I716367

(19)中華民國智慧財產局

(12)發明說明書公告本　**(11)證書號數：TW I716367 B**

(45)公告日： 中華民國 110 (2021) 年 01 月 21 日

(21)申請案號：104135960　(22)申請日： 中華民國 104 (2015) 年 10 月 30 日

(51)Int. Cl.：　*C12N15/10　(2006.01)*　*C12N15/11　(2006.01)*

A61K45/06　(2006.01)　*A61P35/00　(2006.01)*

(30)優先權：2014/10/31　美國　62/073126

2015/05/26　美國　62/166302

(71)申請人：麻省理工學院 (美國) MASSACHUSETTS INSTITUTE OF TECHNOLOGY　(US)

美國

(72)發明人：盧　蒂莫西　冠　塔 LU, TIMOTHY KUAN-TA (US)；黃　兆麟 WONG, ALAN SIU LUN (HK)；蔡　正姿 CHOI, CHING GEE (HK)

(74)代理人：鄭志玲

(56)參考文獻：

WO　2013176772A1　WO　2014005042A2

WO　2014093694A1

審查人員：張維纓

申請專利範圍項數：22 項　圖式數：22　共 220 頁

(54)名稱

用於常間回文重複序列叢集（ＣＲＩＳＰＲ）之大量平行組合性基因學

(57)摘要

本文所描述的為可以快速產生高序基因元件組合的方法及組成物，其包含 CRISPR 引導序列及支架序列，以及可快速鑑定單一細胞或匯集的細胞群體內所編碼的基因元件的條碼。本文亦描述表觀遺傳基因的抑制劑組成物，以及用於減少細胞增生及/或治療癌症的方法。

Described herein are methods and compositions that enable rapid generation of high-order combinations of genetic elements comprising a CRISPR guide sequence and a scaffold sequence, and a barcode for rapid identification of the combination of genetic elements encoded within a single cell or a pooled population. Also described herein compositions of inhibitors of epigenetic genes and methods for reducing cell proliferation and/or treating cancer.

指定代表圖：

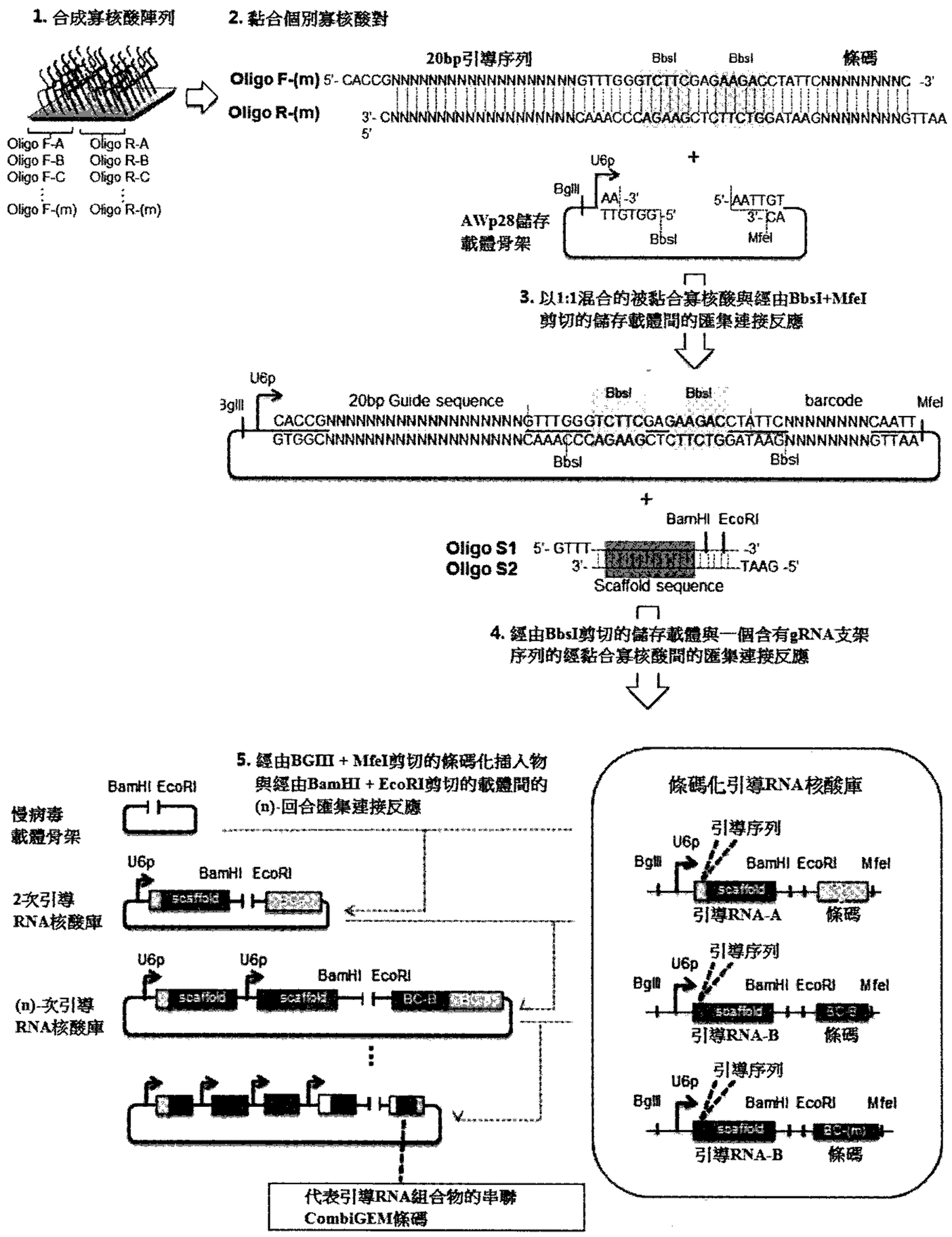

1. 合成寡核酸陣列
2. 黏合個別寡核酸對
Oligo F-A
Oligo F-B
Oligo F-C
Oligo F-(m)
Oligo R-A
Oligo R-B
Oligo R-C
Oligo R-(m)
20bp引導序列
BbsI
BbsI
條碼
Oligo F-(m) 5'- CACCGNNNNNNNNNNNNNNNNNNNNGTTTGGGTCTTCGAGAAGACCTATTCNNNNNNNNC -3'
Oligo R-(m) 3'- CNNNNNNNNNNNNNNNNNNNCAAACCCAGAAGCTCTTCTGGATAAGNNNNNNNNGTTAA 5'
U6p
BglII
AA -3'
TTGTGG-5'
5'-AATTGT
3'- CA
AWp28儲存載體骨架
BbsI
MfeI
3. 以1:1混合的被黏合寡核酸與經由BbsI+MfeI剪切的儲存載體間的匯集連接反應
U6p
3gIII
20bp Guide sequence
BbsI
BbsI
barcode
MfeI
CACCGNNNNNNNNNNNNNNNNNNNNGTTTGGGTCTTCGAGAAGACCTATTCNNNNNNNNCAATT
GTGGCNNNNNNNNNNNNNNNNNNNNCAAACCCAGAAGCTCTTCTGGATAAGNNNNNNNNGTTAA
BbsI
BbsI
BamHI EcoRI
Oligo S1 5'- GTTT -3'
Oligo S2 3'- TAAG -5'
Scaffold sequence
4. 經由BbsI剪切的儲存載體與一個含有gRNA支架序列的經黏合寡核酸間的匯集連接反應
5. 經由BGIII + MfeI剪切的條碼化插入物與經由BamHI + EcoRI剪切的載體間的(n)-回合匯集連接反應
慢病毒載體骨架
BamHI EcoRI
2次引導RNA核酸庫
U6p
BamHI EcoRI
scaffold
(n)-次引導RNA核酸庫
U6p
U6p
BamHI EcoRI
scaffold
scaffold
代表引導RNA組合物的串聯CombiGEM條碼
條碼化引導RNA核酸庫
引導序列
U6p
BglII
BamHI EcoRI MfeI
scaffold
引導RNA-A
條碼
引導序列
U6p
BglII
BamHI EcoRI MfeI
scaffold
引導RNA-B
條碼
引導序列
U6p
BglII
BamHI EcoRI MfeI
scaffold
引導RNA-B
條碼

