中，該2-胺基嘌呤濃度係小於約10 mM。在一些實施例中，該2-胺基嘌呤濃度係小於約5 mM。在一些實施例中，該2-胺基嘌呤濃度係小於約2.25 mM。在一些實

施例中，該2-胺基嘌呤濃度係小於約1.25 mM。在一些實施例中，該2-胺基嘌呤濃度係約1 μ M至約1.25 mM。在一些實施例中，該2-胺基嘌呤濃度係約10 μ M至約1.25 mM。在一些實施例中，該2-胺基嘌呤濃度係約100 μ M至約1.25 mM。在一些實施例中，該2-胺基嘌呤濃度係約1.25 mM。

【0165】 在具體實施例中，本揭露之細胞係在用重組腺病毒感染後約24小時與2-胺基嘌呤接觸。在一些實施例中，本揭露之細胞係在用重組腺病毒感染後約20小時與2-胺基嘌呤接觸。在一些實施例中，本揭露之細胞係在用重組腺病毒感染後約12小時與2-胺基嘌呤接觸。在一些實施例中，本揭露之細胞係在用重組腺病毒感染後約30小時與2-胺基嘌呤接觸。在一些實施例中，本揭露之細胞係在用重組腺病毒感染後約36小時與2-胺基嘌呤接觸。在一些實施例中，本揭露之細胞係在用重組腺病毒感染後約48小時與2-胺基嘌呤接觸。

實例

【0166】 本申請案的下列實例是要進一步說明本申請案的本質。所屬技術領域中具有通常知識者將領會的是，能夠對以上所述的實施例進行變更而不違背其廣義的發明概念。因此，應了解本發明並未受限於揭示之具體實施例，而是意欲涵蓋如本實施方式所定義之屬於本發明之精神及範疇內的修改。

材料

【0167】 細胞：HEK293細胞（American Type Culture Collection (ATCC), Manassas, VA，目錄號CRL-1573）；PEAK-rapid（ATCC, Manassas, VA，目錄號CRL2828）。

【0168】 組織培養基及試劑：OptiMEM培養基（Thermo-fisher, Waltham, MA；目錄號31985-062）；DMEM，高葡萄糖（Thermo-fisher，目錄號10569-010）；DMEM，無酚紅（Thermo-fisher；目錄號A14430-01）；Hyclone透析胎牛血清（Thermo-fisher；目錄號SH30079.03）；96孔TC盤（Corning, Corning NY；目錄號3596）；6孔組織培養盤，清透（Corning，目錄號3516）；培養盤96，不透明白色（PerkinElmer, Waltham, MA；目錄號6005680）；TrypLE選擇細胞解離試劑（Thermo-fisher，目錄號12563-011）；達爾伯克(Dulbecco's)磷酸鹽緩衝鹽水，無鈣，無鎂，D-PBS（Thermo-fisher，目錄號14190-144）；遺傳黴素(Geneticin)(G418) 50 mg/ml（Thermo-fisher，目錄號10131-027）；來自白黑鏈黴菌(*Streptomyces alboniger*)之嘌呤黴素二鹽酸鹽(Sigma Aldrich P9620)；T150組織培養瓶150 mm²（Corning，目錄號CLS430825）；GlutaMax 100x（Thermo-fisher，目錄號35050-061）；經非組織培養處理之6孔培養盤（Corning，目錄號351146）；Hyperflask M容器（Corning，目錄號10030）；2.5% ClonaCell甲基纖維素於DMEM中（沒有麩醯胺酸且含有葡萄糖、丙酮酸鈉、及碳酸氫鈉）（StemCell Technologies, Vancouver, British Columbia, Canada目錄號03899-DI）。

【0169】 轉染試劑：Fugene-HD轉染試劑（Promega, Madison WI，目錄號E2311）；脂染胺3000轉染試劑（Thermo-fisher目錄號L3000008）；去氧核苷酸（Millipore-Sigma, St. Louis, MO，目錄號D7295-2ML）。

【0170】 試管：15 ml錐形管（Corning，目錄號430053）；1.5 ml

螺桿蓋試管（Sarstedt AG & Co. KG, Germany，目錄號72.692.005）。

【0171】純化套組及分析試劑：質體自旋小量製備套組（Qiagen, Hilden, Germany，目錄號27106）；CHROMA SPIN™+TE-1000管柱（Takara Bio USA, Mountainview CA，目錄號636079）；Dual-Glo螢光素酶分析系統（Promega, Madison WI，目錄號E2940）；銀染色套組（Thermo-fisher目錄號24600）；具有Phase-maker tube的Trizol Plus RNA純化套組（Thermo-fisher Catalog Number目錄號A33254）；不含DNA的套組（Thermo-fisher目錄號AM1906）；Nucleospin凝膠及PCR清除套組（Takara Bio USA，目錄號740609.5）。

【0172】酶：SSpe I-HF（New England Biolabs, Ipswich, MA，目錄號R3133S）；來自牛胰臟的DNase I grade II（Sigma-Aldrich，目錄號10104159001）；NEXT Ultra II Q5 Master Mix（New England Biolabs，目錄號M05445S）。

【0173】緩衝液及化學藥品：CutSmart®緩衝液（1X緩衝液組分：50 mM的乙酸鉀、20 mM Tris-乙酸酯、10 mM的乙酸鎂、100 µg/ml BSA、pH 7.9@25℃）（New England Biolabs，目錄號B7204S）；Benzonase核酸酶（Sigma-Aldrich，目錄號E1014-25K）；含有1.5 mM MgCl₂之10x GeneAmp PCR緩衝液I（Thermo-fisher目錄號N8080006）；去氧膽酸鈉（Sigma-Aldrich，目錄號D6750-25g）；1M TRIS-HCL PH8.5（Thermo-fisher，目錄號T1085）；10x GeneAmp PCR緩衝液I（Thermo-Fisher目錄號N8080006）[100 mM Tris-HCl, pH 8.3（在25℃下）；500 mM KCl; 15 mM MgCl₂; 0.01%明膠於高壓滅菌、去離子、經超過濾水]中；10% Pluronic F-68（Thermo-Fisher目錄號

24040-032) ; 剪切鮭魚精子DNA (10 mg/ml) (Thermo-Fisher目錄號AM9680) ; 病毒稀釋緩衝液(VDB) [1x GeneAmp PCR緩衝液I、2 µg/ml的剪切鮭魚精子DNA、及0.05% Pluronic F-68) ; β-巰基乙醇 (Sigma-Aldrich, 目錄號M3148) ; 腺病毒配方緩衝液 (10 mM Tris (pH 7.4)、1 mM MgCl₂、75 mM NaCl、5% 蔗糖、0.02% 聚山梨醇酯80、0.1 mM EDTA、10 mM組胺酸、0.5% EtOH) ; 2-氨基嘌呤，硝酸鹽 (Sigma-Aldrich, 目錄號A2380) , 於DMEM+ 2% FBS中溶解至100 mM。

【0174】 RT-PCR試劑：SuperScript III第一股合成系統 (Thermo-fisher目錄號188080-051) ; Q5熱啟動高保真度2X Master Mix (New England Biolabs, 目錄號M0494S) ; 具有溴化乙錠之1% TAE Mini ReadyAgarose Gel (Bio-RAD, 目錄號1613016) ; Dark Reader藍光透照器 (Clare Chemicals, Dolores, CO, 目錄號DR46B))

【0175】 數位液滴PCR：探針用2X SuperMix (Bio-Rad目錄號186-3026) ; DG32 AutoDG卡匣 (Bio-Rad目錄號1864108) ; 自動液滴產生器油於PBS中 (Bio-Rad目錄號1864110) ; 液滴讀取器油 (Bio-Rad目錄號1863004) ; Eppendorf twin.tec 96孔PCR盤 (目錄號951020346) ; 自動化液滴產生器 (Bio-Rad目錄號186-4101) ; QX200液滴讀取器 (Bio-Rad目錄號186-4003) ; 具有深孔反應模組之C1000Touch熱循環儀 (Bio-Rad, 目錄號185-1197) 。

【0176】 PrimeTime qPCR分析物：這些分析物的20x存料由正向及反向PCR引子 (在18 µM下) 及含有螢光淬滅劑ZEN及Black Hole Quencher 1 (3IABkFQ)及FAM或HEX螢光報導子染料 (在5 µM下) 的5'核酸酶探針所組成。分析物係由Integrated DNA Technologies, Inc.,

Coralville IA合成。用於qPCR分析物之引子及探針序列如下：

mCherry：引子1 (SEQ ID NO:36, 5'-CTGTTCCACGATGGTGTAGTC-3')；引子2 (SEQ ID NO:37, 5'-TGAGGTCAAGACCACCTACA-3')；探針 (SEQ ID NO:38, 5'-FAM-TTGGACATC-ZEN-ACCTCCCACAACGAG-3IABkFQ-3')；

腺病毒外顯子2 (Ad5E2)：引子1 (SEQ ID NO:39, 5'-GGGTGATGCAGTAGAAGGTAAG-3')；引子2 (SEQ ID NO:40, 5'-ATGAAGTTCGGCGGAGATG-3')；探針 (SEQ ID NO:41, 5'-HEX-TC TTGTTCC-Zen-CAGCGGTCCCATC-3IABkFQ-3')；

P5 (AAV之P5啟動子區域)：引子1 (SEQ ID NO:42, 5'-GTGGTCACGCTGGGTATTTA-3')；引子2 (SEQ ID NO:43, 5'-GGGACCTTAATCACAATCTCGT-3')；探針 (SEQ ID NO:44, 5'-FAM-TTTGAAGCG-ZEN-GGAGGTTTGAACGC-3IABkFQ-3')；

AAV REP 基因：引子1 (SEQ ID NO:45, 5'-GTCCGTGAGTGAAGCAGATATT-3')；引子2 (SEQ ID NO:46, 5'-TTCGATCAACTACGCAGACAG-3')；探針 (SEQ ID NO:47, 5'-FAM-TCTGATGCT-ZEN-GTTTCCCTGCAGACA-3IABkFQ-3')；

AAV9 CAP 基因：引子1 (SEQ ID NO:48, 5'-CCGGGTCCAAGGTATTTGTAA-3')；引子2 (SEQ ID NO:49, 5'-CTCAACCCAAGGCAAATCAAC-3')；探針 (SEQ ID NO:50, 5'-FAM-ACATCAAGA-ZEN-CAACGCTCGAGGTCT-3IABkFQ-3')；

及

β內醯胺酶（安比西林抗性）基因：引子1 (SEQ ID NO:51, 5'-

CCAGAAACGCTGGTGAAAGTA-3')；引子2 (SEQ ID NO:52, 5'-CTCAAGGATCTTACCGCTGTTG-3')；探針 (SEQ ID NO:53, 5'-FAM-TGCACGAGT-ZEN-GGGTTACATCGAACT-3IABkFQ-3')。

【0177】 PAGE電解質：4x NuPAGE LDS樣本緩衝液（Thermo-Fisher，目錄號NP0007）；4-12% Bis-Tris PAGE凝膠於1x MOPS運行緩衝液中（Thermo-Fisher，目錄號NP0322PK2）；20x NuPAGE MOPS SDS運行緩衝液（Thermo-Fisher，目錄號NP0001）。

【0178】 AAV純化緩衝液及供應商：0.2 μ m PES的膜過濾器（Thermo-Fisher目錄號567-0020）；0.5 $\times$ 5 cm POROS GoPure層析管柱，用POROS CaptureSelect AAVX樹脂預填充（Thermo-fisher目錄號A36652）；Amicon 15 100kDa MWCO過濾器（Millipore-Sigma目錄號UFC910024）；CIM QA Disk 0.34 ml 體積 (BIA Separations, Slovenia)；緩衝液A (20 mM Tris pH7.5, 400 mM NaCl)；緩衝液B（25 mM Tris pH7.5，40 mM NaCl，及1.5 mM MgCl₂）；緩衝液C（20 mM 檸檬酸鈉pH 2.5，400 mM NaCl）；緩衝液D（100 mM檸檬酸鈉，10 mM Tris，pH 8.0）；緩衝液E (20 mM BTP pH10.0, 0.001% Pluronic F68, 10 mM NaCl)；緩衝液F（20 mM Bis-TRIS丙烷pH 10.0，0.001% Pluronic F68，400 mM NaCl）；Bis-TRIS丙烷(BTP)（Millipore Sigma目錄號B4679）。

【0179】 其他設備：AKTA Explorer FPLC系統 (GE Healthcare Life Sciences, Marlborough, MA)；AKTA純化系統 (GE Healthcare Life Sciences)；Envision 多標記讀取器型號2104 (PerkinElmer, Waltham, MA)。

SR21重組酶之識別及重組表現

【0180】 用SPBeta c2整合酶蛋白(Query, SEQ ID NO:1)作為查詢，在NCBI之非冗餘蛋白資料庫進行BLAST搜索，在沙福芽孢桿菌菌株CCMA-560中鑑定出推定絲胺酸重組酶(Sbjct, SEQ ID NO:2)在蛋白質位準處具有64%序列同一性（圖1）。該推定絲胺酸重組酶或整合酶係推定原噬菌體插入的一部分。將此重組酶命名為SR21（絲胺酸重組酶(Serine Recombinase) 21）。編碼SR21之DNA序列係顯示於SEQ ID No:3。

【0181】 藉由使用來自重組酶編碼區的3'端之CCMA-560 DNA序列作為查詢(SEQ ID NO:58)，在NCBI之序列數據庫進行BLAST搜索，識別出與CCMA-560密切相關且不含有該原噬菌體插入（該「Fairview」菌株）之細菌菌株（圖2）。對應於CCMA-560中之推定原噬菌體插入位點上游及下游序列之Fairview菌株的DNA序列，在本文中稱為「插入前序列(pre-insertion sequence)」，且係顯示於SEQ ID NO:4。使用此序列(SEQ ID NO:4)作為對BLAST的查詢，CCMA-560菌株的基因體序列識別了其他噬菌體-宿主DNA接合點94 kb上游。沙福芽孢桿菌菌株CCMA-560之右及左原噬菌體-宿主DNA接合點的序列分別顯示於SEQ ID NO:5及SEQ ID NO:6中。

【0182】 藉由交換中央相同區域（「ACTGACAAAGCGGT」）(SEQ ID NO:54)上游的序列，分別從該宿主DNA接合點(SEQ ID NO:5)及(SEQ ID NO:6)重構SR21重組酶*attP*及*attB*序列，並選擇最大化該雙對稱之該中央二核苷酸及*att*位點邊界：*attP* (SEQ ID NO:7)；*attB*-CCMA-560 (SEQ ID NO:8)。衍生自沙福芽孢桿菌之Fairview菌株的宿主DNA接合點(SEQ ID NO:4)之*attB*序列(SEQ ID NO:9)含有相對於來自菌株

CCMA-560 (SEQ ID NO:8)之經重構 $attB$ 序列的兩個失配。圖3顯示 $attP$ 與這兩個交替的 $attB$ 序列比對，突顯雙對稱的位置。

測量哺乳動物細胞中的重組酶活性

【0183】 藉由基因合成(GENEWIZ, Plainfield, NJ)建構載體(P175) (SEQ ID NO:10)，以在CMV啟動子的控制下在哺乳動物細胞中表現SR21重組酶，接著是SV40多腺苷酸化信號。該SR21重組酶開讀框係與SEQ ID NO:3相同，除了該細菌轉譯起始密碼子「TTG」係由「ATG」置換，並且引入三個點突變以破壞限制性核酸內切酶識別位點。該開讀框中之此等變化不會導致經編碼SR21重組酶胺基酸序列的任何變化。

【0184】 亦藉由基因合成建構重組酶報導子質體(P41) (GENEWIZ, Plainfield, NJ) (SEQ ID NO:11；圖4)。其編碼由EF1 α 啟動子所驅動的組成性表現之綠色螢光蛋白(GFP)-自切割F2A-海腎螢光素酶(rLUC)融合蛋白。其亦以相對於該CMV啟動子之反義定向編碼由SR21 $attP$ (SEQ ID NO:7)及 $attB$ (SEQ ID NO:9)信號側接之經重組酶活化-mCherry-自切割P2A-螢火蟲螢光素酶(fLUC)報導基因。當SR21重組酶重組該等 $attP$ 及 $attB$ 信號時，該編碼區經反轉成正義定向，並且表現該mCherry-P2A-fLUC蛋白（請參見圖4）。

【0185】 為了測量人類細胞中之SR21重組酶活性，將75,000個HEK293細胞接種至100 μ l高葡萄糖DMEM + 10%胎牛血清的96孔組織培養盤的各孔中。如表1所示，將該重組酶報導子質體(P041) $\pm$ 該SR21重組酶表現質體(P175) +去氧核苷酸（以標準化DNA量）與Fugene-HD轉染試劑於OptiMEM培養基中在室溫下複合15分鐘，並一式三份轉染至所接種細胞的孔中。將盤在37°C下培養48小時。

[表1]轉染條件

樣本	P41重組酶報導子	P175重組酶表現質體	去氧核苷酸	Fugene-HD	OptiMEM
1	4 µl (100 ng)	無	12 µl (300 ng)	1.2 µl	22.8 µl
2	4 µl (100 ng)	4 µl (100 ng)	8 µl (200 ng)	1.2 µl	22.8 µl

【0186】 使用來自Promega的Dual Glo分析套組，依序在轉染孔中分析螢火蟲螢光素酶(fLUC)及海腎螢光素酶(rLUC)。將培養基從該組織培養盤的轉染孔中移除，並添加100 µl之DMEM培養基（無酚紅）與Dual Glo螢光素酶+ fLUC基質的1:1混合物。將盤在室溫下培養10分鐘。將溶解產物轉移至不透明的白色96孔中。使用Envision多標記讀取器測量fLUC活性。接下來，每孔添加50 µl的Stop-and-Glo緩衝液+海腎基質，並將該盤以溫和振盪培養10分鐘。在相同Envision讀取器上讀取該海腎螢光素酶信號。

【0187】 結果：當與該重組酶表現質體共轉染時，該重組酶報導子產生的螢火蟲螢光素酶比與去氧核苷酸共轉染時產生的螢火蟲螢光素酶多1535倍（表2）。此差異不能藉由不同的轉染效率來解釋，因為在單獨的報導子轉染中海腎螢光素酶(rLUC)活性高5倍。此數據證實SR21重組酶在人類細胞中係高度活性的，並且該結果代表三個獨立的實驗。

[表2]在HEK293細胞中的重組酶活性

樣本	說明	fLUC	rLUC	fLUC活性增加倍數
1	僅報導子	4.3E03 ± 1.2E03	1.4E07 ± 2.6E06	
2	報導子+重組酶	6.6E06 ± 3.9E05	3.4E06 ± 1.6E05	1535

建構REP/CAP +轉殖基因質體

【0188】 若該AAV複製(REP)及殼體(CAP)基因可穩定地整合，則哺乳動物細胞中之AAV的大規模生產可係可能的，且之後經誘導以在高

密度培養物中生產AAV。然而，REP蛋白的表現係毒性的，使其難以在表現REP基因的宿主中培養穩定的細胞系，諸如表現該等腺病毒E1基因（諸如HEK293細胞）者。野生型AAV編碼四種具有重疊讀框的REP蛋白，這是由於使用兩種啟動子及交替剪接所致。因此，使用誘導性啟動子來控制REP表現並不容易。先前的工作證實，在人工內含子內部的REP編碼區插入「止擋匣」，可在HEK293細胞中產生穩定細胞系(Qiao *et al.* (2002) *J. Virol.* 76: 13015; Yuan *et al.* (2011) *Hum Gene Therap.* 22: 613-624)。使用藉由腺病毒感染所遞送的Cre重組酶切除止擋匣恢復了REP表現並起始了ITR側接的轉殖基因之AAV複製。在此實例中，在含有ITR側接的轉殖基因的質體中之REP/CAP表現匣的環境下，建構了經重組酶活化的REP基因之經改良版本。

【0189】使用AAV2 REP基因（人類AAV2之bp 190-2202，NC_001401.2）、接著是AAV9 CAP開讀框(AY530579.1)、AAV2多腺苷酸化信號(bp 4411-4466, NC_001401.2)、及AAV2 REP P5啟動子的第二個拷貝(bp 190-313, NC_001401.2)建構了AAV REP/CAP9表現匣(SEQ ID NO:13)。

【0190】使用剪接預測軟體 (NetGene2 at www.cbs.dtu.dk/services/NetGene2/; Brunak, S., Engelbrecht, J., and Knudsen, S.: Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence, *Journal of Molecular Biology*, 1991, 220, 49-65.)以選擇合適的位置，將來自人類 β -肌動蛋白基因的內含子插入至REP編碼區。將內含子插入AAV2 (NC_001401.2)中核苷酸編號1052的下游，該區域係所有四種REP轉錄本共有的區域。該內含子及該插入位置兩者皆

不同於Qiao *et al.* (2002) J. Virol. 76: 13015)所使用者。隨後將止擋匣插入在該 β -肌動蛋白內含子的上游及下游半部兩者之間（分別為SEQ ID 14及SEQ ID 15）。

止擋匣

【0191】 轉錄止擋匣(SEQ ID NO:16)係由下列因子組成：

- SR21 *attP* (SEQ ID NO:7)
- 來自小鼠 HPRT 基因的強剪接受體 (SEQ ID NO:17) (NC_000086.7，核苷酸53001998至53002138)，加上來自人類刺豚鼠傳訊蛋白之29 nt區域（NC_000020.11，核苷酸34262765至34262793）。將此包括在內以預防該止擋匣從初級mRNA轉錄本中剪接出來。
- 新黴素磷酸轉移酶表現匣(SEQ ID NO:18)係藉由哺乳動物啟動子（例如，小鼠磷酸甘油酯激酶1）及細菌(Lac *zya*)啟動子並接著由來自SV40之多腺苷酸化信號所驅動。該基因分別在哺乳動物及細菌細胞中賦予對新黴素及康黴素的抗性。
- 來自編碼自切割RNA模體的多腺苷酸化信號下游的人類 β -球蛋白基因之序列，對於有效的轉錄終止係重要的 (Teixeira *et al.* (2004) Nature 432: 526-30; SEQ ID No:19)。
- SR21 *attB* (SEQ ID NO:8)。

AAV轉殖基因

【0192】 在該AAV REP/CAP區下游之P439載體(SEQ ID NO:12)中編碼AAV反向末端重複(ITR)側接的轉殖基因。130 bp ITR (SEQ ID NO:20)係衍生自3' AAV2 ITR（核苷酸4535至4664，NC_001401.2）且

插入在HPRT-E2A-mCherry轉殖基因上游及該轉殖基因的反向3'。

【0193】 該轉殖基因由人類EF1- α 啟動子（包括外顯子1、內含子1、及內含子2之部分）(SEQ ID NO:21)、編碼mCherry-自剪接E2A連接子-人類HPRT融合基因之序列(SEQ ID No:22)、及來自單純疱疹病毒胸苷激酶基因(Herpes Simplex Virus Thymidine Kinase Gene) (SEQ ID NO:23)之多腺苷酸化信號組成。

絕緣子

【0194】 REP/CAP及ITR-轉殖基因因子係由阻斷基因表現之染色質相關阻遏的基因體因子側接(Kwaks *et al.* (2003) Nature Biotechnology 21: 554-558; Kwaks *et al.*(2003) Nature Biotechnology 21: 822)：人類抗抑制因子40 (AY190756.1, SEQ ID No:24)及小鼠抗抑制因子40 (SEQ ID NO:25)。

質體主鏈

【0195】 該質體主鏈含有下列因子：

- 哺乳動物嘌呤黴素抗性基因匣係建構自疱疹病毒胸苷激酶基因啟動子(SEQ ID NO:26)、嘌呤黴素N-乙酰基轉移酶編碼區(SEQ ID NO:27)、及來自牛生長激素基因之多腺苷酸化信號(SEQ ID NO:28)。
- 一種艾司坦-巴爾病毒(Epstein Barr Virus, EBV) OriP複製起點片段(SEQ ID NO:29)，其代表EBV之「雙對稱」區域及「重複家族」區域之組合
- pUC57載體序列編碼質體複製起點及安比西林(ampicillin)抗性基因(SEQ ID NO:30)。

【0196】 在SEQ ID NO:12中給出完整質體P439的序列。

藉由SR21重組酶移除止擋匣之測試效率

【0197】 為了測試該止擋匣是否可藉由SR21重組酶在人類細胞中精確地移除，根據製造商之說明，使用脂染胺(Lipofectamine) 3000將載體P439 (SEQ ID NO:12)及SR21重組酶表現載體P175 (SEQ ID NO:10)共轉染至PEAK-Rapid細胞中，並且在37℃下於5% CO₂中在含有DMEM及10% FBS的培養基中培養三天。移除培養基，將細胞用D-PBS洗滌一次，然後在37℃下用TrypLE培養5分鐘。將細胞轉移至無菌微量離心管中，藉由離心製粒，用1 ml D-PBS洗滌一次並再次製粒。藉由使用經設計以用於從細菌單離質體的Qiagen Spin Miniprep套組，藉由鹼裂解回收游離的質體。

【0198】 為了破壞未重組的質粒DNA，將等分試樣的回收DNA用Spe I-HF酶在1x CutSmart緩衝液中在37℃下消化1小時並在80℃下消化20分鐘。使用NEXT Ultra II Q5 Master Mix，在下列循環條件下，用引子P349F3 (SEQ ID NO:32)及P349R9 (SEQ ID NO:33)使回收的DNA經受PCR擴增：98℃ 1 min; 35x (98℃ 10s, 72℃ 10s); 5 min 72℃。當在1%瓊脂糖凝膠上進行電泳時，觀察到單一具有預期大小的PCR產物。將PCR產物藉由粒徑篩析層析法使用CHROMA SPIN™+TE-1000管柱純化。使用用於PCR (GeneWiz)之相同引子將PCR產物定序。所得之序列(SEQ ID NO:34)顯示，藉由SR21重組酶透過重組該等attP (SEQ ID NO:7)及attB (SEQ ID NO:8)序列，該止擋匣已從質體P439精確移除，產生attL重組序列(SEQ ID NO:35)。

表現SR21重組酶的重組腺病毒血清型5之構造

【0199】 重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒係藉由在PER.C6細胞中的同源重組程序在Batavia Biosciences (Leiden, the Netherlands)產生(Fallaux *et al.* (1998) Hum Gene Ther. 9: 1909-1917)，如先前所述用於生產E1缺失的載體(Havenga *et al.* (2001) J. Virol 75:3335-3342)，除了使用缺少E3區域之經修飾黏接質體之外（pWE/Ad5.AflII-rITRsp. $\Delta E3$ ，專利US6340595B1）。PER.C6細胞與此黏接質體及含有1至454之Ad5序列（左邊ITR及包裝信號）的質體P321 (SEQ ID NO:31)、含有巨細胞病毒(CMV)啟動子(nt -672至+15)之轉殖表現匣、該SR21重組酶編碼區、猴病毒40 (SV40)多腺苷酸化信號(NC_001669.1, nt 2550至2774)、及第二Ad5序列（範圍從nt 3511至6095）共表現。在PER.C6細胞中的P321 Ad5序列(nt 3511-6095)與黏接質體pWE/Ad5.AflII-rITRsp. $\Delta E3$ 之間的同源重組產生重組腺病毒。藉由兩步驟CsCl-梯度帶化程序獲得經純化的病毒原液，並將經單離之病毒原液透析至腺病毒配方緩衝液（10 mM Tris (pH 7.4)，1 mM MgCl₂，75 mM NaCl，5%蔗糖，0.02%聚山梨醇酯80，0.1 mM EDTA，10 mM組胺酸，0.5% EtOH）。

穩定的細胞系產生

【0200】 根據製造商的說明，使用脂染胺3000將質體P439 (SEQ ID NO:12)轉染至黏附的PEAK-RAPID細胞，並在37°C下將其在T25培養瓶中於DMEM+ 10% FBS+ 0.05 mg/ml遺傳黴素中培養。24小時之後，用TrypLE處理細胞並將其轉移至含有DMEM+ 10% FBS+ 0.05 mg/ml遺傳黴素+ 0.5 μ g/ml嘌呤黴素之T75培養瓶中。將細胞以每週1:10分至相同培養基中連續兩週。在轉染後第三週時，將細胞每週以1:10分在含有DMEM+ 10% FBS+ 0.05 mg/ml遺傳黴素+ 5.0 μ g/ml嘌呤黴素之培養基

中達三週。

【0201】 藉由將細胞稀釋至1% ClonaCell甲基纖維素於DMEM + 30% FBS+ 1x GlutaMax+ 5 µg/ml嘌呤黴素+ 0.05 mg/ml遺傳黴素中，接種至非經組織培養處理之6孔盤中，並在37°C下培養三週來生產單細胞殖株。使用移液器將殖株自甲基纖維素盤轉移至96孔經TC處理之盤中，該盤含有DMEM+ 10% FBS+ 0.05 mg/ml遺傳黴素+5.0 µg/ml嘌呤黴素。藉由標準方法在相同培養基中擴增殖株。

篩選殖株

【0202】 為了篩選生產AAV的殖株，將細胞以一式兩份接種至100 µl DMEM+ 10% FBS中之96孔盤中，並在37°C下培養整夜。將SR21腺病毒稀釋成每ml無血清DMEM中1E8個病毒基因體。用100 µl稀釋的腺病毒置換來自經接種細胞的培養基，並將盤在37°C下培養四天。藉由添加10 µl之下列混合物來裂解細胞：5% 去氧膽酸鹽於PBS中+ 10單位Benzonase。將盤在37°C下培養2小時。將盤以3000 rpm離心5分鐘以沉澱細胞碎屑，且上清液中之AAV病毒藉由數位液滴PCR (ddPCR)定量。

數位液滴PCR (ddPCR)

【0203】 ddPCR 定量係基於 Lock *et al.* (2014) Human Gene Therapy methods 23: 115-125所述之方法。在20 µl含有1x PCR緩衝液+ 20 mM Tris pH 8.5 + 8單位DNAse I的反應中，將二µl的裂解液在37°C下之熱循環儀的96孔盤中用DNAse消化1小時。將2 µl的DNAse消化樣本用98 µl病毒稀釋緩衝液(VDB)稀釋，並且將2 µl的該稀釋液添加至ddPCR反應中，該ddPCR反應含有用於mCherry轉殖基因之1x PCR SuperMix + 1x PCR引子/探針（參見材料章節）。使用Bio-Rad自動化液滴產生器形成

ddPCR液滴。PCR循環如下：95℃ 10 min；42x（94℃ 30s，60℃ 1 min，72℃ 15s，所有三個循環時間為每秒2℃）；98℃ 10min；保持4℃。依據製造商之說明，在Bio-Rad液滴讀取器上偵測到FAM螢光。使產生最高DNAse抗性粒子且偵測為FAM螢光陽性液滴之殖株經歷擴增及進一步篩選。

篩選殖株-第二次分析

【0204】將待篩選之殖株的1.25E6個細胞接種至於3 ml MEM+ 10% FBS中之6孔盤的單孔中，並在37℃下培養2天。將生長培養基用含有5E8個Ad5-SR21病毒粒子之3 ml DMEM+ 10% FBS置換。將盤返回至37℃培養3天。將細胞與培養基轉移至15 ml試管中並使其經歷3個冷凍解凍循環（乾冰/37℃培養），接著在3000 rpm下離心5分鐘以將細胞碎屑製粒。使2 µl的各樣本經歷DNAse消化並用上述的mCherry分析進行ddPCR定量。在用Ad5-SR21病毒感染後，P439C4細胞產生最多AAV，且經選擇以用於進一步表徵（表3）。

[表3]篩選分析中之AAV生產

殖株#	總AAV（DNAse-抗性粒子）
殖株1	4.4E+08 ± 7.6E+07
殖株3	2.5E+09 ± 3.4E+08
殖株4	4.4E+09 ± 5.6E+07
殖株5	1.8E+09 ± 1.9E+08
殖株12	7.5E+08 ± 9.8E+07
殖株18	1.6E+08 ± 2.6E+07
殖株20	9.8E+07 ± 1.4E+07
殖株25	1.2E+09 ± 6.4E+07
殖株28	1.3E+08 ± 1.2E+07
殖株32	1.5E+09 ± 1.3E+08

植株36	2.8E+08 ± 3.0E+07
植株41	1.0E+09 ± 8.9E+07

時間進程實驗

【0205】 進行新的實驗以判定在兩個不同生長溫度下培養基中之AAV生產及分泌之動力學。將兩mL非酶解離溶液添加至T150培養瓶中經PBS-洗滌之P439-C4細胞單層，並且將燒瓶在37℃下培養5分鐘。將培養瓶用8 ml的DMEM+10% FBS洗滌，並且將細胞轉移至50 ml的離心管中。將細胞以1500 rpm離心5分鐘，，並將團塊再懸浮於DMEM+2% FBS中。將細胞以相同培養基稀釋至每毫升1.25E6個細胞。將四ml的細胞接種至四種6孔盤的各孔中。將1 ml (2E8 vp)於DMEM+ 2% FBS中的Ad5-CMV-SR21腺病毒添加至孔中。將兩盤在37℃下培養，且將兩盤在32℃下在5% CO2下培養。每天使用細胞刮刀回收細胞及培養基持續8天，以移除附接的細胞，並將樣本轉移至15 ml錐形管中。將試管以3000 rpm旋轉5分鐘，並將等分試樣轉移至1.5 ml螺旋蓋管中並在-20℃下冷凍直到ddPCR分析。

【0206】 將樣本如上所述以一式兩份進行處理，並且將各經DNAse處理的樣本在VDB中製成三份連續稀釋液。在ddPCR反應中將樣本定量，該反應含有1x PCR Master Mix + 1x mCherry-FAM分析物+ 1x Ad5E2-HEX分析物（請參見材料章節）。如上所述進行ddPCR。

結果：

【0207】 在細胞培養基中的腺病毒及AAV在8天時程期間增加（表4）。腺病毒複製在32℃下較慢，導致較高的AAV生產，這可能是由於腺病毒的細胞病變作用延遲所致。每個細胞在32℃下的AAV產量超過14,000個基因組拷貝。

[表4]經過8天時間進程之AAV及腺病毒產率

AAV GC/細胞			Ad5 GC/細胞	
感染後天數	32℃	37℃	32℃	37℃
1	4 ± 1	22 ± 0	8 ± 1	161 ± 4
2	11 ± 2	174 ± 26	223 ± 14	1319 ± 292
3	2009 ± 137	803 ± 110	2028 ± 155	4941 ± 75
4	4275 ± 274	975	2600 ± 168	5806 ± 159
5	7406 ± 309	1672 ± 49	5041 ± 446	12000 ± 442
6	7469 ± 75	4109 ± 214	4781 ± 610	33344 ± 1781
7	10313 ± 619	5034 ± 75	8344 ± 663	20594 ± 1547
8	14031 ± 221	5563 ± 88	13313 ± 177	22875 ± 442

Hyperflask培養物

【0208】 將8.3E07個P439C4細胞接種至550 ml的DMEM + 10% FBS+ 0.5 μg/mL嘌呤黴素+50.0 μg/mL G418中的兩個Hyperflask M容器中，並在37℃下培養3天。3天生長後的密度估計為每瓶3.6E8個細胞。藉由在550 ml的DMEM10% FBS中稀釋病毒，將培養瓶以40 MOI (1.4E10 vp)或20 MOI (7.2E09 vp)感染，並用稀釋的病毒置換在hyperflask中的培養基。將細胞在32℃下在5% CO2下培養7天。7天後從感染物中收集上清液，並藉由通過0.2 μm PES的膜過濾器來提純。

AAVX純化

【0209】 將0.5 × 5 cm的POROS GoPure層析管柱（用POROS CaptureSelect AAVX樹脂預填充至床體積為1 mL，附接至AKTA Explorer FPLC系統），用10倍管柱體積(CV)緩衝液A (20 mM Tris pH7.5, 400 mM NaCl)在3 ml/min之流率下平衡。4.5 mL/min之流速裝載病毒懸浮液，接著添加10 mL的緩衝液A以洗去未結合樣本。藉由用5 ml的低鹽benzonase緩衝液、緩衝液B（25 mM Tris pH7.5，40 mM NaCl，