圖 1

I716367

公告本

發明摘要

※ 申請案號：104135960

※ 申請日： 104年10月30日 ※IPC 分類：*C12N 15/10* (2006.01)
C12N 15/11 (2006.01)
A61K 45/06 (2006.01)
A61P 35/00 (2006.01)

【發明名稱】（中文/英文）用於常間回文重複序列叢集(CRISPR)之大量平行組合性基因學/MASSIVELY PARALLEL COMBINATORIAL GENETICS FOR CRISPR

【中文】本文所描述的為可以快速產生高序基因元件組合的方法及組成物，其包含CRISPR引導序列及支架序列，以及可快速鑑定單一細胞或匯集的細胞群體內所編碼的基因元件的條碼。本文亦描述表觀遺傳基因的抑制劑組成物，以及用於減少細胞增生及/或治療癌症的方法。

【英文】

Described herein are methods and compositions that enable rapid generation of high-order combinations of genetic elements comprising a CRISPR guide sequence and a scaffold sequence, and a barcode for rapid identification of the combination of genetic elements encoded within a single cell or a pooled population. Also described herein compositions of inhibitors of epigenetic genes and methods for reducing cell proliferation and/or treating cancer.

【代表圖】

【本案指定代表圖】：第（ 1 ）圖。

【本代表圖之符號簡單說明】：無

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

發明專利說明書

（本說明書格式、順序，請勿任意更動）

【發明名稱】（中文/英文）

用於常間回文重複序列叢集(CRISPR)之大量平行組合性基因學

/MASSIVELY PARALLEL COMBINATORIAL GENETICS FOR CRISPR

【相關申請案】本專利申請主張根據35 U.S.C. § 119(e)而於2014年10月31日提交的第62/073,126號美國臨時專利申請以及2015年5月26日提交的第62/166,302號美國臨時專利申請的優先權，其全部內容藉由引用而併入本文中。

【政府資金】本發明係根據美國國家衛生研究院所核准的計畫案號OD008435由政府出資而申請。政府對本發明具有一定的權利。

【技術領域】

【0001】 本發明關於用於快速產生包含CRISPR 引導序列及支架序列的高序基因元件組合物以及鑑定該基因元件的方法及組成物。本發明亦關於靶向表觀遺傳基因而可抑制細胞增生的抑制劑組成物及相關方法。

【先前技術】

【0002】 叢集且有規律間隔的短回文重複序列(CRISPR)/Cas系統最初發現於細菌和古細菌物種中做為對抗外源遺傳物質(例如：質體及噬菌體)的一種防禦機制。天然存在的CRISPR / Cas系統依靠三個成分的表現：互補於標的序列的引導RNA序列、協助吸引第三成分，核酸內切酶，到適當位點的支架RNA。雖然在許多細菌和古細菌物種中，CRISPR / Cas系統是用來使外源遺傳物質降解，該系統已適用於各種原核和真核生物，

並已用於許多方法中，包括基因敲除、突變以及基因表現激活或抑制(Hsu, et al. *Cell* (2014) 157(6):1262-1278)。在經由基因工程處理的CRISPR / CAS系統中，三個獨立成分的需求可藉由表現一個小引導RNA(sgRNA)而規避，該sgRNA含有用於與標的序列結合的CRISPR引導RNA序列以及支架RNA兩者，且一起模擬由個別的引導RNA序列以及支架序列形成的結構，並足以吸引核酸內切酶至適當的標的位點上(Jinek, et al. *Science* (2012) 337(6096):816-821)。

【發明內容】

【0003】 產生用於表現包含多於sgRNA的多重CRISPR系統的載體及基因元件是很費力的，且採用傳統選殖方法建構CRISPR系統庫的複雜性非常有限。本文所述方法可產生包含多重sgRNA的載體，其各個皆包含一個CRISPR 引導序列及一個支架序列，以及可被偵測且可作為該CRISPR引導序列特性指標的串聯條碼元件。該方法亦可簡易且快速地產生載體的高度複合基因庫。

【0004】 本發明一方面提供包含一第一DNA元件的基因構築體，該第一DNA元件包含一CRISPR引導序列和支架序列；在該第一DNA元件兩側的一第一相容端元件及一第二相容端元件，其中該第一及第二相容端元件能夠互相黏合；一條碼元件；在該條碼元件兩側的一第三相容端元件及一第四相容端元件，其中該第三及第四相容端元件能夠互相黏合，但不能與該第一或第二相容端元件黏合；以及一位於該第四相容端元件及第一相容端元件之間的分隔位點，其中該DNA元件、第一相容端元件及第二相容端元件係在該分隔位點一側，而該條碼元件、第三相容端元件及第四相容端

元件則位於該分隔位點的另一側。

【0005】 在某些實施例中，基因構築體進一步包含一個位於該第一DNA元件上游的啟動子元件。

【0006】 一些方面提供包含本文所提供的基因構築體中之任一個的載體。

【0007】 其他方面提供基因構築體，其包含複數個DNA元件，其中該複數個DNA元件中的各個DNA元件包含一CRISPR引導序列及一支架序列；在該複數個DNA元件兩側的一第一相容端元件和一第二相容端元件，其中該第一及第二相容端元件能夠互相黏合；複數個條碼元件；在該複數個條碼元件兩側的一第三相容端元件和一第四相容端元件，其中該第三和第四相容元件能夠互相黏合，但不會與該第一或第二相容端元件黏合；以及一個位於該複數個DNA元件及該複數個條碼元件之間的分隔位點。

【0008】 其他方面提供包含本文所提供的任何基因構築體以及位於各個該CRISPR引導序列上游的啟動子序列的載體。

【0009】 另外方面提供用於產生組合性載體的方法，其包含(a)提供一種含有一種包含CRISPR引導序列的第一基因構築體；一個第二相容端元件及一個在該CRISPR引導序列側邊且針對一個第一限制酶的第一識別位點；一個條碼元件；以及一個第三相容端及一個在該條碼元件側邊且針對一個第二限制酶的第二識別位點；(b)在該第一識別位點上剪切該第一基因構築體，而產生一個第五相容端元件，且在該第二識別位點上剪切該載體而產生一個第六相容端元件；(c)提供一種支架元件，其包含一個支架序列；一個包含一個第一相容端元件及一個第四相容端元件的分隔位點；在該支

架元件兩側的一個第七相容端元件及一個第八相容端元件，其中該第七相容端元件能夠黏合於該第五相容端元件，而該第八相容端元件則能夠黏合於該第六相容端元件；以及(d)將該支架元件黏合於該經剪切的第一基因構築體，其中該黏合係發生在該載體和該支架元件內能夠互相黏合的相容端元件處，且其中，在黏合後，該支架元件係被整合於該CRISPR 引導序列和條碼元件之間，且其中，該分隔位點係位於該支架序列和該條碼元件之間，進而創建出一個組合性載體。

【0010】 在某些實施例中，該方法進一步包含(a)提供本文所述的任何組合性載體；(b)在該支架元件內的分隔位點上剪切該載體，而產生一第一相容端元件及一個第四相容端元件；(c)提供一第二基因構築體，其包含一CRISPR 引導序列；一支架序列；一條碼元件；以及在該第二基因構築體兩側的一第二相容端元件和一第三相容端元件，其中該第二基因構築體的第二相容端元件能夠與該載體的第一相容端元件黏合，而該第二基因構築體的第三相容端元件則能夠與該載體的第四相容端元件黏合；以及(d)使該第二基因構築體黏合於該經剪切的載體上，其中該黏合係發生在該第二基因構築體與載體內且能夠互相黏合的相容端元件處，且其中，在黏合後，該第二基因構築體係被整合於該載體，進而創建出一個包含串聯的條碼元件和串聯的CRISPR引導及支架序列的組合性載體。

【0011】 在某些實施例中，該組合性載體進一步包含一或多個位於該CRISPR 引導序列上游的啟動子。在某些實施例中，該方法是迭代的。在某些實施例中，該第一識別位點及該第二識別位點具有相同的識別位點序列，且該第一限制酶及該第二限制酶為相同的限制酶。

【0012】 本發明其他方面提供基因構築體，其包含至少兩個CRISPR引導序列；一條碼序列；以及一介於各個CRISPR 引導序列之間，且介於該條碼元件和該CRISPR 引導序列之間最靠近該條碼元件的限制酶識別位點。

【0013】 其他方面提供基因構築體，其包含複數個DNA元件，各包含一CRISPR 引導序列以及一支架序列；一條碼元件；以及一位於該複數個DNA元件的各DNA元件上游的啟動子序列。在某些實施例中，該條碼序列係位於該基因構築體的5’端。在某些實施例中該條碼序列係位於該基因構築體的3’端。

【0014】 其他方面提供包含本文所述任何基因構築體的載體。

【0015】 又其他方面提供用於產生一種組合性載體的方法，其包含(a)提供一種載體，其包含：複數個CRISPR 引導序列；一條碼元件，其中該條碼元件係位於該複數個CRISPR 引導序列的下游；視情況可選擇的一位於該複數個CRISPR 引導序列中至少一個的上游的啟動子序列；以及複數個針對複數個限制酶的識別位點，其中各個該複數個識別位點係位於該複數個CRISPR 引導序列中之一個的下游；(b)以至少一個的該複數個限制酶在至少一個的該複數個識別位點上剪切該載體，而產生一第一相容端元件及一第二相容端元件；(c)提供一第一支架元件，其包含：一支架序列、視情況可選擇的一啟動子序列、以及在該第一支架元件兩側的一第三相容端元件及一第四相容端元件，其中該第三相容端元件能夠黏合於該經剪切載體的第一相容端元件，且該第四相容端元件能夠黏合於該經剪切載體的第二相容端元件；以及(d)使該第一支架元件黏合於該經剪切的載體，其中

該黏合係發生在該第一支架元件與經剪切載體內的相容端元件處，且其中，在黏合後，該第一支架元件係被整合於該複數個CRISPR引導序列中之一個的下游，因而產生一個組合性載體。在某些實施例中，該方法為反覆進行的。

【0016】 其他方面提供用於產生一種組合性載體的方法，其包含(a)提供一載體，其包含：複數個CRISPR 引導序列、一條碼元件，其中該條碼元件係位於該複數個CRISPR 引導序的上游；視情況可選擇的一位於該複數個CRIPSR引導序列中至少一個的上游的啟動子序列；以及複數個針對複數個限制酶的識別位點，其中各個該複數個識別位點係位於該複數個CRISPR引導序列中之一個的上游；(b)以至少一個的該複數個限制酶在至少一個的該複數個識別位點上剪切該載體，而產生一第一相容端元件及一第二相容端元件；(c)提供一第一支架元件，其包含視情況可選擇的一支架序列、一啟動子序列、以及在該第一支架元件兩側的一第三相容端元件及一第四相容端元件，其中該第三相容端元件能夠黏合於該經剪切載體的第一相容端元件，且該第四相容端元件能夠黏合於該經剪切載體的第二相容端元件；(d)使該第一支架元件黏合於該經剪切的載體，其中該黏合係發生在該第一支架元件與經剪切載體內的相容端元件處，且其中，在黏合後，該第一支架元件係被整合於該複數個CRISPR引導序列中之一個的上游，因而產生一個組合性載體。在某些實施例中，該方法為迭代的。

【0017】 本發明的多個方面提供包含二或多個靶向二或多個選自表2中所述之表觀遺傳基因組合物的表觀遺傳基因的抑制劑的組成物。在某些實施例中，各個該二或多個抑制劑可減少或防止一表觀遺傳基因的表現，

或者減少或防止由該表觀遺傳基因所編碼的一種蛋白質的活性。在某些實施例中，各個該抑制劑係選自由CRISPR 引導序列及支架序列；shRNA；以及小分子所組成的群組中。在某些實施例中，至少一個該抑制劑為CRISPR 引導序列及支架序列；且該組成物進一步包含或者編碼一個Cas9核酸內切酶。在某些實施例中，該CRISPR 引導序列或shRNA係由一個重組表現載體表現而得。在某些實施例中，該表觀遺傳基因的組合物包含BRD4和KDM4C或BRD4和KDM6B。在某些實施例中，BRD4的抑制劑為JQ1((6S)-4-(4-氯苯基)-2,3,9-三甲基-6H-噻吩並[3,2-f][1,2,4]三唑並[4,3-a][1,4]二氮雜-6-乙酸1,1-二甲基乙基酯)。在某些實施例中，KDM4C的抑制劑為SD70(N-(呋喃-2-基(8-羥基喹啉-7-基)甲基)異丁醯胺)。在某些實施例中，KDM6B的抑制劑為GSK-J4 (3-((6-(4,5-二氫-1H-苯並[d]氮雜-3(2H)-基)-2-(吡啶-2-基)嘧啶-4-基)氨基)丙酸乙酯，單鹽酸鹽)。

【0018】 其他方面提供用於減少細胞增生的方法，其包含使該細胞與一種靶向二或多個選自表2中所述之表觀遺傳基因組合物的表觀遺傳基因的二或多個抑制劑組合物接觸。在某些實施例中，該細胞係癌細胞。在某些實施例中，該癌細胞係卵巢癌細胞。在某些實施例中，各個該二或多個抑制劑可減少或防止表觀遺傳基因的表現，或者減少或防止由該表觀遺傳基因編碼的一種蛋白質的活性。在某些實施例中，各個該抑制劑係選自由CRISPR 引導序列及支架序列；shRNA；以及小分子所組成的群組中。在某些實施例中，至少一個該抑制劑為CRISPR 引導序列及支架序列；且該組成物進一步包含或編碼Cas9核酸內切酶。在某些實施例中，該CRISPR 引導序列或shRNA係由重組表現載體表現而得。在某些實施例中，該表觀遺

傳基因的組合物包含BRD4和KDM4C或BRD4和KDM6B。在某些實施例中，BRD4的抑制劑為JQ1 ((6S)-4-(4-氯苯基)-2,3,9-三甲基-6H-噻吩並[3,2-f][1,2,4]三唑並[4,3-a][1,4]二氮雜-6-乙酸1,1-二甲基乙基酯)。在某些實施例中，KDM4C的抑制劑為SD70 (N-(呋喃-2-基(8-羥基喹啉-7-基)甲基)異丁醯胺)。在某些實施例中，KDM6B的抑制劑為GSK-J4 (3-((6-(4,5-二氫-1H-苯並[d]氮雜-3(2H)-基)-2-(吡啶-2-基)嘧啶-4-基)氨基)丙酸乙酯，單鹽酸鹽)。