及1.5 mM MgCl₂) 平衡管柱、接著用含有250單位/ml的Benzonase之15 ml緩衝液B裝載管柱來進行管柱上的DNA消化。然後將管柱在室溫下培養30分鐘，接著用緩衝液A的緩衝液進行15 CV洗滌。用15 CV的緩衝液C (20 mM檸檬酸鈉pH 2.5, 400 mM的NaCl) 以0.5 mL流份洗提病毒，該等流份立即用25 µL的500 mM Bis-TRIS丙烷pH 10.0中和。觀察到單峰洗提。收集曲線下的所有流份、濃縮，並使用15 100kDa MWCO (Cat# UFC910024, Fisher)用三輪緩衝液添加/離心，將緩衝液交換成緩衝液D (100 mM檸檬酸鈉，10 mM Tris，pH 8.0)。使經緩衝液交換及濃縮的親和力層析產物進行陰離子交換層析，以進一步將AAV純化遠離空殼體。

陰離子交換層析法

【0210】 將親和力層析產物（病毒懸浮液）於緩衝液E (20 mM BTP pH10.0, 0.001% Pluronic F68, 10 mM NaCl)中稀釋至45 mL，並在AKTA純化系統(GE Healthcare Life Sciences)上以2 ml/min之流速裝載至CIM QA碟（BIA分離，0.34 ml體積）上。用10 CV的無菌過濾緩衝液E (20 mM BTP pH10.0, 0.001% Pluronic F68, 10 mM NaCl)洗滌管柱。在60 CV梯度下自100%緩衝液E至100%緩衝液F (20 mM Bis-TRIS丙烷pH 10.0, 0.001% Pluronic F68, 400 mM NaCl) 洗提病毒；收集0.5 mL流份。收集曲線下的所有流份，並使用Amicon 15 100kDa MWCO (cat#: UFC910024, Fisher)以2000 × g離心5分鐘進行濃縮，並將緩衝液更換成緩衝液D (100 mM檸檬酸鈉，10 mM Tris，pH 8.0)。

蛋白質視覺化

【0211】 在補充有5% β-巯基乙醇的NuPage LDS樣本緩衝液(4x)

中，將2 μ L濃縮洗出液加熱變性（95℃達10分鐘），並在1x MOPS運行緩衝液中在4至12% Bis-Tris PAGE凝膠上進行電泳。根據製造商的說明，使凝膠經受銀染色。

ddPCR

【0212】 使用上述的mCherry分析，藉由數位液滴PCR測量病毒濃度。

結果

【0213】 當以20及40 MOI感染時，Hyperflask容器中的P439C4細胞之感染及生長分別產生了1.9E13及7.0E13個基因體拷貝(GC)。20及40 MOI的感染分別對應於每個細胞5.2E4及1.9E5個GC。透過PAGE電泳及銀染色檢查病毒樣本的純度。僅對應於三個AAV9殼體異構體（VP1、VP2、及VP3）的三條帶係可見的（圖5）。該等殼體蛋白（VP1 (87 kDa)、VP2 (72 kDa)、及VP3 (62 kDa)係以大約1:1:10的預期化學計量存在，如先前針對其他重組AAV載體所記述者(Daya and Berns (2008) Clin Microbiol Rev. 21: 583-593)。

測量錯誤包裝的DNA位準

【0214】 編碼AAV REP或CAP基因的序列及衍生自在生產過程中使用的質體載體之原核序列可非特異性地包裝至AAV粒子中，在用於基因治療時存在潛在的安全隱患（請參見例如，Schnodt and Buning, *Hum Gene Ther Methods.*, 2017; 28(3):101-108）。風險包括透過同源重組產生具有復制能力的AAV、殼體基因表現觸發細胞毒性T淋巴細胞反應、及對原核序列的免疫系統識別，從而導致炎症反應及/或基因靜默。在藉由三重轉染或從生產的細胞系中純化的重組AAV製劑中，分別報導了2%、

0.4%至1.0%、及1.3%至6.3%的殼體化rep、cap、及原核生物序列(Nony *et al.*(2003) J. Virology 77: 776-781; Gao *et al.* (2008) Molecular Therapy 16: S105; Chaudeuf *et al.* (2005) Molecular therapy 12: 744-753)。

【0215】 為了判定與上述生產系統關聯的錯誤包裝位準，藉由ddPCR判定該轉染載體（ITR側接的轉殖基因之外部）的四種序列之豐度：a) P5啟動子；b) AAV REP基因；c) AAV9 CAP基因；及d) β-內醯胺酶（安比西林抗性）基因。將來自前述之20及40 MOI hyperflask培養物的純化病毒製劑一式三份以DNAse消化、在VDB中連續稀釋、並使其經受ddPCR。含有這些序列之病毒粒子的濃度表示為含有mCherry轉殖基因的AAV粒子的百分比（表5）。0.04%之最高的殼體化率係P5啟動子在病毒A prep（由感染重組腺病毒的20 MOI產生）中之殼體化率。然而，在prep B (40 MOI)中（其中AAV產率高出許多），P5的殼體化率僅0.007%。REP、CAP、及安比西林基因序列係相同或更低。CAP位準係0.007%至0.009%，其低於先前報導的用於血友病B基因治療試驗所生產的重組AAV2的四種臨床批次之0.016%至0.021% cap殼體化率。(2009) Molecular Therapy 17: 144-152)。因此，此處描述對於生產及純化重組AAV的方法導致錯誤包裝DNA的發生率非常低，與臨床基因治療計劃可能需要的一致。

[表5]包裝在經純化病毒中之非轉殖基因序列的豐度

	P5啟動子	REP	CAP	安比西林
病毒A (20 MOI)	0.0398% ± 0.0023%	0.0078% ± 0.0003%	0.0092% ± 0.0002%	0.0051% ± 0.005%
病毒B (40 MOI)	0.0074% ± 0.0032%	0.0005% ± 0.0002%	0.007% ± 0.0001%	0.0003% ± 0.0001%

【0216】 顯示四種探針相對於含有 mCherry 轉殖基因的粒子之 DNase 抗性粒子平均百分比 $\pm$ 標準偏差，用於分析兩種 AAV 載體製劑。

在止擋匣切除之後，REP 基因的 RNA 剪接之 RT-PCR 分析

【0217】 為了判定當止擋匣切除時是否正確地剪接插入至建構體 P439 中的 REP 基因中的內含子，進行了 RT-PCR 實驗。

【0218】 將來自 PEAK-RAPID 細胞中 P439 之穩定池的一千萬個細胞藉由離心沉澱，並將其再懸浮於 15 ml 的 DMEM+2% FBS + 1E9 Ad5-CMV-SR21 病毒顆粒中。將細胞接種至 T75 培養瓶中然後在 37°C 下培養四十八小時。使用細胞刮刀分離細胞。將培養基及細胞轉移至 15 ml 離心管中，並以 1500 rpm 離心 10 分鐘以沉澱細胞。使用具有 Phase-maker 管的 Trizol Plus RNA 純化套組自細胞團塊純化 RNA。

【0219】 為了移除任何污染的 DNA，將 31 μ g 的 RNA 用來自不含 DNA 套組之 1 μ l 的 DNase 在 1x 消化緩衝液中在 37°C 下處理 30 分鐘。添加 5 μ l 的 DNase 無活化漿液，並且在 2 分鐘的培養期間將樣本倒置數次。將 RNA 樣本以 10,000 $\times$ g 離心 5 分鐘並將 RNA 轉移至新的無菌管。

【0220】 用 SuperScript III 第一股合成系統將 RNA 反轉錄。將 80 μ g RNA、1 μ l 50 μ M 寡聚-DT、及 1 μ l 10 mM dNTPS 在無菌管中混合並在 65°C 下培養 5 分鐘及在冰上培養 2 分鐘。添加 10 μ l 的 2x 混合物（2x RT 緩衝液，10 mM MgCl₂，20 mM 二硫蘇糖醇，0.5 μ l RNase-out，及 0.5 μ l 反轉錄酶）。模擬 RT 反應係相同的，除了用水置換反轉錄酶。將反應在 50°C 下培養 50 分鐘並在冰上培養 2 分鐘。添加 1 μ l RNase H 並將樣本在 37°C 下培養 20 分鐘。

【0221】 五十 μ l PCR 反應含有 1 μ l 的反轉錄 RNA、25 μ l Q5 熱啟動

高保真2X Master Mix、及0.5 μ M的兩種引子。使反應進行熱循環如下：

【0222】 98°C 1 min；35個循環之(98°C 10s, 69°C 10s, 72°C 36s)；5 min 72°C。

【0223】 將五 μ l的反應在1xTAE緩衝液及溴化乙錠中的1%瓊脂糖凝膠上離析。在Dark Reader透照器上在藍光照明下將條帶可視化。使用Nucleospin凝膠及PCR清潔套組從切除的帶中回收DNA。使用循環定序及染劑-終止子化學法，將DNA在GeneWiz (South Plainfield, NJ)用PCR引子定序。

結果：

【0224】 來自模擬RT模板的PCR反應沒有產生可檢測的產物，表示基因體DNA已從RNA樣本中消除。從P439中切除止擋匣後，使用引子AAVRT-F1 (SEQ ID NO:62)及P349R9 (SEQ ID NO:63)進行PCR，產生兩種具有類似螢光強度的PCR產物，這些產物均衍生自剪接的轉錄本。一種產物是由在經工程改造之 β -肌動蛋白剪接供體及受體位點處剪接而得到（分別為SEQ ID NO:14及SEQ ID NO:15；圖8）。第二產物是由5' REP基因(SEQ ID NO:64)供體位點與下游的 β -肌動蛋白受體之間的剪接得到（圖8）。預測此剪接事件相對於野生型AAV2移除REP編碼序列的64 bp，將產生框移並產生截頭REP蛋白。此意味著突變該上游剪接供體位點可增加活性REP蛋白的豐度，並使AAV生產更加有效。

更新AAV建構體：P600 (SEQ ID NO:70)

【0225】 對質體P439 (SEQ ID NO:12)進行幾處改變，得到建構體P600 (SEQ ID NO:70)。首先，該止擋匣上游的REP基因之該剪接供體係經突變的。簡而言之，將在5' REP序列(SEQ ID NO:64)中所識別之剪接

供體位點的核苷酸GT突變成AT (SEQ ID NO:65)。預測此突變以消除在此位點的剪接而不改變該REP蛋白質序列。

【0226】 為了降低止擋匣切除後可將該REP/CAP基因包裝至AAV殼體中的可能性，設計2 kb隨機序列(SEQ ID NO:66)並插入在該attB序列下游及該肌動蛋白剪接序列上游，以增加經工程改造的內含子的大小。使用NetGene 2軟體（上述的）鑑別潛在的剪接位點，並將其移除。此插入將REP/CAP基因的大小從4.3 kb增加至6.4 kb，其遠高於5.0 kb AAV包裝的限制。

【0227】 基於這樣的假設，相鄰AAV ITR的序列亦可能在從基因體中進行轉殖基因拯救期間被擴增，並且可能被錯誤包裝成AAV殼體（參見例如，Schnodt and Buning, *Hum Gene Ther Methods.*, 2017; 28(3):101-108），兩個隨機2 kb非編碼間隔子因子係經設計以側接該轉殖基因，以減小錯誤包裝DNA的潛在影響。在左AAV ITR的上游插入一個因子(SEQ ID NO:67)，並在右AAV ITR的下游插入第二個(SEQ ID NO:68)經置換的小鼠抗抑制因子40 (SEQ ID NO:25)。

【0228】 此外，該cap基因係AAV9變體（參見例如，Hinderer *et al.*, *Hum Gene Ther.* 2018; 29(3):285-298）。

【0229】 最後，P439中之ITR側接的轉殖基因編碼序列係由編碼mCherry-IRES-SEAP（分泌型鹼性磷酸酶）蛋白之SEQ ID NO:69置換。

【0230】 所得建構體P600的完整序列在SEQ ID NO:70中揭示，並且將該質體之繪示顯示於圖9中。

在穩定彙集中來自P600的AAV生產

【0231】 將建構體P600轉染至Peak-RAPID細胞中，並且透過選擇

0.5 $\mu\text{g/ml}$ 嘌呤黴素來產生穩定池，基本上如上文對於P439細胞所述。在分析AAV生產之前，使細胞以1:10通過達6週。

【0232】 將 2.5E6 個 P600-PEAK-Rapid p6 細胞接種至 5 ml 的 DMEM+ 10% FBS 中的三個T25燒瓶中，並在37°C下培養三天。在感染當天，將該等燒瓶中之一者的細胞密度判定為6.6E6個活細胞，其中用 TrypLE 回收細胞並使用錐蟲藍排除法計數。將兩個餘留培養瓶中之培養基用含有2.6E8個Ad5-CMV-SR21病毒粒子之11 ml的DMEM+ 2% FBS置換。將培養瓶在32°C下培養24小時。將一ml於DMEM+ 2% FBS中之1.25 mM的2-胺基嘌呤添加至一個燒瓶中。將一ml的DMEM+ 2% FBS添加至另一培養瓶，並且將兩個培養瓶在32°C下額外培養7天。將培養基自培養瓶中回收，以3,000 rpm離心5分鐘以沈澱細胞及與碎屑。使2 μl 的各樣本上清液經歷DNase消化並使用上述的mCherry分析進行ddPCR定量。

【0233】 AAV生產位準係示於表6中。當用Ad5-CMV-SR21感染時，P600穩定池活躍於產生AAV。在2-胺基嘌呤（一種經報導可阻斷腺病毒誘導的CAP依賴性mRNA轉譯的藥物）的存在下，AAV病毒生產量增加了2.5倍（參見例如，Zhang and Schneider (1994) J. Virology 68: 2544-2555; and Huang and Schneider (1990) PNAS 87: 7115-7119）。儘管已報導在感染後1至2小時10 mM 2-AP治療阻斷了腺病毒感染的細胞病變作用並且至少三天無毒（參見例如，Zhang and Schneider (1994) J. Virology 68: 2544-2555; and Huang and Schneider (1990) PNAS 87: 7115-7119)），我們發現濃度在1.25 mM以上且添加時間早於24小時，對於我們的AAV生產細胞系統中的AAV生產具有抑制作用。這些數據意味著抑制晚期腺病毒基因計劃（尤其是cap-依賴性mRNA轉譯的關閉）係一

種在生產細胞系中增加AAV產量的有用策略。

[表6]

樣本	介質	AAV GC/細胞
1	DMEM+ 2% FBS	23,826 ± 2990
2	DMEM+ 2% FBS+ 1.25 mM 2-AP	59, 356 ± 11,026

【0234】 在人類細胞中重組的腺相關病毒(AAV)的製造需要AAV複製(REP)及殼體(CAP)基因、腺病毒基因、及AAV-可包裝轉殖基因的表現，該轉殖基因由AAV反向末端重複(ITR)側接的表現匣組成。所有三種組分都可仰賴單獨的質體遞送至細胞以產生AAV，但是現有的轉染方法難以擴展成大規模培養。將一些該等因子摻入宿主細胞系中可使AAV生產更有效率，然而一些AAV及腺病毒基因係細胞生長抑制性或細胞毒性，因而限制了此方法。本發明描述使AAV REP基因可逆失活的方式，使得AAV REP、CAP、及可包裝轉殖基因可整合至合適的宿主細胞中並擴增。藉由表現重組酶之複製缺陷型重組腺病毒（例如，ΔE1/ΔE3）感染這些細胞會重新活化該等REP基因並誘導AAV複製及包裝。

【序列表】

<110> 美商健生生物科技公司(JANSSEN BIOTECH, INC.)

<120> 用於生物製造之合成遺傳因子

<130> 14620-192-185

<140>

<141>

<150> US 62/877,577

<151> 2019-07-23

<150> US 62/877,561

<151> 2019-07-23

<150> US 62/877,551

<151> 2019-07-23

<150> US 62/877,540

<151> 2019-07-23

<150> US 62/877,532

<151> 2019-07-23

<150> US 62/877,524

<151> 2019-07-23

<150> US 62/877,516

<151> 2019-07-23

<150> US 62/877,508

<151> 2019-07-23

<160> 73

<170> PatentIn version 3.5

<210> 1

<211> 545

<212> PRT

<213> 人工序列

<220>

<223> SPBetac2整合酶(NP_046553.1)蛋白質序列

<400> 1

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

Lys Asp Tyr Ser Tyr Thr Ala Ile Ala Ser His Leu Thr Asn Leu Gln
210 215 220

Ile Pro Thr Pro Ser Gly Lys Lys Arg Trp Asn Gln Tyr Thr Ile Lys
225 230 235 240

Ala Ile Leu Gln Asn Glu Val Tyr Ile Gly Thr Val Lys Tyr Lys Val
245 250 255

Arg Glu Lys Thr Lys Asp Gly Lys Arg Thr Ile Arg Pro Glu Lys Glu
260 265 270

Gln Ile Val Val Gln Asp Ala His Ala Pro Ile Ile Asp Lys Glu Gln
275 280 285

Phe Gln Gln Ser Gln Val Lys Ile Ala Asn Lys Val Pro Leu Leu Pro
290 295 300

Asn Lys Asp Glu Phe Glu Leu Ser Glu Leu Ala Gly Val Cys Thr Cys
305 310 315 320

Ser Lys Cys Gly Glu Pro Leu Ser Lys Tyr Glu Ser Lys Arg Ile Arg
325 330 335

Lys Asn Lys Asp Gly Thr Glu Ser Val Tyr His Val Lys Ser Leu Thr
340 345 350

Cys Lys Lys Asn Lys Cys Thr Tyr Val Arg Tyr Asn Asp Val Glu Asn
355 360 365

Ala Ile Leu Asp Tyr Leu Ser Ser Leu Asn Asp Leu Asn Asp Ser Thr
370 375 380

Leu Thr Lys His Ile Asn Ser Met Leu Ser Lys Tyr Glu Asp Asp Asn
385 390 395 400

Ser Asn Met Lys Thr Lys Lys Gln Met Ser Glu His Leu Ser Gln Lys
405 410 415

Glu Lys Glu Leu Lys Asn Lys Glu Asn Phe Ile Phe Asp Lys Tyr Glu
420 425 430

Ser Gly Ile Tyr Ser Asp Glu Leu Phe Leu Lys Arg Lys Ala Ala Leu
435 440 445

Asp Glu Glu Phe Lys Glu Leu Gln Asn Ala Lys Asn Glu Leu Asn Gly
450 455 460

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

Ile Asn Lys Ile Ile Asp Gln Tyr His Ile Glu Ser Ser Ser Glu Lys
485 490 495

Lys Asn Glu Leu Leu Arg Met Val Leu Lys Asp Val Ile Val Asn Met
500 505 510

Thr Gln Lys Arg Lys Gly Pro Ile Pro Ala Gln Phe Glu Ile Thr Pro
515 520 525

Ile Leu Arg Phe Asn Phe Ile Phe Asp Leu Thr Ala Thr Asn Ser Phe
530 535 540

His
545

<210> 2
<211> 535
<212> PRT
<213> 人工序列

<220>
<223> SR21重組酶(WP_029708089.1)蛋白質序列

<400> 2

Met Glu Leu Lys His Ile Val Asn Ser Tyr Asn Val Thr Lys Ile Ile
1 5 10 15

Gly Tyr Leu Arg Arg Ser Arg Gln Asp Val Glu Arg Glu Lys Arg Thr
 20 25 30

Gly Glu Asp Thr Leu Thr Glu Gln Lys Glu Leu Met Asn Lys Ile Leu
 35 40 45

Thr Gly Ile Glu Ile Pro Tyr Glu Thr Arg Thr Glu Ile Gly Ser Gly
 50 55 60

Glu Ser Ile Glu Gly Arg Pro Val Phe Lys Ser Cys Leu Ala Asp Leu
65 70 75 80

Arg Ser Gly Lys Phe Gln Ala Ile Ala Val Lys Glu Ile Thr Arg Leu
 85 90 95

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

Glu Lys Arg Ile Ile Ile Ile Thr Pro Tyr Lys Ile Tyr Asp Pro Arg
 115 120 125

Asn Pro Val Asp Ala Arg Gln Ile Arg Phe Glu Leu Phe Met Ala Arg
 130 135 140

Glu Glu Phe Glu Met Thr Arg Glu Arg Met Asn Gly Ala Lys Phe Thr
145 150 155 160

Tyr Ala Ala Gln Gly Lys Trp Ile Ser Gly Leu Ala Pro Phe Gly Tyr
 165 170 175

Lys Leu Asn Lys Arg Thr Ser Arg Leu Glu Pro Ser Asp Glu Asp Lys
 180 185 190

Val Val Val Lys Leu Ile Phe Asp Ile Phe Leu Asn Gly Leu Asp Gly
 195 200 205

Lys Asp Leu Ser Tyr Thr Ala Ile Ala Thr His Leu Ser Lys Leu Gln
210 215 220

Phe Thr Thr Pro Arg Gly Gly Lys Arg Trp Ser Lys Asp Thr Val Arg
225 230 235 240

Lys Ile Leu Gln Asn Glu Ala Tyr Met Gly Arg Val Arg Tyr Lys Ala
245 250 255

Arg Glu Thr Thr Lys Asp Gly Lys Lys Val Phe Arg Pro Glu Ser Glu
260 265 270

His Ile Val Val Asp Asp Ala His Glu Pro Ile Ile Asn Lys Glu Asp
275 280 285

Phe Glu Ala Val Gln Glu Lys Ile Lys Asn Lys Val Pro Leu Leu Pro
290 295 300

Val Val Thr Ser Tyr Glu Pro Asn Glu Leu Ala Gly Ile Cys Val Cys
305 310 315 320

Ser Val Cys Gly Lys Ser Leu Gln Lys Phe Glu Ser Glu Tyr Asn Arg
325 330 335

Lys Asn Lys Asp Gly Thr Ser Ser Tyr Phe His Val Lys Leu Leu Ile
340 345 350

Cys Lys Ile Asn Lys Cys Thr Ser Val Arg Tyr Glu Tyr Val Glu Glu
355 360 365

Ala Ile Leu Glu Tyr Leu Glu Gln Leu Ile Ala Leu Glu Asn Asn Lys
370 375 380

Leu Lys Ala Ile Ile Glu Lys Ser Met Glu Ala Ala Glu Thr Asn Asn
385 390 395 400

Ser Glu Lys Thr Ser Glu Gln Met Leu Val Gln Ala Asn Gln Lys Gln
405 410 415

Lys Glu Leu Glu Asn Lys Leu Thr Phe Ile Phe Glu Lys Phe Glu Ser
420 425 430

Gly Ile Tyr Thr Asp Glu Met Phe Leu Gln Arg Lys Ala Ala Ile Glu
435 440 445

Lys Glu Val Ala Asp Ile Lys Lys Leu Lys Gln Glu Leu Ser Met Thr
450 455 460

Phe Glu Val Lys Glu Lys Asp Val Asn Glu Phe Arg Val Asn Ile Ser
465 470 475 480

Asp Val Val Lys Phe Tyr Lys Glu Ser Lys Ser Arg Gly Leu Lys Asn
485 490 495

Glu Lys Leu Arg Ser Ile Phe Asp Phe Ile Val Leu Glu Met Thr Glu
500 505 510

Lys Arg Arg Gly Pro Ile Pro Ala Lys Phe Asn Ile Tyr Pro Val Leu
515 520 525

Arg Ile Pro Ser Val Val Asn
530 535

<210> 3
<211> 1608
<212> DNA
<213> 人工序列

<220>
<223> SR21重組酶DNA序列

<400> 3
ttggaattaa aacacattgt caattcctat aatgtgacta aaattattgg ctatttgaga 60

cgttctagac aagatgttga aagagagaag cgtactgggtg aagatacttt gacagaacaa 120

aaagaactta tgaataaaat attgactggg attgaaatac cgtacgaaac aagaactgag 180

atcggttctg gtgaaagcat tgaaggaaga ccagtgttta aaagttgctt agctgatctg	240
aggtctggaa aatttcaagc aattgctgta aaagaaataa ctcgctcttag tcgtggtagt	300
tacagtgatg ctggtgaaat cgtcaatctt ctaaataaaa aaaggattat cataattaca	360
ccatataaaa tatatgatcc aagaaatcct gttgatgctc gtcaaattag atttgagcta	420
tttatggcta gagaagaatt tgaaatgaca agagaaagaa tgaatggtgc caagtttact	480
tatgctgcac aaggaaaatg gatttctggt ttagcaccat ttggctataa attaaataag	540
cgtacatctc gcctcgaacc atcggatgaa gacaaagttg tggtaaagtt gatatttgac	600
attttcttaa atggccttga tggtaaagat ttaagttata cagctattgc aactcatctt	660
tctaaacttc aattcactac cccaagggga ggaaaaagat ggagtaagga cactgttaga	720
aaaattcttc aaaatgaagc ctatatggga agggttaggt ataaagcaag agaaaccaca	780
aaagatggaa aaaaagtctt ccgacctgaa agtgaacata ttgtttgtaga tgatgcccac	840
gaaccaataa taaataaaga ggactttgaa gctgttcaag aaaagattaa aaataaagtg	900
cctctgcttc cagtcgtaac atcttatgaa cctaataaat tagcgggaat atgtgtatgc	960
tctgtctgtg gaaaatcact tcaaaaattt gaatccgaat acaataggaa aaacaaagat	1020
gggacttcca gctattttca cgtgaagtta ttgatttgca aaatcaacaa gtgcacatca	1080
gtgagatatg agtatgttga agaagcaatt ttggagtatc ttgagcagtt gattgcttta	1140
gaaaataata aactaaaggc tataattgaa aaatctatgg aagcagcaga aactaacaat	1200
tcagaaaaga caagtgagca aatgctagta caagcaaacc aaaaacagaa agagctcgaa	1260
aacaaattga catttatctt tgaaaaattt gaatctggaa ttatatacaga tgaaatgttt	1320
ttgcaaagga aagcagctat tgaaaaagaa gtagctgaca ttaaaaaatt gaaacaagaa	1380
cttagcatga catttgaagt taaagagaaa gatgttaatg aatttagggt taacatatca	1440
gatgtagtta aattttataa agaattctaaa agtagaggat taaaaaatga aaagttgaga	1500
tccatttttg actttatagt attagaaatg acagagaaaa gacgaggccc catccctgct	1560
aagttcaata ttatatccgt tttgcgaatt ccgagtgtag ttaattaa	1608

<210> 4
<211> 188
<212> DNA
<213> 人工序列

<220>
<223> 沙福芽孢桿菌(Bacillus safensis) Fairview菌株插入前基因座
(gil640668059:c464939-464752

<400> 4
agataatccg ttcgaagctg ttcaaacgaa tatcatcggc gggcagcacg tcattgaagc 60

cgcaattgaa catgaagtga gtcatgtcgt caacatctcc actgacaaag cggtttctcc 120

aacgaatgcc atgggtgcaa caaaattgat ttcagaaaaa ctatttttcc aagcaaacga 180

aagtattc 188

<210> 5
<211> 120
<212> DNA
<213> 人工序列

<220>
<223> 右原噬菌體-宿主DNA連接點，沙福芽孢桿菌株
CCMA-560

<400> 5
tttgcgaatt ccgagtgtag ttaattaagt atcatgataa actgacaaag cggtttctcc 60

aacgaatgcc atgggtgcaa caaaattgat ttcagaaaaa ctatttttcc aagcaaacga 120

<210> 6
<211> 158
<212> DNA
<213> 人工序列

<220>
<223> 左原噬菌體-宿主DNA連接點，沙福芽孢桿菌株
CCMA-560

<400> 6
tatcatcggc gggcagcacg tcattgaagc cgctctcgaa catgaagtga gtcatgtcgt 60

caacatttca actgacaaag cggtatatca agatacttaa tacatgataa caaaaaaaga 120

agcctacata ttggcttctt ttataattat atttgcta 158

<210>	7	
<211>	50	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	由沙福芽孢桿菌菌株CCMA-560（人工序列）重建之SR21重組酶AttP	
<400>	7	
aattaagtat catgataaac tgacaaagcg gtatatcaag atacttaata		50
<210>	8	
<211>	40	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	由沙福芽孢桿菌菌株CCMA-560重建之SR21重組酶AttB	
<400>	8	
cgtcaacatt tcaactgaca aagcggtttc tccaacgaat		40
<210>	9	
<211>	40	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	由沙福芽孢桿菌菌株Fairview重建之SR21重組酶AttB	
<400>	9	
cgtcaacatc tccactgaca aagcggtttc tccaacgaat		40
<210>	10	
<211>	4532	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	P175- SR21哺乳動物表現載體	
<400>	10	
taactataac ggtcctaagg tagcgaagct cttcagatgg acagtcagac tgaagagcct		60

ctcttaaggt agctcgagga gcttggccca ttgcatacgt tgtatccata tcataatatg	120
tacatttata ttggctcatg tccaacatta ccgccatgtt gacattgatt attgactagt	180
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt	240
acataactta cggtaaattgg cccgcctggc tgaccgcca acgacccccg cccattgacg	300
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	360
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt	420
acgcccccta ttgacgtcaa tgacggtaaa tggcccgctt ggcattatgc ccagtacatg	480
accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	540
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt	600
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac	660
tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg	720
tgggaggtct atataagcag agctcgttta gtgaaccgtc ggcgcgccgc caccatggaa	780
ttaaaacaca ttgtcaattc ctataatgtg actaaaatta ttggctattt gagacgttcc	840
agacaagatg ttgaaagaga gaagcgtact ggtgaagata ctttgacaga acaaaaagaa	900
cttatgaata aaatattgac tgggattgaa ataccgtacg aaacaagaac tgagatcgg	960
tctggtgaaa gcattgaagg aagaccagtg tttaaaagtt gcttagctga tctgaggtct	1020
ggaaaatttc aagcaattgc tgtaaaagaa ataactcgtc ttagtcgtgg tagttacagt	1080
gatgctggtg aaatcgtcaa tcttctaaat gaaaaaagga ttatcataat tacaccatat	1140
aaaatatatg atccaagaaa tcctgttgat gctcgtcaaa ttagatttga gctatttatg	1200
gctagagaag aatttgaaat gacaagagaa agaatgaatg gtgccaaagt tacttatgct	1260
gcacaaggaa aatggatttc tggtttagca ccatttggct ataaattaaa taagcgtaca	1320
tctcgctcg aaccatcgga tgaagacaaa gttgtggtaa agttgatatt tgacattttc	1380
ttaaattggcc ttgatggtaa agatttaagt tatacagcta ttgcaactca tctttctaaa	1440
cttcaattca ctaccccaag gggaggaaaa agatggagta aggacactgt tagaaaaatt	1500
cttcaaaatg aagcctatat gggaagggtt aggtataaag caagagaaac cacaaaagat	1560

ggaaaaaaag	tcttccgacc	tgaagtgaa	catattgttg	tagatgatgc	ccacgaacca	1620
ataataaata	aagaggactt	tgaagctgtt	caagaaaaga	ttaaaaataa	agtgcctctg	1680
cttccagtcg	taacatctta	tgaacctaat	gaattagcgg	gaatatgtgt	atgctctgtc	1740
tgtggaaaat	cacttcaaaa	atttgaatcc	gaatacaata	ggaaaaacaa	agatgggact	1800
tccagctatt	ttcacgtgaa	gttattgatt	tgcaaaatca	acaagtgcac	atcagtgaga	1860
tatgagtatg	ttgaagaagc	aattttggag	tatcttgagc	agttgattgc	tttagaaaat	1920
aataaactaa	aggctataat	tgaaaaatct	atggaagcag	cagaaactaa	caattcagaa	1980
aagacaagtg	agcaaatgct	agtacaagca	aaccaaaaac	agaaagaact	cgaaaacaaa	2040
ttgacattta	tttttgaaaa	atttgaatct	ggaatttata	cagatgaaat	gtttttgcaa	2100
aggaaagcag	ctattgaaaa	agaagtagct	gacattaaaa	aattgaaaca	agaacttagc	2160
atgacatttg	aagttaaaga	gaaagatgtt	aatgaattta	gggttaacat	atcagatgta	2220
gttaaatttt	ataaagaatc	taaaagtaga	ggattaaaaa	atgaaaagtt	gagatccatt	2280
tttgacttta	tagtattaga	aatgacagag	aaaagacgag	gccccatccc	tgctaagttc	2340
aatattttatc	cggttttgcg	aatccccagt	gtagttaatt	aatctagaga	tacattgatg	2400
agtttgacaa	aaccacaact	agaatgcagt	gaaaaaaatg	ctttatttgt	gaaatttgtg	2460
atgctatttg	tttatttgta	accattataa	gctgcaataa	acaagttaac	aacaacaatt	2520
gcattcatth	tatgtttcag	gttcaggggg	aggtgtggga	ggttttttta	agcaagtaaa	2580
acctctacaa	atgtggtatg	gctgattatg	atcgcgggcg	caataaagat	cccgggtagg	2640
gataacaggg	taatgctctt	cagatggaca	gtcagactga	agagctggca	aacagctatt	2700
atgggtatta	tgggtgacgt	caggtggcac	ttttcgggga	aatgtgcgcg	gaacccttat	2760
ttgtttatth	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	aaccctgata	2820
aatgcttcaa	taatattgaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	2880
tattccctth	tttgcgcat	tttgccctcc	tgthtttgct	caccagaaa	cgttggtgaa	2940
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	3000
cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	tgagcactth	3060

taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg	3120
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca	3180
tettacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa	3240
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt	3300
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc	3360
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa	3420
actattaact ggcgaaactac ttactctagc ttcccggcaa caattaatag actggatgga	3480
ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggtggct ggtttattgc	3540
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga	3600
tggtaaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga	3660
acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga	3720
ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat	3780
ctaggtgaag atcctttttg ataatctcat gaccaaaac ccttaacgtg agttttcgtt	3840
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct	3900
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtggt tttgtttgcc	3960
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc	4020
aaatactggt cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc	4080
gcctacatac ctcgctctgc taatcctggt accagtggct gctgccagtg gcgataagtc	4140
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg	4200
aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata	4260
cctacagcgt gagctatgag aaagcgccac gcttcccga gggagaaagg cggacaggta	4320
tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc	4380
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg	4440
atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt	4500
cctggccttt tgctggcctt ttgctcacat gt	4532

<210> 11
<211> 9555
<212> DNA
<213> 人工序列

<220>
<223> P41- SR21重組酶報導子質體

<400> 11	
taactataac ggtcctaagg tagcgaagct cttcagatgg acagtcagac tgaagagcct	60
ctcttaaggt agctcgagga gcttggccca ttgcatacgt tgtatccata tcataatatg	120
tacatttata ttggctcatg tccaacatta ccgccatggt gacattgatt attgactagt	180
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt	240
acataactta cggtaaattgg cccgcctggc tgaccgccca acgacccccg cccattgacg	300
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	360
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt	420
acgcccccta ttgacgtcaa tgacggtaaa tggccccgct ggcattatgc ccagtacatg	480
accttatggg actttcctac ttggcagtac atctacgtat tagtcategc tattaccatg	540
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt	600
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac	660
tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg	720
tgggaggtct atataagcag agctcgttta gtgaaccgtc ggcgcgcccc tcaacatctc	780
cactgacaaa gcggtttctc caacgaatgg atccttacac ggcgatcttg ccgcccttct	840
tggccttaat gagaatctcg cggatcttgc gggcgtccaa cttgccggtc agtcctttag	900
gcacctcgtc cacgaacaca acaccaccgc gcagcttctt ggcggttgta acctggctgg	960
ccacatagtc cacgatctcc ttctcgggtca tggttttacc gtgttccagc acgacgactg	1020
cggcggggcag ctcgccggca tcgtcgtcgg gcaggccggc gacccccggc tcgaagatgt	1080
tggggtgttg cagcaggatg ctctccagtt cggctggggc tacctggtag cccttgtatt	1140
tgatcaggct cttcagccgg tccacgatga agaagtgctc gtctcgtcc cagtaggcga	1200

tgtcgccgct	gtgcagccag	ccgtcccttgt	cgatgagagc	gtttgtagcc	tcggggttgt	1260
taacgtagcc	gctcatgate	atggggccac	ggacgcacag	ctcgccgcgc	tggttcacac	1320
ccagtgtctt	accggtgtcc	aagtccacca	ccttagcctc	gaagaagggc	accaccttgc	1380
ctactgcgcc	aggcttgteg	tccccctcgg	gggtgatcag	aatggcgctg	gttgtttctg	1440
tcaggccgta	gccctggcgg	atgcctggta	ggtggaagcg	tttggccacg	gcctcaccta	1500
cctcccttgc	gagcggcgcc	ccgccgctgg	cgatctcgtg	caagttgctt	aggtcgtact	1560
tgtcgatgag	agtgctctta	gcgaagaagc	taaatagtgt	gggcaccagc	agggcagatt	1620
gaatcttata	gtcttgcaag	ctgcgcaaga	atagctcctc	ctcgaagcgg	tacatgagca	1680
cgacccgaaa	gccgcagatc	aagtagccca	gcgtggtgaa	catgccgaag	ccgtggtgaa	1740
atggcaccac	gctgaggata	gcggtgtcgg	ggatgatctg	gttgccgaag	atggggtcgc	1800
gggcatgact	gaatcggaca	caagcgggtc	ggtgcggtag	ggctacgccc	ttgggcaatc	1860
cggtagctgc	actactgttc	atgatcaggg	cgatggtttt	gtcccggteg	aagctctcgg	1920
gcacgaagtc	gtactcgttg	aagccgggtg	gcaaatggga	agtcacgaag	gtgtacatgc	1980
tttggaagcc	ctggtagteg	gtcttgctat	ccatgatgat	gatcttttgt	atgatcggta	2040
gcttcctttg	cacgttgagg	atcttttgca	gccctttcct	gtcacgaat	acgacggtgg	2100
gctggctgat	gcccattgct	ttcagcagct	cgcgctcggt	gtagatgtcg	ttagctgggg	2160
ccacagccac	accgatgaac	agggcaccca	acacgggcat	gaagaactgc	aagctattct	2220
cgctgcacac	cacgatccga	tggtttgtat	tcagcccata	gcgcttcata	gcttctgcca	2280
gccgaacgct	catctcgaag	tactcggcgt	aggtaatgtc	cacctcgata	tgtgcgtcgg	2340
taaaggcgat	ggtgccgggc	accagggcgt	agcgcttcat	ggctttgtgc	agctgctcgc	2400
cggcgggtccc	gtcttcgagt	gggtagaatg	gcgctggacc	cttcttaatg	tttttgcat	2460
cttccatagg	tccagggttg	gactccacgt	ctcccgccaa	cttgagaagg	tcaaaattca	2520
aagtctgttt	cactccgctt	cccttgtaaa	gtcgttccat	gccgccggtg	gagtggcggc	2580
cctcggcgcg	ttcgtactgt	tccacgatgg	tgtagtcctc	gttggtgggag	gtgatgtcca	2640
acttgatgtt	gacgttgtag	gcgccgggca	gctgcacggg	cttctttggcc	ttgtaggtgg	2700

tcttgacctc	agcgtcgtag	tggccgccgt	ccttcagctt	cagcctctgc	ttgatctcgc	2760
ccttcagggc	gccgtcctcg	gggtacatcc	gctcggagga	ggcctcccag	cccatgggtct	2820
tcttctgcat	tacggggccg	tcggagggga	agttgggtgcc	gcgcagcttc	accttgtaga	2880
tgaactcgcc	gtcctgcagg	gaggagtccct	gggtcacggt	caccacgccg	ccgtcctcga	2940
agttcatcac	gcgtcccccac	ttgaagccct	cggggaagga	cagcttcaag	tagtcgggga	3000
tgtcggcggg	gtgcttcacg	taggccttgg	agccgtacat	gaactgaggg	gacaggatgt	3060
cccaggcgaa	gggcaggggg	ccacccttgg	tcaccttcag	cttggcggtc	tgggtgccct	3120
cgtagggggcg	gccctcgccc	tcgccctega	tctcgaactc	gtggccgttc	acggagccct	3180
ccatgtgcac	cttgaagcgc	atgaactcct	tgatgatggc	catgttatcc	tcctcgccct	3240
tgtccaccat	ggtggcgtcg	actattaagt	atcttgatat	accgctttgt	cagtttatca	3300
tgatacttaa	tttctagaga	tacattgatg	agtttggaca	aaccacaact	agaatgcagt	3360
gaaaaaatg	ctttattttgt	gaaatttgtg	atgctattgc	tttatttgta	accattataa	3420
gctgcaataa	acaagttaac	aacaacaatt	gcattcatTT	tatgttttcag	gttcaggggg	3480
aggtgtggga	ggtttttttaa	agcaagtaaa	acctctacaa	atgtggtatg	gctgatttatg	3540
atcgcggccg	ccgtgaggct	ccggtgcccc	tcagtgggca	gagcgcacat	cgcccacagt	3600
ccccgagaag	ttgggggggag	gggtcggcaa	ttgaaccggt	gcctagagaa	ggtggcgcgg	3660
ggtaaactgg	gaaagtgatg	tcgtgtactg	gtctccgcctt	tttcccgagg	gtggggggaga	3720
accgtatata	agtgcagtag	tcgccgtgaa	cgttcttttt	cgcaacgggt	ttgccgccag	3780
aacacaggta	agtgccgtgt	gtggttcccc	cgggccttggc	ctcttttacgg	gttatggccc	3840
ttgcgtgcct	tgaattactt	ccacgccccct	ggctgcagta	cgtgatttctt	gatcccgagc	3900
ttcgggttgg	aagtgggtgg	gagagttcga	ggccttgccgc	ttaaggagcc	ccttcgcctc	3960
gtgcttgagt	tgaggccttg	cctgggcgct	ggggccgccg	cgtgcgaatc	tggtggcacc	4020
ttcgcgccctg	tctcgctgct	ttcgataagt	ctctagccat	ttaaaatttt	tgatgacctg	4080
ctgcgacgct	ttttttcttg	caagatagtc	ttgtaaatgc	gggccaaagat	ctgcacactg	4140
gtatttcggt	ttttggggcc	gcgggcggcg	acggggcccc	tgcgtcccag	cgtcatgtt	4200

cggcgaggcg	gggcctgcga	gcgcggccac	cgagaatcgg	acgggggtag	tctcaagctg	4260
gccggcctgc	tctggtgcct	ggcctcgcgc	cgccgtgtat	cgccccgccc	tgggcggcaa	4320
ggctggcccc	gtcggcacca	gttgctgag	cggaaagatg	gccgttccc	ggccctgctg	4380
cagggagctc	aaaatggagg	acgcggcgt	cgggagagcg	ggcgggtgag	tcacccacac	4440
aaaggaaaag	ggcctttccg	tcctcagccg	tcgtttcatg	tgactccacg	gagtaccggg	4500
cgccgtccag	gcacctcgat	tagttctcga	gcttttggag	tacgtcgtct	ttaggttggg	4560
gggaggggtt	ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	4620
cagcttggca	cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttgg	4680
tcattctcaa	gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	4740
ccaccatgga	gagcgacgag	agcggcctgc	ccgccatgga	gacgagtg	cgcacaccg	4800
gcacctgaa	cggcgtggag	ttcgagctgg	tgggcggcgg	agagggcacc	cccgagcagg	4860
gccgcatgac	caacaagatg	aagagcacca	aaggcgcctt	gaccttcagc	ccctacctgc	4920
tgagccacgt	gatgggctac	ggcttctacc	acttcggcac	ctaccccagc	ggctacgaga	4980
accccttcc	gcacgccatc	aacaacggcg	gctacaccaa	cacccgcac	gagaagtacg	5040
aggacggcgg	cgtgctgcac	gtgagcttca	gctaccgcta	cgaggccggc	cgcgtgatcg	5100
gcgacttcaa	ggtgatgggc	accggcttcc	ccgaggacag	cgtgatcttc	accgacaaga	5160
tcatccgcag	caacgccacc	gtggagcacc	tgcaccccat	gggcgataac	gatctggatg	5220
gcagcttcac	ccgcaccttc	agcctgcgcg	acggcggcta	ctacagctcc	gtggtggaca	5280
gccacatgca	cttcaagagc	gccatccacc	ccagcatcct	gcagaacggg	ggacctatgt	5340
tcgccttccg	ccgcgtggag	gaggatcaca	gcaacaccga	gctgggcatc	gtggagtacc	5400
agcacgcctt	caagaccccc	gatgcagatg	ccggtgaaga	aggaagcgga	gctactaact	5460
tcagcctgct	gaagcaggct	ggagacgtgg	aggagaacce	tggacctatg	gcttccaagg	5520
tgtacgacce	cgagcaacgc	aaacgcatga	tactgggcc	tcagtgtgg	gtcgtgca	5580
agcaaatgaa	cgtgctggac	tccttcatca	actactatga	ttccgagaag	cacgccgaga	5640
acgccgtgat	ttttctgcat	ggtaacgctg	cctccagcta	cctgtggagg	cacgtcgtgc	5700

ctcacatcga gcccgtggct agatgcatca tccctgatct gatcggaatg ggtaagtccg	5760
gcaagagcgg gaatggctca tategcctcc tggatcacta caagtacctc accgcttggt	5820
tcgagctgct gaaccttcca aagaaaatca tctttgtggg ccacgactgg ggggcttgtc	5880
tggcctttca ctactcctac gaggaccaag acaagatcaa ggccatcgtc catgctgaga	5940
gtgtcgtgga cgtgatcgag tcctgggacg agtggcctga catcgaggag gacatcgccc	6000
tgatcaagag cgaagagggc gagaaaatgg tgcttgagaa taactttctc gtcgagacca	6060
tgctcccaag caagatcatg cggaaactgg agcctgagga gttcgctgcc tacctggagc	6120
cattcaagga gaagggcgag gttagacggc ctacctctc ctggcctcgc gagatccctc	6180
tcgttaaggg aggcaagccc gatgtcgtcc agattgtccg caactacaac gcctaccttc	6240
gggccagcga cgatctgcct aagatgttca tcgagtcga cctgggttc ttttccaacg	6300
ctattgtcga gggagctaag aagtcccta acaccgagtt cgtgaagggtg aagggcctcc	6360
acttcagcca ggaggacgct ccagatgaaa tgggtaagta catcaagagc ttcgtggagc	6420
gcgtgctgaa gaacgagcag taagctcact gcccatgatg cagagctttc aaggataggc	6480
tttattctgc aagcaatcaa ataataaatc tattctgcta agagatcaca catggttgtc	6540
ttcagttctt tttttatgtc tttttaaata tatgagccac aaagggtttt atgttgaggg	6600
atgtgtttat gtgtatttat acatggctat gtgtgtttgt gtcatgtgca cactccacac	6660
ttttttgttt acgttagatg tgggttttga tgagcaaata aaagaactag gcaataaaga	6720
aacttgtaca tgggagttct gcaagtggga gtaaaagggtg caggagaaat ctggttggaa	6780
gaaagacctc tataggacag gactcctcag aaacagatgt tttggaagag atgggggaaag	6840
gttcagtgaa gggggctgaa ccccttccc tggattgcag cacagcagcg aggaaggggc	6900
tcaacgaaga aaaagtgttc ctagctttag gaagtcaagg tttaggcagg gatagccatt	6960
ctattttatt aggggcaata ctatttccaa cggcatctgg cttttctcag cccttgtag	7020
gctctacagg gaggttgagg tgttagagat cagagcagga aacaggtttt tctttccacg	7080
gtaactacaa tgaagtgate ctactttac taaggaactt ttcattttaa gtgttgacgc	7140
atgcctaaag aggtgaaatt aatcccatac ccttaagtct acagactggg cacagcattt	7200

caaggaggag	acctcattgt	aagctactag	ggaggtgggg	acttaggtga	aggaaatgag	7260
ccagcagaag	ctcacaagtc	agcatcagcg	tgtcatgtct	cagcagcaga	acagcacggt	7320
cagatgaaaa	tatagtgtga	agaatttgta	taacattaat	tgagaaggca	gattcactgg	7380
agttcttata	taattgaaag	ttaatgcacg	ttaataagca	agagtttagt	ttaatgtgat	7440
ggtgttatga	acttaacgct	tgtgtctcca	gaaaattcac	atgctgaatc	cccaactccc	7500
aattggctcc	atttgtgggg	gaggctttgg	aaaagtaatc	aggttttagag	gggctcatga	7560
gagcagatcc	ccatcataga	attattttcc	tcatcagaag	cagagagatt	agccatttct	7620
cttccttctg	gtgaggacac	agtgggaagt	cagccacctt	agggataaca	gggtaatgct	7680
cttcagatgg	acagtcagac	tgaagagctg	gcaaacagct	attatgggta	ttatgggtga	7740
cgtcaggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	tttttctaaa	7800
tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	7860
gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttatcccc	ttttttgcgg	7920
cattttgcct	tcctgttttt	gtcacccag	aaacgctgg	gaaagtaaaa	gatgctgaag	7980
atcagttggg	tgcacgagtg	ggttacatcg	aactggatct	caacagcgg	aagatccttg	8040
agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaagtt	ctgctatgtg	8100
gcgcggtatt	atcccgtatt	gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	8160
ctcagaatga	cttggttgag	tactcaccag	tcacagaaaa	gcattcttac	gatggcatga	8220
cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	8280
ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgcacaac	atgggggatc	8340
atgtaactcg	ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	8400
gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	8460
tacttactct	agcttcccgg	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	8520
gaccacttct	gcgctcggcc	cttcgggctg	gctggtttat	tgctgataaa	tctggagccg	8580
gtgagcgtgg	gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	ccctcccgta	8640
tcgtagttat	ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	8700

ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata	8760
tacttttagat tgatttaaaa cttcattttt aattttaaag gatctaggtg aagatccttt	8820
ttgataatct catgaccaaa atcccttaac gtgagttttc gtccactga gcgtcagacc	8880
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct	8940
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa	9000
ctctttttcc gaaggtaact ggcttcagca gagegcagat accaaatact gttcttctag	9060
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc	9120
tgetaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg	9180
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca	9240
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat	9300
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggtg agcggcaggg	9360
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcttggtat ctttatagtc	9420
ctgtcggggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc	9480
ggagcctatg gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttgctggc	9540
cttttgctca catgt	9555

- <210>

12
- <211>

14985
- <212>

DNA
- <213>

人工序列
- <220>
- <223>

P439，游離型REP/CAP + 轉殖基因質體

- <220>
- <221>

misc_feature
- <222>

(2103)..(2137)
- <223>

SEQ ID NO:14 β -肌動蛋白內含子之上游半部
- <220>
- <221>

misc_feature
- <222>

(2144)..(2398)
- <223>

止擋匣之5'端

<220>
<221> misc_feature
<222> (4069)..(4287)
<223> 止擋匣之3'端

<220>
<221> misc_feature
<222> (4295)..(4370)
<223> SEQ ID NO:15 β-肌動蛋白內含子之下游半部

<220>
<221> misc_feature
<222> (7984)..(8124)
<223> 左反向末端重複(ITR)

<220>
<221> misc_feature
<222> (11067)..(11207)
<223> 右反向末端重複(ITR)

<400> 12	
taactataac ggtcctaagg tagcgaagat atccgatcaa gaaagcactc cgggctccag	60
aaggagcctt ccaggccagc tttgagcata agctgctgat gagcagtgag tgtcttgagt	120
agtgttcagg gcagcatgtt accattcatg cttgacttct agccagtgtg acgagaggct	180
ggagtcaggt ctctagagag ttgagcagct ccagccttag atctcccagt cttatgcggt	240
gtgcccattc gctttgtgtc tgcagtcccc tggccacacc cagtaacagt tctgggatct	300
atgggagtag cttccttagt gagctttccc ttcaaatact ttgcaaccag gtagagaagt	360
ttggagtgaa ggttttgttc ttcgtttctt cacaatatgg atatgcatct tcttttgaaa	420
atgttaaagt aaattacctc tcttttcaga tactgtcttc atgcgaactt ggtatcctgt	480
ttccatccca gccttctata acccagtaac atcttttttg aaaccagtgg gtgagaaaga	540
cacctggtea ggaacgcgga ccacaggaca actcaggctc acccacggca tcagactaaa	600
ggcaaacaag gactctgtat aaagtaccgg tggcatgtgt attagtggag atgcagcctg	660
tgtctctgcag acagggagtc acacagacac ttttctataa tttcttaagt gctttgaatg	720
ttcaagtaga aagtctaaca ttaaatttga ttgaacaatt gtatatcat ggaatathtt	780
ggaacggaat accaaaaaat ggcaatagtg gttctttctg gatggaagac aaacttttct	840

tctttaaaat aaattttatt ttatatat	gaggttgacc acatgacctt aaggatacat	900
atagacagta aactgggttac tacagtgaag	caaattaaca tatctaccat cgtacatagt	960
tacatTTTTT tgtgtgacag gaacagctaa	aatctacgta ttttaacaaaa ctccctaaaga	1020
caatacatTTT ttattaacta tagccctcat	gatgtacatt agatccctctc ttaaggtagc	1080
ccgaataaaaa atagaaaaac aacaaatg	tt aattttctat ttttatTTTgg tgtcgtactc	1140
aatccataat cgtcacaagc gttatagttt	aagtacaaca atgtcgtcac cttgtttggtg	1200
taattaggtt tacgccaaca gggtgataac	aggcgcgcgg gtccctgtatt agaggtcacg	1260
tgagtgtttt gcgacatTTT gcgacaccat	gtggtcacgc tgggtattta agcccgagtg	1320
agcacgcagg gtctccatTTT tgaagcggga	ggtttgaacg cgcagccgcc atgccgggg	1380
tttacgagat tgtgattaag gtccccagcg	accttgacga gcctctgccc ggcatttctg	1440
acagctttgt gaactgggtg gccgagaagg	aatgggagtt gccgccagat tctgacatgg	1500
atctgaatct gattgagcag gcacccctga	ccgtggccga gaagctgcag cgcgactttc	1560
tgacggaatg gcgccgtgtg agtaaggccc	cggaggccct tttctttgtg caatttgaga	1620
agggagagag ctacttccac atgcacgtgc	tcttggaaac caccggggtg aaatccatgg	1680
ttttgggacg tttcctgagt cagattcgcg	aaaaactgat tcagagaatt taccgcggga	1740
tcgagccgac tttgccaaac tggttcgcg	tcacaaagac cagaaatggc gccggaggcg	1800
ggaacaaggt ggtggatgag tgctacatcc	ccaattactt gtcceccaaa acccagcctg	1860
agctccagtg ggcgtggact aatatggaac	agtatttaag cgctgtttg aatctcacgg	1920
agcgtaaacg gttggtggcg cagcatctga	cgcacgtgtc gcagacgcag gagcagaaca	1980
aagagaatca gaatcccaat tctgatgcgc	cggtgatcag atcaaaaact tcagccaggt	2040
acatggagct ggtcgggtgg ctctggaca	aggggattac ctccgagaag cagtggatcc	2100
aggtgggtgt ctttcctgcc tgagctgacc	tgggcagacg cgtaattaag tatcatgata	2160
aactgacaaa gcggtatata aagatactta	atacatatgg tcctgtaaaa gtttaatgtg	2220
taagaagtat ttgttataaa agataaatat	tcagaatctt ctttttaatt cctgatttta	2280
tttctatagg actgaaagac ttgctcgaga	tgtcatgaag gagatgggag gccatcacat	2340

tctcacttca	ctgcagctca	acgtcccca	ttctaccggg	taggggaggc	gcttttccca	2400
aggcagtctg	gagcatgcgc	tttagcagcc	ccgctgggca	cttggcgcta	cacaagtggc	2460
ctctggcctc	gcacacattc	cacatccacc	ggtaggcgcc	aaccggctcc	gttcctttggt	2520
ggcccccttcg	cgccaccttc	tactcctccc	ctagtcagga	agttcccccc	cgccccgcag	2580
ctcgcgtcgt	gcaggacgtg	acaaatggaa	gtagcacgtc	tcactagtct	cgtgcagatg	2640
gacagcaccg	ctgagcaatg	gaagcgggta	ggcctttggg	gcagcggcca	atagcagctt	2700
tgtctcttcg	ctttctgggc	tcagaggctg	ggaaggggtg	ggtccggggg	cgggctcagg	2760
ggcgggctca	ggggcggggc	gggcgcccga	aggtcctccg	gaggccccggc	attctgcacg	2820
cttcaaaagc	gcacgtctgc	cgcgtgttc	tcctcttctt	catctccggg	cctttcgacc	2880
tagcgggcag	tgagcgcaac	gcaattaatg	tgagttagct	cactcattag	gcaccccagg	2940
ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	tgtgagcgga	taacaatttc	3000
acacaggaaa	cagctgccac	catgattgaa	caagatggat	tgcacgcagg	ttctccggcc	3060
gcttgggtgg	agaggctatt	cggctatgac	tgggcacaa	agacaatcgg	ctgctctgat	3120
gccgccgtgt	tccggctgtc	agcgcagggg	cgcccggttc	tttttgta	gaccgacctg	3180
tccggtgccc	tgaatgaact	gcaagacgag	gcagcgcggc	tatcgtggct	ggccacgacg	3240
ggcgttcctt	gcgcagctgt	gctcgacgtt	gtcactgaag	cgggaaggga	ctggctgcta	3300
ttgggcgaag	tgccggggca	ggatctcctg	tcatctcacc	ttgctcctgc	cgagaaagta	3360
tccatcatgg	ctgatgcaat	gcggcggctg	catacgtttg	atccggctac	ctgcccattc	3420
gaccaccaag	cgaaacatcg	catcgagcga	gcacgtactc	ggatggaagc	cggctcttgtc	3480
gatcaggatg	atctggacga	agagcatcag	gggctcgcgc	cagccgaact	gttcgccagg	3540
ctcaaggcga	gcatgccccga	cggcgaggat	ctcgtcgtga	cccatggcga	tgcctgcttg	3600
ccgaatatca	tggtggaaaa	tggccgcttt	tctggattca	tcgactgtgg	ccggctgggt	3660
gtggcggacc	gctatcagga	catagcgttg	gctacccgtg	atattgctga	agagcttggc	3720
ggcgaatggg	ctgaccgctt	cctcgtgctt	tacggtatcg	ccgtccccga	ttcgcagcgc	3780
atcgctttct	atcgctttct	tgacgagttc	ttctgagata	cattgatgag	tttggacaaa	3840

ccacaactag aatgcagtga aaaaaatgct ttattttgtga aattttgtgat gctattgctt	3900
tattttgtaac cattataagc tgcaataaac aagttaacaa caacaattgc attcatttta	3960
tgtttcgggt tcagggggag gtgtgggagg ttttttaaag caagtaaaac ctctacaaat	4020
gtggtatggc tgattatgat cgcatagtgt taccatcaac caccttaact tcatttttct	4080
tattcaatac ctaggtaggt agatgctaga ttcttgaaat aaaatatgag tctcaagtgg	4140
tccttgtcct ctctcccagt caaattctga atctagttgg caagattctg aaatcaaggc	4200
atataatcag taataagtga tgatagaagg gtatatagaa ggctagccgt caacatctcc	4260
actgacaaag cggtttctcc aacgaattat cgattcggct gtggggtcct gtggtgtgtg	4320
gggagctgtc acatccaggg tcctcactgc ctgtccccct cctcctcag gaggaccagg	4380
cctcatacat ctccctcaat gcggcctcca actcgcggtc ccaaattaag gctgccttgg	4440
acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg gtgggccagc	4500
agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta aacgggtacg	4560
atccccaata tgcggcttcc gtctttctgg gatggggccac gaaaaagttc ggcaagagga	4620
acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg gaggccatag	4680
cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt cccttcaacg	4740
actgtgtcga caagatgggtg atctgggtggg aggaggggaa gatgaccgcc aaggtcgtgg	4800
agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa tgcaagtcct	4860
cggcccagat agacccgact cccgtgacg tcacctccaa caccaacatg tgcgccgtga	4920
ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg atgttcaaat	4980
ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag gaagtcaaag	5040
actttttccg gtggggcaaag gatcacgtgg ttgaggtgga gcatgagttc tacgtcaaaa	5100
agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag cccaaacggg	5160
tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgac aactacgcag	5220
acaggtacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg tttccctgca	5280
gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga cagaaagact	5340

gttttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc	5400
agaaactgtg ctacattcat catatcatgg gaaagggtgcc agacgcttgc actgcctgcg	5460
atctgggtcaa tgtggatttg gatgactgca tctttgaaca ataaatgatt taaatcaggt	5520
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga agggattcgc	5580
gagtgggtggg ctttgaaacc tggagccccct caaccaagg caaatcaaca acatcaagac	5640
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac	5700
aagggggagc cggteaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac	5760
cagcagctca aggccggaga caacccttac ctcaagtaca accacgccga cgccgagttc	5820
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag	5880
gccaaaaaga ggcttcttga acctcttggg ctggttgagg aagcggctaa gacggctcct	5940
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggatttggc	6000
aaatcgggtg cacagccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag	6060
tcagteccag acctcaacc aatcggagaa cctcccgag cccctcagg tgtgggatct	6120
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga	6180
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac	6240
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc	6300
tccaacagca catctggagg atcttcaaat gacaacgcct atttcggcta cagcaccccc	6360
tgggggtatt ttgacttcaa cagattccac tgccattct caccacgtga ctggcagcga	6420
ctcatcaaca acaactgggg attccggcct aagcgactca atttcaagct cttcaacatt	6480
cagggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc	6540
acggteccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggetcac	6600
gagggctgcc tcccgcggtt cccagcggac gttttcatga ttcttcagta cgggtatctg	6660
acgcttaatg atggaagcca ggccgtgggt cgttcgctct tttactgcct ggaatatattc	6720
cgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta	6780
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc	6840

gaccaatact	tgtactatct	ctcaaagact	attaacggtt	ctggacagaa	tcaacaaacg	6900
ctaaaattca	gtgtggccgg	accagcaac	atggctgtcc	agggaagaaa	ctacatacct	6960
ggaccagct	accgacaaca	acgtgtctca	accactgtga	ctcaaaacaa	caacagcgaa	7020
tttgcttggc	ctggagcttc	ttcttgggct	ctcaatggac	gtaatagctt	gatgaatcct	7080
ggacctgcta	tggccagcca	caaagaagga	gaggaccgtt	tctttccttt	gtctggatct	7140
ttaatttttg	gcaaacaagg	aactggaaga	gacaacgtgg	atgcggacaa	agtcatgata	7200
accaacgaag	aagaaattaa	aactactaac	ccggtagcaa	cggagtccta	tggacaagtg	7260
gccacaaacc	accagagtgc	ccaagcacag	gcgcagaccg	gctgggttca	aaaccaagga	7320
atacttccgg	gtatggtttg	gcaggacaga	gatgtgtacc	tgcaaggacc	catttgggcc	7380
aaaattcctc	acacggacgg	caactttcac	ccttctccgc	tgatgggagg	gtttggaatg	7440
aagcacccgc	ctcctcagat	cctcatcaaa	aacacacctg	tacctgcgga	tcctccaacg	7500
gccttcaaca	aggacaagct	gaactctttc	atcacccagt	attctactgg	ccaagtcagc	7560
gtggagatcg	agtgggagct	gcagaaggaa	aacagcaagc	gctggaacce	ggagatccag	7620
tacacttcca	actattacaa	gtctaataat	gttgaatttg	ctgttaatac	tgaaggtgta	7680
tatagtgaac	cccgccccat	tggcaccaga	tacctgactc	gtaatctgta	attgcttgtt	7740
aatcaataaa	ccgtttaatt	cgtttcagtt	gaacttttgt	ctctgcggtt	taaacggtec	7800
tgtattagag	gtcacgtgag	tgttttgcga	cattttgcga	caccatgtgg	tcacgctggg	7860
tatttaagcc	cgagtgagca	cgcagggtct	ccattttgaa	gcgggagggt	tgaacgcgcg	7920
cggccgcccc	aataaaaata	gaaaaacaac	aatgttaat	tttctatttt	tatttggctc	7980
agacctgcag	gcagctgcgc	gctcgctcgc	tactgaggc	cgcccgggca	aagcccgggc	8040
gtcgggcgac	ctttggctgc	ccggcctcag	tgagcgagcg	agcgcgcaga	gagggagtgg	8100
ccaactccat	cactaggggt	tccttctaga	cgtgaggctc	cggtgcccgt	cagtgggcag	8160
agcgcacatc	gcccacagtc	cccgagaagt	tggggggagg	ggtcggcaat	tgaaccggtg	8220
cctagagaag	gtggcgcggg	gtaaaactggg	aaagtgatgt	cgtgtactgg	ctccgccttt	8280
ttcccgaggg	tgggggagaa	ccgtatataa	gtgcagtagt	cgccgtgaac	gttccttttc	8340

gcaacgggtt	tgccgccaga	acacaggtaa	gtgccgtgtg	tggttcccgc	gggcctggcc	8400
tctttacggg	ttatggccct	tgcgtgccct	gaattacttc	cacgcccctg	gctgcagtac	8460
gtgattcttg	atcccgagct	tcgggttgga	agtgggtggg	agagttcgag	gccttgcgct	8520
taaggagccc	cttcgcctcg	tgcttgagtt	gaggcctggc	ctgggcgctg	gggccgccgc	8580
gtgcgaatct	ggtggcacct	tcgcgcctgt	ctcgtctgct	tcgataagtc	tctagccatt	8640
taaaattttt	gatgacctgc	tcgcacgctt	tttttctggc	aagatagtct	tgtaaatgcg	8700
ggccaagatc	tgcacactgg	tatttcggtt	tttggggccg	cgggcggcga	cggggcccgt	8760
gcgtcccagc	gctcatgttc	ggcgaggcgg	ggcctgcgag	cgcggccacc	gagaatcgga	8820
cgggggtagt	ctcaagctgg	ccggcctgct	ctggtgcctg	gcctcgcgcc	gccgtgtatc	8880
gccccgccct	gggcggcaag	gctggcccgg	tcggcaccag	ttgcgtgagc	ggaaagatgg	8940
ccgcttcccc	gccctgctgc	agggagctca	aaatggagga	cgcggcgctc	gggagagcgg	9000
gcgggtgagt	cacccacaca	aaggaaaagg	gcctttccgt	cctcagccgt	cgtttcatgt	9060
gactccacgg	agtaccgggc	gccgtccagg	cacctcgatt	agttctcgag	cttttggagt	9120
acgtcgtctt	taggttgggg	ggaggggttt	tatgcgatgg	agtttcccca	cactgagtgg	9180
gtggagactg	aagttaggcc	agcttggcac	ttgatgtaat	tctccttgga	atttgccctt	9240
tttgagtttg	gatcttggtt	cattctcaag	cctcagacag	tggttcaaag	tttttttctt	9300
ccatttcagg	tgtcgtgagc	caccatggtg	agcaagggcg	aggaggataa	catggccatc	9360
atcaaggagt	tcatgcgctt	caaggtgcac	atggagggct	ccgtgaacgg	ccacgagttc	9420
gagatcgagg	gcgagggcga	gggccgcccc	tacgagggca	cccagaccgc	caagctgaag	9480
gtgaccaagg	gtggccccct	gcccttcgcc	tgggacatcc	tgteccctca	gttcatgtac	9540
ggctccaagg	cctacgtgaa	gcaccccgcc	gacatccccg	actacttgaa	gctgtccttc	9600
cccgagggct	tcaagtggga	gcgcgtgatg	aacttcgagg	acggcggcgt	ggtgaccgtg	9660
accagggact	cctccctgca	ggacggcgag	ttcatctaca	aggtgaagct	gcgcggcacc	9720
aacttccccct	ccgacggccc	cgtaatgcag	aagaagacca	tgggctggga	ggcctcctcc	9780
gagcggatgt	accccagga	cggcgccttg	aagggcgaga	tcaagcagag	gctgaagctg	9840

aaggacggcg gccactacga cgctgaggtc aagaccacct acaaggccaa gaagcccgtg	9900
cagctgcccc gcgcctacaa cgtcaacatc aagttaggaca tcacctccca caacgaggac	9960
tacaccatcg tggaacagta cgaacgcgcc gagggccgcc actccaccgg cggcatggac	10020
gagctgtaca agggaagcgg acagtgtact aattatgctc tcttgaaatt ggctggagat	10080
gttgagagca accctggacc tatggcgacc cgcagccctg gcgtcgtgat tagtgatgat	10140
gaacccggtt atgaccttga tttattttgc atacctaatc attatgctga ggatttggaa	10200
agggtgttta ttccatcatgg actaattatg gacaggactg aacgtcttgc tcgagatgtg	10260
atgaaggaga tgggaggcca tcacattgta gccctctgtg tgctcaaggg gggctataaa	10320
ttctttgctg acctgctgga ttacatcaaa gcactgaata gaaatagtga tagatccatt	10380
cctatgactg tagattttat cagactgaag agctattgta atgaccagtc aacagggggac	10440
ataaaagtaa ttggtggaga tgatctctca actttaactg gaaagaatgt cttgattgtg	10500
gaagatataa ttgacactgg caaaacaatg cagactttgc ttcccttggt caggcagtat	10560
aatccaaaga tggteaaggt cgctagcttg ctggtgaaaa ggaccccacg aagtgttgga	10620
tataagccag actttgttgg atttgaaatt ccagacaagt ttgtttagg atatgccctt	10680
gactataatg aatacttcag ggatttgaat catgtttgtg tcattagtga aactggaaaa	10740
gcaaaataca aagcctaaac gcgtggggga ggctaactga aacacggaag gagacaatac	10800
cggaaggAAC ccgcgctatg acggcaataa aaagacagaa taaaacgcac ggtgttgggt	10860
cgtttgttca taaacgcggg gtteggtecc agggctggca ctctgtcgat accccaccga	10920
gaccccatlg gggccaatac gcccgcttt cttecttttc cccacccac cccccaagtt	10980
cgggtgaagg cccagggctc gcagccaacg tcggggcggc aggcctgcc atagccatct	11040
atgtcgggtg cggagaaaga ggtaatagga acccctagtg atggagttgg ccactccctc	11100
tctgcgcgt ctctcgctca ctgaggccgg gcgaccaaag gtcgcccga ccccggtt	11160
tgcccgggcg gcctcagtga gcgagcgagc gcgcagctgc ctgcaggtag ggataacagg	11220
gtaatacttg aattttcaag tataaagtct agtgctaaat ttaatttgaa caactgtata	11280
gtttttgctg gttgggggaa ggaaaaaaaa tggtggcagt gtttttttca gaattagaag	11340

tgaaatgaaa	acttgttgtg	tgtgaggatt	tctaattgcca	tgtggtggtt	gcatactgag	11400
tgaagccggt	gagcattctg	ccatgtcacc	ccctcgtgct	cagtaatgta	ctttacagaa	11460
atcctaaact	caaaagattg	atataaacca	tgcttcttgt	gtatatccgg	tctcttctct	11520
gggtagtctc	actcagcctg	catttctgcc	atcccttgtc	ctctggcccc	cgtcgcccag	11580
atatctggca	aacagctatt	atgggtatta	tgggtttcag	ggagtggcgc	agctgcttca	11640
tccccgtggc	ccgttgctcg	cgtttgctgg	cggtgtcccc	ggaagaaata	tatttgcatg	11700
tctttagttc	tatgatgaca	caaacccegc	ccagcgtctt	gtcattggcg	aaaacacgca	11760
gatgcagtcg	gggcggcgcg	gtcccaggtc	cacttcgcat	attaaggtga	cgcgtgtggc	11820
ctcgaacaca	gagcgacggc	caccatgacc	gagtacaagc	ccacggtgcg	cctcgccacc	11880
cgcgacgacg	tcccccgggc	cgtacgcacc	ctcgccgccg	cgttcgccga	ctaccccgcc	11940
acgcgccaca	ccgttgaccc	ggaccgccac	atcgagcggg	tcaccgagct	gcaagaactc	12000
ttcctcacgc	gcgtcgggct	cgacatcggc	aaggtgtggg	tcgcggacga	cggcgcgcgcg	12060
gtggcggtct	ggaccacgcc	ggagagcgtc	gaagcggggg	cggtgttcgc	cgagatcggc	12120
ccgcgcatgg	ccgagttgag	cggttcccgg	ctggccgcgc	agcaacagat	ggaaggcctc	12180
ctggcgccgc	accggcccaa	ggagcccgcg	tggttcctgg	ccaccgtcgg	cgtctcgccc	12240
gaccaccagg	gcaagggtct	gggcagcgcg	gtcgtgtctc	ccggagtgga	ggcggccgag	12300
cgcgccgggg	tgcccgcttt	cctggagacc	tccgcgcccc	gcaacctccc	cttctacgag	12360
cggctcggct	tcaccgtcac	cgccgacgtc	gaggtgcccc	aaggaccgcg	cacctggtgc	12420
atgacccgca	agccccgtgc	ctgatgtgcc	ttctagttag	cagccatctg	ttgtttgccc	12480
ctcccccgctg	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	cctaataaaa	12540
tgaggaaatt	gcategcatt	gtctgagtag	gtgtcattct	attctggggg	gtgggggtggg	12600
gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	atgcggtggg	12660
ctctatggtc	ccccttatta	accctaaacg	ggtagcatat	gcttccccggg	tagtagtata	12720
tactatccag	actaacceta	attcaatagc	atatgttacc	caacgggaag	catatgctat	12780
cgaattaggg	ttagtaaaag	ggtcctaagg	aacagcgatc	tggatagcat	atgctatcct	12840

aatctatatac tgggtagcat atgctatcct aatctatatac tgggtagcat aggctatcct	12900
aatctatatac tgggtagcat atgctatcct aatctatatac tgggtagtat atgctatcct	12960
aatttatatac tgggtagcat aggctatcct aatctatatac tgggtagcat atgctatcct	13020
aatctatatac tgggtagtat atgctatcct aatctgtatac cgggtagcat atgctatcct	13080
catgcatata cagtcagcat atgataccca gtagtagagt gggagtgcta tcctttgcat	13140
atgccgccac ctcccaagga ttatgggtga cgtcaggtgg cacttttcgg ggaaatgtgc	13200
gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac	13260
aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt	13320
tccgtgtcgc ccttatcccc ttttttgcgg cattttgcct tcctgttttt gctcaccag	13380
aaacgctggg gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg	13440
aactggatct caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa	13500
tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc	13560
aagagcaact cggtcgcgcg atacactatt ctcagaatga cttgggtgag tactcaccag	13620
tcacagaaaa gcactttacg gatggcatga cagtaagaga attatgcagt gctgccataa	13680
ccatgagtga taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc	13740
taaccgcttt tttagcacaac atgggggatc atgtaactcg ctttgatcgt tgggaaccgg	13800
agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa	13860
caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa	13920
tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg	13980
gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag	14040
cactgggggc agatggtaag cctccccgta tcgtagttat ctacacgacg gggagtcagg	14100
caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt	14160
ggtaactgtc agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt	14220
aattttaaag gatctaggtg aagatccttt ttgataatct catgaccaaa atcccctaac	14280
gtgagttttc gtccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttccttgag	14340

atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg	14400
tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca	14460
gagcgcagat accaaatact gttctttctag tgtagccgta gttaggccac cacttcaaga	14520
actctgtagc accgcctaca tacctcgtct tgctaatect gttaccagtg gctgctgcca	14580
gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc	14640
agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca	14700
ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa	14760
aggcggacag gtatccggtt agcggcaggg tcggaacagg agagcgcacg agggagcttc	14820
cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc	14880
gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg	14940
cctttttacg gttcctggcc ttttgctggc cttttgctca catgt	14985