【0019】 其他方面提供用於治療受試者中癌症的方法，其包含給藥予該受試者一種靶向二或多個選自表2中所述之表觀遺傳基因組合的表觀遺傳基因的二或多個抑制劑組合物，其中各個該二或多個抑制劑係以一個有效量進行給藥。在某些實施例中，各個該抑制劑係選自由CRISPR 引導序列、shRNA、以及小分子所組成的群組中。在某些實施例中，在該組合物中所施用的各個該二或多個抑制劑的有效量較在該組成物中不施用該抑制劑的有效量為少。在某些實施例中，各個該二或多個抑制劑可減少或防止一種表觀遺傳基因的表現，或者減少或防止由該表觀遺傳基因編碼的一種蛋白質的活性。在某些實施例中，各個該抑制劑係選自由CRISPR 引導序列及支架序列；shRNA；以及小分子所組成的群組中。在某些實施例中，至少一個該抑制劑為CRISPR 引導序列及支架序列；且該組成物進一步包含或編碼Cas9核酸內切酶。在某些實施例中，該CRISPR 引導序列或shRNA係由重組表現載體表現而得。在某些實施例中，該表觀遺傳基因的組合物包含BRD4和KDM4C或BRD4和KDM6B。在某些實施例中，BRD4的抑制劑為JQ1 ((6S)-4-(4-氯苯基)-2,3,9-三甲基-6H-噻吩並[3,2-f][1,2,4]三唑並[4,3-a][1,4]二氮雜-6-乙酸1,1-二甲基乙基酯)。在某些實施例中，KDM4C的

抑制劑為SD70 (N-(呋喃-2-基(8-羥基喹啉-7-基)甲基)異丁醯胺)。在某些實施例中，KDM6B的抑制劑為GSK-J4 (3-((6-(4,5-二氫-1H-苯並[d]氮雜-3(2H)-基)-2-(吡啶-2-基)嘧啶-4-基)氨基)丙酸乙酯，單鹽酸鹽)。

【0020】 其他方面提供用於鑑定可減少細胞增生的一種表觀遺傳基因抑制劑組合物的方法，其包含使一種第一群細胞和一種第二群細胞與二或多個CRISPR引導序列及支架序列以及一個Cas9核酸內切酶的複數個組合物接觸；培養該第一群細胞和第二群細胞，使得該第二群細胞的培養時程較該第一群細胞為長；鑑定該第一群細胞中的該二或多個CRISPR 引導序列及支架序列的組合物，與該第二群細胞中的該二或多個CRISPR 引導序列及支架序列的組合物；比較該第一群細胞中各個該二或多個CRISPR 引導序列及支架序列的組合物的豐度，與該第二群細胞中各個該二或多個CRISPR 引導序列及支架序列的組合物的豐度；以及將在該第二群細胞中缺乏或豐度降低但在該第一群細胞中存在或豐度增加的二或多個CRISPR 引導序列及支架序列的組合物鑑定為可減少細胞增生的CRISPR 引導序列及支架序列的組合物。

【0021】 又其他方面提供用於鑑定一種將被抑制而減少細胞增生的基因組合物的方法，其包含使一種第一群細胞和一種第二群細胞與二或多個CRISPR引導序列及支架序列以及一個Cas9核酸內切酶的多種組合物接觸；培養該第一群細胞和第二群細胞，使得該第二群細胞的培養時程較該第一群細胞為長；鑑定該第一群細胞中的該二或多個CRISPR 引導序列及支架序列的組合物，與該第二群細胞中的該二或多個CRISPR 引導序列及支架序列的組合物；比較該第一群細胞中各個該二或多個CRISPR 引導序

列及支架序列的組合物的豐度，與該第二群細胞中各個該二或多個CRISPR引導序列及支架序列的組合物的豐度；以及將在該第二群細胞中缺乏或豐度降低，但在該第一群細胞中存在或豐度增加的二或多個CRISPR 引導序列及支架序列的組合物鑑定為可減少增生的被抑制基因組合物。

【0022】 本發明的這些和其他方面，以及其各種實施例，將由引用本發明的附圖及詳細說明而變得明顯。

【0023】 本發明的各種限制，可以涵蓋本發明的各種實施例。因此，可預期的是，本發明中涉及任一元件或元件組合的各種限制可包括於本發明的各方面中。本發明不受其應用於以下說明書所述或附圖所說明的建構細節及元件安排所限制。本發明具有其他實施例且能以各種方式執行或實現。

【圖式簡單說明】

【0024】 附圖並非意欲按比例繪製。為了清楚起見，並非每個元件皆被標記在每個附圖中。在附圖中：

【0025】 圖1表示一個描繪本發明非限制實施例的示意圖。步驟1和2中，合成一個寡核苷酸庫，且使相對應的寡核苷酸黏合在一起。各個寡核苷酸含有一個CRISPR引導序列、兩個BbsI限制酶識別位點、以及一個條碼元件。步驟3中，在一個單鍋連接反應中，寡核苷酸被連接至一個儲存載體上，而產生一個包含該核苷酸的載體。步驟4中，在該載體的BbsI限制酶識別位點上進行消化，以插入一個支架序列以及一個由BamHI和EcoRI限制酶識別位點所形成的分隔位點。該條碼引導RNA核酸庫可在該分隔位點上被迭代地消化，而插入包含一個CRISPR引導序列、支架序列、分隔位點及條

碼元件的額外元件，進而產生一個帶有串聯的條碼元件的複合引導RNA核酸庫。該序列，從頭到尾，係對應於SEQ ID Nos: 364-366。

【0026】 圖2A及2B表示一個描繪本發明非限制實施例的示意圖。圖2A的步驟1中，合成寡核苷酸，各個包含複數個CRISPR序列和一個位於CRISPR引導序列下游的單一條碼元件。限制酶識別位點(RE)存在於各個CRISPR引導序列之後。步驟2中，以單鍋連接反應法將匯集的合成寡核苷酸連接於標的載體中。如步驟3所示，可使用不同的限制酶，接續地在各個CRISPR引導序列之後的各個限制酶識別位點上消化該載體，以插入一個支架元件，而在某些情況下，可插入一個啟動子元件，以啟動一個下游CRISPR引導序列的表現，進而產生一個編碼了複數個CRISPR引導序列和支架序列並含一個單一條碼元件的條碼組合性引導RNA核酸庫。在圖2B的步驟1中，合成寡核苷酸，各個包含複數個CRISPR序列和一個位於CRISPR引導序列上游的單一條碼元件。限制酶識別位點(RE)存在於各個CRISPR引導序列之後。步驟2中，以單鍋連接反應法將匯集的合成寡核苷酸連接於標的載體中。如步驟3所示，可使用不同的限制酶，接續地在各個CRISPR引導序列之後的各個限制酶識別位點上消化該載體，以插入一個支架元件，而在某些情況下，可插入一個啟動子元件，以啟動一個下游CRISPR引導序列的表現，進而產生一個編碼了複數個CRISPR引導序列和支架序列並含一個單一條碼元件的條碼組合性引導RNA核酸庫。

【0027】 圖3A-3D顯示生產一種高覆蓋組合式gRNA核酸庫以及有效地將該核酸庫輸送至人類細胞中。圖3A表示由大腸桿菌所純化出的質體庫中一回gRNA核酸庫條碼讀取值的累積分佈狀態，結果說明對於所有預期的

組合而言可全部覆蓋。圖3B表示在質體庫與被慢病毒感染的OVCAR8-ADR-Cas9細胞庫兩者中的二回gRNA核酸庫的情況，結果說明其對於所有預期的組合而言可接近全部覆蓋。最為條碼化的gRNA核酸庫可在每個組合中的平均條碼讀取值的5倍範圍內側得(以陰影區域強調，並以箭頭指示)。圖3C顯示在質體庫與被感染的OVCAR8-ADR-Cas9細胞庫內條碼表示值間的高關聯性(標準化的條碼計數log_2數值)，其結果說明慢病毒可有效地將二回核酸庫輸送至人類細胞中。圖3D顯示，以二回gRNA核酸庫感染後培養5天時，OVCAR8-ADR-Cas9細胞中兩個生物性複製體間的條碼表示值的高度在線性。R是皮爾森相關係數(Pearson correlation coefficient)。

【0028】 圖4A-4C顯示使用高通量篩選法所得可抑制癌細胞增生的gRNA組合物的鑑定結果。圖4A顯示高通量篩選示意圖，其中以條碼化的二回gRNA核酸庫感染OVCAR8-ADR-Cas9，並培養15或20天。以Illumina HiSeq鑑定並定量細胞庫內的條碼表示值，並進行兩個細胞庫間的比較。圖4B(右欄)顯示在培養20天及15天的細胞中表準化條碼計數數值間，被發現可調控細胞增生的二回gRNA組合物，以log_2比值排序的結果。圖4B(左欄)顯示相同的gRNA與對照組gRNA配對的結果。含對照組的gRNA對組合物以空心三角形表示。在其他生物性複製體中可確認gRNA組合物的抗增生效果，並以空心圓表示(參見圖12)。進一步驗證經標記的gRNA組合物。圖4C表示可調控癌細胞增生的二回組合物的驗證結果。以指定的二回gRNA組合物感染OVCAR8-ADR-Cas9細胞群，並培養15天。將等量的細胞重新塗盤並培養一段額外的時間，如圖所示。使用MTT檢測法測定細胞活性，並以吸收值(OD_{570} - OD_{650}) (n = 3)確認。數據以平均值±標準差(SD)表示。

【0029】 圖 5A-5D顯示KDM4C和BRD4或KDM6B和BRD4抑制人類卵巢癌細胞生長的組合式抑制效果。圖5A顯示，相對於經由表現對照gRNA的慢病毒感染的細胞而言，經由表現指定的單一或組合式gRNA的慢病毒感染OVCAR8-ADR-Cas9細胞的細胞活性的倍率改變值。培養細胞15天，然後將等量的被感染細胞重新塗盤並額外培養5天。圖5B顯示，相對於經由表現對照shRNA的慢病毒感染的細胞而言，經由表現指定的shRNA共同感染的OVCAR8-ADR細胞的細胞活性倍率改變值。培養細胞9天，然後將等量的被感染細胞重新塗盤並額外培養4天。圖5C顯示，相對於沒有接受藥物的對照組細胞而言，以指定濃度的SD70和JQ1處理5天後，OVCAR8-ADR細胞的細胞生長抑制百分比。亦顯示經由預測的Bliss independence及HAS模型針對SD70和JQ1的組合物計算後的過度抑制情況(中間及右欄)。圖5D顯示，相對於沒有接受藥物的對照組細胞而言，以指定濃度的GSK-J4和JQ1處理7天後，OVCAR8-ADR細胞的細胞生長抑制百分比。亦顯示經由預測的Bliss independence及HAS模型針對GSK-J4和JQ1的組合物計算後的過度抑制情況(中間及右欄)。以MTT檢測法測定細胞活性。數據以平均值±SD表示(圖5A中，n =3；圖5B-5D中n = 6)。星號(*$P < 0.05$)及米號(#$P < 0.05$)分別代表指定樣本間，以及經藥物處理與無藥物的對照組樣本間的顯著差異。

【0030】 圖6A-6E 顯示以慢病毒輸送組合式gRNA表現構築體可提供有效的標的基因的表現。圖6A表示，人類細胞中組合式慢病毒gRNA表現構築體的試驗策略示意圖。製作慢病毒，其包含有分別經由UBC和CMV啟動子所控制的而可編碼出RFG和GFP的基因，以及可標定RFP(RFP-sg1或RFP-sg2)和GFP(GFP-sg1)且由U6啟動子重複序列起始的gRNA表現區域。使

用慢病毒感染OVCAR8-ADR或OVCAR8-ADR-Cas9細胞，並使用流式細胞儀與螢光顯微鏡分析GFP和RFP的表現。圖6B顯示，經由編碼了指定gRNA表現構築體的慢病毒感染細胞4天後，細胞中GFP和RFP表現量的流式細胞儀散佈圖。編碼組合式gRNA表現構築體的慢病毒使OVCAR8-ADR-Cas9細胞中RFP和GFP螢光值表示陽性的細胞百分比減少，而OVCAR8-ADR細胞則沒有。圖6C表示，經由編碼了指定的gRNA表現構築體的慢病毒感染後4天的GFP(左欄)和RFP(右欄)陽性細胞的百分比。圖6D表示，經由編碼了指定的gRNA表現構築體的慢病毒感染後8天的GFP(左欄)和RFP(右欄)陽性細胞的百分比。在靶向RFP和GFP的gRNA間可測得有限的交叉反應性。圖6B中的數據代表以流式細胞儀測量經感染4天後的細胞的測量值，而圖6C和6D的定量結果以平均值±標準差(n=3)表示。圖6E表示螢光顯微鏡分析結果，其說明在感染後3天，組合式gRNA表現構築體可有效地抑制OVCAR8-ADR-Cas9細胞，但非OVCAR8-ADR細胞中的RFP和GFP螢光值。

【0031】 圖 7A-7C顯示OVCAR8-ADR-Cas9細胞中標的基因的gRNA的剪切效率。圖7A表示，使用Surveyor檢測法檢測OVCAR8-ADR-Cas9細胞中的插入刪除百分比的彙整表。以8種由篩選基因庫中隨機選定的不同gRNA感染細胞8或12天。由Surveyor檢測法偵測到的未被剪切及被剪切的PCR產物的預期大小，以鹼基表列。圖7B和7C表示瓊脂糖膠體測定結果，其顯示以Surveyor檢測法測定未經感染或以指定的gRNA感染8或12天的OVCAR8-ADR-Cas9細胞中DNA剪切效率的結果。

【0032】 圖8A和8B顯示OVCAR8-ADR-Cas9細胞中標的基因的二回gRNA表現構築體的剪切效率。圖8A表示，由Surveyor檢測法偵測到的未被

剪切及被剪切的PCR產物的預期大小，以鹼基表列(上半部)。瓊脂糖膠體顯示，以Surveyor檢測法檢測經由指定的單一或二回gRNA表現構築體感染12天的OVCAR8-ADR-Cas9細胞中所測得的插入刪除百分比(下半部)。圖8B是免疫墨點分析結果，其顯示經由對照組在體或指定的單一或二回gRNA構築體感染15天的OVCAR8-ADR-Cas9中的蛋白質含量。

【0033】 圖9A-9C表示經由二回gRNA表現構築體感染的單細胞衍生OVCAR8-ADR-Cas9殖株中標的等位基因的DNA比對結果。圖9A顯示經由表現靶向BMI1和PHF8的sgRNA的慢病毒感染的OVCAR8-ADR-Cas9細胞中所得序列的比對結果。該些序列，從頭到尾，係分別對應於 SEQ ID NOs: 203、204、203、203、204、204、203、205、206、206、203、203、204、207、208、209、204、210、208、208、210、210、203、219、206、206、208、208、220、220、221、222、204、204、223、224、204、204、203、203、204、和204。圖9B顯示經由表現靶向BRD4和KDM4C的gRNA的慢病毒感染的OVCAR8-ADR-Cas9細胞中所得序列的比對結果。OVCAR8-ADR-Cas9 細胞係經由慢病毒感染12天，並塗盤成單細胞群。萃取各個單細胞擴增殖株的基因組DNA。以PCR擴增標的等位基因，並以TA選殖法將其插入TOPO載體中而用於Sanger定序。顯示各殖株的兩個等位基因的序列。粗體代表核苷酸突變和插入，而刪除則以「-」表示。顯示針對標的基因的野生型(WT)序列做為參考，以底線表示20 bp gRNA標的，而PAM序列則以粗斜體表示。該些序列，從頭到尾，係分別對應於SEQ ID NOs: 211、212、211、211、213、213、211、211、214、214、211、215、212、216、211、211、217、218、211、211、212、212、211、211、225、226、