- <210> 13
- <211> 4409
- <212> DNA
- <213> 人工序列

- <220>
- <223> 在插入止擋匣之前的REP2/CAP9序列

<400> 13	
ggtcctgtat tagaggtcac gtgagtgttt tgcgacattt tgcgacacca tgtggtcacg	60
ctgggtattt aagcccgagt gaggacgcag ggtctccatt ttgaagcggg aggtttgaac	120
gcgcagccgc catgccgggg ttttacgaga ttgtgattaa ggtccccagc gaccttgacg	180
agcatctgcc cggcattttct gacagctttg tgaactgggt ggccgagaag gaatgggagt	240
tgccgccaga ttctgacatg gatctgaatc tgattgagca ggcaccctg accgtggccg	300
agaagctgca gcgcgacttt ctgacggaat ggcgccgtgt gagtaaggcc ccggaggccc	360
ttttctttgt gcaatttgag aagggagaga gctacttcca catgcacgtg ctcgtggaaa	420
ccaccgggggt gaaatccatg gttttgggac gtttcctgag tcagattcgc gaaaaactga	480
ttcagagaat ttaccgcggg atcgagccga ctttgccaaa ctggttcgcg gtcacaaaga	540

ccagaaatgg cgccggaggc gggaacaagg tggatggatga gtgctacatc cccaattact	600
tgtccccaa aaccacagcct gagctccagt gggcgtggac taatatggaa cagtatttaa	660
gcgcctgttt gaatctcacg gagcgtaaac ggttggatggc gcagcatctg acgcacgtgt	720
cgcagacgca ggagcagaac aaagagaatc agaateccaa ttctgatgcg ccggtgatca	780
gatcaaaaac ttcagccagg tacatggagc tggtcgggtg gctcgtggac aaggggatta	840
cctcggagaa gcagtggatc caggaggacc aggcctcata catctccttc aatgcggcct	900
ccaactcgcg gtcccaaatt aaggctgcct tggacaatgc gggaaagatt atgagcctga	960
ctaaaaccgc ccccgactac ctggtgggcc agcagcccgt ggaggacatt tccagcaatc	1020
ggatttataa aattttggaa ctaaacgggt acgatcccca atatgcggct tccgtctttc	1080
tgggatgggc cacgaaaaag ttcggcaaga ggaacaccat ctggctgttt gggcctgcaa	1140
ctaccgggaa gaccaacatc gcggaggcca tagcccacac tgtgcccttc tacgggtgcg	1200
taaactggac caatgagaac tttcccttca acgactgtgt cgacaagatg gtgatctggt	1260
gggaggaggg gaagatgacc gccaaagtcg tggagtcggc caaagccatt ctccgaggaa	1320
gcaaggtgcg cgtggaccag aaatgcaagt cctcggccca gatagaccgc actcccgtga	1380
tcgtcacctc caacaccaac atgtgcgccg tgattgacgg gaactcaacg accttcgaac	1440
accagcagcc gttgcaagac cggatgttca aatttgaact caccgcctgt ctggatcatg	1500
actttgggaa ggtcaccaag caggaagtca aagacttttt ccggtgggca aaggatcacg	1560
tggttgaggt ggagcatgag ttctacgtca aaaaggggtg agccaagaaa agacccgccc	1620
ccagtgacgc agatataagt gagcccaaac gggatgcgcga gtcagttgcg cagccatcga	1680
cgtcagacgc ggaagcttcg atcaactacg cagacaggta caaaacaaa tgttctcgtc	1740
acgtgggcat gaatctgatg ctgtttccct gcagacaatg cgagagaatg aatcagaatt	1800
caaatatctg cttcactcac ggacagaaag actgtttaga gtgctttccc gtgtcagaat	1860
ctcaaccctg ttctgtcgtc aaaaaggcgt atcagaaact gtgctacatt catcatatca	1920
tgggaaaggt gccagacgt tgcactgcct gcgatctggt caatgtggat ttggatgact	1980
gcactcttga acaataaatg atttaaatca ggtatggctg ccgatggtta tcttccagat	2040

tggctcgagg	acaaccttag	tgaagggatt	cgcgagtgg	gggctttgaa	acctggagcc	2100
cctcaaccca	aggcaaata	acaacatcaa	gacaacgctc	gaggtcttgt	gcttccgggt	2160
tacaaatacc	ttggacccgg	caacggactc	gacaaggggg	agccggtcaa	cgcagcagac	2220
gcggcggccc	tcgagcacga	caaggcctac	gaccagcagc	tcaaggccgg	agacaacccg	2280
tacctcaagt	acaaccaagc	cgacgccgag	ttccaggagc	ggctcaaaga	agatacgtct	2340
tttgggggca	acctcgggcg	agcagtcttc	caggccaaaa	agaggcttct	tgaacctctt	2400
ggtctggttg	aggaagcggc	taagacggct	cctggaaaga	agaggcctgt	agagcagtct	2460
cctcaggaac	cggactcctc	cgcgggtatt	ggcaaatcgg	gtgcacagcc	cgtaaaaag	2520
agactcaatt	tcggtcagac	tggcgacaca	gagtcagtcc	cagaccctca	accaatcgga	2580
gaacctcccg	cagccccctc	aggtgtggga	tctcttacia	tggttcagg	tggtggcgca	2640
ccagtggcag	acaataacga	aggtgccgat	ggagtgggta	gttcctcggg	aaattggcat	2700
tgcgattccc	aatggctggg	ggacagagtc	atcaccacca	gcacccgaac	ctgggccctg	2760
cccacctaca	acaatcaact	ctacaagcaa	atctccaaca	gcacatctgg	aggatcttca	2820
aatgacaacg	cctacttcgg	ctacagcacc	ccctgggggt	attttgactt	caacagattc	2880
cactgccact	tctcaccacg	tgactggcag	cgactcatca	acaacaactg	gggattccgg	2940
cctaagcgac	tcaacttcaa	gctcttcaac	attcaggtea	aagaggttac	ggacaacaat	3000
ggagtcaaga	ccatcgccaa	taaccttacc	agcacggtec	aggtcttcac	ggactcagac	3060
tatcagctcc	cgtacgtgct	cgggtcggct	cacgagggtc	gcctcccgcc	gttcccagcg	3120
gacgttttca	tgattcctca	gtacgggtat	ctgacgctta	atgatggaag	ccaggccgtg	3180
ggtcgttcgt	ccttttactg	cctggaatat	ttcccgtcgc	aaatgctaag	aacgggtaac	3240
aacttccagt	tcagctacga	gtttgagaac	gtacctttcc	atagcagcta	cgtcacagc	3300
caaagcctgg	accgactaat	gaatccactc	atcgaccaat	acttgtacta	tctctcaaag	3360
actattaacg	gttctggaca	gaatcaacaa	acgctaaaat	tcagtgtggc	cggaccacgc	3420
aacatggctg	tccagggaag	aaactacata	cctggaccca	gtaccgaca	acaacgtgtc	3480
tcaaccactg	tgactcaaaa	caacaacagc	gaatttgctt	ggcctggagc	ttcttcttgg	3540

gctctcaatg gacgtaatag cttgatgaat cctggacctg ctatggccag ccacaaagaa	3600
ggagaggacc gtttctttcc tttgtctgga tctttaatit ttggcaaaca aggaactgga	3660
agagacaacg tggatgcgga caaagtcatg ataaccaacg aagaagaaat taaaactact	3720
aacccggtag caacggagtc ctatggacaa gtggccacaa accaccagag tgcccaagca	3780
caggcgcaga ccggctgggt tcaaaaccaa ggaatacttc cgggtatggt ttggcaggac	3840
agagatgtgt acctgcaagg acccatttgg gccaaaattc ctcacacgga cggcaacttt	3900
cacccttctc cgctgatggg agggtttgga atgaagcacc cgcctcctca gatcctcatc	3960
aaaaacacac ctgtacctgc ggatcctcca acggccttca acaaggacaa gctgaactct	4020
ttcatcacc c agtattctac tggccaagtc agcgtggaga tcgagtggga gctgcagaag	4080
gaaaacagca agcgttgga cccggagatc cagtacactt ccaactatta caagtctaat	4140
aatgttgaat ttgctgttaa tactgaaggt gtatatagtg aaccccgccc cattggcacc	4200
agatacctga ctcgtaatct gtaattgctt gttaatcaat aaaccgttta attcgtttca	4260
gttgaacttt ggtctctgcg gtttaaacgg tcttgtatta gaggtcacgt gagtgttttg	4320
cgacattttg cgacaccatg tggtcacgct gggtatttta gcccgagtga gcacgcaggg	4380
tctccatttt gaagcgggag gtttgaacg	4409

<210> 14
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> β -肌動蛋白內含子之上游半部

<400> 14	
gtgggtgtct ttctgcctg agctgacctg ggcag	35

<210> 15
<211> 76
<212> DNA
<213> 人工序列

<220>
<223> β -肌動蛋白內含子之下游半部

<400> 15
tcggctgtgg ggtcctgtgg tgtgtgggga gctgtcacat ccagggtcct cactgcctgt 60

ccccctccct cctcag 76

<210> 16
<211> 2144
<212> DNA
<213> 人工序列

<220>
<223> 轉錄止擋匣

<400> 16
aattaagtat catgataaac tgacaaagcg gtatatcaag atacttaata catatggtec 60

tgtaaaagtt taatgtgtaa gaagtatttg ttataaaaga taaatattca gaatcttctt 120

tttaattcct gattttattt ctataggact gaaagacttg ctcgagatgt catgaaggag 180

atgggaggcc atcacattct cacttcactg cagctcaacg tcccgaattc taccgggtag 240

gggaggcgct ttteccaagg cagtctggag catgcgcttt agcagccccg ctgggcactt 300

ggcgctacac aagtggcctc tggcctcgca cacattccac atccaccggt aggcgccaac 360

cggtccggtt ctttgggtggc cccttcgcgc caccttctac tcctccccta gtcaggaagt 420

tccccccgc cccgcagctc gcgtcgtgca ggacgtgaca aatggaagta gcacgtctca 480

ctagtctcgt gcagatggac agcaccgctg agcaatggaa gcgggtaggc ctttggggca 540

gcggccaata gcagctttgc tccttcgctt tctgggctca gaggctggga aggggtgggt 600

ccggggggcgg gctcaggggc gggctcaggg gcgggggcggg cgcccgaagg tcctccggag 660

gcccggcatt ctgcacgctt caaaagcgca cgtctgccgc gctgttctcc tcttctcat 720

ctccgggcct ttcgacctag cgggcagtga gcgcaacgca attaatgtga gttagctcac 780

tcattaggca ccccaggctt tacactttat gcttccggct cgtatgttgt gtggaattgt 840

gagcggataa caatttcaca caggaaacag ctgccaccat gattgaacaa gatggattgc 900

acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 960

caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 1020

ttgtcaagac cgacctgtcc ggtgccctga atgaactgca agacgaggca gcgcggctat	1080
cgtaggctggc cacgacgggc gtcccttgcg cagctgtgct cgacgttgtc actgaagcgg	1140
gaagggactg gctgctattg ggcgaaagtgc cggggcagga tctcctgtca tctcaccttg	1200
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc	1260
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga	1320
tggaagccgg tcttgtcgat caggatgac tggacgaaga gcatcagggg ctgcgccag	1380
ccgaactgtt cgccaggctc aaggcgagca tgcccgcgg cgaggatctc gtcgtgacct	1440
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg	1500
actgtggccg gctgggtgtg gcggaccgt atcaggacat agcgttggct acccgtgata	1560
ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg	1620
ctcccgattc gcagcgcac gccctctatc gccctcttga cgagttcttc tgagatacat	1680
tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat	1740
ttgtgatgct attgctttat ttgtaacct tataagctgc aataaacaag ttaacaacaa	1800
caattgcatt cattttatgt ttcgggttca gggggaggtg tgggaggttt tttaaagcaa	1860
gtaaaacctc taaaatgtg gtatggctga ttatgatcgc atagtgttac catcaaccac	1920
cttaacttca tttttcttat tcaataccta ggtaggtaga tgctagattc tggaaataaa	1980
atatgagtct caagtgtcc ttgtcctctc tcccagtcaa attctgaatc tagttggcaa	2040
gattctgaaa tcaaggcata taatcagtaa taagtgatga tagaagggta tatagaaggc	2100
tagccgtcaa catctccact gacaaagcgg tttctccaac gaat	2144

- <210> 17
- <211> 170
- <212> DNA
- <213> 人工序列

- <220>
- <223> 強的剪接受體

<400> 17	
gtcctgtaaa agtttaatgt gtaagaagta tttgttataa aagataaata ttcagaatct	60

tctttttaat	tcctgatttt	atttctatag	gactgaaaga	cttgctcgag	atgtcatgaa	120
ggagatggga	ggccatcaca	ttctcacttc	actgcagctc	aacgtcccca		170
<210>	18					
<211>	1672					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	新黴素/康黴素抗性基因					
<400>	18					
attctaccgg	gtaggggagg	cgcttttccc	aaggcagtct	ggagcatgcg	cttttagcagc	60
cccgtgggc	acttggcgct	acacaagtgg	cctctggcct	cgcacacatt	ccacatccac	120
cggtaggcgc	caaccggctc	cgttcttttg	tggccccctt	gcgccacctt	ctactcctcc	180
cctagtcagg	aagttccccc	ccgccccgca	gctcgcgtcg	tgcaggacgt	gacaaatgga	240
agtagcacgt	ctcactagtc	tcgtgcagat	ggacagcacc	gctgagcaat	ggaagcgggt	300
aggccttttg	ggcagcggcc	aatagcagct	ttgctccttc	gctttctggg	ctcagaggct	360
gggaaggggt	gggtccgggg	gcgggctcag	gggcgggctc	aggggcgggg	cgggcgcccc	420
aaggtcctcc	ggaggccccg	cattctgcac	gcttcaaaag	cgcacgtctg	ccgcgtgttt	480
ctcctcttcc	tcattctcgg	gcctttcgac	ctagcgggca	gtgagcgcaa	cgcaattaat	540
gtgagttagc	tcactcatta	ggcaccccag	gctttacact	ttatgcttcc	ggctcgtatg	600
ttgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctgcca	ccatgattga	660
acaagatgga	ttgcacgcag	gttctccggc	cgcttgggtg	gagaggctat	tcggctatga	720
ctgggcacaa	cagacaatcg	gctgctctga	tgccgccgtg	ttccggctgt	cagcgcaggg	780
gcgcccggtt	ctttttgtca	agaccgacct	gtccggtgcc	ctgaatgaac	tgcaagacga	840
ggcagcgcgg	ctatcgtggc	tggccacgac	gggcgttcct	tgcgcagctg	tgctcgacgt	900
tgteactgaa	gcgggaaggg	actggctgct	attgggcgaa	gtgccggggc	aggatctcct	960
gtcatctcac	cttgctcctg	ccgagaaagt	atccatcatg	gctgatgcaa	tgcggcggct	1020
gcatacgctt	gatccggcta	cctgcccatt	cgaccaccaa	gcgaaacatc	gcatcgagcg	1080

agcacgtact cggatggaag ccggtcttgt cgatcaggat gatctggacg aagagcatca	1140
ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg agcatgcccg acggcgagga	1200
tctcgtcgtg acccatggcg atgcctgctt gccgaatatc atggtggaaa atggccgctt	1260
ttctggattc atcgactgtg gccggctggg tgtggcggac cgctatcagg acatagcgtt	1320
ggctaccCGT gatattgctg aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct	1380
ttacgggtatc gccgctcccg attcgcagcg catcgccctc tatcgccctc ttgacgagtt	1440
cttctgagat acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc	1500
tttatttgtg aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa	1560
caagttaaca acaacaattg cattcatttt atgtttcggg ttcaggggga ggtgtgggag	1620
gtttttttaa gcaagtaaaa cctctacaaa tgtggtatgg ctgattatga tc	1672

<210> 19
<211> 200
<212> DNA
<213> 人工序列

<220>
<223> 來自人類β 球蛋白之轉錄終止序列

<400> 19 gcatagtgtt accatcaacc accttaactt catTTTTcTT attcaatacc taggtaggta	60
gatgctagat tctggaaata aaatatgagt ctcaagtggT ccttgTcctc tctcccagtc	120
aaattctgaa tctagtTggc aagattctga aatcaaggca tataatcagt aataagtgat	180
gatagaaggg tatatagaag	200

<210> 20
<211> 130
<212> DNA
<213> 人工序列

<220>
<223> ITR，正向

<400> 20 ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt	60
---	----

ggtcgccccg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttcct	130
<210> 21	
<211> 1188	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人類EF1- α 啓動子、外顯子1、內含子1、及外顯子2（部分）	
<400> 21	
cgtgaggctc cggtgccccgt cagtgggcag agcgcacatc gcccacagtc cccgagaagt	60
tggggggagg ggtcggcaat tgaaccggtg cctagagaag gtggcgcggg gtaaactggg	120
aaagtgatgt cgtgtactgg ctccgccttt ttcccagagg tgggggagaa ccgtatataa	180
gtgcagtagt cgccgtgaac gttctttttc gcaacgggtt tgccgccaga acacaggtaa	240
gtgccgtgtg tggttcccgc gggcctggcc tctttacggg ttatggccct tgcgtgcctt	300
gaattacttc cacgcccctg gctgcagtac gtgattcttg atcccagact tcgggttgga	360
agtgggtggg agagttcgag gccttgcgct taaggagccc cttcgcctcg tgcttgagtt	420
gaggcctggc ctgggcgctg gggccgccgc gtgcgaatct ggtggcacct tcgcgcctgt	480
ctcgctgctt tcgataagtc tctagccatt taaaattttt gatgacctgc tgcgacgctt	540
tttttctggc aagatagtct tgtaaatgcg ggccaagatc tgcacactgg tatttcggtt	600
tttggggccg cgggcggcga cggggcccgt gcgtcccagc gctcatgttc ggcgaggcgg	660
ggcctgcgag cgcggccacc gagaatcgga cgggggtagt ctcaagctgg ccggcctgct	720
ctggtgcctg gcctcgcgc gccgtgtatc gccccgccct gggcggcaag gctggcccgg	780
tcggcaccag ttgcgtgagc ggaaagatgg ccgcttcccc gccctgctgc agggagctca	840
aatggagga cgcggcgctc gggagagcgg gcgggtgagt caccacaca aaggaaaagg	900
gcctttccgt cctcagccgt cgcttcatgt gactccacgg agtaccgggc gccgtccagg	960
cacctcgatt agttctcgag cttttggagt acgtcgtctt taggttgggg ggaggggttt	1020
tatgcgatgg agtttcccca cactgagtgg gtggagactg aagttaggcc agcttggcac	1080

ttgatgtaat tctccttgga atttgccctt tttagatttg gatcttggtt cattctcaag	1140
cctcagacag tggttcaaag tttttttctt ccatttcagg tgtcgtga	1188
<210> 22	
<211> 1434	
<212> DNA	
<213> 人工序列	
<220>	
<223> mCherry-E2A-人類HPRT融合基因	
<400> 22	
atggtgagca agggcgagga ggataacatg gccatcatca aggagttcat gcgcttcaag	60
gtgcacatgg agggctccgt gaacggccac gagttcgaga tcgagggcga gggcgagggc	120
cgccccctacg agggcaccca gaccgccaag ctgaagggtga ccaagggtgg cccccgccc	180
ttcgctggg acatcctgtc cctcagttc atgtacggct ccaaggccta cgtgaagcac	240
cccgccgaca tccccgacta cttgaagctg tccttccccg agggcttcaa gtgggagcgc	300
gtgatgaact tcgaggacgg cggcgtggtg accgtgacce aggactcctc cctgcaggac	360
ggcgagttca tctacaaggt gaagctgcgc ggcaccaact tccccccga cggccccgta	420
atgcagaaga agaccatggg ctgggaggcc tcctccgagc ggatgtacce cgaggacggc	480
gccctgaagg gcgagatcaa gcagaggctg aagctgaagg acggcggcca ctacgacgt	540
gaggtaaga ccacctaaa ggccaagaag cccgtgcagc tgccccggcg ctacaacgtc	600
aacatcaagt tggacatcac ctcccacaac gaggactaca ccategtgga acagtacgaa	660
cgcgccgagg gccgccactc caccggcggc atggacgagc tgtacaaggg aagcggacag	720
tgtactaatt atgctctctt gaaattggct ggagatgttg agagcaacce tggacctatg	780
gcgacccgca gccctggcgt cgtgattagt gatgatgaac ccggttatga ccttgattta	840
ttttgcatac ctaatcatta tgctgaggat ttggaaaggg tgttttattcc tcatggacta	900
attatggaca ggactgaacg tcttgctcga gatgtgatga aggagatggg aggccatcac	960
attgtagccc tctgtgtgct caaggggggc tataaattct ttgctgacct gctggattac	1020
atcaaagcac tgaatagaaa tagtgataga tccattccta tgactgtaga ttttatcaga	1080

ctgaagagct attgtaatga ccagtcaca ggggacataa aagtaattgg tggagatgat	1140
ctctcaactt taactggaaa gaatgtcttg attgtggaag atataattga cactggcaaa	1200
acaatgcaga ctttgctttc cttggtcagg cagtataatc caaagatggc caaggtcgct	1260
agcttgctgg tgaaaaggac cccacgaagt gttggatata agccagactt tgttggattt	1320
gaaattccag acaagtttgt tgtaggatat gcccttgact ataataata cttcagggat	1380
ttgaatcatg tttgtgtcat tagtgaaact ggaaaagcaa aatacaaagc ctaa	1434

<210> 23
<211> 272
<212> DNA
<213> 人工序列

<220>
<223> 來自單純疱疹病毒胸苷激酶基因之多腺苷酸化信號

<400> 23	
gggggaggct aactgaaaca cggaaggaga caataccgga aggaacccgc gctatgacgg	60
caataaaaag acagaataaa acgcacggtg ttgggtcggt tgttcataaa cgcgggggttc	120
ggccccaggg ctggcactct gtcgataccc caccgagacc ccattggggc caatacggcc	180
gcgtttcttc cttttcccca cccaccccc caagttcggg tgaaggccca gggctcgcag	240
ccaacgtcgg ggcggcaggc cctgccatag cc	272

<210> 24
<211> 1031
<212> DNA
<213> 人工序列

<220>
<223> 人類抗抑制因子40 (AY190756.1)

<400> 24	
gatcaagaaa gcactccggg ctccagaagg agccttccag gccagctttg agcataagct	60
gctgatgagc agtgagtgtc ttgagtagtg ttcagggcag catgttacca ttcatgcttg	120
acttctagcc agtgtgacga gaggctggag tcaggtctct agagagttga gcagctccag	180
ccttagatct cccagtttta tgcggtgtgc ccattcgctt tgtgtctgca gtcccctggc	240

cacacccagt aacagttctg ggatctatgg gagtagcttc cttagtgage tttcccttca	300
aatactttgc aaccaggtag agaagtttgg agtgaaggtt ttgtttctcg tttcttcaca	360
atatggatat gcatcttctt ttgaaaatgt taaagtaaata taccctctctt ttcagatact	420
gtcttcatgc gaacttggtg tcctgtttcc atcccagcct tctataaccc agtaacatct	480
tttttgaaac cagtgggtga gaaagacacc tggtcaggaa cgcggaccac aggacaactc	540
aggctcacc caggcatcag actaaaggca aacaaggact ctgtataaag taccggtggc	600
atgtgtatta gtggagatgc agcctgtgct ctgcagacag ggagtcacac agacactttt	660
ctataatttc ttaagtgcct tgaatgttca agtagaaagt ctaacattaa atttgattga	720
acaattgtat attcatggaa tattttggaa cggaatacca aaaaatggca atagtggttc	780
tttctggatg gaagacaaac tttcttctt taaaataaat tttattttat atatttgagg	840
ttgaccacat gaccttaagg atacatatag acagtaaact ggttactaca gtgaagcaaa	900
ttaacatata taccatcgta catagttaca tttttttgtg tgacaggaac agctaaaatc	960
tacgtattta acaaaactcc taaagacaat acatttttat taactatagc cctcatgatg	1020
tacattagat c	1031

<210> 25
<211> 354
<212> DNA
<213> 人工序列

<220>
<223> 小鼠抗抑制因子40

<400> 25 acttgaatth tcaagtataa agtctagtgc taaatttaata ttgaacaact gtatagthth	60
tgctggthtg gggaaggaaa aaaaatggtg gcagtgttht tttcagaatt agaagtgaah	120
tgaaaacttg ttgtgtgtga ggattttctaa tgccatgtgg tggthgcata ctgagtgaag	180
ccggtgagca ttctgccatg tcacccccctc gtgctcagta atgtacttta cagaaatcct	240
aaactcaaaa gattgatata aaccatgctt cthgtgtata tccggtctct tctctgggta	300
gtctcactca gcctgcattt ctgccatccc ttgtctctctg gccccgctc ccca	354

<210>	26	
<211>	222	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	疱疹病毒胸苷激酶基因啓動子	
<400>	26	
	ttcagggagt ggcgcagctg cttcatcccc gtggcccgtt gctcgcgttt gctggcggtg	60
	tccccggaag aaatatatTT gcatgtcttt agttctatga tgacacaaac cccgcccagc	120
	gtcttgtcat tggcgaaaac acgcagatgc agtcggggcg gcgcgggtccc aggtccactt	180
	cgcataTTaa ggtgacgcgt gtggcctcga acacagagcg ac	222
<210>	27	
<211>	600	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	嘌呤黴素N-乙酰基轉移酶編碼區域	
<400>	27	
	atgaccgagt acaagcccac ggtgcgcctc gccacccgcg acgacgtccc ccgggccgta	60
	cgcaccctcg ccgccgcgtt cgccgactac cccgccacgc gccacaccgt tgacccggaC	120
	cgccacatcg agcgggtcac cgagctgcaa gaactcttcc tcacgcgcgt cgggctcgac	180
	atcggaagg tgtgggtcgc ggacgacggc gccgcggtgg cggctctggac cacgccggag	240
	agcgtcgaag cggggggcgtt gttcgccgag atcggcccgC gcatggccga gttgagcgtt	300
	tcccggctgg ccgcgcagca acagatggaa ggcctcctgg cgccgcaccg gcccAaggag	360
	cccgctgggt tcctggccac cgtcggcgtc tcgcccgacc accagggcaa gggctctgggc	420
	agcgccgtcg tgctccccgg agtggaggcg gccgagcgcg ccgggggtgcc cgccttcctg	480
	gagacctccg cgccecgcaa cctcccccttC tacgagcggc tcggcttcac cgtcaccgcc	540
	gacgtcgagg tgcccgaagg accgcgcacc tggtgcatga cccgcaagcc cggtgcctga	600

<210> 28
<211> 224
<212> DNA
<213> 人工序列

<220>
<223> 來自牛生長荷爾蒙基因之多腺苷酸化信號

<400> 28
tgtgccttct agttgccagc catctgttgt ttgccccctcc cccgtgcctt ccttgaccct 60

ggaaggtgcc actcccactg tcttttcccta ataaaatgag gaaattgcat cgcattgtct 120

gagtaggtgt cattctattc tggggggtgg ggtggggcag gacagcaagg gggaggattg 180

ggaagacaat agcaggcatg ctggggatgc ggtgggctct atgg 224

<210> 29
<211> 490
<212> DNA
<213> 人工序列

<220>
<223> 艾司坦-巴爾病毒(Epstein Barr Virus) OriP複製因子

<400> 29
cccccttatt aaccctaaac gggtagcata tgcttcccgg gtagtagtat atactatcca 60

gactaaccct aattcaatag catatgttac ccaacgggaa gcatatgcta tcgaattagg 120

gttagtaaaa gggtcctaag gaacagcgat ctggatagca tatgctatcc taatctatat 180

ctgggtagca tatgctatcc taatctatat ctgggtagca taggctatcc taatctatat 240

ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatttatat 300

ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat 360

ctgggtagta tatgctatcc taatctgtat ccgggtagca tatgctatcc tcatgcatat 420

acagtcagca tatgataccc agtagtagag tgggagtgct atcctttgca tatgccgcca 480

cctcccaagg 490

<210> 30
<211> 1815
<212> DNA
<213> 人工序列

<220>

<223> 編碼質體複製起點及安比西林(ampicillin)抗性基因之
pUC57載體序列

<400> 30

cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa	60
tacattcaaa tatgtatcgg ctcatgagac aataaccctg ataaatgctt caataatatt	120
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc cttattccc ttttttgcg	180
cattttgcct tcctgttttt gtcacccag aaacgctggt gaaagtaaaa gatgctgaag	240
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg	300
agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg	360
gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt	420
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcattcttac gatggcatga	480
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac	540
ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatac	600
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc	660
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac	720
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag	780
gaccacttct gcgctcggcc ctccggctg gctggtttat tgctgataaa tctggagccg	840
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta	900
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg	960
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata	1020
tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt	1080
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc	1140
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct	1200
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa	1260
ctctttttcc gaaggtaaact ggcttcagca gagegcagat accaaatact gttcttctag	1320

tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca taccctcgctc	1380
tgctaatcct gttaccagtg gctgctgcc a gtggcgataa gtcgtgtctt accgggttgg	1440
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg gggttcgtgca	1500
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat	1560
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggt a agcggcaggg	1620
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc	1680
ctgtcggggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc	1740
ggagcctatg gaaaaacgcc agcaacgcgg ccttttttacg gttcctggcc ttttgctggc	1800
cTTTTGCTCA catgt	1815

<210> 31
<211> 7671
<212> DNA
<213> 人工序列

<220>
<223> 用於產生表現SR21重組酶之重組腺病毒的
P321質體

<400> 31	
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt	60
ggattgaagc caatatgata atgaggggggt ggagtttgtg acgtggcgcg gggcgtggga	120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca	180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt	240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga	300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt	360
actcatagcg cgtaatatTT gtctagggcc gcggggactt tgaccgttta cgtggagact	420
cgcccagggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat	480
tatagtcagt acgtaccagt gcactggcct aggcctctctt aaggtagctc gaggagcttg	540
gcccattgca tacgttgtat ccatatcata atatgtacat ttatatTggc tcatgtccaa	600
cattaccgcc atgttgacat tgattattga ctagttatta atagtaatca attacggggt	660

cattagttca tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc	720
ctggctgacc gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag	780
taacgccaat agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc	840
acttggcagt acatcaagtg tatcatatgc caagtacgcc cctatitgac gtcaatgacg	900
gtaaatggcc cgctggcat tatgcccagt acatgacctt atgggacttt cctacttggc	960
agtacatcta cgtattagtc atcgtatatta ccatggtgat gcggttttgg cagtacatca	1020
atgggcgtgg atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca	1080
atgggagttt gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactccg	1140
ccccattgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata agcagagctc	1200
gttttagtgaa ccgtcggcgc gccgccacca tggaattaaa acacattgtc aattcctata	1260
atgtgactaa aattattggc tatttgagac gticcagaca agatgttgaa agagagaagc	1320
gtactggtga agatactttg acagaacaaa aagaacttat gaataaaata ttgactggga	1380
ttgaaatacc gtacgaaaca agaactgaga tcggttctgg tgaaagcatt gaaggaagac	1440
cagtgtttaa aagttgctta gctgatctga ggtctggaaa atttcaagca attgctgtaa	1500
aagaaataac tcgtcttagt cgtggtagtt acagtgatgc tggtgaaatc gtcaatcttc	1560
taaatgaaaa aaggattatc ataattacac catataaaat atatgatcca agaaatcctg	1620
ttgatgctcg tcaaattaga tttgagctat ttatggctag agaagaattt gaaatgacaa	1680
gagaaagaat gaatggtgcc aagtttactt atgctgcaca aggaaaatgg atttctggtt	1740
tagcaccatt tggctataaa ttaaataagc gtacatctcg cctcgaacca tcggatgaag	1800
acaaagttgt ggtaaagttg atatttgaca ttttcttaaa tggccttgat ggtaaagatt	1860
taagttatac agctattgca actcatcttt ctaaacttca attcactacc ccaaggggag	1920
gaaaaagatg gagtaaggac actgttagaa aaattcttca aaatgaagcc tatatgggaa	1980
gggttaggta taaagcaaga gaaaccacaa aagatggaaa aaaagtcttc cgacctgaaa	2040
gtgaacatat tgttgtagat gatgcccacg aaccaataat aaataaagag gactttgaag	2100
ctgttcaaga aaagattaaa aataaagtgc ctctgcttcc agtcgtaaca tcttatgaac	2160

cta	atgaatt	agc	gggaata	tgt	gtatgct	ctgtctgtgg	aaa	atcactt	caaaa	atttg	2220
aat	ccgaata	caat	aggaaa	aaca	aagatg	ggacttccag	ctatttt	tcac	gtga	agttat	2280
tga	tttgcaa	aat	caacaag	tgc	acatcag	tgagatatga	gtatgttgaa	gaag	caattt		2340
tgg	agtatct	tgag	cagttg	att	gctttag	aaaataataa	actaa	aggct	ataatt	gaaa	2400
aat	ctatgga	agc	agcagaa	acta	acaatt	cagaaaagac	aagt	gagcaa	atgct	tagtac	2460
aag	caaacca	aaa	acagaaa	gaac	tgc	gaaa	acaattgac	atttattttt	gaaaa	atttg	2520
aat	ctggaat	ttata	cagat	gaa	atgtttt	tgcaaaggaa	agc	agctatt	gaaaa	agaag	2580
tag	ctgacat	taaaa	attg	aaaca	agaac	ttagcatgac	atttga	agtt	aaaga	gaaag	2640
atg	ttaatga	atttag	gggtt	aacata	tcag	atgtagttaa	attttataaa	gaatc	taaaa		2700
gtag	aggatt	aaaaa	atgaa	aagt	tgagat	ccatttttga	ctttatagta	ttaga	aatga		2760
caga	gaaaag	acg	aggcccc	atccc	tgc	ta	gttcaatat	ttatcc	ggtt	ttgcgaatcc	2820
cgag	ttagt	taatt	agtct	agag	atacat	tgatgagttt	ggacaaacca	caact	tagaat		2880
gcag	tgaaaa	aaat	gcttta	ttt	gtgaaat	ttgtgatgct	attgctttat	ttgt	aacct		2940
tata	agctgc	aataa	acaag	tta	aca	acaa	caattgcatt	cattttatgt	ttc	aggttca	3000
ggggg	aggtg	tggg	aggttt	ttta	aagcaa	gtaaa	accctc	tacaa	atgtg	gtatggctga	3060
ttat	gatcgc	ggccg	cctag	atct	gtactg	aaatgtgtgg	gcgtggctta	agggt	gggaa		3120
aga	atatata	aggt	gggggt	cttat	gtagt	tttgtatctg	ttttgcagca	gccg	ccgccg		3180
ccat	gagcac	caact	cgttt	gat	ggaagca	ttgtgagctc	atatttgaca	acgc	gcatgc		3240
cccc	atgggc	cgggg	tgcgt	caga	atgtga	tgggctccag	cattgatggt	cgc	ccccgtcc		3300
tgc	ccgcaa	ctct	actacc	ttg	acctacg	agaccgtgtc	tggaacgccg	ttgg	agactg		3360
cag	cctccgc	cgc	cgcttca	gcc	gtgcag	ccaccgccccg	cgggattgtg	act	gactttg		3420
cttt	cctgag	ccc	gcttgca	agc	agtgcag	cttcccgttc	atccgccccg	gat	gacaagt		3480
tgac	ggctct	ttt	ggc	acaa	ttggattctt	tgacccggga	acttaatgtc	gttt	ctcagc		3540
aget	gttgga	tct	gcgccag	cagg	tttctg	ccctgaaggc	ttcctccccct	ccca	atgcgg		3600
ttta	aaacat	aaata	aaaaaa	ccag	actctg	tttggatttg	gatcaagcaa	gtgt	ctttgct		3660

gtctttatftt	aggggtftftg	cgcgcgcggt	aggcccggga	ccagcggctct	cggtcgttga	3720
gggtcctgtg	tatftttttcc	aggacgtggt	aaaggtgact	ctggatgttc	agatacatgg	3780
gcataagccc	gtctctgggg	tggaggtagc	accactgcag	agcttcatgc	tgcgggggtgg	3840
tgttgtagat	gatccagtcg	tagcaggagc	gctgggcgtg	gtgcctaaaa	atgtctttca	3900
gtagcaagct	gattgccagg	ggcaggccct	tggtgtaagt	gtttacaaag	cggttaagct	3960
gggatgggtg	catacgtggg	gatatgagat	gcatacttga	ctgtatfttt	aggttggcta	4020
tgttcccagc	catatccctc	cggggattca	tgttggtgcag	aaccaccagc	acagtgtatc	4080
cggtgcactt	gggaaatfttg	tcatgtagct	tagaaggaaa	tgcgtggaag	aacttggaga	4140
cgcccttggt	acctccaaga	ttttccatgc	attcgtccat	aatgatggca	atgggcccac	4200
gggcggcggc	ctgggcgaag	atattttctgg	gatacctaac	gtcatagttg	tgttccagga	4260
tgagatcgtc	ataggccatt	tttaciaaagc	gcgggcggag	ggtgccagac	tgcggtataa	4320
tggttccatc	cggcccaggg	gcgtagttaac	cctcacagat	ttgcattttcc	cacgctttga	4380
gttcagatgg	ggggatcatg	tctacctgcg	gggcgatgaa	gaaaacggtt	tccggggtag	4440
gggagatcag	ctgggaagaa	agcaggtttcc	tgagcagctg	cgacttaccg	cagccggtgg	4500
gcccgtaaat	cacacctatt	accggctgca	actggtagtt	aagagagctg	cagctgccgt	4560
catccctgag	cagggggggcc	acttcgttaa	gcatgtccct	gactcgcata	ttttccctga	4620
ccaaatccgc	cagaaggcgc	tgcgcccca	gcgatagcag	ttcttgcaag	gaagcaaagt	4680
ttttcaacgg	tttgagaccg	tccgccgtag	gcatgctttt	gagcgtttga	ccaagcagtt	4740
ccaggcggtc	ccacagctcg	gtcacctgct	ctacggcatac	tcatccagc	atatctctctc	4800
gtttcgcggg	ttggggcggc	tttcgctgta	cggcagtagt	cggtgctcgt	ccagacgggc	4860
cagggtcatg	tctttccacg	ggcgcaggggt	cctcgtcagc	gtagtctggg	tcacggtgaa	4920
ggggtgcgct	ccgggctgcg	cgttggccag	ggtgcgcttg	aggctgggtcc	tgctgggtgct	4980
gaagcgtgc	cggtcttcgc	cctgcgcgtc	ggccaggtag	catttgacca	tgggtgcata	5040
gtccagcccc	tccgcggcgt	ggcccttggc	gcgcagcttg	cccttgagg	aggcgcgcga	5100
cgaggggcag	tgcagacttt	tgagggcgta	gagcttgggc	gcgagaaata	ccgattccgg	5160

ggagtaggca tccgcgccgc agggcccgca gacggtctcg cattccacga gccaggtgag	5220
ctctggccgt tcgggggtcaa aaaccagggt tcccccatgc tttttgatgc gtttcttacc	5280
tctggtttcc atgagccggt gtccacgctc ggtgacgaaa aggcctgtccg tgtccccgta	5340
tacagacttg agaggcctgt cctcgagcgg tgttccgcgg tcctcctcgt atagaaactc	5400
ggaccactct gagacaaagg ctcgcgctcca ggccagcacg aaggaggcta agtgggaggg	5460
gtagcggtcg ttgtccacta ggggggtccac tcgctccagg gtgtgaagac acatgtcgcc	5520
ctcttcggca tcaaggaagg tgattggttt gtaggtgtag gccacgtgac cgggtgttcc	5580
tgaagggggg ctataaaagg ggggtgggggc gcgttcgtcc tcactctctt ccgcatecgt	5640
gtctgcgagg gccagctggt ggggtgagtc gacgcgaggc tggatggcct tccccattat	5700
gattcttctc gcttccggcg gcategggat gcccgcgttg caggccatgc tgtccaggca	5760
ggtagatgac gaccatcagg gacagcttca aggccagcaa aaggccagga accgtaaaaa	5820
ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc aaaaaaatcg	5880
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc	5940
tggaagctcc ctctgtcgct ctctgtttcc gacctgccg cttaccggat acctgtccgc	6000
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc	6060
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg	6120
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc	6180
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga	6240
gttcttgaag tgggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc	6300
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaac	6360
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaagg	6420
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc	6480
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa	6540
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttgggt ctgacagtta	6600
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt	6660

tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag	6720
tgctgcaatg ataccgcgag acccacgctc accggtcca gatttatcag caataaacca	6780
gccagccgga agggccgagc gcagaagtgg tectgcaact ttatccgctt ccatccagtc	6840
tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt	6900
tgttgccatt gctgcaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag	6960
ctccggttcc caacgatcaa ggcgagttac atgatcccc atgtttgtgca aaaaagcggc	7020
tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt taccactcat	7080
ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt	7140
gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc	7200
ttgcccggcg tcaacacggg ataataccgc gccacatagc agaactttaa aagtgcctcat	7260
cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag	7320
ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt	7380
ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg	7440
gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta	7500
ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc	7560
gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt	7620
aacctataaa aataggcgta tcacgaggcc ctttcgtctt caagaattgt t	7671

<210> 32
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 引子P349F3

<400> 32	
cgtaaacggt tgggtggcgca	20

<210> 33
<211> 24

<212> DNA
<213> 人工序列

<220>
<223> 引子P349R9

<400> 33
gtagaagggc acagtgtggg ctat 24

<210> 34
<211> 606
<212> DNA
<213> 人工序列

<220>
<223> 含有SEQ ID NO:14 β -肌動蛋白內含子之上游半部及
SEQ ID NO:15 β -肌動蛋白內含子之下游半部的序列

<220>
<221> misc_feature
<222> (144)..(178)
<223> SEQ ID NO:14 β -肌動蛋白內含子之上游半部

<220>
<221> misc_feature
<222> (237)..(312)
<223> SEQ ID NO:15 β -肌動蛋白內含子之下游半部

<400> 34
cgcagacgca ggagcagaac aaagagaatc agaateccaa ttctgatgcg ccggtgatca 60

gatcaaaaac ttcagccagg tacatggagc tggtcgggtg gctcgtggac aaggggatta 120

cctcggagaa gcagtggatc caggtgggtg tctttcctgc ctgagctgac ctgggcagac 180

gcgtaattaa gtatcatgat aaactgacaa agcggtttct ccaacgaatt atcgattcgg 240

ctgtggggtc ctgtggtgtg tggggagctg tcacatccag ggtcctcact gcctgtcccc 300

ttccctcctc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcgg 360

tcccaaatta aggetgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc 420

cccgactacc tgggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa 480

atttttggaac taaacgggta cgatccccaa tatgcggcct ccgtctttct gggatgggcc 540

acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcttgcaac taccgggaag 600

accaac	606
<210> 35 <211> 45 <212> DNA <213> 人工序列 <220> <223> AttL經重組序列 <400> 35 aattaagtat catgataaac tgacaaagcg gtttctccaa cgaat	45
<210> 36 <211> 21 <212> DNA <213> 人工序列 <220> <223> mCherry引子1 <400> 36 ctgttccacg atggtgtagt c	21
<210> 37 <211> 20 <212> DNA <213> 人工序列 <220> <223> mCherry引子2 <400> 37 tgaggtcaag accacctaca	20
<210> 38 <211> 24 <212> DNA <213> 人工序列 <220> <223> mCherry探針，在5'端之FAM，在鹼基9與鹼基10之間的ZEN， 在3'端之3IABkFQ <400> 38	

ttggacatca cctcccacaa cgag		24
<210> 39		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 腺病毒外顯子2 (Ad5E2)引子1		
<400> 39		
gggtgatgca gtagaaggta ag		22
<210> 40		
<211> 19		
<212> DNA		
<213> 人工序列		
<220>		
<223> 腺病毒外顯子2 (Ad5E2)引子2		
<400> 40		
atgaagttcg gcggagatg		19
<210> 41		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 腺病毒外顯子2 (Ad5E2)探針，在5'端之HEX，在鹼基9與鹼基10之間的Zen， 在3'端之3IABkFQ		
<400> 41		
tcttgttccc agcggtccea tc		22
<210> 42		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> P5（AAV之P5啓動子區域）引子1		
<400> 42		
gtggtcacgc tgggtattta		20

<210>	43		
<211>	22		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	P5（AAV之P5啓動子區域）引 子2		
<400>	43		
gggaccttaa tcacaatctc gt		22	
<210>	44		
<211>	23		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	P5（AAV之P5啓動子區域）探針，在5'端之FAM， 在鹼基9與鹼基10之間的ZEN， 在3'端之31ABkFQ		
<400>	44		
tttgaagcgg gaggtttgaa cgc		23	
<210>	45		
<211>	22		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV REP基因引 子1		
<400>	45		
gtccgtgagt gaagcagata tt		22	
<210>	46		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV REP基因引 子2		
<400>	46		
ttcgatcaac tacgcagaca g		21	

<210>	47		
<211>	24		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV REP基因探針，在5'端之FAM，在鹼基9與鹼基10之間的ZEN，在3'端之3IABkFQ		
<400>	47		
tctgatgctg tttccctgca gaca			24
<210>	48		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV9 CAP基因引子1		
<400>	48		
ccgggtccaa ggtatttgta a			21
<210>	49		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV9 CAP基因引子2		
<400>	49		
ctcaacccaa ggcaaatcaa c			21
<210>	50		
<211>	24		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV9 CAP基因探針，在5'端之FAM，在鹼基9與鹼基10之間的ZEN，在3'端之3IABkFQ		
<400>	50		
acatcaagac aacgctcgag gtct			24

<210>	51	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	β 內醯胺酶（安比西林抗性）基因引子1	
<400>	51	
ccagaaacgc tggtgaaagt a		21
<210>	52	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	β 內醯胺酶（安比西林抗性）基因引子2	
<400>	52	
ctcaaggatc ttaccgctgt tg		22
<210>	53	
<211>	24	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	β 內醯胺酶（安比西林抗性）基因探針，在5'端之FAM， 在鹼基9與鹼基10之間的ZEN， 在3'端之3IABkFQ	
<400>	53	
tgcacgagtg gggttacatcg aact		24
<210>	54	
<211>	14	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	沙福芽孢桿菌菌株CCMA-560	
<400>	54	
actgacaaag cggt		14

<210> 55
<211> 863
<212> DNA
<213> 人工序列

<220>
<223> 在所例示之建構體中止擋匣之AAV2 rep
基因上游之5'部分

<400> 55
ggtcctgtat tagaggtcac gtgagtgttt tgcgacattt tgcgacacca tgttggtcacg 60

ctgggtathtt aagcccgagt gaggacgcag ggtctccatt ttgaagcggg aggtttgaac 120

gcgcagccgc catgccgggg ttttacgaga ttgtgattaa ggtccccagc gaccttgacg 180

agcatctgcc cggcattttct gacagctttg tgaactgggt ggccgagaag gaatgggagt 240

tgccgccaga ttctgacatg gatctgaatc tgattgagca ggcacccctg accgtggccg 300

agaagctgca gcgcgacttt ctgacggaat ggcgccgtgt gagtaaggcc ccggaggccc 360

ttttctttgt gcaatttgag aagggagaga gctacttcca catgcacgtg ctctgtgaaa 420

ccaccgggggt gaaatccatg gttttgggac gtttcctgag tcagattcgc gaaaaactga 480

ttcagagaat ttaccgcggg atcgagccga ctttgccaaa ctggttcgcg gtcacaaaga 540

ccagaaatgg cgcgggaggc gggaacaagg tggatgatga gtgctacatc cccaattact 600

tgctccccaa aaccagcct gagctccagt gggcgtggac taatatggaa cagtatttaa 660

gcgcctgttt gaatctcacg gacgtaaac ggttggtggc gcagcatctg acgcacgtgt 720

cgcagacgca ggagcagaac aaagagaatc agaatcccaa ttctgatgcg ccggtgatca 780

gatcaaaaac ttcagccagg tacatggagc tggtcgggtg gctcgtggac aaggggatta 840

cctcggagaa gcagtggatc cag 863

<210> 56
<211> 1150
<212> DNA
<213> 人工序列

<220>
<223> 在所例示之建構體中止擋匣之AAV2 rep
基因下游之3'部分

<400>	56	
gaggaccagg cctcatacat ctctttcaat gcggcctcca actcgcggtc ccaaattaag	60	
gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg	120	
gtgggccagc agcccgtgga ggacatttcc agcaatcgga ttataaaaat tttggaacta	180	
aacgggtacg atccccaata tgcggcttcc gtctttcttg gatgggccac gaaaaagttc	240	
ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg	300	
gaggccatag cccacactgt gccctttctac ggggtgcgtaa actggacca a tgagaacttt	360	
cccttcaacg actgtgtcga caagatggtg atctgggtgg aggaggggaa gatgaccgcc	420	
aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa	480	
tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctcaa caccaacatg	540	
tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg	600	
atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag	660	
gaagtcaaag acttttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgagttc	720	
tacgtcaaaa aggggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag	780	
cccaaacggg tgcgcgagtc agttgcgcag ccategacgt cagacgcgga agcttcgata	840	
aactacgcag acaggtacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg	900	
tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga	960	
cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa	1020	
aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc	1080	
actgcctgcg atctgggtcaa tgtggatttg gatgactgca tctttgaaca ataaatgatt	1140	
taaatcaggt	1150	

- <210> 57
- <211> 2399
- <212> DNA
- <213> 人工序列

- <220>
- <223> 在所例示之建構體中包括的AAV9 cap基因，其包括具有GenBank NC_001401.2序列的核苷

酸編號4411至4466之多腺苷酸化信號，以及
具有下列序列的核苷酸190之AAV2 rep P5啟動子

<400> 57	
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga agggattcgc	60
gagtgggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac	120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac	180
aagggggagc cggtcacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac	240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc	300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcctccag	360
gccaaaaaga ggcttcttga acctcttggc ctggttgagg aagcggctaa gacggctcct	420
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggatttggc	480
aaatcgggtg cacagccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag	540
tcagtcccag acctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct	600
cttacaatgg cttcaggtgg tggcgacca gtggcagaca ataacgaagg tgccgatgga	660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac	720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaac	780
tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc	840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga	900
ctcatcaaca acaactgggg attccggcct aagcgactca attcaagct cttcaacatt	960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc	1020
acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac	1080
gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcttcagta cgggtatctg	1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc	1200
ccgtcgcaaa tgctaagaac gggtacaac ttccagttca gctacgagtt tgagaacgta	1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc	1320
gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg	1380

ctaaaattca	gtgtggccgg	accagcaac	atggctgtcc	agggaagaaa	ctacatacct	1440
ggaccagct	accgacaaca	acgtgtctca	accactgtga	ctcaaaacaa	caacagcgaa	1500
tttgcttggc	ctggagcttc	ttcttgggct	ctcaatggac	gtaatagctt	gatgaatcct	1560
ggacctgcta	tggccagcca	caaagaagga	gaggaccgtt	tctttccttt	gtctggatct	1620
ttaatttttg	gcaaacaagg	aactggaaga	gacaacgtgg	atgcggacaa	agtcatgata	1680
accaacgaag	aagaaattaa	aactactaac	ccggtagcaa	cggagtccta	tggacaagtg	1740
gccacaaacc	accagagtgc	ccaagcacag	gcgcagaccg	gctgggttca	aaaccaagga	1800
atacttccgg	gtatggtttg	gcaggacaga	gatgtgtacc	tgcaaggacc	catttggggc	1860
aaaattcctc	acacggacgg	caactttcac	ccttctccgc	tgatgggagg	gtttggaatg	1920
aagcaccgcg	ctcctcagat	cctcatcaaa	aacacacctg	tacctgcgga	tcctccaacg	1980
gccttcaaca	aggacaagct	gaactctttc	atcaccagct	attctactgg	ccaagtcagc	2040
gtggagatcg	agtgggagct	gcagaaggaa	aacagcaagc	gctggaaccc	ggagatccag	2100
tacacttcca	actattacaa	gtctaataat	gttgaatttg	ctgttaatac	tgaaggtgta	2160
tatagtgaac	cccgccccat	tggcaccaga	tacctgactc	gtaatctgta	attgcttggt	2220
aatcaataaa	ccgtttaatt	cgtttcagtt	gaacttttgt	ctctgcggtt	taaacggtec	2280
tgtattagag	gtcacgtgag	tgttttgcga	cattttgcga	caccatgtgg	tcacgctggg	2340
tatttaagcc	cgagtgagca	cgcagggtct	ccattttgaa	gcgggaggtt	tgaacgcgc	2399

<210> 58
<211> 600
<212> DNA
<213> 人工序列

<220>
<223> 來自圖2之CCMA-560 DNA序列(query)

<400> 58	
tgactttata	gtattagaaa
tgacagagaa	aagacgaggc
cccatccctg	ctaagttcaa
	60
tatttatccg	gttttgcgaa
ttccgagtgt	agttaattaa
gtatcatgat	aaactgacaa
	120
agcggtttct	ccaacgaatg
ccatgggtgc	aacaaaattg
atttcagaaa	aactatTTTT
	180

ccaagcaaac gaaagtattc cgaatcaaaa aacgaggttt tgctctgtac gctttggcaa	240
tgtgcttgga tctagaggtt ccgttattcc gatcatgctc cagcagctat taaatgaaaa	300
acctttgacc gtgactgata ctcatatgac acgttttttt atgtccattg aagaggctgt	360
ttccctcaca ctccaagcag caatcatgat gaaaggcggc gaaaccttca ttctcaagat	420
ggagtcctta cagcttgccg atctcctaaa agcgtttcat gaatatgccg ctcaaatcaa	480
tgctaaatct ccggatattc ttgtagtcgg aaaaagacct ggcgaaaagc ttcacgagga	540
gtcacatttt ccgcacgaag cagatgcact gtttgaacat gaacaatttt atgccatttt	600

<210> 59
<211> 488
<212> DNA
<213> 人工序列

<220>
<223> 來自圖2之沙福芽孢桿菌菌株Fairview contig56_1 (Sbjct)的全基因體散彈槍序列之核苷酸464352至464839

<400> 59 actgacaaag cggtttctcc aacgaatgcc atgggtgcaa caaaattgat ttcagaaaaa	60
ctatttttcc aagcaaacga aagtattcca aataaaaaaa ccaagttttg ctctgtacgc	120
tttggcaatg tgcttggatc tagaggttcc gtgattccga tcatgctcca gcagctatta	180
aatgaaaaac ctttgaccgt gactgacct catatgacac gtttttttat gtccattgaa	240
gaggctgttt ccctcacact tcaagcagca atcatgatga aaggcggcga aaccttcatt	300
ctcaagatgg agtccttaca gcttgccgat ctccctaaaag cgtttcatga atatgccgct	360
caaatcaatg ctaaattctcc ggatattctt gtagtcggaa aaagacctgg cgaaaagctt	420
cacgaggagc tcacatttcc gcacgaagca gatgcactgt ttgaacatga acaattttat	480
gccatttt	488

<210> 60
<211> 225
<212> DNA
<213> 人工序列

<220>
<223> SV40多腺苷酸化信號

<400> 60
gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt 60

gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 120

acaacaacaa ttgcattcat tttatgtttc gggttcaggg ggaggtgtgg gaggtttttt 180

aaagcaagta aaacctctac aaatgtggta tggctgatta tgate 225

<210> 61
<211> 4281
<212> DNA
<213> 人工序列

<220>
<223> 經修飾rep基因，其包括5'rep、內含子5'、止擋匣、內含子3'、及3' rep

<400> 61
ggtcctgtat tagaggtcac gtgagtgttt tgcgacattt tgcgacacca tgttggtcacg 60

ctgggtattt aagcccgagt gagcacgcag ggtctccatt ttgaagcggg aggtttgaac 120

gcgcagccgc catgccgggg ttttacgaga ttgtgattaa ggtccccagc gaccttgacg 180

agcatctgcc cggcattttct gacagctttg tgaactgggt ggccgagaag gaatgggagt 240

tgccgccaga ttctgacatg gatctgaatc tgattgagca ggcaccctg accgtggccg 300

agaagctgca gcgcgacttt ctgacggaat ggcgccgtgt gagtaaggcc ccggaggccc 360

ttttctttgt gcaatttgag aaggagaga gctacttcca catgcacgtg ctcgtggaaa 420

ccaccggggt gaaatccatg gttttgggac gtttctgag tcagattcgc gaaaaactga 480

ttcagagaat ttaccgcggg atcgagccga ctttgccaaa ctggttcgcg gtcacaaaga 540

ccagaaatgg cgccggaggc gggaacaagg tggtggatga gtgctacatc cccaattact 600

tgctcccaa aaccagcct gagctccagt gggcgtggac taatatggaa cagtatttaa 660

gcgcctgttt gaatctcacg gacgtaaac ggttggtggc gcagcatctg acgcacgtgt 720

cgcagacgca ggagcagaac aaagagaatc agaatcccaa ttctgatgcg ccggtgatca 780

gatcaaaaac ttcagccagg tacatggagc tggtcgggtg gctcgtggac aaggggatta 840

cctcggagaa gcagtggatc caggtgggtg tctttcctgc ctgagctgac ctgggcagac	900
gcgtaattaa gtatcatgat aaactgacaa agcgggtatat caagatactt aatacatatg	960
gtcctgtaaa agtttaatgt gtaagaagta tttgtttataa aagataaata ttcagaatct	1020
tctttttaat tcctgatttt atttctatag gactgaaaga cttgctcgag atgtcatgaa	1080
ggagatggga ggccatcaca ttctcacttc actgcagctc aacgtcccca attctaccgg	1140
gtaggggagg cgcttttccc aaggcagtct ggagcatgcg ctttagcagc cccgctgggc	1200
acttggcgct acacaagtgg cctctggcct cgcacacatt ccacatccac cggtaggcgc	1260
caaccggctc cgttctttgg tggccccctc gcgccacctt ctactcctcc cctagtcagg	1320
aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt	1380
ctcactagtc tcgtgcagat ggacagcacc gctgagcaat ggaagcgggt aggcctttgg	1440
ggcagcggcc aatagcagct ttgctccttc gctttctggg ctcagaggct gggaaggggt	1500
gggtccgggg gcgggctcag gggcgggctc aggggcgggg cgggcgcccc aaggtcctcc	1560
ggaggeccgg cattctgcac gcttcaaaag cgcacgtctg ccgcgctgtt ctccctctcc	1620
tcatctccgg gcctttcgac ctagcgggca gtgagcgcaa cgcaattaat gtgagttagc	1680
tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa	1740
ttgtgagcgg ataacaattt cacacaggaa acagctgcca ccatgattga acaagatgga	1800
ttgcacgcag gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa	1860
cagacaatcg gctgctctga tgccgcctg ttccggctgt cagcgcaggg gcgccccggtt	1920
ctttttgtca agaccgacct gtccggtgcc ctgaatgaac tgcaagacga ggcagcgcgg	1980
ctatcgtggc tggccacgac gggcgttctt tgcgcagctg tgctcgacgt tgtcactgaa	2040
gcgggaaggg actggctgct attgggcgaa gtgccggggc aggatctctt gtcatctcac	2100
cttgctcctg ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt	2160
gateccggcta cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact	2220
cggatggaag ccggtctttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg	2280
ccagccgaac tgttcgccag gctcaaggcg agcatgcccc acggcgagga tctcgtcgtg	2340

acccatggcg atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc	2400
atcgactgtg gccggctggg tgtggcggac cgctatcagg acatagcggt ggctaccgt	2460
gatattgctg aagagcttgg cggcgaatgg gctgaccgt tctctgtgct ttacggtatc	2520
gccgtcccc attcgcagcg catgccttc tatgccttc ttgacgagtt cttctgagat	2580
acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg	2640
aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa caagttaaca	2700
acaacaattg cattcatttt atgtttcggg ttcaggggga ggtgtgggag gttttttaaa	2760
gcaagtaaaa cctctacaaa tgtggtatgg ctgattatga tcgcatagtg ttaccatcaa	2820
ccaccttaac ttcatTTTTT ttattcaata cctaggtagg tagatgctag attctggaaa	2880
taaaatatga gtctcaagtg gtctttgtcc tctctccag tcaaattctg aatctagttg	2940
gcaagattct gaaatcaagg catataatca gtaataagt atgatagaag ggtatataga	3000
aggctagccg tcaacatctc cactgacaaa gcggtttctc caacgaatta tcgattcggc	3060
tgtggggtec tgtggtgtgt ggggagctgt cacatccagg gtctcactg cctgtccct	3120
tccctcctca ggaggaccag gcctcataca tctccttcaa tgcggcctcc aactcgcgt	3180
cccaaattaa ggctgccttg gacaatgcgg gaaagattat gagcctgact aaaaccgccc	3240
cggactacct ggtgggccag cagcccgtag aggacatttc cagcaatcgg atttataaaa	3300
ttttggaact aaacgggtac gatccccaat atgcggcttc cgtctttctg ggatgggcca	3360
cgaaaaagtt cggcaagagg aacaccatct ggctgtttgg gcctgcaact accgggaaga	3420
ccaacatcgc ggaggccata gccacactg tgcccttcta cgggtgcgta aactggacca	3480
atgagaactt tcccttcaac gactgtgtcg acaagatggt gatctggtgg gaggagggga	3540
agatgaccgc caaggtcgtg gagtcggcca aagccattct cggaggaagc aaggtgcgcg	3600
tggaccagaa atgcaagtec tcggcccaga tagaccgcac tcccgtagc gtcacctcca	3660
acaccaacat gtgcgccgtg attgacggga actcaacgac cttegaacac cagcagccgt	3720
tgcaagaccg gatgttcaaa tttgaactca cccgccgtct ggatcatgac tttgggaagg	3780
tcaccaagca ggaagtcaaa gactttttcc ggtgggcaaa ggatcacgtg gttgaggtgg	3840

agcatgagtt ctacgtcaaa aagggtggag ccaagaaaag acccgcccc agtgacgcag	3900
atataagtga gccc aaacgg gtgcgcgagt cagttgcgca gccatcgacg tcagacgcgg	3960
aagcttcgat caactacgca gacaggtacc aaaacaaatg ttctcgtcac gtgggcatga	4020
atctgatgct gtttccctgc agacaatgcg agagaatgaa tcagaattca aatatctgct	4080
tcactcacgg acagaaagac tgtttagagt gctttcccg tgcagaatct caacccgttt	4140
ctgtcgtcaa aaaggcgtat cagaaactgt gctacattca tcatatcatg ggaaagggtgc	4200
cagacgcttg cactgcctgc gatctggtea atgtggattt ggatgactgc atctttgaac	4260
aataaatgat ttaaatcagg t	4281

<210> 62
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> AAVRT-F1

<400> 62
tgatgcgccg gtgatcagat 20

<210> 63
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> P349R9

<400> 63
gtagaagggc acagtgtggg ctat 24

<210> 64
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 在5'REP中之剪接供體位點

<400> 64

cttcagccag gtacatggag	20
<210> 65	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 在5'REP中之經突變剪接供體位點	
<400> 65	
cttcagccag atacatggag	20
<210> 66	
<211> 2000	
<212> DNA	
<213> 人工序列	
<220>	
<223> 內含子間隔子	
<400> 66	
ttacgtataa tacctccacg ttgagaccgt acgagcgcac aacggaagcg ttgttgagac	60
ccagtcgcct gtgataccgc cttggttaag gcccgtagcc tagcttaagg caacgaaaca	120
agacgactca ggagaatcat tatgaacagg gcgcgcggga ggtcgtggat atggccggta	180
ttttacttat gggaggatct tttggccgcg agacttccga gaaaccatca acagttcgat	240
ttcctccttt tatccttaac atagcacaga gttgccacat tgataggggg gcaggagatc	300
gtcacatgaa ccgacggtcg ggacatattc gatgccgtcc cctctcgaca accggattcc	360
tcgtttccaa tcgaatacca cgccccgggg atcgtgcgca cgataagcac aataagcgtc	420
actgcgggggt caagctgtgg cttgcaggat gctaactcgt aacgacatta agacagcaga	480
gcaatgggcg acccaaaaag tttaagcctt tacctcgggg tggcaacata agtgtgccga	540
ctacgacagg aacctatgac tcctttgggg cgttctaata ggccgggtct gctgttaacc	600
agctgatggg gcaagtggaa tagagcatgc cctacgcgtc attaaggaat gcgtggctac	660
ttggctgtgc gtacttgtct aatgtatggc aacaaacaac ctaatctttg tcggcaacgc	720
gataatctcg ccacgttaat cggctacttg catggtagga ttcgatttta cgtatgtacc	780

acggattcta ttgcacgcta cagcatcaat gtgcccgatg actactagct gcccagaggg	840
gataaatcat gtgtaatttg gcgtgcattc gagttattgg cattaatttc tctatcaagt	900
gcagtccta gcttcgagta agctatgcgt tccccccgt acattttaat cccataggga	960
acggcacccg cacctattaa gagagcgcca tagctctaaa cgaccacgg tccaatgctt	1020
ataatttcta aaatttaagg tcgcattgcc ttgaccgtta gtccccctca cgtttgaggg	1080
cataatgttc ctgcgccctt tacaattagt ctaattctac ttaaaatcga cgcaggattt	1140
tttctgttgg ctccgctgcc atgggaagct gtctatggag ggctcggta tccctggctt	1200
cttactatgt aatctacett tacctttgct aattacgtgt acgtgaccag aaactacacg	1260
gaacgagtgg gacccataag ctgagcgatg gctagtgaga ccgcctgata tgactcacgg	1320
ggtaaccgtg caacgtgcat atatctaatz agtacggcac ctttccactt tgcttatatg	1380
agagtacagg ctacttagcc cgacgtgtac gccagatctg ggcacccgcc agcaggctcc	1440
actaggccgg cctgccaggc aaatgacctg agtgccgcga ctagcctcgg ccagacagct	1500
gactagcgat tcacaagtgc cgacttagta gttcttttagc aatagaaata tagacatagt	1560
tccctcgaac tacagggaaa tacggcctgt ggttgcaaga taagaggctt ttaacactga	1620
cgtagtttac acactctggc cggaccatga atttcgtctg ctcccacgaa ccacattatc	1680
agtacttcct cttaagtgtt ttattaaatc gaataatcta caaatagttc taaaggagcg	1740
agttagagag aagttccacg ctccataact tcgagaccgt tagcgaagtt ttccagcaca	1800
cggctacggg accctgcact cagcttcatg gttttaggca cgagaaccgc atatagtgga	1860
gatttcgccg tggaagcaag tctgtcttag agagcctagg tgcttaggtt ttcggaactc	1920
tectgattat attcaatttc cateccattc tggatcatca agtgcccgcg gatgacctac	1980
gggcaaattg ccccagataa	2000