211、211、226、227、211、211、212、212、211、228、226、229、211、211、229、226、211、211、230、和226。圖9C是文氏圖(Venn diagram)，其顯示經單一及二回基因修飾的細胞頻率。以FACS將攜帶指定的二回gRNA表現構築體的OVCAR8-ADR-Cas9細胞塗盤成單細胞群。以Illumina MiSeq，由40個經由全染色體擴增的單細胞群中進行標的等位基因的定序。有75% (即30/40) 和80% (即32/40)的單細胞群分別攜帶至少一種位於標的BMI1和PHF8基因位點的突變等位基因。有62.5% (即25/40)的單細胞群在BMI1和PHF8基因中含有至少一種突變等位基因。各個單細胞群的兩個等位基因的序列如表6所示。經由Sanger定序法(圖9B)測定單細胞衍生殖株以及以Illumina MiSeq定序經全染色體擴增的單細胞群(圖9C)，可觀察到相似的突變等位基因頻率。

【0034】 圖 10A和10B顯示，針對組合式gRNA篩選的生物性複製體間條碼定量的高度再現性。圖10A表示，經由二回gRNA核酸庫感染後並培養15天的OVCAR8-ADR-Cas9細胞中，兩個生物性複製體間的條碼表示值(標準化條碼計數的log_2數值)的比較結果。圖10B表示 ，經由二回gRNA核酸庫感染後並培養20天的OVCAR8-ADR-Cas9細胞中，兩個生物性複製體間的條碼表示值(標準化條碼計數的log_2數值)的比較結果。R是皮爾森相關係數。

【0035】 圖11顯示在表現構築體內以不同次序排列的相同gRNA組合物中條碼定量數值的一致性倍率變化。針對各個不同排序的二回gRNA組合物(亦即：sgRNA-A + sgRNA-B and sgRNA-B + sgRNA-A)測定其變異係數(CV；定義為SD/培養OVCAR8-ADR-Cas9細胞培養20天後的條碼計數值，

相對於培養15天之數值的倍率變化的平均值)。在細胞增生篩選中，分別有超過82%的二回gRNA組合物具有<0.2的CV，及有95%的二回gRNA組合物具有 <0.4的CV。

【0036】 圖 12A-12C顯示以組合式篩選而鑑定出可抑制癌細胞增生的gRNA對的生物性複製體。圖12A顯示經由與圖4B中相同的二回gRNA核酸庫感染OVCAR8-ADR-Cas9細胞所得的log_2倍率變化。以針對細胞培養20天對比於15天所得的標準化條碼計數值的log_2比率排序可調控增生的引導RNA對組合物(右欄)以及其gRNA+對照組(即，基因靶向gRNA+對照組gRNA；左欄)。以空心圓標示在另一個生物性複製體中確認的gRNA組合物對於抗增生的效果(圖4B)，而以空心三角形標示以對照組RNA對的組合物。進一步於圖4C中驗證經標記的gRNA組合物。圖12B為經由二回gRNA核酸庫感染的OVCAR8-ADR-Cas9細胞的生物性複製體之間，細胞培養20天對比於15天所得的標準化條碼計數值的log_2比率散佈圖。圖12C顯示在細胞庫篩選中gRNA組合物的log_2比率的頻率分佈。由兩個生物性複製體的平均值計算圖中所顯示的log_2比率。

【0037】 圖13顯示在細胞庫篩選中的個別數值與在驗證數值間的高度一致性。針對各個二回gRNA組合物，由細胞庫篩選數據(‘篩選表型’)中所得的細胞培養20天對比於15天的標準化條碼計數值的倍率改變，與其相對的細胞活性進行作圖，並與由個別細胞增生試驗所測得的載體對照組進行比較(‘驗證表型’) (R = 0.932)。篩選表型的數據係以兩個生物性複製體的平均值表示；個別的驗證表型代表三個獨立試驗的平均。的細胞 R是皮爾森相關係數。

【0038】 圖14A-14F顯示在OVCAR8-ADR 細胞中，由shRNA介導的標的基因敲除。圖14A顯示表現了對照組shRNA或靶向KDM4C的shRNA的OVCAR8-ADR 細胞中KDM4C的相對mRNA含量。圖14B顯示表現了對照組shRNA或靶向BRD4的shRNA的OVCAR8-ADR 細胞中BRD4的相對mRNA含量。圖14C顯示表現了對照組shRNA或靶向KDM6B的shRNA的OVCAR8-ADR 細胞中KDM6B的相對mRNA含量。以qRT-PCR定量mRNA含量，並以肌動蛋白mRNA的含量進行標準化。數據以平均值±SD表示(n=3)。圖14D-14F顯示表現了對照組shRNA或靶向KDM6B、BRD4或KDM6B的shRNA的OVCAR8-ADR 細胞中相對蛋白質含量的西方墨點分析結果。將測得的蛋白質含量標準化至肌動蛋白含量，且標準化至對照組shRNA樣本，並以相對的蛋白質含量作圖如下。星號(*$P < 0.05$)代表表現基因靶向shRNA與對照組shRNA的細胞間mRNA或蛋白質含量的顯著差異。

【0039】 圖15顯示條碼化組合式gRNA核酸庫的組裝策略。合成、黏合條碼化的gRNA寡核酸對，並將其以匯集形式選殖入儲存載體中。將含有gRNA支架序列的寡核酸插入該匯集的儲存載體庫中，藉以創建出條碼化的sgRNA核酸庫。詳細的組裝步驟如圖1所示。使用CombiGEM策略於建構該組合式gRNA核酸庫。將由BglII和MfeI進行剪切的sgRNA核酸庫所製得的匯集條碼化sgRNA插入物係經由在標的載體中以BamHI和EcoRI剪切而得的相容突出端進行連接。迭代的單鍋連接反應可產生具有對應於在一端串聯的gRNA的特定條碼的(n)-重gRNA核酸庫，因而，可藉由次世代定序法追蹤匯集群體中的個別組合式成員。

【0040】 圖16A-16E顯示OVCAR8-ADR-Cas9細胞中標的基因組位點

上針對gRNA的插入刪除分析的深度定序結果。以指定的sgRNA感染細胞15天，然後進行深度定序。圖16A顯示插入刪除頻率。圖16B顯示架構轉移及架構內突變的百分比。圖16C顯示插入刪除大小的分佈。圖16D顯示，以一回sgRNA(KDM4C或BRD4)，上圖，或者二回gRNA表現構築體(KDM4和BRD4)，下圖，進行感染15天後，進行OVCAR8-ADR-Cas9細胞中標的基因組位點的深度定序所分析得的插入刪除分佈。圖16E顯示，以一回sgRNA(PHF8或BMI1)，上圖，或者二回gRNA表現構築體(PHF8和BMI1)，下圖，進行感染15天後，進行OVCAR8-ADR-Cas9細胞中標的基因組位點的深度定序所分析得的插入刪除分佈。