<210> 67
<211> 2072
<212> DNA
<213> 人工序列

<220>
<223> 間隔子2

<400> 67	
gtagtcccgc agtacgcggc taggcagatc ggcccccttc aaatccgtgt catcggaaga	60
ccatgttggg agtttcataa tggattttat tgattgctat gtaacctgaa cgacggtatt	120
aatgggcttc ttaacatgaa cgtgagtaac ggaaacatta cgcttttcgg tcgcgaatga	180
cggcgaactc tgggccttgt gggattatag atacgtcagt gtccctcggg aacggacaaa	240
gtaggaaata tcgacgacat ggatcatgcc aatatatcaa cgcgggaaga tgaccattga	300
tgcagtggga actggtgtcc cacaaatata tagggcttgt gcagcggtea agggctgtgt	360
tccacctggg acgccagccc aattatacgc cagaaacgct tgattttact caggtagaaa	420
gggggaaaat acactacgcg accaactaat aattccttcc gcaaacttaa tggctgctaa	480
ttttcctcaa gaccggctta atgcctacta cacttaaggc gcggatggcg agatacgacg	540
cgggcacata ctaggcacgt catcggccag gcgttctaata ttaagttcac gcataaataa	600
gcgtctctcg agactgcaga tccccgtctt aaatttagta cacaacgctg tttaccaccg	660
gggcgttatg acttgcagac cegtgcgtga atatcctgta tcattcatca ctgctgggac	720
taccccactt tacatatcta agggattcaa caaacttcgg gtggttaatg tcaagtgcgt	780
tattggcgta taacagggtg aatgcacatt gatggaatgc ctcaaccaa gctgtggcgt	840
cacgccagcc gaacacgttt tacccaaact tagccctttt gctgcgaccc tgctatctat	900
tggctatact ctatatgcaa tgggtgcaga tcggagaccg cgccattgtc accaccccaa	960
ggagggggaga ttcacgctcg ctgggattag ctccccgat catttateaa aactggaaat	1020
actetaagct aagttcacct tcagatcctt gtcaagagtt gtcggcccga ttagtcgttg	1080
cgcacacgtg gcgtaattac agcacgcgag gtacattcga gtttatcctc tgggctccat	1140
tcaggcaata tgctgtcagc acgaagagtg tgacctggcg ctattgttgg cgacttgacg	1200
ccacggaata ccagtgtata tacatcctag gagaactgca agacttcggg cgtgtatcgt	1260
aggttatggg agtcactggg atttgcacct gcgcacaatc tgtttggtaa agtaaaagcg	1320
atacttcgcg actcccaata agtaacgata ctgcgcggca tcccttagcg ctctcatacc	1380
gagcgagata cagtctatat actcaacaga gcacgtttgg aacctgtgag gttcagcacc	1440

aatcattcat ctataaactc atcgagaaaa gtatggaggc cacgtgggct tcttcggttc	1500
agctaagcag actcatgagg gtgtggtctc ggctttggag gcagatattg tccggaattg	1560
taattgcgac tgcgcggcag tatagagtgt ctctatcatt gtaggactca cgcacccatc	1620
ggacaaccac gtgatggccc cttagttata tcgagggata ggggctatgc tcaagttagt	1680
cttaatttcg cgtgcttgca aggccttatcg ccggtccctc ctaccgtacc cggcctttca	1740
gggccatgat ctttgcctgc ttccctgttt atctacacca atattgggag acttgatatga	1800
cgagcgagcc actaaagagg acgattgtca acagggtgct ttatcacata atagcctgtt	1860
tctccaacag tgcttcaatt acctagcatc acagaatgcg ataagcccac cacgtgtttc	1920
ccaatggcgg ctattcatta atcgggacga caccctatt cgtggagaga gcctcctatt	1980
gagtccgagc tgttagcgca tgattacggt gggccagaca gattcgacgg aatgccgcta	2040
ctacgctggc gcgtaatcag tcggaaccag aa	2072

<210> 68
<211> 2091
<212> DNA
<213> 人工序列

<220>
<223> 間隔子3

<400> 68 tgtacacaat gatacgatcc cgtttctgcg acgtactcgg gtgtagctgg gaccctgtcg	60
catttaatcg ccggtcaata ggtacaccgt gctgatcttg gatactcctg atcgcgatgt	120
aacccgagcg gcgctactaa tgggcgtcct aacacgagcg cgaataacca cttgcgctag	180
gcaggattta cgatggtcat ctggaacgga cgggtacacc atgccaaaat ttgctcctag	240
ttcacaaggt caaacggtga cccggaggcc agtccactcc ttgtcctcgt ccttacgtta	300
ccgtaacctc tgacctggag cctaaggctg ctccggcatt ttaaaaacca cacaataag	360
gcttcacact attcctatta tcttgtgtcc tacctaagag gtagtggagc gagtaggtag	420
acacaggcac tcaccacega tgttccgccc atgggccatg gttaaggtcc gaatgaaacc	480
cgtatgcaag ttaccggcct tatattaata aggttggtat acatcctaaa agctctctcg	540

gacgtgacca aatacaggca tccacggagc actactgaat gggcacgtgg accggcgagt	600
gagaaatcga gcagcattaa gtcattctta gtaagccgag atgggtctgc gttatgtacg	660
tctagtgggg aaacaagttg gtggacaaac gttacgccta aggagcta atttcggtea	720
ggttgcgctg ggtcaaatea accggtcctg gtaacggagc gatggttagg tgtattacac	780
tagcgatggg ccgtgccgga tggtaagcca cgttctccga gccagtgaat ttggatgaga	840
actcgtaaat ctctcgttgc aaacagcttc ctcaatatat gacctggcca ctgatgagag	900
ggtaacactt aattcctaga atgtctgacg acatcacctt taagcagaca cactgcgatg	960
tgtggacgct aggggtgtct tttccacgtt tgacttagcc catgatgtct tgactaagat	1020
ggcagggtat ttccccttat tcgtgtatgt accactatta ccgtcgattg accgtcgccg	1080
cttatacgtc gtgccctgtg ccttggcaat tttgtccctg gataaccgcg cttgcttcta	1140
tcgaggtcct tgaataagtt aatgcaactt gtggaggaac atagggaatg tggcgtgctg	1200
cttagatatg gaggcaaga agtcctctgg gttgtagttt cactgttacg tcttagccag	1260
tgaagcactg gcatgcagct ttatctgagg attccactca ctgcttaca gaatttgcca	1320
agctgggtgcc gatatgacct caacgctatc cagagtttaa gttgtcatct ctctccatca	1380
taattgccac agtgtgttgt gtctacgaat agctattaaa ataccagccc ctgagaacga	1440
tgcccgacac atagggcagt gcaggcagca taacgtgcag atcaacataa tttccactaa	1500
gtgttgatat agtttactgg ctgagagccg ccatattgga acgacgacga gagcgcgtat	1560
cctgcccac cagaaatacg tgtccaatcc atataactga ccaaagctgc ggacagtttc	1620
ctctccacct acaacgggat tagcgcactt gcttgtagtg aggcaggat aaccacggc	1680
acgtaacacc tttagacgtt cggttccgca atcgcaaat gtactactat tagagaacta	1740
tagcaaggaa cttatagcgt atcgaaagtt catecctgcg gccatagcc gtaaccaccg	1800
tggggttatt ttaaacttta agttgtacct tcaggcaagc gaagaccggt actttaccta	1860
catggggccg ctttccgagc cagccgcac cccaacatat tctacgcgc tgtaataatg	1920
aacataagag aaaatggagt tgggtgacag cacatatcga ctacctagc actaccgcaa	1980
cacaaagtag atcccgccgg tatectctt tcatgcagga cgggttacct aggaatgcgt	2040

ttgtttgact ggtgcccgt gaggcgtga atatctcgcg atagcgtgta c	2091
<210> 69	
<211> 2794	
<212> DNA	
<213> 人工序列	
<220>	
<223> mCherry-IRES-SEAP編碼序列	
<400> 69	
atggtgagca agggcgagga ggataacatg gccatcatca aggagtteat gcgcttcaag	60
gtgcacatgg agggctccgt gaacggccac gagttcgaga tcgagggcga gggcgagggc	120
cgcccctacg agggcaccca gaccgccaag ctgaaggtga ccaaggggtg ccccctgccc	180
ttcgcctggg acatcctgtc ccttcagttc atgtacggt ccaaggccta cgtgaagcac	240
cccgccgaca tccccgacta cttgaagctg tccttccccg agggcttcaa gtgggagcgc	300
gtgatgaact tcgaggacgg cggcgtggtg accgtgaccc aggactcctc cctgcaggac	360
ggcgagtcca tctacaaggt gaagctgcgc ggcaccaact tcccctccga cggccccgta	420
atgcagaaga agaccatggg ctgggaggcc tcctccgagc ggatgtaccc cgaggacggc	480
gccctgaagg gcgagatcaa gcagaggctg aagctgaagg acggcggcca ctacgacgt	540
gaggtcaaga ccacctacaa ggccaagaag cccgtgcagc tgcccggcgc ctacaacgtc	600
aacatcaagt tggacatcac ctcccacaac gaggactaca ccatcgtgga acagtacgaa	660
cgcgccgagg gccgccactc caccggcggc atggacgagc tgtacaagta agtcgactac	720
tggccgaagc cgcttggaat aaggccggtg tgcgtttgtc tatatgtgat tttccaccat	780
attgccgtct tttggcaatg tgagggcccg gaaacctggc cctgtcttct tgacgagcat	840
tcctaggggt ctttcccctc tcgccaaagg aatgcaaggt ctgttgaatg tcgtgaagga	900
agcagttcct ctggaagctt cttgaagaca aacaacgtct gtagcgaccc tttgcaggca	960
gcggaacccc ccacctggcg acaggtgcct ctgcggccaa aagccacgtg tataagatac	1020
acctgcaaag gcggcacaaac cccagtgcc a gtttgtgagt tggatagttg tggaaagagt	1080
caaatggctc tcctcaagcg tattcaacaa ggggctgaag gatgcccaga aggtacccca	1140

ttgtatggga	tctgatctgg	ggcctcggtg	cacatgcttt	acatgtgttt	agtcgaggtt	1200
aaaaaacgtc	taggcccccc	gaaccacggg	gacgtggttt	tcctttgaaa	aacacgatcg	1260
tacggccacc	atgtggggag	cctgcttgct	gctgctgggc	ctcagcctac	aggtgtgccc	1320
gagcgtcatc	ccagtggagg	aggagaatcc	agcctttctg	aaccgaaagg	cagccgaggc	1380
cctggatgct	gccaagaagc	tgaagcccat	tcagacatca	gctaagaacc	tcgttatcct	1440
catgggagat	gggatggggg	tgtctacagt	aaccgctacc	cggatcttaa	aggggcagca	1500
gcaaggccat	ctgggacctg	agacacagtt	agctatggat	cgctttccac	acatggctct	1560
gtccaagaca	tacaacacag	acaagcagat	cccggacagc	gcaggcacag	gtacagcctt	1620
tctctgcggg	gtcaaaacca	acatgaaggt	cattggcttg	agtgcagctg	cacgcttcaa	1680
ccagtgcaac	acgacatggg	gtaacgaggt	cgtctcggtg	atgcaccgtg	ctaagaaagc	1740
aggaaagtct	gtgggagtgg	tgaccaccac	gtcggtgcac	catgcttctc	cggccggcac	1800
ctacgcgcac	actgtgaacc	gtggttggtg	ctcggatgca	caaatgcctg	cctcagcgct	1860
acaggatggc	tgcaaagaca	tctctactca	gtcctctccc	aacatggaca	ttgatgtgat	1920
cctcggtggt	ggccgcaagt	tcatgtttcc	caaggggaca	ccagaccagg	aatatccaac	1980
tgacaccaag	caggctggaa	ccaggctgga	tggacgcaac	ctagttcaag	agtggctggc	2040
aaagcaccag	ggagcccggg	atgtttggaa	ccgctcagag	ctgattcagg	catccctgaa	2100
ccgatctgtg	acacacctca	tgggcctctt	tgagcctaac	gacatgaaat	atgaaatcca	2160
ccgagaccct	gctcaggacc	cttccttggc	cgagatgacg	gaggtggctg	tgcgcatgct	2220
cagcaggaac	cccaaaggct	tctatctctt	tgtggaaggg	ggtcgcatcg	accatggcca	2280
ccatgaaact	gtagcctatc	gtgccctgac	tgaggctgtc	atgttcgact	cggctgttga	2340
caaggcagac	aagctcacca	gtgagcagga	cacgatgac	cttgtcactg	ccgaccactc	2400
tcacgtcttc	tcttttggtg	gttacacaca	gagaggggct	tccatctttg	gactggctcc	2460
cttcaaggct	gaggatggca	aatcctttac	ctcgatacta	tacggcaacg	gtcctggtta	2520
caagctccat	aatggcgccc	gggctgatgt	cactgaggaa	gagagcagca	accccacgta	2580
ccagcagcag	gctgctgtac	ccctgtcgtc	agagaccac	agcggggagg	acgtggcaat	2640

attcgcgcgt ggcccacagg cgcagctggt gcacggagtt caggagcaga actacatcgc	2700
gcacgtcatg gccttcgcag cctgcctgga gccctaacct gactgcggcc tggcateccc	2760
tgctggccag agcagtgcag tgagcccggg ctga	2794

<210> 70
<211> 22126
<212> DNA
<213> 人工序列

<220>
<223> P600

<400> 70 taactataac ggtcctaagg tagcgaagat atccgatcaa gaaagcactc cgggctccag	60
aaggagcctt ccaggccagc tttgagcata agctgctgat gagcagtgag tgtcttgagt	120
agtgttcagg gcagcatggt accattcatg cttgacttct agccagtgtg acgagaggct	180
ggagtcaggt ctctagagag ttgagcagct ccagccttag atctcccagt cttatgcggt	240
gtgcccattc gctttgtgtc tgcagtcccc tggccacacc cagtaacagt tctgggatct	300
atgggagtag cttccttagt gagctttccc ttcaaatact ttgcaaccag gtagagaagt	360
ttggagtgaa ggttttgttc ttcgtttctt cacaatatgg atatgcatct tcttttgaaa	420
atgttaaagt aaattacctc tcttttcaga tactgtcttc atgcgaactt ggtatcctgt	480
ttccatccca gccttctata acccagtaac atcttttttg aaaccagtgg gtgagaaaga	540
cacctgggtca ggaacgcgga ccacaggaca actcaggctc acccacggca tcagactaaa	600
ggcaaacaag gactctgtat aaagtaccgg tggcatgtgt attagtggag atgcagcctg	660
tgctctgcag acagggagtc acacagacac ttttctataa tttcttaagt gctttgaatg	720
ttcaagtaga aagtctaaca ttaaatttga ttgaacaatt gtatatteat ggaatathtt	780
ggaacggaat accaaaaaat ggcaatagtg gttctttctg gatggaagac aaacttttct	840
tctttaaaat aaatthttatt ttatatatth gaggttgacc acatgacctt aaggatacat	900
atagacagta aactggttac tacagtgaag caaattaaca tatctacat cgtacatagt	960
tacatthttt tgtgtgacag gaacagctaa aatctacgta tttaacaaaa ctccataaga	1020

caatacattt ttattaacta tagccctcat gatgtacatt agatccctctc ttaaggtagc	1080
ccgaataaaa atagaaaaac aacaaatgtt aattttctat ttttatittg tgtcgtactc	1140
aatccataat cgtcacaagc gttatagttt aagtacaaca atgtcgtcac cttgttggtg	1200
taattaggtt tacgccaaca gggtgataac aggcgcgccg gtccctgtatt agaggtcacg	1260
tgagtgtttt gcgacatttt gcgacaccat gtggtcacgc tgggtattta agcccgagtg	1320
agcacgcagg gtctccattt tgaagcggga ggtttgaacg cgcagccgcc atgccggggt	1380
tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc ggcatttctg	1440
acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat tctgacatgg	1500
atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag cgcgactttc	1560
tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg caatttgaga	1620
agggagagag ctacttccac atgcacgtgc tcgtggaaac caccggggtg aaatccatgg	1680
ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt taccgcggga	1740
tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc gccggaggcg	1800
ggaacaaggt ggtggatgag tgctacatcc ccaattactt gtcceccaaa acccagcctg	1860
agctccagtg ggcgtggact aatatggaac agtatttaag cgctgtttg aatctcacgg	1920
agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag gagcagaaca	1980
aagagaatca gaatcccaat tctgatgcgc cggtgatcag atcaaaaact tcagccagat	2040
acatggagct ggtcgggtgg ctcttggaac aggggattac ctcggagaag cagtggatcc	2100
aggtgggtgt ctttcctgcc tgagctgacc tgggcagacg cgtaattaag tatcatgata	2160
aactgacaaa gcggtatata aagatactta atacatatgg tcctgtaaaa gtttaatgtg	2220
taagaagtat ttgttataaa agataaatat tcagaatctt ctttttaatt cctgatttta	2280
tttctatagg actgaaagac ttgctcgaga tgtcatgaag gagatgggag gccatcacat	2340
tctcacttca ctgcagctca acgtccceaa ttctaccggg taggggaggc gcttttccca	2400
aggcagtcctg gagcatgcgc tttagcagcc ccgctgggca cttggcgcta cacaagtggc	2460
ctctggcctc gcacacattc cacatccacc ggtaggcgcc aaccggctcc gttctttggt	2520

ggcccccttcg cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag	2580
ctcgcgtcgt gcaggacgtg acaaatggaa gtagcacgtc tcactagtct cgtgcagatg	2640
gacagcaccg ctgagcaatg gaagcgggta ggccctttggg gcagcggcca atagcagctt	2700
tgtctcttcg ctttctgggc tcagaggctg ggaaggggtg ggtccggggg cgggctcagg	2760
ggcgggctca ggggcggggc gggcgcccga aggtcctccg gaggcccggc attctgcacg	2820
cttcaaaagc gcacgtctgc cgcgtgttc tcctcttcct catctccggg cctttcgacc	2880
tagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg	2940
ctttacactt tatgtctccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc	3000
acacaggaaa cagctgccac catgattgaa caagatggat tgcacgcagg ttctccggcc	3060
gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat	3120
gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg	3180
tccggtgccc tgaatgaact gcaagacgag gcagcgcggc tategtggct ggccacgacg	3240
ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	3300
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta	3360
tccatcatgg ctgatgcaat gcggcggtg catacgttg atccggtac ctgcccattc	3420
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	3480
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	3540
ctcaaggcga gcatgcccga cggcgaggat ctcgtcgtga cccatggcga tgcctgcttg	3600
ccgaatatca tgggtggaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt	3660
gtggcggacc gctatcagga catagcgttg gctacccgtg atattgctga agagcttggc	3720
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttgcgagcgc	3780
atcgccttct atcgccttct tgacgagttc ttctgagata cattgatgag tttggacaaa	3840
ccacaactag aatgcagtga aaaaaatgct ttatttgtga aatttgtgat gctattgctt	3900
tatttgtaac cattataagc tgcaataaac aagttaacaa caacaattgc attcatttta	3960
tgtttcgggt tcagggggag gtgtgggagg ttttttaaag caagtaaaac ctctacaaat	4020

gtggtatggc tgattatgat cgcatagtgt taccatcaac caccttaact tcatttttct	4080
tattcaatac ctaggtaggt agatgctaga ttctggaaat aaaatatgag tctcaagtgg	4140
tccttgtcct ctctcccagt caaattctga atctagttgg caagattctg aaatcaaggc	4200
atataatcag taataagtga tgatagaagg gtatatagaa ggctagccgt caacatctcc	4260
actgacaaag cggtttctcc aacgaattat cgattttacgt ataatacctc cacgttgaga	4320
ccgtacgagc gcataacgga agcgttgttg agaccacgtc gcctgtgata ccgccttggt	4380
taaggcccggt agcctagctt aaggcaacga aacaagacga ctcaggagaa tcattatgaa	4440
cagggcgcgc gggaggtcgt ggatatggcc ggtattttac ttatgggagg atcttttggc	4500
cgcgagactt ccgagaaacc atcaacagtt cgatttccctc cttttatcct taacatagca	4560
cagagttgcc acattgatag gggggcagga gatcgtcaca tgaaccgacg gtcgggacat	4620
attcgatgcc gtccctctct gacaaccgga ttctctgttt ccaatcgaat accacgcccc	4680
ggggatcgtg cgcacgataa gcacaataag cgtcactgcg gggtcraagct gtggcttgca	4740
ggatgctaac tcgtaacgac attaagacag cagagcaatg ggcgacccaa aaagtttaag	4800
cctttacctc ggggtggcaa cataagtgtg ccgactacga caggaaccca tgactccttt	4860
ggggcgttct aatgggcggg gtctgctgtt aaccagctga tggggcaagt ggaatagagc	4920
atgccctacg cgctattaaag gaatgcgtgg ctacttggct gtgcgtactt gctcaatgta	4980
tggcaacaaa caacctaate tttgtcggca acgcgataat ctgccacgt taatcggcta	5040
cttgcattgt aggattcgat ttacgtatg taccacggat tctattgcac gctacagcat	5100
caatgtgccc gatgactact agctgcccag aggggataaa tcatgtgtaa tttggcgtgc	5160
attcgagtta ttggcattaa tttctctatc aagtgcagtc cctagcttcg agtaagctat	5220
gcgttccccc ccgtacattt taatcccata gggaacggca cccgcaccta ttaagagagc	5280
gccatagctc taaacgaccc acggtccaat gcttataatt tctaaaattt aaggtcgcat	5340
tgccttgacc gttagtcccc ctcacgtttg agggcataat gttcctgcgc cctttacaat	5400
tagtctaatt ctacttaaaa tcgacgcagg attttttctg ttggetccgc tgccatggga	5460
agctgtctat ggagggcctg gttatccctg gcttcttact atgtaatcta cctttacctt	5520

tgctaattac	gtgtacgtga	ccagaaacta	cacggaacga	gtgggacceca	taagctgagc	5580
gatggctagt	gagaccgcct	gatatgactc	acggggtaac	cgtgcaacgt	gcatatatct	5640
aatgagtacg	gcacctttcc	actttgctta	tatgagagta	caggctactt	agccccgacgt	5700
gtacgccaga	tctgggcacc	cgccagcagg	tcccactagg	ccggcctgcc	aggcaaatga	5760
cctgagtgcc	gcgactagcc	tcggccagac	agctgactag	cgattcaca	gtgccgactt	5820
agtagttctt	tagcaataga	aatatagaca	tagttccctc	gaactacagg	gaaatacggc	5880
ctgtggttgc	aagataagag	gcttttaaca	ctgacgtagt	ttacacactc	tggccggacc	5940
atgaatttcg	tctgctccca	cgaaccacat	tatcagtact	tcttcttaag	tgttttatta	6000
aatcgaataa	tctacaaata	gttctaaagg	agcgagttag	agagaagttc	cacgctccat	6060
aacttcgaga	ccgttagcga	agttttccag	cacacggcta	cgggaccctg	cactcagctt	6120
catggtttta	ggcacgagaa	ccgcatatag	tggagatttc	gccgtggaag	caagtctgtc	6180
ttagagagcc	taggtgctta	ggttttcgga	actctcctga	ttatatcca	tttccatccc	6240
attctggatc	atcaagtgcc	cgcggatgac	ctacgggcaa	attgccccag	ataaatcggg	6300
ctgtggggtc	ctgtggtgtg	tggggagctg	tcacatccag	ggtcctcact	gcctgtcccc	6360
ttccctcctc	aggaggacca	ggcctcatac	atctccttca	atgcggcctc	caactcgcgg	6420
tcccaaatta	aggctgcctt	ggacaatgcg	ggaaagatta	tgagcctgac	taaaaccgcc	6480
cccgactacc	tgggtgggcca	gcagcccgtg	gaggacattt	ccagcaatcg	gattttataa	6540
attttggaac	taaacgggta	cgatcccca	tatgcggctt	ccgtctttct	gggatgggcc	6600
acgaaaaagt	tgggcaagag	gaacaccatc	tggctgtttg	ggcctgcaac	taccgggaag	6660
accaacatcg	cggaggccat	agcccacact	gtgcccttct	acgggtgcgt	aaactggacc	6720
aatgagaact	ttcccttcaa	cgactgtgtc	gacaagatgg	tgatctggtg	ggaggagggg	6780
aagatgaccg	ccaaggtcgt	ggagtcggcc	aaagccattc	tcggaggaag	caaggtgcgc	6840
gtggaccaga	aatgcaagtc	ctcggcccag	atagacccca	ctcccgtgat	cgtcacctcc	6900
aacaccaaca	tgtgcgcctg	gattgacggg	aactcaacga	ccttcgaaca	ccagcagccg	6960
ttgcaagacc	ggatgttcaa	atttgaactc	acccgccgtc	tggatcatga	ctttgggaag	7020

gtcaccaagc aggaagtcaa agacttttttc cgggtgggcaa aggatcacgt ggttgaggtg	7080
gagcatgagt tctacgtcaa aaaggggtgga gccaagaaaa gacccgcccc cagtgacgca	7140
gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg	7200
gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg	7260
aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc	7320
ttcactcacg gacagaaaga ctgttttagag tgctttcccg tgtcagaatc tcaacccgtt	7380
tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg	7440
ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa	7500
caataaatga tttaaatcag gtatggctgc cgatggttat ctccagatt ggctcgagga	7560
caaccttagt gaagggattc gcgagtggtg ggctttgaaa cctggagccc ctcaacceaa	7620
ggcaaatcaa caacatcaag acaacgctcg aggtcttgtg ctccgggtt acaaatacct	7680
tggacccggc aacggactcg acaaggggga gccggtcaac gaagcagacg cggcggccct	7740
cgagcacgac aaggcctacg accagcagct caaggccgga gacaacccgt acctcaagta	7800
caaccacgcc gacgccgagt tccaggagcg gctcaaagaa gatacgtctt ttgggggcaa	7860
cctcggggcga gcagtcttcc aggccaaaaa gaggtctctt gaacctcttg gtctggttga	7920
ggaagcggct aagacggctc ctggaaagaa gaggcctgta gagcagtcct ctcaggaacc	7980
ggactcctcc gtgggtattg gcaaatcggg tgcacagccc gctaaaaaga gactcaattt	8040
cggtcagact ggcgacacag agtcagtccc agacctcaa ccaatcggag aacctcccgc	8100
agccccctca ggtgtgggat ctcttacaat ggcttcaggt ggtggcgcac cagtggcaga	8160
caataacgaa ggtgccgatg gagtgggtag ttctctggga aattggcatt gcgattccca	8220
atggctgggg gacagagtca tcaccaccag caccogaacc tgggccctgc ccacctaaa	8280
caatcacctc tacaagcaaa tctccaacag cacatctgga ggatcttcaa atgacaacgc	8340
ctacttcggc tacagcacc cctgggggta ttttgacttc aacagattcc actgccactt	8400
ctcaccacgt gactggcagc gactcatcaa caacaactgg ggattccggc ctaagcgact	8460
caacttcaag ctcttcaaca ttcaggtcaa agaggttacg gacaacaatg gagtcaagac	8520

catcgccaat aacettacca gcacgggtcca ggtctttcacg gactcagact atcagctccc	8580
gtacgtgctc gggtcgggctc acgagggctg cctccccgccg ttcccagcgg acgtttttcat	8640
gattcctcag tacgggtatc tgacgcttaa tgatggaagc caggccgtgg gtcgttcgtc	8700
cttttactgc ctggaatatt tcccgtcgca aatgctaaga acgggtaaca acttccagtt	8760
cagctacgag tttgagaacg tacctttcca tagcagctac gctcacagcc aaagcctgga	8820
ccgactaatg aatccactca tcgaccaata cttgtactat ctctcaaaga ctattaacgg	8880
ttctggacag aatcaacaaa cgctaaaatt cagtgtggcc ggaccacagca acatggctgt	8940
ccagggaaga aactacatac ctggaccag ctaccgacaa caacgtgtct caaccactgt	9000
gactcaaaac aacaacagcg aatttgcttg gcctggagct tcttcttggg ctctcaatgg	9060
acgtaatagc ttgatgaatc ctggacctgc tatggccagc cacaagaag gagaggaccg	9120
tttctttcct ttgtctggat ctttaatttt tggcaaacaa ggaactggaa gagacaacgt	9180
ggatgcggaac aaagtcatga taaccaacga agaagaaatt aaaactacta acccggtagc	9240
aacggagtcc tatggacaag tggccacaaa ccaccagagt gcccagcac aggcgcagac	9300
cggctgggtt caaaaccaag gaatacttcc gggtatggtt tggcaggaca gagatgtgta	9360
cctgcaagga cccatttggg ccaaattcc tcacacggac ggcaacttcc acccttctcc	9420
gctgatggga gggtttggaa tgaagcacc gcctcctcag atcctcatca aaaacacacc	9480
tgtacctgcg gatcctccaa cggccttcaa caaggacaag ctgaactctt tcataccca	9540
gtattctact ggccaagtca gcgtggagat cgagtgggag ctgcagaagg aaaacagcaa	9600
gcgctggaac ccggagatcc agtacacttc caactattac aagtctaata atgttgaatt	9660
tgctgttaat actgaaggtg tatatagtga accccgcccc attggcacca gatacctgac	9720
tcgtaatctg taattgcttg ttaatcaata aaccgtttaa ttcgtttcag ttgaactttg	9780
gtctctgcgg tttaaacggt cctgtattag aggtcacgtg agtgttttgc gacattttgc	9840
gacaccatgt ggtcacgctg ggtatttaag cccgagtgag cacgcagggt ctccattttg	9900
aagcgggagg tttgaacgcg cgcggccgcg tagtccccga gtacgcggct aggcagatcg	9960
gccccttcca aatccgtgtc atcggaagac catgttggtg gtttcataat ggattttatt	10020

gattgctatg taacctgaac gacggtatta atgggcttct taacatgaac gtgagtaacg	10080
gaaacattac gcttttcggt cgcgaatgac ggcgaactct gggccttgtg ggattataga	10140
tacgtcagtg ttccctcggga acggacaaag taggaaatat cgacgacatg gatcatgcca	10200
atatatcaac gcgggaagat gaccattgat gcagtgggaa ctggtgtccc acaaatatct	10260
agggcttgtg cagcgggtcaa gggctgtgtt ccacctggga cgcagccca attatacgcc	10320
agaaacgctt gattttactc aggtagaaag ggggaaaata cactacgcga ccaactaata	10380
attccittcg caaacttaat ggctgctaatt ttccctcaag accggcttaa tgcctactac	10440
acttaaggcg cggatggcga gatacgacgc gggcacatac taggcacgtc atcggccagg	10500
cgttctaatt taagttcacg cataaataag cgtctctcga gactgcagat ccccgcttta	10560
aatttagtac acaacgctgt ttaccaccgg ggcgttatga ctgcagacc cgtgcgtgaa	10620
tatcctgtat cattcatcac tgctgggact accccacttt acatatctaa gggattcaac	10680
aaacttcggg tggttaatgt caagtgcgtt attggcgtat aacagggatga atgcacattg	10740
atggaatgcc ttcaaccaag ctgtggcgtc acgccagccg aacacgtttt acccaaactt	10800
agcccttttg ctgcgacctt gctatctatt ggctatactc tatatgcaat gggatgcagat	10860
cggagaccgc gccattgtca ccacccaag gaggggagat tcacgctcgc tcggattagc	10920
ttcccggatc atttatcaaa actggaaata ctctaagcta agttcacctt cagatccttg	10980
tcaagagttg tcggcccgat tagtcgttgc gcacacgtgg cgtaattaca gcacgcgagg	11040
tacattegag tttatcctct gggctccatt caggcaatat gctgtcagca cgaagagtgt	11100
gacctggcgc tattgttggc gacttgacgc cacggaatac cagtgtatat acatcctagg	11160
agaactgcaa gacttcgggc gtgtatcgta ggttatggta gtcactggga tttgcacctg	11220
cgcacaatct gtttggtaaa gtaaaagcga tacttcgcga ctcccaataa gtaacgatac	11280
tgcgcggcat cccttagcgc tctcataccg agcgagatac agtctatata ctcaacagag	11340
cacgtttgga acctgtgagg ttcagcacca atcattcadc tataaactca tcgagaaaag	11400
tatggaggcc acgtgggctt ctteggttca gctaagcaga ctcatgaggg tgtggtctcg	11460
gctttggagg cagatattgt ccggaattgt aattgcgact gcgcggcagt atagagtgtc	11520

tctatcattg taggactcac gcacccatcg gacaaccacg tgatggcccc ttagttatat	11580
cgagggatag gggctatgct caagttagtc ttaatttcgc gtgcttgcaa ggcttatcgc	11640
cggtcacctc taccgtaacc ggcttttcag ggccatgac tttgcctgct tccctgttta	11700
tctacaccaa tattgggaga cttgtatgac gagcgagcca ctaaagagga cgattgtcaa	11760
cagggtgctt tatcacataa tagcctgttt ctccaacagt gcttcaatta cctagcatca	11820
cagaatgcga taagcccacc acgtgtttcc caatggcggc tattcattaa tcgggacgac	11880
acccctattc gtggagagag catcctattg agtccgagct gttagcgcac gattacggtg	11940
ggccagacag attcgacgga atgccgctac tacgctggcg cgtaatcagt cggaaccaga	12000
attaattaac ctgcaggcag ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc	12060
ccggggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg	12120
gagtggccaa ctccatcact aggggttcct tctagacgtg aggcctccgt gcccgtcagt	12180
gggcagagcg cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa	12240
ccggtgccta gagaaggtag cgcggggtaa actgggaaaag tgatgtcgtg tactggctcc	12300
gcctttttcc cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc	12360
tttttcgcaa cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc	12420
ctggcctctt tacgggttat ggcccttgcg tgccttgaat tacttccacg cccctggctg	12480
cagtacgtga ttcttgatcc cgagcttcgg gttggaagtg ggtgggagag ttcgaggcct	12540
tgcgcttaag gagccccctc gcctcgtgct tgagttgagg cctggcctgg gcgctggggc	12600
cgccgcgtgc gaatctggtg gcaccttcgc gcctgtctcg ctgctttcga taagtctcta	12660
gccatttaaa atttttgatg acctgctgcg acgctttttt tctggcaaga tagtcttgta	12720
aatgcggggc aagatctgca cactggtatt tcggtttttg gggccgcggg cggcgacggg	12780
gcccgtgcgt cccagcgctc atgttcggcg aggcggggcc tgcgagcgcg gccaccgaga	12840
atcggaaggg ggtagtctca agctggccgg cctgctctgg tgccctggcct cgcgccgccg	12900
tgtatcgccc cgccctgggc ggcaaggctg gcccggtcgg caccagttgc gtgagcggaa	12960
agatggccgc ttcccggccc tgctgcaggg agctcaaaat ggaggacgcg gcgctcggga	13020

gagcgggcgg	gtgagtcacc	cacacaaagg	aaaagggcct	ttccgtcctc	agccgtcgt	13080
tcattgtgact	ccacggagta	ccgggcgccg	tccaggcacc	tcgattagtt	ctcgagcttt	13140
tggagtacgt	cgtctttagg	ttggggggag	gggttttatg	cgatggagtt	tccccacact	13200
gagtgggtgg	agactgaagt	taggccagct	tggcacttga	tgtaattctc	cttggaattt	13260
gccccctttg	agtttggatc	ttggttcatt	ctcaagcctc	agacagtggg	tcaaagtttt	13320
tttcttccat	ttcaggtgtc	gtgagccacc	atggtgagca	agggcgagga	ggataacatg	13380
gccatcatca	aggagttcatt	gcgcttcaag	gtgcacatgg	agggcctcgt	gaacggccac	13440
gagttcgaga	tcgagggcga	gggcgagggc	cgcacctacg	agggcaccca	gaccgccaag	13500
ctgaagggtga	ccaagggtgg	ccccctgccc	ttcgccctggg	acatccctgtc	ccctcagttc	13560
atgtacggct	ccaaggccta	cgtgaagcac	cccgcgcaca	tccccgacta	cttgaagctg	13620
tccttccccg	agggtttcaa	gtgggagcgc	gtgatgaact	tcgaggacgg	cggcgtgggtg	13680
accgtgacct	aggactcctc	cctgcaggac	ggcgagttca	tctacaaggt	gaagctgcgc	13740
ggcaccaaact	tccccctcga	cggccccgta	atgcagaaga	agaccatggg	ctgggaggcc	13800
tcctccgagc	ggatgtacct	cgaggacggc	gccctgaagg	gcgagatcaa	gcagaggctg	13860
aagctgaagg	acggcggcca	ctacgacgct	gaggtcaaga	ccacctacaa	ggccaagaag	13920
cccgtgcagc	tgccccggcg	ctacaacgtc	aacatcaagt	tggacatcac	ctcccacaa	13980
gaggactaca	ccatcgtgga	acagtacgaa	cgcgccgagg	gccgccactc	caccggcggc	14040
atggacgagc	tgtacaagta	agtcgactac	tggccgaagc	cgtttggaat	aaggccggtg	14100
tgcgtttgtc	tatatgtgat	tttccaccat	attgccgtct	tttggcaatg	tgagggcccg	14160
gaaacctggc	cctgtcttct	tgacgagcat	tcctaggggt	ctttcccctc	tcgccaaagg	14220
aatgcaaggt	ctgttgaatg	tcgtgaagga	agcagttcct	ctggaagctt	cttgaagaca	14280
aacaacgtct	gtagcgacct	tttgcaggca	gcggaacccc	ccacctggcg	acaggtgcct	14340
ctgcggccaa	aagccacgtg	tataagatac	acctgcaaag	gcggcacaa	cccagtgcc	14400
cgttgtgagt	tggatagtgt	tggaaagagt	caaattggctc	tcctcaagcg	tattcaacaa	14460
ggggctgaag	gatgcccgag	aggtacccca	ttgtatggga	tctgatctgg	ggcctcggtg	14520