【0041】 圖 17A-17C顯示在混合細胞群中前期增生性gRNA及抗增生性gRNA頻率的數學模型。圖17A顯示在一個具有不同率(即，2、5或10%)的一開始含有抗增生性gRNA(f_s)和前期增生性(f_f) gRNA的細胞群中，前期增生性gRNA即抗增生性gRNA的相對頻率的刺激結果。相對頻率係定義為相對於起始時間點而言，在給定時間下的條碼豐度值。在此實例中，因CRISPR-Cas9系統而產生的基因擾動(p)所得具有修正生長速率的細胞分率設定為1.0(即，100%)，而抗增生性殖株的倍增時間($T_{doubling,m}$)為48小時。圖17B和17C顯示套用數學模型後，在混合細胞群中因參數：p、$T_{doubling,m}$、f_s、及 f_f的變異而得到的抗增生性gRNA相對頻率。在圖17B和17C各圖中，線條代表，從頂部到底部，p=0.2、p=0.4、p=0.6、p=0.8、和p=1.0。圖17A-17C中，前期增生性殖株的倍增時間設定為12小時。詳細的定義於實例3中描述。

【0042】 圖18顯示個別gRNA組合物的匯集篩選及驗證數據。針對各個gRNA組合物，將由匯集篩選數據(‘篩選表型’)所得的而在細胞培養20天

對比於15天的標準化條碼計數值上的倍率改變與其相對的細胞活性進行作圖，並與由個別細胞增生試驗所測得的載體對照組進行比較('驗證表型')。各個個別sgRNA的篩選表型係由含有各個三個對照組sgRNA的相對應sgRNA對的倍率改變取平均而得。篩選數據的數據為兩個生物性複製體的平均值，而各個驗證數據則以平均值 ± SD ($n \geq 3$)表示。

【0043】 圖19 呈現靶向KDM4C、KDM6B、及BRD4的gRN在對靶及脫靶的插入缺失序列產生率上的測量結果。每一列代表一個對應於一個20 bp引導序列(黑色)接著一個3 bp的PAM序列(灰色)的基因組位點。粗黑字體的序列代表gRNA的對靶基因組序列。對於KDM4C-sg1、KDM6B-sg2、及BRD4-sg3而言，在每條虛線下方的序列為使用CRIPSR設計(Ran, et al. *Nature Protocols* (2013) 8:2281-2308)及CCTop(Stemmer, *PLoS One* (2015) 10:e0124633)工具所鑑定出的所有預測性外顯子脫靶基因組序列。亦評估五個針對BRD4-sg2而預測出的外顯子/內含子脫靶位點。以底線標示的核苷酸強調脫靶序列與對靶序列中的差異性。以PCR從~10,000個細胞中進行每個基因組位點的擴增，並以>4.2百萬個讀值進行深度定序。「n.d.」代表因無法提供特定的擴增子進行定序，使得進行基因序列PCR時產生失敗。該序列，從頭到尾，對應於 SEQ ID NOs: 368-393。

【0044】 圖20A-20B顯示在攜帶了KDM4C及BRD4架構轉移突變的OVCAR8-ADR-Cas9細胞中生長能力降低的現象。圖20A描述攜帶了KDM4C及BRD4架構轉移突變的OVCAR8-ADR-Cas9細胞，經由單細胞擴增而得到的突變殖株(即，得自圖9B所示的#3殖株)的細胞生長結果。在進行MTT測定之前，將等量的細胞接種並培養5天。數據以相對於對照組

OVCAR8-ADR-Cas9細胞的吸光測量值的平均值 ± SD (n = 3)表示 。星號(*P < 0.01)代表在對照組和突變細胞之間的顯著差異。圖20B呈現在圖20A的對照組和突變細胞中蛋白質含量的免疫墨點分析結果。

【0045】 圖21A-21B顯示以gRNA表現構築體感染OVCAR8-ADR-Cas9細胞後所進行的RNA定序分析結果。圖21A顯示代表性熱圖，其顯示出分別以單一或雙重gRNA感染OVCAR8-ADR-Cas9細胞後，其每個樣本(欄)中的每個基因轉錄體(列)的相對表現量。以指定的gRNA感染OVCAR8-ADR-Cas9的細胞，經鑑定後所表現出的轉錄體含量與載體對照組相比具有顯著差異時，則包含於該熱圖中。數值為使用RNA-Seq進行測量並以該轉錄體作為中心平均值的經log_2-轉換的FPKM。根據皮爾森相關法進行轉錄體和樣本的分級叢集。圖21B顯示生物程序中的前10大富集基因集，其係當與載體對照組比較時(Q-值<0.05)，在經由指定的gRNA感染的OVCAR8-ADR-Cas9細胞中所鑑定出的差異化表現基因。在表中，以灰色陰影表示與基因集(y軸)相關聯的差異化表現基因子集(x軸)。

【0046】 圖22A-22B顯示KDM4C和BRD4以及KDM6B和BRD4對於額外癌細胞株的細胞生長效果。圖22A顯示，組合性gRNA基因表現構築體可有效地抑制乳癌MDA-MB231-Cas9細胞及胰腺癌Bx-PC3-Cas9細胞中靶向的螢光基因。在含有或不含由U6啟動子重複序列起始的靶向RFP及GFP的gRNA表現區域的情況下，將含有由組合型啟動子所表現的RFP及GFP的慢病毒載體輸送至MDA-MB231-Cas9和Bx-PC3-Cas9細胞中，並藉由流式細胞儀，分析GFP和RFP的表現。詳細的策略說明於圖6中。在感染後4天，編碼了組合性gRNA基因表現構築體的慢病毒可減少RFP和GFP螢光呈現陽性

的細胞百分比。圖22B顯示，將以表現了指定的單一或組合性gRNA的慢病毒感染MDA-MB231-Cas9和Bx-PC3-Cas9細胞培養14天。然後，將等量的被感染細胞重新接種並額外培養5天。以MTT測定法測定相對於對照組sgRNA的細胞活性。數據以由生物性複製體所得的平均值± SD (n =6)表示。星號(*P < 0.05)代表在指定的樣本之間的顯著差異。這些結果說明靶向表觀遺傳基因的組合性gRNA可因不同的細胞情況而具有各種不同的表型。

【實施方式】

【0047】 產生用於表現包含多於一個sgRNA(引導序列及支架序列)的多重CRISPR系統的載體及基因元件是很費力的，且採用傳統選殖方法建構CRISPR系統庫的複雜性非常有限。本文所述方法可產生帶有串聯式條碼及CRISPR引導序列及支架序列的載體。該方法在建構用於組合性基因篩選的大型基因庫具有高度潛力，且利用此一事實，進而可容易地列印大量寡核苷酸，且因該引導序列及條碼元件的長度短，因此帶有標的特異性決定區的引導序列可被列印至這些寡核苷酸上。

【0048】 如本文所述，利用大量平行組合性基因學方法產生CRISPR構築體和載體的方式，可快速地產生包含有可標靶宿主中核酸的CRISPR系統元件(CRISPR 引導序列及支架序列)的組合性基因構築體組。該方法亦能進行多重組合次序的匯集篩選(例如：可同時匯集並篩選二次、三次與n次組合)，並鑑定出特定應用層面所需的最小組合。組合性基因構築體組，例如使用本文所述方法而生產者，可用於鑑定能協同地交互作用的基因及基因路徑，以調節細胞程序或表型，例如癌細胞的生長。本文亦描述使用本文所述的組合性CRISPR構築體鑑定出新穎的表觀遺傳基因組合物，當一起進