cacatgcttt	acatgtgttt	agtcgaggtt	aaaaaacgtc	taggcccccc	gaaccacggg	14580
gacgtggttt	tcctttgaaa	aacacgatcg	tacggccacc	atgtggggag	cctgcttgct	14640
gctgctgggc	ctcagcctac	aggtgtgccc	gagcgtcatc	ccagtggagg	aggagaatcc	14700
agccttctgg	aaccgaaagg	cagccgaggc	cctggatgct	gccaaagaagc	tgaagcccat	14760
tcagacatca	gctaagaacc	tcgttatcct	catgggagat	gggatggggg	tgtctacagt	14820
aaccgctacc	cggatcttaa	aggggcagca	gcaaggccat	ctgggacctg	agacacagtt	14880
agctatggat	cgctttccac	acatggctct	gtccaagaca	tacaacacag	acaagcagat	14940
cccggacagc	gcaggcacag	gtacagcctt	tctctgcggg	gtcaaaacca	acatgaaggt	15000
cattggcttg	agtcagctg	cacgcttcaa	ccagtgcaac	acgacatggg	gtaacgaggt	15060
cgtctcggtg	atgcaccgtg	ctaagaaagc	aggaaagtct	gtgggagtgg	tgaccaccac	15120
gtcggtgcac	catgcttctc	cggccggcac	ctacgcgcac	actgtgaacc	gtggttggtg	15180
ctcggatgca	caaatgcctg	cctcagcgtc	acaggatggc	tgcaaagaca	tctctactca	15240
gtcatctcc	aacatggaca	ttgatgtgat	cctcggtggt	ggccgcaagt	tcatgtttcc	15300
caaggggaca	ccagaccagg	aatatccaac	tgacaccaag	caggctggaa	ccaggctgga	15360
tggacgcaac	ctagttcaag	agtggctggc	aaagcaccag	ggagccccgt	atgtttggaa	15420
ccgctcagag	ctgattcagg	catccctgaa	ccgatctgtg	acacacctca	tgggcctctt	15480
tgagcctaac	gacatgaaat	atgaaatcca	ccgagaccct	gtcaggacc	cttccttggc	15540
cgagatgacg	gaggtggctg	tgcgcatgct	cagcaggaac	cccaaagget	tctatctctt	15600
tgtggaaggg	ggtcgcatcg	accatggcca	ccatgaaact	gtagcctatc	gtgccctgac	15660
tgaggctgtc	atgttcgact	cggctgttga	caaggcagac	aagctcacca	gtgagcagga	15720
cacgatgate	cttgtcactg	ccgaccactc	tcacgtcttc	tcttttggtg	gttacacaca	15780
gagaggggct	tccatctttg	gactggctcc	cttcaaggct	gaggatggca	aatcctttac	15840
ctcgatacta	tacggcaacg	gtcctggtta	caagctccat	aatggcgccc	gggctgatgt	15900
cactgaggaa	gagagcagca	accccacgta	ccagcagcag	gtgctgttac	ccctgtcgtc	15960
agagaccac	agcggggagg	acgtggcaat	attcgcgcgt	ggcccacagg	cgcagctggt	16020

gcacggagtt caggagcaga actacatcgc gcacgtcatg gccttcgcag cctgcctgga	16080
gccctacact gactgcggcc tggcatcccc tgctggccag agcagtgcag tgagcccggg	16140
ctgaacgcgt gggggaggct aactgaaaca cggaaggaga caataccgga aggaacccgc	16200
gctatgacgg caataaaaag acagaataaa acgcacggtg ttgggtcgtt tgttcataaa	16260
cgcggggttc ggtcccaggg ctggcactct gtcgataccc caccgagacc ccattggggc	16320
caatacgcgc gcgtttcttc cttttcccca cccaccccc caagttcggg tgaaggccca	16380
gggctcgcag ccaacgtcgg ggcggcaggc cctgccatag ccatctatgt cgggtgcgga	16440
gaaagaggta ataggaaccc ctagtgatgg agttggccac tccctctctg cgcgctcgt	16500
cgtcactga ggccgggcga ccaaaggtcg cccgacgccc gggctttgcc cgggcggcct	16560
cagtgagcga gcgagcgcgc agctgcctgc aggttaatta acctgcaggt gtacacaatg	16620
atacgatccc gtttctgcga cgtactcggg tgtagctggg acctgtcgc atttaatcgc	16680
cggtcfaatag gtacaccgtg ctgatcttgg atactcctga tcgcgatgta acccgagcgg	16740
cgtactaat gggcgtccta acacgagcgc gaataaccac ttgcgctagg caggatttac	16800
gatggtcac tggaacggac gggtacacca tgccaaaatt tgctcctagt tcacaaggtc	16860
aaacggtgac ccggaggcca gtccactcct tgtcctcgtc cttacgttac cgtaacctct	16920
gacctggagc ctaaggctgc tccggcattt taaaaaccac acaaataagg cttcacacta	16980
ttcctattat cttgtgtcct acctaaagg tagtggagcg agtaggtaga cacaggcact	17040
caccaccgat gttcgcgcca tgggccatgg ttaagggtccg aatgaaaccc gtatgcaagt	17100
taccggcctt atattaataa ggttgggtata catcctaaaa gctctctcgg acgtgaccaa	17160
atacaggcat ccacggagca ctactgaatg ggcacgtgga ccggcgagtg agaaatcgag	17220
cagcattaag tcattcttag taagccgaga tgggtctgcg ttatgtacgt ctagtgggga	17280
aacaagttgg tggacaaacg ttacgcctaa ggagctaata tttcggtcag gttgcgctgg	17340
gtcaaatcaa ccggtcctgg taacggagcg atggttaggt gtattacact agcgatgggc	17400
cgtgccggat ggtacgccac gttctccgag ccagtgaaat tggatgagaa ctcgtaaate	17460
tctcgttgca aacagcttcc tcaatatatg acctggccac tgatgagagg gtaacactta	17520

attcctagaa	tgtctgacga	catcaccctt	aagcagacac	actgcgatgt	gtggacgcta	17580
gggtgtctct	ttccacgttt	gacttagccc	atgatgtctt	gactaagatg	gcagggtatt	17640
tcccccttatt	cgtgtatgta	ccactattac	cgtcgattga	ccgtcgccgc	ttatacgtcg	17700
tgccctgtgc	cttggcaatt	ttgtccctgg	ataacccgcc	ttgcttctat	cgaggtcctt	17760
gaataagtta	atgcaacttg	tggaggaaca	tagggaatgt	ggcgtgctgc	ttagatatgg	17820
agtgcaagaa	gtcctctggg	ttgtagtttc	actgtttacgt	cctagccagt	gaagcactgg	17880
catgcagctt	tatctgagga	ttccactcac	tgcttacaag	aatttgccaa	gctggtgccg	17940
atatgacccc	aacgctatcc	agagttttaag	ttgtcatctc	tctccatcat	aattgccaca	18000
gtgtgtttgtg	tctacgaata	gctattaaaa	taccagcccc	tgagaacgat	gcccgaacaca	18060
tagggcagtg	caggcagcat	aacgtgcaga	tcaacataat	ttccactaag	tgtttgatata	18120
gtttactggc	tgagagccgc	catattggaa	cgacgacgag	agcgcgtatc	ctgcccgaacc	18180
agaaatacgt	gtccaatcca	tataactgac	caaagctgcg	gacagtttcc	tctccaccta	18240
caacgggatt	agcgcacttg	cttgtagtga	ggcaggtata	accacaggca	cgtaaacacct	18300
ttagacgttc	ggttccgcaa	tcgcacaatg	tactactatt	agagaactat	agcaaggaac	18360
ttatagcgta	tcgaaagttc	atccctgcgg	cccatagccg	taaccaccgt	gggggttattt	18420
taaactttaa	gttgtaccct	caggcaagcg	aagaccggtg	ctttacctac	atggggccgc	18480
tttccgagcc	acgccgcacc	ccaacatatt	cctacgcgct	gtaataatga	acataagaga	18540
aatggagtt	gggtgacagc	acatatcgac	tacctagcga	ctaccgcaac	acaaagtaga	18600
tcccgccggt	atcctccttt	catgcaggac	gggttacctg	ggaatgcgtt	tgtttgactg	18660
gtgcccgcctg	aggcgcctgaa	tatctcgcg	tagcgtgtac	cgatcaagaa	agcactccgg	18720
gatatctggc	aaacagctat	tatgggtatt	atgggtttca	gggagtggcg	cagctgcttc	18780
atccccgtgg	cccgttgctc	gcgtttgctg	gcggtgtccc	cggaagaaat	atatttgcat	18840
gtcttttagtt	ctatgatgac	acaaaccccc	cccagcgtct	tgtcattggc	gaaaacacgc	18900
agatgcagtc	ggggcggcgc	ggtcccaggt	ccacttcgca	tattaaggtg	acgcgtgtgg	18960
cctcgaacac	agagcgacgg	ccaccatgac	cgagtacaag	cccacggtgc	gcctcgccac	19020

ccgcgacgac	gtcccccggg	ccgtacgcac	cctcgccgcc	gcgttcgccg	actaccccgc	19080
cacgcgccac	accgttgacc	cggaccgccca	catcgagcgg	gtcaccgagc	tgcaagaact	19140
cttcctcacg	cgcgtcgggc	tcgacatcgg	caaggtgtgg	gtcgcggacg	acggcgccgc	19200
ggtggcggtc	tggaccacgc	cggagagcgt	cgaagcgggg	gcggtgttcg	ccgagatcgg	19260
cccgcgcatg	gccgagttga	gcggttcccc	gctggccgcg	cagcaacaga	tggaaggcct	19320
cctggcgccg	caccggccca	aggagcccgc	gtggttcctg	gccaccgtcg	gcgtctcgcc	19380
cgaccaccag	ggcaagggtc	tgggcagcgc	cgtcgtgctc	cccggagtgg	aggcggccga	19440
gcgcgccggg	gtgccccgct	tcctggagac	ctccgcgcc	cgcaacctcc	ccttctacga	19500
gcggctcggc	ttcaccgtea	ccgccgacgt	cgaggtgccc	gaaggaccgc	gcacctggtg	19560
catgaccgcg	aagcccggtg	cctgatgtgc	cttctagttg	ccagccatct	gttgtttgcc	19620
cctcccccg	gccttccttg	acctggaag	gtgccactcc	cactgtcctt	tcctaataaa	19680
atgaggaaat	tgcatcgcat	tgtctgagta	ggtgtcattc	tattctgggg	ggtggggtgg	19740
ggcaggacag	caagggggag	gattgggaag	acaatagcag	gcatgctggg	gatgcggtgg	19800
gctctatggt	cccccttatt	aaccctaaac	gggtagcata	tgcttcccgg	gtagtagtat	19860
atactatcca	gactaaccct	aattcaatag	catatgttac	ccaacgggaa	gcatatgcta	19920
togaattagg	gttagtaaaa	gggtcctaag	gaacagcgat	ctggatagca	tatgctatcc	19980
taatctatat	ctgggtagca	tatgctatcc	taatctatat	ctgggtagca	taggctatcc	20040
taatctatat	ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	20100
taatttatat	ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	20160
taatctatat	ctgggtagta	tatgctatcc	taatctgtat	ccgggtagca	tatgctatcc	20220
tcatgcatat	acagtcagca	tatgataccc	agtagtagag	tgggagtgct	atcctttgca	20280
tatgccgccca	cctcccaagg	attatgggtg	acgtcagggtg	gcacttttcg	gggaaatgtg	20340
cgcggaaccc	ctatttgttt	atTTTTctaa	atacattcaa	atatgtatcc	gctcatgaga	20400
caataaccct	gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	20460
ttccgtgtcg	cccttattec	cttttttgcg	gcattttgcc	ttcctgtttt	tgtcaccca	20520

gaaacgctgg	tgaaagtaaa	agatgctgaa	gacagtttg	gtgcacgagt	gggttacatc	20580
gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgtttttcca	20640
atgatgagca	cttttaaagt	tctgctatgt	ggcgcggtat	tatcccgtat	tgacgccggg	20700
caagagcaac	tcggtcgccg	catacactat	tctcagaatg	acttggttga	gtactcacca	20760
gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	aattatgcag	tgttgccata	20820
accatgagtg	ataaacttgc	ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	20880
ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	20940
gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	21000
acaacgttgc	gcaaactatt	aactggcgaa	ctacttactc	tagcttcccg	gcaacaatta	21060
atagactgga	tggaggcgga	taaagttgca	ggaccacttc	tgcgctcggc	ccttccggct	21120
ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	ggtctcgcgg	tatcattgca	21180
gcactggggc	cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	21240
gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	21300
tggtaaactgt	cagaccaagt	ttactcatat	atactttaga	ttgatttaaa	acttcatttt	21360
taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	tcatgaccaa	aatcccttaa	21420
cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	agatcaaagg	atcttcttga	21480
gacccTTTTT	ttctgcgcgt	aatctgctgc	ttgcaaacaa	aaaaaccacc	gctaccagcg	21540
gtggtttggt	tgccggatca	agagctacca	actctttttc	cgaaggtaac	tggttcagc	21600
agagcgcaga	taccaaatac	tgttcttcta	gtgtagccgt	agttaggcca	ccacttcaag	21660
aactctgtag	caccgcctac	atacctcgct	ctgctaatcc	tgttaccagt	ggctgctgcc	21720
agtggcgata	agtcgtgtct	taccgggttg	gactcaagac	gatagttacc	ggataaggcg	21780
cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca	gcttggagcg	aacgacctac	21840
accgaactga	gataacctaca	gcgtgagcta	tgagaaagcg	ccacgcttcc	cgaagggaga	21900
aaggcggaca	ggtatccggt	aagcggcagg	gtcggaacag	gagagcgcac	gagggagctt	21960
ccagggggaa	acgcctggta	tctttatagt	cctgtcgggt	ttcgccacct	ctgacttgag	22020

cgtcgatttt	tgtgatgctc	gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	22080
gcctttttac	ggttcctggc	cttttgctgg	ccttttgctc	acatgt		22126
<210>	71					
<211>	19					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	β -肌動蛋白剪接供體					
<400>	71					
gtggatccag	gtgggtgtc					19
<210>	72					
<211>	19					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	β -肌動蛋白剪接受體					
<400>	72					
ccctcctcag	gaggaccag					19
<210>	73					
<211>	863					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	在所例示之建構體中止擋匣之AAV2 rep 基因上游之經突變5'部分					
<400>	73					
ggtcctgtat	tagaggtcac	gtgagtgttt	tgcgacattt	tgcgacacca	tgtggtcacg	60
ctgggtat	tt aagcccgagt	gagcacgcag	ggtctccatt	ttgaagcggg	aggtttgaac	120
gcgcagccgc	catgccgggg	ttttacgaga	ttgtgattaa	ggtccccagc	gaccttgacg	180
agcatctgcc	cggcatttct	gacagctttg	tgaactgggt	ggccgagaag	gaatgggagt	240
tgccgccaga	ttctgacatg	gatctgaatc	tgattgagca	ggcacccctg	accgtggccg	300
agaagctgca	gcgcgacttt	ctgacggaat	ggcgccgtgt	gagtaaggcc	ccggaggccc	360

ttttctttgt gcaatttgag aagggagaga gctacttcca catgcacgtg ctcgtggaaa	420
ccaccggggt gaaatccatg gttttgggac gtttcttgag tcagattcgc gaaaaactga	480
ttcagagaat ttaccgcggg atcgagccga ctttgccaaa ctggttcgcg gtcacaaaga	540
ccagaaatgg cgccggaggc gggaacaagg tggatgatga gtgctacatc cccaattact	600
tgtccccaa aaccagcct gagctccagt gggcgtggac taatatggaa cagtatttaa	660
gcgcctgttt gaatctcacg gacgtaaac ggttggatgc gcagcatctg acgcacgtgt	720
cgcagacgca ggagcagaac aaagagaatc agaatcccaa ttctgatgcg ccggtgatca	780
gatcaaaaac ttcagccaga tacatggagc tggtcgggtg gctcgtggac aaggggatta	840
cctcggagaa gcagtggatc cag	863

【發明申請專利範圍】

【請求項1】

一種非天然存在的核酸分子，其包含經修飾腺相關病毒(AAV) *rep*基因，該基因具有編碼四種Rep蛋白Rep78、Rep68、Rep52、及Rep40之AAV *rep*基因及插入至由該等四種Rep蛋白共享之該*rep*基因的編碼序列中的人工內含子，其中該人工內含子包含止擋匣(stop cassette)，該止擋匣係插入在該人工內含子之5'剪接位點下游及分支位點上游，且該止擋匣以5'至3'的順序包含：

- (a) 具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attP*位點，較佳的是具有SEQ ID NO:7之核苷酸序列的*attP*位點；
- (b) 剪接受體；
- (c) 終止子；及
- (d) 具有與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的*attB*位點，較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點。

【請求項2】

如請求項1之非天然存在的核酸分子，其中該剪接受體包含SEQ ID NO:17之核苷酸序列。

【請求項3】

如請求項1或2之非天然存在的核酸分子，其中該終止子包含多腺苷酸化信號。

【請求項4】

如請求項3之非天然存在的核酸分子，其中該終止子進一步包含SEQ ID NO:19之核苷酸序列。

【請求項5】

如請求項1至4中任一項之非天然存在的核酸分子，其中該止擋匣包含編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶(neomycin phosphotransferase)表現匣。

【請求項6】

如請求項1至5中任一項之非天然存在的核酸分子，其中該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:15之核苷酸序列。

【請求項7】

如請求項1至6中任一項之非天然存在的核酸分子，其中該AAV *rep*基因包含AAV1至AAV8中之一者、或其混合之*rep*基因。

【請求項8】

如請求項7之非天然存在的核酸分子，其中該AAV *rep*基因包含人類AAV2之該*rep*基因，該基因具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至2202。

【請求項9】

如請求項8之非天然存在的核酸分子，其中該人工內含子係插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號996至1905之間。

【請求項10】

如請求項9之非天然存在的核酸分子，其中該人工內含子係直接插入在 GenBank 存取號 NC_001401.2 之核苷酸序列的核苷酸編號 1052、1061、1712、1906、1022、1112、1475、1514、1700、1742、1784 或 1340、較佳的是核苷酸編號 1052 的下游。

【請求項11】

一種非天然存在的核酸分子，其包含經修飾 AAV *rep* 基因，該基因以 5' 至 3' 的順序包含：

- (a) 具有 SEQ ID NO:55 之核苷酸序列的 AAV *rep* 基因之 5' 部分；
- (b) 人工內含子，其以 5' 至 3' 的順序包含：
 - (i) 具有 SEQ ID NO:14 之核苷酸序列的 5' 內含子片段；
 - (ii) 止擋匣，其以 5' 至 3' 的順序包含：
 - (1) 具有 SEQ ID NO:7 之核苷酸序列的 *attP* 位點；
 - (2) 具有 SEQ ID NO:17 之核苷酸序列的剪接受體；
 - (3) 具有 SEQ ID NO:18 之核苷酸序列的新黴素磷酸轉移酶表現匣；
 - (4) 具有 SEQ ID NO:19 之核苷酸序列的終止子；及
 - (5) 具有 SEQ ID NO:8 或 SEQ ID NO:9 之核苷酸序列的 *attB* 位點；及
 - (iii) 具有 SEQ ID NO:15 之核苷酸序列的 3' 內含子片段；及
- (c) 具有 SEQ ID NO:56 之核苷酸序列的該 AAV *rep* 基因之 3' 部分。

【請求項12】

一種非天然存在的核酸分子，其包含經修飾 AAV *rep* 基因，該基因

以5'至3'的順序包含：

- (a) 具有SEQ ID NO:73之核苷酸序列的AAV *rep*基因之5'部分；
- (b) 人工內含子，其以5'至3'的順序包含：
 - (i) 具有SEQ ID NO:14之核苷酸序列的5'內含子片段；
 - (ii) 止擋匣，其以5'至3'的順序包含：
 - (1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；
 - (2) 具有SEQ ID NO:17之核苷酸序列的剪接受體；
 - (3) 具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；
 - (4) 具有SEQ ID NO:19之核苷酸序列的終止子；及
 - (5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及
 - (iii) 具有SEQ ID NO:66之核苷酸序列的3'內含子片段；及
- (c) 具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。

【請求項13】

如請求項11或12之非天然存在的核酸分子，其中該止擋匣包含SEQ ID NO:16之核苷酸序列。

【請求項14】

如請求項1至13中任一項之非天然存在的核酸分子，其進一步包含編碼三種殼體蛋白VP1、VP2及VP3之AAV *cap*基因。

【請求項15】

如請求項14之非天然存在的核酸分子，其中該AAV *cap*基因包含

AAV1至AAV9及AAVDJ中之一者、或其混合之*cap*基因。

【請求項16】

如請求項15之非天然存在的核酸分子，其中該AAV *cap*基因包含具有GenBank存取號AY530579.1之核苷酸序列的人類AAV9之*cap*基因。

【請求項17】

如請求項14至16中任一項之非天然存在的核酸分子，其中該AAV *cap*基因進一步包含多腺苷酸化信號（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號4411至4466之AAV2的多腺苷酸化信號）及強化子（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至313之AAV2 *rep* P5啟動子），其中該多腺苷酸化信號及該強化子兩者皆係該*cap*基因之該編碼序列的下游。

【請求項18】

如請求項14至17中任一項之非天然存在的核酸分子，其進一步包含由該AAV *cap*基因下游的一對AAV反向末端重複(ITR)側接之轉殖基因。

【請求項19】

如請求項18之非天然存在的核酸分子，其進一步包含該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：

- (a) 具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；
- (b) 具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；
- (c) 具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；

(d) 具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因子06；

(e) 具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07；

(f) 具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12；

(g) 具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13；

(h) 具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35；

(i) 具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36；

(j) 具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因子52；

(k) 具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53；及

(l) 來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子，

更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及SEQ ID NO:25之核苷酸序列。

【請求項20】

如請求項19之非天然存在的核酸分子，其中該非天然存在的核酸分子包含該經修飾AAV *rep*基因上游的該第一絕緣子，並且進一步分別包含

該轉殖基因上游及下游的第一間隔序列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自由下列所組成之群組：

- a) SEQ ID NO:67之核苷酸序列；及
- b) SEQ ID NO:68之核苷酸序列。

【請求項21】

如請求項18至20中任一項之非天然存在的核酸分子，其中該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；較佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列。

【請求項22】

一種非天然存在的核酸分子，其以5'至3'的順序包含：

(A) 第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；

(B) 經修飾AAV *rep*基因，其以5'至3'的順序包含：

(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:55之核苷酸序列；

(ii) 人工內含子，其以5'至3'的順序包含：

(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；

(b) 止擋匣，其以5'至3'的順序包含：

(1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；

(2) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17之

核苷酸序列；

(3) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；

(4) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及

(5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及

(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:15之核苷酸序列；

(iii) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；

(C) AAV *cap*基因，較佳地該AAV *cap*基因包含SEQ ID NO:57之核苷酸序列；

(D) 由一對AAV ITR側接之轉殖基因，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及

(E) 第二絕緣子，較佳地該第二絕緣子具有SEQ ID NO:25之核苷酸序列。

【請求項23】

一種非天然存在的核酸分子，其以5'至3'的順序包含：

(A) 第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核

苷酸序列；

(B) 經修飾AAV *rep*基因，其以5'至3'的順序包含：

(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:73之核苷酸序列；

(ii) 人工內含子，其以5'至3'的順序包含：

(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；

(b) 止擋匣，其以5'至3'的順序包含：

(1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；

(2) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；

(3) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；

(4) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及

(5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及

(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:66之核苷酸序列；

(iii) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；

(C) AAV *cap*基因；

(D) 由下列側接之轉殖基因：

(1) 一對AAV ITR，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及

(2) 一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列。

【請求項24】

一種載體，其包含如請求項1至22中任一項之非天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:12之核苷酸序列。

【請求項25】

一種載體，其包含如請求項1至21或23中任一項之非天然存在的核酸分子；較佳地，該載體係質體；更佳地，該質體包含SEQ ID NO:70之核苷酸序列。

【請求項26】

一種製造如請求項1至23中任一項之非天然存在的核酸分子之方法。

【請求項27】

一種製造如請求項24或25之載體的方法。

【請求項28】

一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾腺相關病毒(AAV) *rep* 基因，該基因具有編碼四種Rep蛋白Rep78、Rep68、Rep52、及Rep40之AAV *rep* 基因及插入至由該等四種Rep蛋白共

享之該 *rep* 基因的編碼序列中的人工內含子，其中該人工內含子包含止擋匣，該止擋匣係插入在該人工內含子之5'剪接位點下游及分支位點上游，且該止擋匣以5'至3'的順序包含：

- (a) 具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 *attP* 位點，較佳的是具有SEQ ID NO:7之核苷酸序列的 *attP* 位點；
- (b) 剪接受體；
- (c) 終止子；及
- (d) 具有與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 *attB* 位點，較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 *attB* 位點。

【請求項29】

如請求項28之細胞，其中該剪接受體包含SEQ ID NO: 17之核苷酸序列。

【請求項30】

如請求項28或29之細胞，其中該終止子包含多腺苷酸化信號。

【請求項31】

如請求項30之細胞，其中該終止子進一步包含SEQ ID NO:19之核苷酸序列。

【請求項32】

如請求項28至31中任一項之細胞，其中該止擋匣包含編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移

酶表現匣。

【請求項33】

如請求項28至32中任一項之細胞，其中該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:15之核苷酸序列。

【請求項34】

如請求項28至32中任一項之細胞，其中該人工內含子以5'至3'的順序包含SEQ ID NO:14之核苷酸序列、該止擋匣、及SEQ ID NO:66之核苷酸序列。

【請求項35】

如請求項28至33中任一項之細胞，其中該AAV *rep*基因包含AAV1至AAV8中之一者、或其混合之*rep*基因。

【請求項36】

如請求項35之細胞，其中該AAV *rep*基因包含人類AAV2之該*rep*基因，該基因具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至2202。

【請求項37】

如請求項36之細胞，其中該人工內含子係插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號996至1905之間。

【請求項38】

如請求項37之細胞，其中該人工內含子係直接插入在GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號1052、1061、1712、1906、1022、1112、1475、1514、1700、1742、1784或1340、較佳的是核苷酸

編號1052的下游。

【請求項39】

一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾 AAV *rep* 基因，該基因以5'至3'的順序包含：

- (a) 具有SEQ ID NO:55之核苷酸序列的AAV *rep* 基因之5'部分；
- (b) 人工內含子，其以5'至3'的順序包含：
 - (i) 具有SEQ ID NO:14之核苷酸序列的5'內含子片段；
 - (ii) 止擋匣，其以5'至3'的順序包含：
 - (1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；
 - (2) 具有SEQ ID NO:17之核苷酸序列的剪接受體；
 - (3) 具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；
 - (4) 具有SEQ ID NO:19之核苷酸序列的終止子；及
 - (5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及
 - (iii) 具有SEQ ID NO:15之核苷酸序列的3'內含子片段；及
- (c) 具有SEQ ID NO:56之核苷酸序列的該AAV *rep* 基因之3'部分。

【請求項40】

一種細胞，其包含非天然存在的核酸分子，該核酸分子包含經修飾 AAV *rep* 基因，該基因以5'至3'的順序包含：

- (a) 具有SEQ ID NO:73之核苷酸序列的AAV *rep* 基因之5'部分；
- (b) 人工內含子，其以5'至3'的順序包含：

- (i) 具有SEQ ID NO:14之核苷酸序列的5'內含子片段；
- (ii) 止擋匣，其以5'至3'的順序包含：
 - (1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；
 - (2) 具有SEQ ID NO:17之核苷酸序列的剪接受體；
 - (3) 具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；
 - (4) 具有SEQ ID NO:19之核苷酸序列的終止子；及
 - (5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及
- (iii) 具有SEQ ID NO:66之核苷酸序列的3'內含子片段；及
- (c) 具有SEQ ID NO:56之核苷酸序列的該AAV *rep*基因之3'部分。

【請求項41】

如請求項39或40之細胞，其中該止擋匣包含SEQ ID NO:16之核苷酸序列。

【請求項42】

如請求項28至41中任一項之細胞，其進一步包含編碼三種殼體蛋白VP1、VP2及VP3之AAV *cap*基因。

【請求項43】

如請求項42之細胞，其中該AAV *cap*基因包含AAV1至AAV9及AAVDJ中之一者、或其混合之*cap*基因。

【請求項44】

如請求項43之細胞，其中該AAV *cap*基因包含具有GenBank存取號

AY530579.1之核苷酸序列的人類AAV9之*cap*基因。

【請求項45】

如請求項43之細胞，其中AAV *cap*基因包含AAV9的混合之*cap*基因。

【請求項46】

如請求項42至45之細胞，其中該AAV *cap*基因進一步包含多腺苷酸化信號（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號4411至4466之AAV2的多腺苷酸化信號）及強化子（較佳的是具有GenBank存取號NC_001401.2之核苷酸序列的核苷酸編號190至313之AAV2 *rep* P5啟動子），其中該多腺苷酸化信號及該強化子兩者皆係該*cap*基因之該編碼序列的下游。

【請求項47】

如請求項42至46中任一項之細胞，其進一步包含由該AAV *cap*基因下游的一對AAV反向末端重複(ITR)側接之轉殖基因。

【請求項48】

如請求項47之細胞，其進一步包含該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：

- (a) 具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；
- (b) 具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；
- (c) 具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；
- (d) 具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因

子06；

(e) 具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因

子07；

(f) 具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因

子12；

(g) 具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因

子13；

(h) 具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因

子35；

(i) 具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因

子36；

(j) 具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因

子52；

(k) 具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因

子53；及

(l) 來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的
球蛋白基因座之雞HS4絕緣子，

更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及
SEQ ID NO:25之核苷酸序列。

【請求項49】

如請求項48之細胞，其中該細胞包含該經修飾AAV *rep*基因上游的
該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第一間隔序
列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自

由下列所組成之群組：

- a) SEQ ID NO:67之核苷酸序列；及
- b) SEQ ID NO:68之核苷酸序列。

【請求項50】

如請求項47至49中任一項之細胞，其中該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；較佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列。

【請求項51】

一種細胞，其包含非天然存在的核酸分子，該核酸分子以5'至3'的順序包含：

(A) 第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；

(B) 經修飾AAV *rep*基因，其以5'至3'的順序包含：

(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:55之核苷酸序列；

(ii) 人工內含子，其以5'至3'的順序包含：

(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；

(b) 止擋匣，其以5'至3'的順序包含：

(1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；

(2) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17

之核苷酸序列；

(3) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；

(4) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及

(5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及

(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:15之核苷酸序列；

(iii) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；

(C) AAV *cap*基因，較佳地該AAV *cap*基因包含SEQ ID NO:57之核苷酸序列；

(D) 由一對AAV ITR側接之轉殖基因，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及

(E) 第二絕緣子，較佳地該第二絕緣子具有SEQ ID NO:25之核苷酸序列。

【請求項52】

一種細胞，其包含非天然存在的核酸分子，該核酸分子以5'至3'的順序包含：

(A) 第一絕緣子，較佳地該第一絕緣子具有SEQ ID NO:24之核苷酸序列；

(B) 經修飾AAV *rep*基因，其以5'至3'的順序包含：

(i) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因之該5'部分具有SEQ ID NO:73之核苷酸序列；

(ii) 人工內含子，其以5'至3'的順序包含：

(a) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；

(b) 止擋匣，其以5'至3'的順序包含：

(1) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；

(2) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；

(3) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；

(4) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及

(5) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及

(c) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:66之核苷酸序列；

(iii) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；

(C) AAV *cap*基因；及

(D) 由下列側接之轉殖基因：

(i) 一對AAV ITR，較佳地，該AAV ITR具有SEQ ID NO:20之核苷酸序列，且該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及

(ii) 一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列。

【請求項53】

如請求項28至51中任一項之細胞，其中該非天然存在的核酸分子係游離型(episomal)，其具有SEQ ID NO:12之核苷酸序列。

【請求項54】

如請求項28至50或52中任一項之細胞，其中該非天然存在的核酸分子係游離型，其具有SEQ ID NO:70之核苷酸序列。

【請求項55】

如請求項52至54中任一項之細胞，其進一步包含編碼重組酶之核酸分子，該重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的胺基酸序列；較佳地，該核酸包含與SEQ ID NO: 3之核苷酸序列至少85%、至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%同一的核苷酸序列；更佳地，該細胞包含重組 Δ E1/ Δ E3腺病毒血清型5 (Ad5)病毒，該病毒編碼具有SEQ ID NO:2之胺基酸序列的重組酶。

【請求項56】

如請求項52至55中任一項之細胞，其進一步包含腺病毒E1A及E1B基因，較佳地該細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6細胞。

【請求項57】

一種生產包含轉殖基因之重組AAV的方法，其包含：

(A) 獲得第一宿主細胞，其包含：

(i) 經修飾AAV *rep*基因，其以5'至3'的順序包含：

(a) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因具有SEQ ID NO:55之核苷酸序列；

(b) 人工內含子，其以5'至3'的順序包含：

(1) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；

(2) 止擋匣，其以5'至3'的順序包含：

(aa) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；

(bb) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；

(cc) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；

(dd) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及

(ee) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列

的*attB*位點；及

(3) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:15之核苷酸序列；

(c) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:56之核苷酸序列；

(ii) AAV *cap*基因，較佳地該AAV *cap*基因包含SEQ ID NO:57之核苷酸序列；及

(iii) 由一對AAV ITR側接之該轉殖基因，較佳地，該ITR具有SEQ ID NO:20之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；

(B) 將該第一宿主細胞用重組腺病毒感染，該重組腺病毒包含編碼具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%序列同一性的重組酶之重組酶基因，以獲得進一步含有該重組酶基因的第二宿主細胞；

(C) 在生產包含該轉殖基因之該重組AAV的條件下使該第二宿主細胞生長；及

(D) 可選地收集該重組AAV。

【請求項58】

一種生產包含轉殖基因之重組AAV的方法，其包含：

(A) 獲得第一宿主細胞，其包含：

- (i) 經修飾AAV *rep*基因，其以5'至3'的順序包含：
 - (a) AAV *rep*基因之5'部分，較佳地該AAV *rep*基因具有SEQ ID NO:73之核苷酸序列；
 - (b) 人工內含子，其以5'至3'的順序包含：
 - (1) 5'內含子片段，較佳地該5'內含子片段具有SEQ ID NO:14之核苷酸序列；
 - (2) 止擋匣，其以5'至3'的順序包含：
 - (aa) 具有SEQ ID NO:7之核苷酸序列的*attP*位點；
 - (bb) 剪接受體，較佳地該剪接受體具有SEQ ID NO:17之核苷酸序列；
 - (cc) 編碼可選擇標記之基因，較佳的是具有SEQ ID NO:18之核苷酸序列的新黴素磷酸轉移酶表現匣；
 - (dd) 終止子，較佳地該終止子具有SEQ ID NO:19之核苷酸序列；及
 - (ee) 具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的*attB*位點；及
 - (3) 3'內含子片段，較佳地該3'內含子片段具有SEQ ID NO:66之核苷酸序列；
 - (c) 該AAV *rep*基因之3'部分，較佳地該AAV *rep*基因之該3'部分具有SEQ ID NO:66之核苷酸序列；
- (ii) AAV *cap*基因；及
- (iii) 該轉殖基因係由下列側接：
 - (a) 一對AAV ITR，較佳地，該ITR具有SEQ ID NO:20

之核苷酸序列，該轉殖基因包含可操作地連接至編碼序列的啟動子，且該編碼序列係可操作地連接的多腺苷酸化信號；更佳地，該啟動子具有SEQ ID NO:21之核苷酸序列且該多腺苷酸化信號具有SEQ ID NO:23之核苷酸序列；及

(b) 一對間隔序列，較佳地，該間隔序列具有SEQ ID NO:67及SEQ ID NO:68之核苷酸序列；

(B) 將該第一宿主細胞用重組腺病毒感染，該重組腺病毒包含編碼具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%序列同一性的重組酶之重組酶基因，以獲得進一步含有該重組酶基因的第二宿主細胞；

(C) 在生產包含該轉殖基因之該重組AAV的條件下使該第二宿主細胞生長；及

(D) 可選地收集該重組AAV。

【請求項59】

如請求項57或58之方法，其中該第一宿主細胞進一步包含該經修飾AAV *rep*基因上游的第一絕緣子及可選地由該等ITR側接之該轉殖基因下游的第二絕緣子，較佳地，該第一絕緣子及該第二絕緣子係獨立地選自由下列所組成之群組：

- (a) 具有SEQ ID NO:24之核苷酸序列的人類抗抑制因子40；
- (b) 具有SEQ ID NO:25之核苷酸序列的小鼠抗抑制因子40；
- (c) 具有GenBank存取號AY190749.1之核苷酸序列的抗抑制因子04；

(d) 具有GenBank存取號AY190750.1之核苷酸序列的抗抑制因子06；

(e) 具有GenBank存取號AY190751.1之核苷酸序列的抗抑制因子07；

(f) 具有GenBank存取號AY190752.1之核苷酸序列的抗抑制因子12；

(g) 具有GenBank存取號AY190753.1之核苷酸序列的抗抑制因子13；

(h) 具有GenBank存取號AY190754.1之核苷酸序列的抗抑制因子35；

(i) 具有GenBank存取號AY190755.1之核苷酸序列的抗抑制因子36；

(j) 具有GenBank存取號AY190757.1之核苷酸序列的抗抑制因子52；

(k) 具有GenBank存取號AY190758.1之核苷酸序列的抗抑制因子53；及

(l) 來自在二或更多個拷貝中具有AY040835.1之核苷酸序列的球蛋白基因座之雞HS4絕緣子，

更佳地，該第一絕緣子及該第二絕緣子分別具有SEQ ID NO:24及SEQ ID NO:25之核苷酸序列。

【請求項60】

如請求項59之方法，其中該第一宿主細胞包含該經修飾AAV *rep*基因上游的該第一絕緣子，並且進一步分別包含該轉殖基因上游及下游的第

一間隔序列及第二間隔序列，其中該第一間隔序列及該第二間隔序列係獨立地選自由下列所組成之群組：

- a) SEQ ID NO:67之核苷酸序列；及
- b) SEQ ID NO:68之核苷酸序列。

【請求項61】

如請求項57至59中任一項之方法，其中該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、及該第二絕緣子。

【請求項62】

如請求項61之方法，其中該第一宿主細胞係藉由將核酸分子引入至該細胞中而獲得，該核酸分子以5'至3'的順序包含該第一絕緣子、該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、及該第二絕緣子、較佳的是包含SEQ ID NO:12之核苷酸序列的質體。

【請求項63】

如請求項57、58、或60之方法，其中該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、該第一間隔序列、及該第二間隔序列。

【請求項64】

如請求項63之方法，其中該第一宿主細胞係藉由將一或多種核酸分子引入至細胞中而獲得，該等核酸分子包含該經修飾AAV *rep*基因、該

AAV *cap*基因、由該等ITR側接之該轉殖基因、該第一絕緣子、該第一間隔序列、及該第二間隔序列、較佳的是包含SEQ ID NO:70之核苷酸序列的質體。

【請求項65】

如請求項57至62中任一項之方法，其中該重組腺病毒係重組 Δ E1/ Δ E3腺病毒血清型5 (Ad5)病毒，該病毒包含SEQ ID NO:3之核苷酸序列。

【請求項66】

如請求項57至65中任一項之方法，其中該宿主細胞包含腺病毒E1A及E1B基因，較佳地該宿主細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6細胞。

【請求項67】

如請求項57至66中任一項之方法，其中用於使該第二宿主細胞生長之該等條件包含用2-胺基嘌呤培養該第二細胞。

【請求項68】

如請求項67之方法，其中該2-胺基嘌呤濃度係小於約1.25 mM。

【請求項69】

如請求項67或68之方法，其中該2-胺基嘌呤濃度係約1 μ M至約1.25 mM。

【請求項70】

如請求項67或68之方法，其中該2-胺基嘌呤濃度係約10 μ M至約1.25 mM。

【請求項71】

如請求項67或68之方法，其中該2-胺基嘌呤濃度係約100 μM 至約1.25 mM。

【請求項72】

如請求項67或68之方法，其中該2-胺基嘌呤濃度係約1.25 mM。

【請求項73】

如請求項67至72中任一項之方法，其中用2-胺基嘌呤培養該第二細胞係在用重組腺病毒感染該第一宿主細胞後約24小時起始。

【請求項74】

一種組成物，其包含如請求項55之細胞、及2-胺基嘌呤。

【請求項75】

如請求項74之組成物，其中該2-胺基嘌呤濃度係小於約1.25 mM。

【請求項76】

如請求項74之組成物，其中該2-胺基嘌呤濃度係約1 μM 至約1.25 mM。

【請求項77】

如請求項74之組成物，其中該2-胺基嘌呤濃度為約10 μM 至約1.25 mM。

【請求項78】

如請求項74之組成物，其中該2-胺基嘌呤濃度為約100 μM 至約1.25 mM。

【請求項79】

如請求項74之組成物，其中該2-胺基嘌呤濃度係約1.25 mM。

【請求項80】

一種非天然存在的核酸分子，其包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。

【請求項81】

如請求項74之非天然存在的核酸分子，其包含與SEQ ID NO:3之核苷酸序列具有至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核苷酸序列。

【請求項82】

一種載體，其包含如請求項80或81之非天然存在的核酸。

【請求項83】

如請求項82之載體，其進一步包含啟動子，較佳的是可操作地連接至編碼該絲胺酸重組酶之核苷酸序列的巨細胞病毒(CMV)啟動子。

【請求項84】

如請求項82或83之載體，其進一步包含多腺苷酸化信號，諸如猿猴病毒40 (SV40)多腺苷酸化信號，其可操作地連接至編碼該絲胺酸重組酶的該核苷酸序列。

【請求項85】

如請求項82至84中任一項之載體，其係DNA質體。

【請求項86】

如請求項82至85中任一項之載體，其係重組腺病毒載體。

【請求項87】

如請求項86之載體，其係重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列在CMV啟動子的控制下編碼具有SEQ ID NO:2之胺基酸序列的絲胺酸重組酶，其中該核苷酸序列係進一步可操作地連接至SV40多腺苷酸化信號（NC_001669.1，nt 2550至2774）。

【請求項88】

一種細胞，其包含非天然存在的核酸分子，該核酸分子包含編碼絲胺酸重組酶之核苷酸序列，該絲胺酸重組酶具有與SEQ ID NO:2之胺基酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的胺基酸序列。

【請求項89】

如請求項88之細胞，其包含與SEQ ID NO: 3之核苷酸序列至少85%、諸如至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的核苷酸序列。

【請求項90】

一種細胞，其包含載體，該載體包含如請求項88或89之非天然存在的核酸。

【請求項91】

如請求項90之細胞，其進一步包含啟動子，較佳的是可操作地連接至編碼該絲胺酸重組酶之該核苷酸序列的巨細胞病毒(CMV)啟動子。

【請求項92】

如請求項90或91之細胞，其進一步包含多腺苷酸化信號，諸如猿猴病毒40 (SV40)多腺苷酸化信號，其可操作地連接至編碼該絲胺酸重組酶的該核苷酸序列。

【請求項93】

如請求項90至92之細胞，其中該載體係DNA質體。

【請求項94】

如請求項90至93中任一項之細胞，其中該載體係重組腺病毒載體。

【請求項95】

如請求項94之細胞，其中該重組腺病毒載體係重組 $\Delta E1/\Delta E3$ 腺病毒血清型5 (Ad5)病毒，該病毒包含核苷酸序列，該核苷酸序列在CMV啟動子的控制下編碼具有SEQ ID NO:2之胺基酸序列的絲胺酸重組酶，其中該核苷酸序列係進一步可操作地連接至SV40多腺苷酸化信號 (NC_001669.1，nt 2550至2774)。

【請求項96】

如請求項88至95之細胞，其包含腺病毒E1A及E1B基因，較佳地該細胞係911細胞、pTG6559細胞、GH329細胞、N52.E6細胞、HeLa-E1細胞、UR細胞、VLI-293細胞、HEK293細胞、或PER.C6細胞。

【請求項97】

一種在細胞中進行位點特異性重組的方法，其包含：

(a) 獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attP$ 位點（較佳的是具有SEQ ID

NO:7之核苷酸序列的 $attP$ 位點) 及與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attB$ 位點(較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 $attB$ 位點)；

(b) 將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；及

(c) 在使該絲胺酸重組酶催化在該等 $attP$ 與 $attB$ 位點之間該位點特異性重組的條件下使該細胞生長。

【請求項98】

一種藉由在細胞中進行位點特異性重組之程序所生產的產物，該程序包含：

(a) 獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attP$ 位點(較佳的是具有SEQ ID NO:7之核苷酸序列的 $attP$ 位點) 及與SEQ ID NO:8或SEQ ID NO:9至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 $attB$ 位點(較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 $attB$ 位點)；

(b) 將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同

一性的絲胺酸重組酶；及

(c) 在使該絲胺酸重組酶催化在該等 *attP* 與 *attB* 位點之間該位點特異性重組的條件下使該細胞生長。

【請求項99】

一種用於從細胞中獲得產物之程序，其包含：

(a) 獲得包含核酸分子的細胞，該核酸分子具有與SEQ ID NO:7 至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 *attP* 位點（較佳的是具有SEQ ID NO:7之核苷酸序列的 *attP* 位點）及與SEQ ID NO:8或SEQ ID NO:9 至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一的核苷酸序列的 *attB* 位點（較佳的是具有SEQ ID NO:8或SEQ ID NO:9之核苷酸序列的 *attB* 位點）；

(b) 將非天然存在的核酸分子引入至該細胞中，該核酸分子編碼具有與SEQ ID NO:2 至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的絲胺酸重組酶；

(c) 在使該絲胺酸重組酶催化在該等 *attP* 與 *attB* 位點之間該位點特異性重組的條件下生長該細胞；及

(d) 從該細胞中生產並回收產物。

【請求項100】

一種非天然存在的系統，其包含：

用於經AAV介導之重組之手段，其中該手段可選地包含轉殖基因因子。

【請求項101】

一種用於轉移如請求項100之非天然存在的系統之手段。

【請求項102】

一種非天然存在的系統，其包含：

用於重組如請求項100之系統的重組手段，其中該重組手段包括在催化期間使用至少一種絲胺酸殘基。

【請求項103】

一種用於轉移如請求項102之非天然存在的系統之手段。

【請求項104】

一種用於製造分子之手段，其中該用於製造分子之手段包含如請求項100至103中任一項之手段並且能夠複製。

【請求項105】

一種用於AAV介導之位點特異性重組之程序，其包含：

(a) 用於執行獲得細胞的功能之步驟，其包含如請求項100之手段；

(b) 在催化期間使用至少一種絲胺酸殘基以使位點特異性重組的條件下，用於執行使該細胞生長的功能之步驟。

【請求項106】

如請求項105之用於AAV介導之位點特異性重組之程序，其包含獲得產物，其中可選地該產物係治療性產物。

【發明圖式】

得分	預期	方法	識別	陽性	間隙
697位元(1798)	0.0	組成基質調整	341/531(64%)	415/531(78%)	2/531(0%)

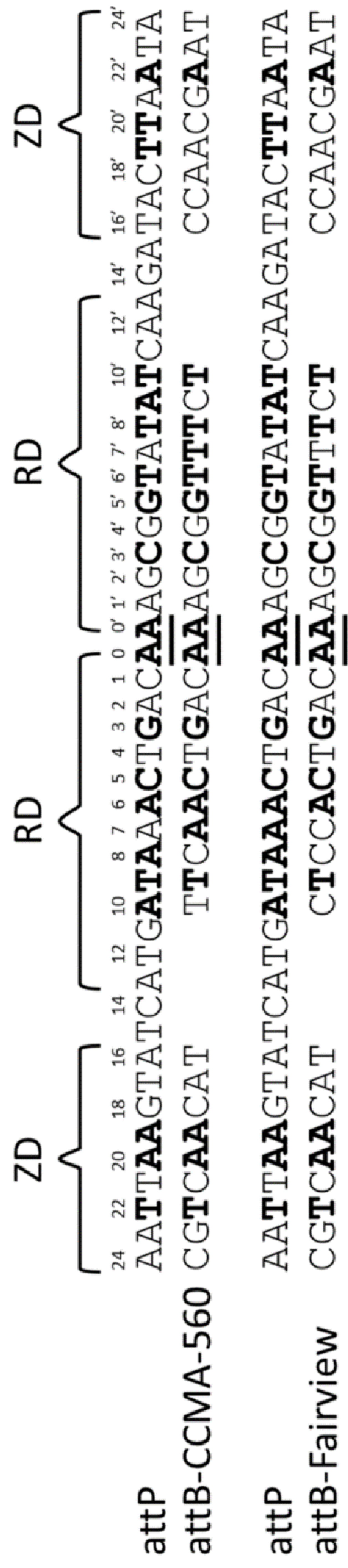
Query	1	MELKNIVNSYNITNILGYLRRSRQDMEREKRTGEDTLTEQKELMNKILTAEIIPYELKME	60
Sbjct	1	MELK+IVNSYN+T I+GYLRRSRQD+EREKRTGEDTLTEQKELMNKILT IEIPYE + E	60
Query	61	IGSGESIDGRPVFKECLKDLEEGKYQAIHAVKEITRLSRGSYSYDAGQIVNLLQSKRLIIIT	120
Sbjct	61	IGSGESI+GRPVPK CL DL GK+QAIHAVKEITRLSRGSYSYDAG+IVNLL KR+IIIT	120
Query	121	PYKVYDPRNPVDMRQIRFELFMAREEFEMTRERMTGAKYTYAAQGWISGLAPYGYQLNK	180
Sbjct	121	PYK+YDPRNPVD RQIRFELFMAREEFEMTRERM GAK+TYAAQGWISGLAP+GY+LNK	180
Query	181	KTSKLDPVEDEAKVVQLIFNIFLNGLNGKDYSYTAIASHLTNLQIPTPSGKKRWNQYTIK	240
Sbjct	181	+TS+L+P +++ VV+LIF+IFLNGL+GKD SYTAIA+HL+ LQ TP G KRW++ T++	240
Query	241	AILQNEVYIGTVKYKVREKTKDGKRTIRPEKEQIVVQDAHAPIIDKEQFQQSQVKIANKV	300
Sbjct	241	ILQNE Y+G V+YK RE TKDGK+ RPE E IVV DAH PII+KE F+ Q KI NKV	300
Query	301	PLLPNKDEFELSELAGVCTCSKCGEPLSKYESKRIRKNKDGTESVYHVKSLTCKKNKCTY	360
Sbjct	301	PLLP +E +ELAG+C CS CG+ L K+ES+ RKNKDGT S +HVK L CK NKCT	360
Query	361	VRYNDVENAILDYLSSLNDLNDSTLTKHINSMLSKYEDDNSNMKTKKQMSEHLSQKEKEL	420
Sbjct	361	VRY VE AIL+YL L L ++ L I + E +NS KT +QM +QK+KEL	419
Query	421	KNKENFIFDKYESGIYSDELFLKRKAALDEEFKELQNAKNELNGLQDTQSEIDSNTVRNN	480
Sbjct	420	+NK FIF+K+ESGIY+DE+FL+RKAA+++E +++ K EL+ + + E D N R N	478
Query	481	INKIIDQYHIESSEKKNELLRMVLDVIVNMTQKRKGPIPAQFEITPILR	531
Sbjct	479	I+ ++ Y S KNE LR + +++ MT+KR+GPIPA+F I P+LR	529

【圖1】

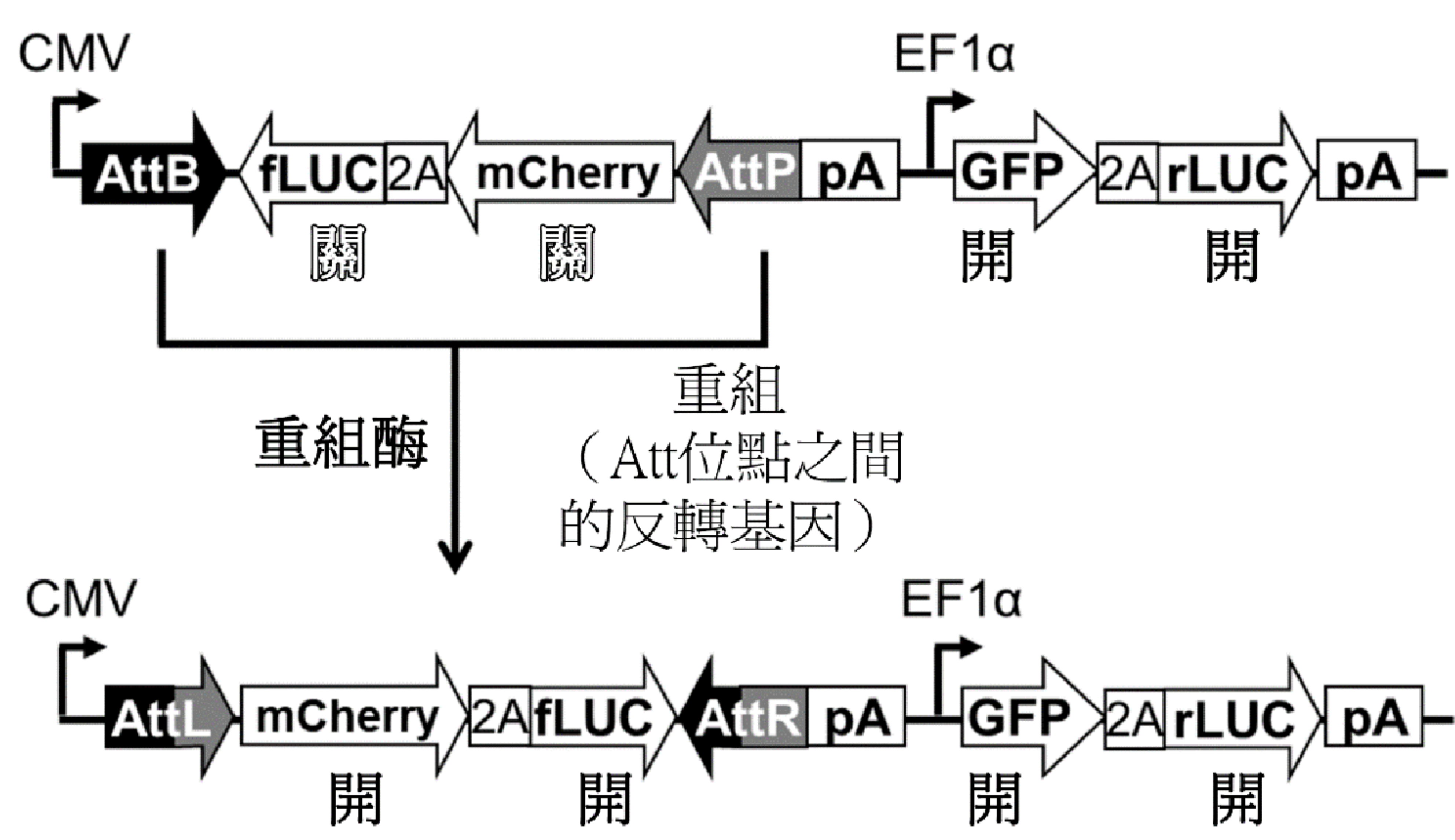
得分	預期	識別	間隙	股
874 位元(473)	0.0	483/488(99%)	0/488(0%)	正/負

Query	113	ACTGACAAAGCGGTTTCTCCAACGAATGCCATGGGTGCAACAAAATTGATTTTCAGAAAAA	172
Sbjct	464839	ACTGACAAAGCGGTTTCTCCAACGAATGCCATGGGTGCAACAAAATTGATTTTCAGAAAAA	464780
Query	173	CTATTTTTCGAAGCAAACGAAAGTATTCCGAATCAAAAAACGAGGTTTGCTCTGTACGC	232
Sbjct	464779	CTATTTTTCGAAGCAAACGAAAGTATTCCAAATAAAAAACCAAGTTTGCTCTGTACGC	464720
Query	233	TTTGGCAATGTGCTTGGATCTAGAGGTTCCGTTATTCCGATCATGCTCCAGCAGCTATTA	292
Sbjct	464719	TTTGGCAATGTGCTTGGATCTAGAGGTTCCGTGATTCCGATCATGCTCCAGCAGCTATTA	464660
Query	293	AATGAAAAACCTTTGACCGTGACTGATCCTCATATGACACGtttttttATGTCCATTGAA	352
Sbjct	464659	AATGAAAAACCTTTGACCGTGACTGATCCTCATATGACACGTTTTTTTATGTCCATTGAA	464600
Query	353	GAGGCTGTTTCCCTCACACTTCAAGCAGCAATCATGATGAAAGGCGGCGAAACCTTCATT	412
Sbjct	464599	GAGGCTGTTTCCCTCACACTTCAAGCAGCAATCATGATGAAAGGCGGCGAAACCTTCATT	464540
Query	413	CTCAAGATGGAGTCCTTACAGCTTGCCGATCTCCTAAAAGCGTTTCATGAATATGCCGCT	472
Sbjct	464539	CTCAAGATGGAGTCCTTACAGCTTGCCGATCTCCTAAAAGCGTTTCATGAATATGCCGCT	464480
Query	473	CAAATCAATGCTAAATCTCCGGATATTCTTGTAAGTCGGAAAAAGACCTGGCGAAAAGCTT	532
Sbjct	464479	CAAATCAATGCTAAATCTCCGGATATTCTTGTAAGTCGGAAAAAGACCTGGCGAAAAGCTT	464420
Query	533	CACGAGGAGCTCACATTTCCGCACGAAGCAGATGCACTGTTTGAACATGAACAATTTTAT	592
Sbjct	464419	CACGAGGAGCTCACATTTCCGCACGAAGCAGATGCACTGTTTGAACATGAACAATTTTAT	464360
Query	593	GCCATTTT 600	
Sbjct	464359	GCCATTTT 464352	

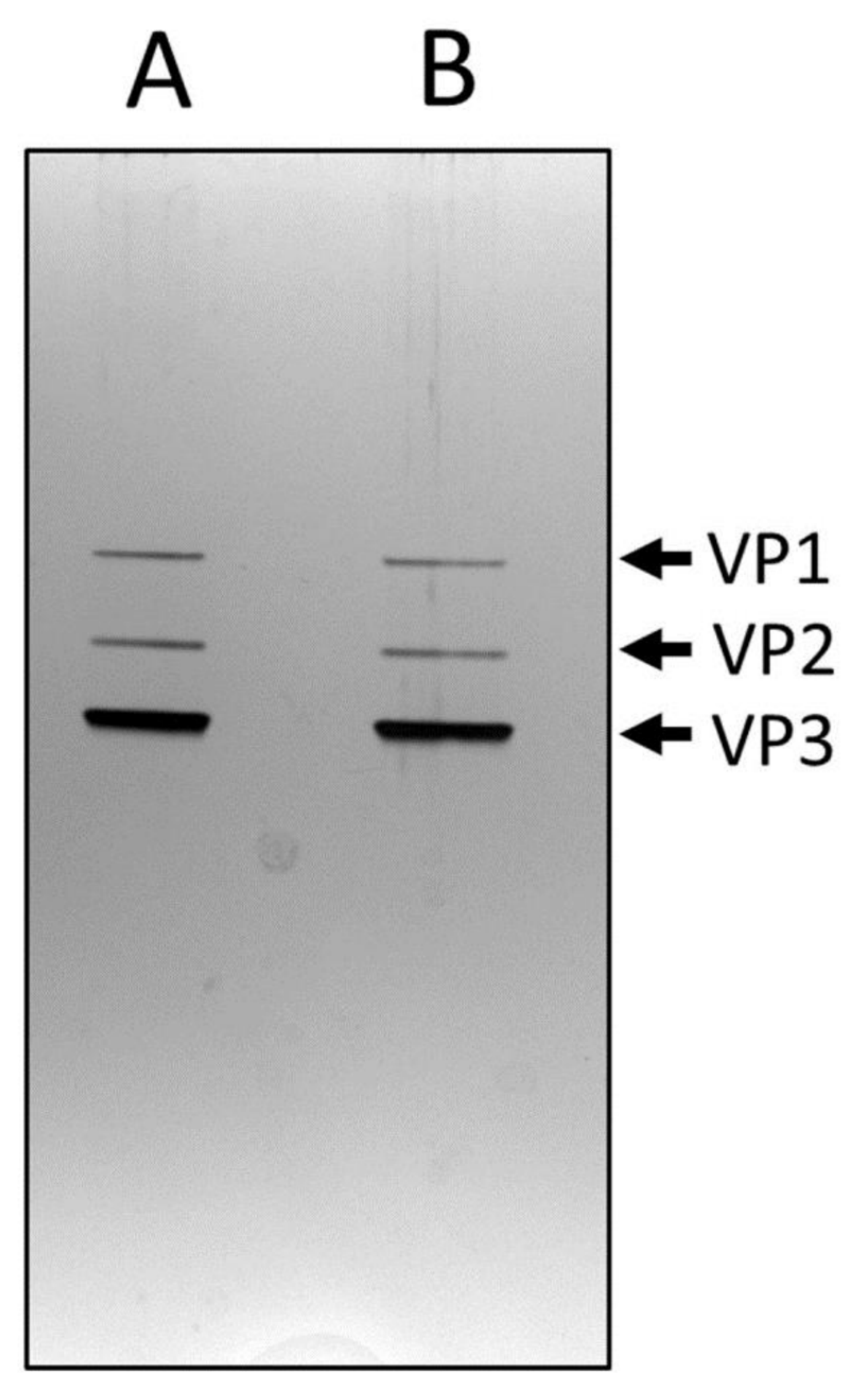
【圖2】



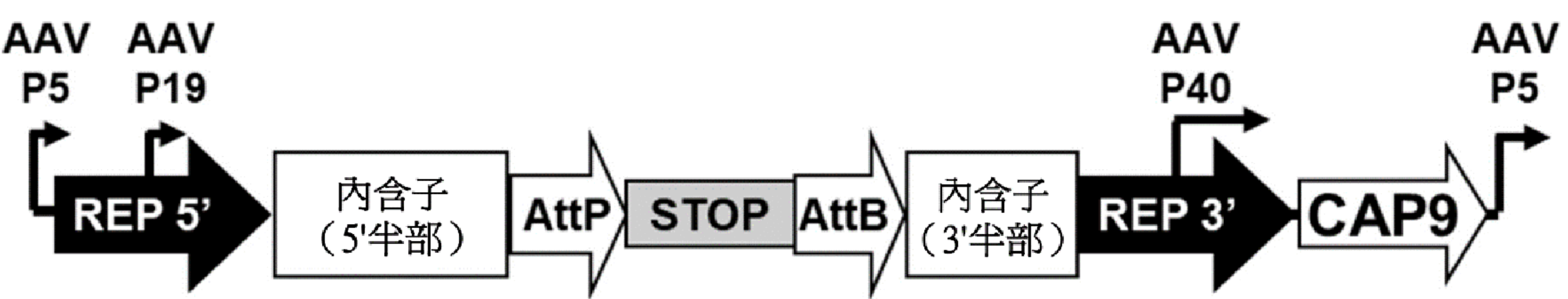
【圖3】



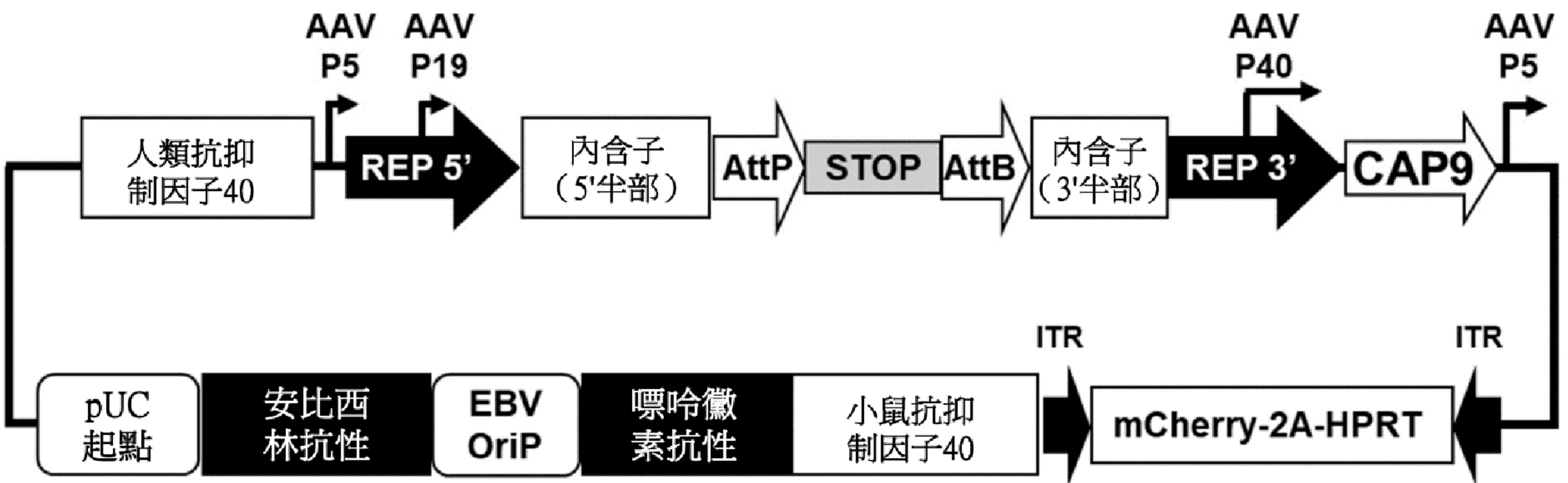
【圖4】



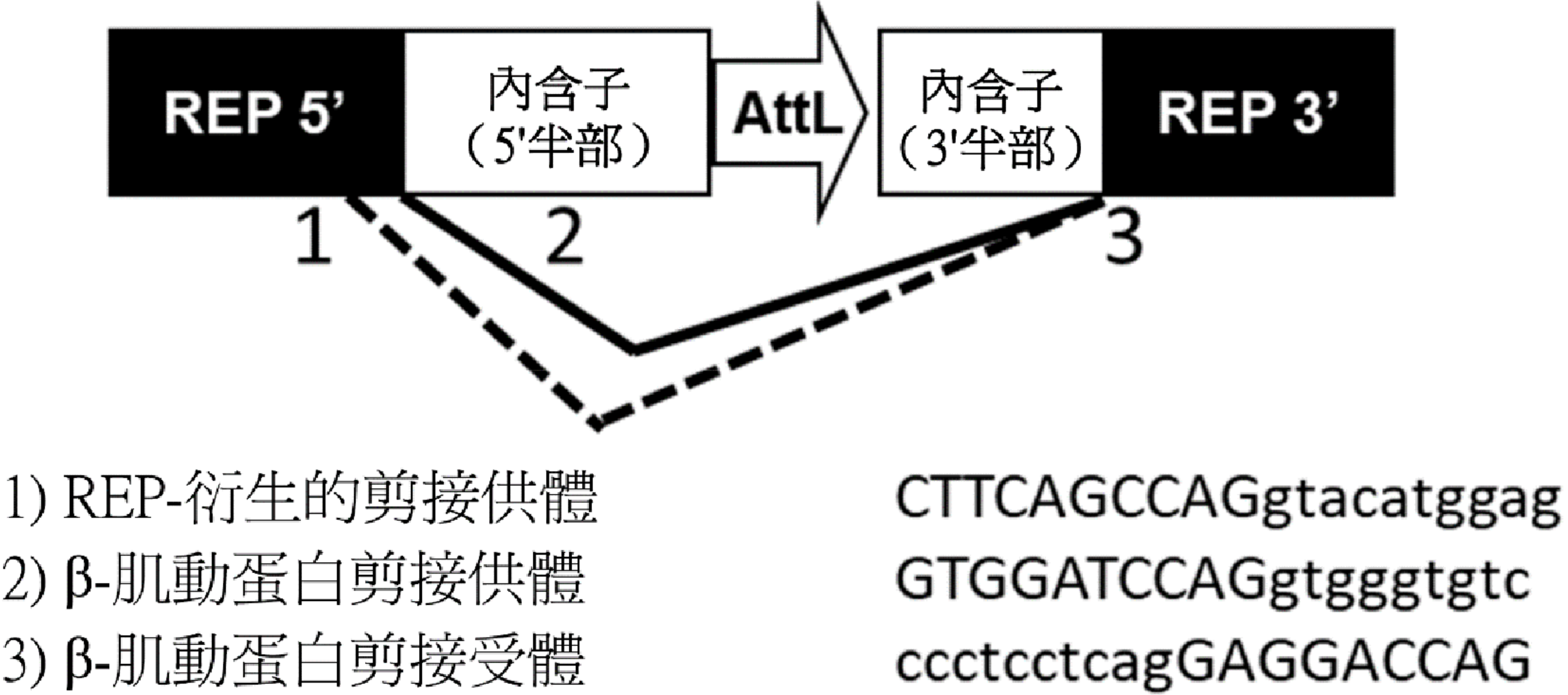
【圖5】



【圖6】

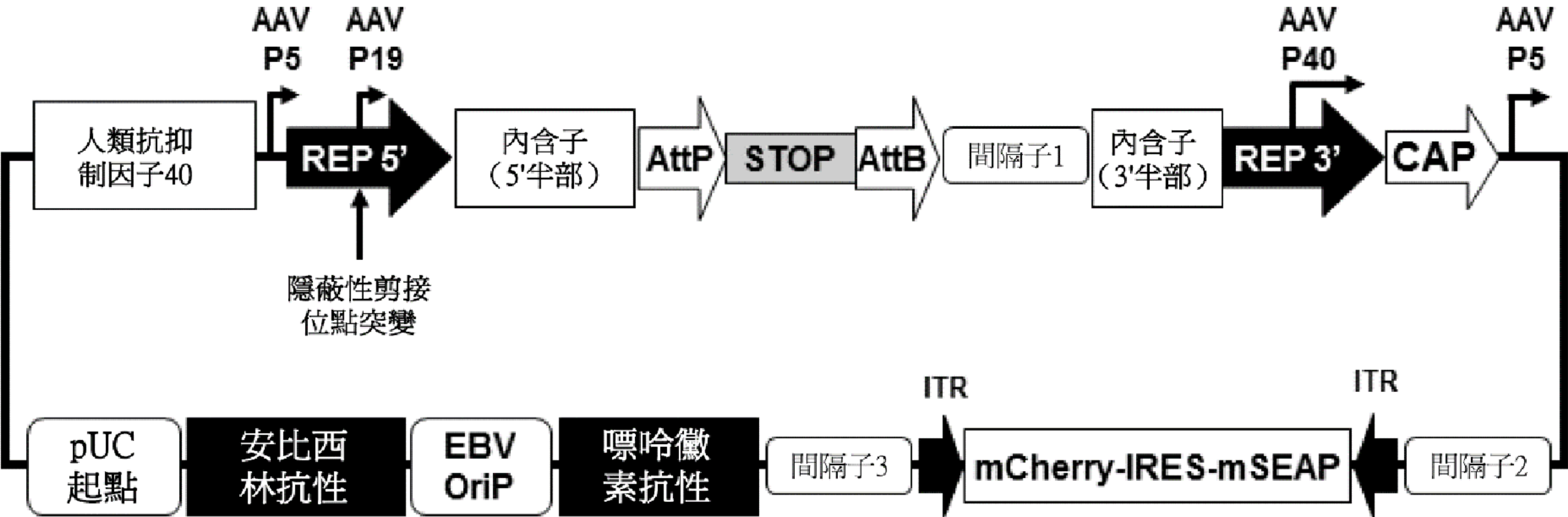


【圖7】



【圖8】

P600



【圖9】