行抑制時，具有抗癌效果，例如減少細胞的增生。

【0049】 本發明多個方面關於基因構築體、包含基因構築體的載體、組合性載體、以及使用大量平行組合性基因學途徑產生組合性載體的方法，如PCT公開號WO2014/00542中所描述，其係以引用方式納入本文中。如本文所述，一個「基因構築體」係指一或多個包含CRISPR引導序列及支架序列以及條碼序列的DNA元件，而每個DNA元件係關聯於一個條碼元件。如本文所述，特定DNA元件與條碼元件之間的關聯意味著特定DNA元件與條碼元件總會包含在相同的基因構築體中。因此，在一個基因構築體中存有或可偵測到一個特定的條碼元件意指該關聯的特定DNA元件亦存在相同的基因構築體中。

【0050】 宿主細胞中，包含一個CRISPR 引導序列及一個支架序列的DNA元件被轉錄，並形成一個CRISPR小型引導RNA (sgRNA)，其功能為吸引核酸內切酶至宿主細胞中特異的標靶核酸上，而產生位點特異性CRISPR活性。如本文所述，「CRISPR 引導序列」係指與宿主中標靶核酸序列互補的的核酸序列。該CRISPR 引導序列使sgRNA靶向一個標靶核酸序列，其亦稱為標靶位點。與標靶核酸互補的該CRISPR 引導序列長度可介於15-25個核苷酸、18-22個核苷酸、或19-21個核苷酸之間。在某些實施例中，與標靶核酸互補的該CRISPR 引導序列的長度為15、16、17、18、19、20、21、22、23、24、或25個核苷酸。在某些實施例中，與標靶核酸互補的該CRISPR 引導序列的長度為20個核苷酸。

【0051】 據瞭解，若一個CRISPR 引導序列能與宿主細胞中的一個標靶核酸雜交，則該CRISPR 引導序列係互補於該標靶核酸。在某些實施例

中，該CRISPR 引導序列至少有 50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、或至少有100%與一個標靶核酸互補(亦參見美國專利案號8,697,359，其所指導的一種CRISPR引導序列與一種標靶多核苷酸序列間的互補性係以引用方式併入本文中)。已證明，CRISPR引導序列及鄰近該標靶核酸3'端的標靶核酸之間的錯誤配對可破壞核酸酶的裂解活性(Upadhyay, et al. *Genes Genome Genetics* (2013) 3(12):2233-2238)。在某些實施例中，該CRISPR 引導序列至少有50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、或至少有100%與該標靶核酸3'端互補(例如：至少有5、6、7、8、9、或10個該標靶核酸3'端的核苷酸)。

【0052】 CRISPR 引導序列可由本領域已知的任何來源獲得。舉例而言，該CRISPR 引導序列可以是存在於宿主細胞核酸中的任何指定長度的核酸序列(例如：基因組核酸及/或 基因組外核酸)。在某些實施例中，CRISPR引導序列可被設計並合成以靶向所需的核酸，例如編碼了轉錄因子、訊息蛋白、運輸蛋白等的核酸。在某些實施例中，該CRISPR引導序列被設計並合成以靶向表觀遺傳基因。舉例而言，該CRISPR引導序列可被設計以靶向靶向表2所述的任何表觀遺傳基因的組合物。在某些實施例中，該CRISPR引導序列包含表1中所提供的任何CRISPR 引導序列的實例。

【0053】 本文所使用的術語「支架序列」亦指一種 tracrRNA，其係指將一種核酸內切酶吸引至一個結合(雜交)於互補CRISPR引導序列的標靶核酸的核酸序列。任何包含至少一個莖環結構(stem loop structure)並可吸引核酸內切酶的支架序列皆可用於本文所述的基因組元件及載體中。對本領

域技術人員而言，示例性支架序列將是顯而易見的，並且可在例如Jinek, et al. *Science* (2012) 337(6096):816-821、Ran, et al. *Nature Protocols* (2013) 8:2281-2308、PCT申請案號WO2014/093694、以及PCT申請案號WO2013/176772中發現。

【0054】 術語「標靶核酸」、「標靶位點」和「標靶序列」可在本文中交替使用，並指稱任何在宿主細胞中可靶向本文所述的該CRISPR引導序列的核酸序列。該標靶核酸位於一個前間區序列鄰近基序(PAM)下游側邊(在3'側)，其可與核酸內切酶作用，並進一步參與將核酸內切酶活性靶向該標靶核酸。普遍認為，在該標靶核酸側邊的PAM序列取決於核酸內切酶並源自該核酸內切酶的來源。舉例而言，對於來自化膿性鏈球菌(*Streptococcus pyogenes*)的Cas9核酸內切酶而言，PAM序列為NGG。對於來自金黃色葡萄球菌(*Staphylococcus aureus*) 的Cas9核酸內切酶而言，PAM序列為NNGRRT。對於來自腦膜炎雙球菌(*Neisseria meningitidis*) 的Cas9核酸內切酶而言，PAM序列為NNNNGATT。對於來自嗜熱鏈球菌(*Neisseria meningitidis*) 的Cas9核酸內切酶而言，PAM序列為NNAGAA。對於來自齒垢密螺旋體(*Treponema denticola*) 的Cas9核酸內切酶而言，PAM序列為NAAAAC。 對於Cpf1核酸酶而言，PAM序列為TTN。

【0055】 在某些實施例中，該CRISPR 引導序列及支架序列係表現為個別的轉錄體。在這樣的實施例中，該CRISPR 引導序列進一步包含一個與該支架序列一部分互補且可與該支架序列結合(雜交)並吸引核酸內切酶至該標靶核酸的額外序列。在其他實施例中，該CRISPR引導序列及支架序列係表現為單一轉錄體、可指稱為單一引導RNA(sgRNA)的嵌合RNA。

sgRNA具有結合(雜交)於標靶核酸與吸引核酸內切酶至該標靶核酸的雙重功能。在此實施例中，該支架序列可進一步包含一個連接環序列。

【0056】 條碼元件可作為鑑定基因構築體之用，且可說明在載體或基因組元件中存在一或多個特異性CRISPR 引導序列。一組條碼元件的成員具有足夠特異的核酸序列，使得每個條碼元件可容易地與該組中的其他條碼元件有所區別。條碼元件可為任何長度的核苷酸，但較佳長度為少於30個核苷酸。在某些實施例中，該條碼元件長度為6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、25、26、27、28、29、或30或者更多核苷酸。偵測條碼元件以及測定一個條碼元件或複數個條碼元件的核酸序列可作為決定一個基因構築體的關聯DNA元件是否存在。本文所述的條碼元件可以任何本領域所習知的方法進行偵測，包括定序或微陣列方法。

【0057】 圖1顯示許多本發明基因構築體的非限制性實例的示意圖。圖1步驟4中，包含以「引導序列」表示的CRISPR 引導序列和以「支架」表示的支架序列的DNA元件，其兩側有一個以「BamHI」表示的第一相容端元件，以及一個以「BglII」表示的第二相容端元件，其能夠互相黏合。該基因構築體亦包含一個以「條碼」表示的條碼元件，其兩側有一個以「EcoRI」表示的第三相容端元件，以及一個以「MfeI」表示的第四相容端元件，其能夠互相黏合，但不能黏合於該第一及第二相容端元件。該基因構築體亦包含一個分隔位點，使得該條碼元件係位於該分隔位點的一側，而該DNA元件則位於該分隔位點的另一側。圖1亦描繪一個在該DNA元件上游(相對於5'端)的啟動子元件，其可用於表現(轉錄)該DNA元件。若圖1所描

繪的DNA元件是在該條碼元件上游(相對於5'端)，則這樣的配置亦可以逆轉。

【0058】 可以用多種不同方法創造相容端，這對於本領域技術人員而言是顯而易見的，且這種相容端可以由多種不同的序列構成。如本文所述，「相容端元件」係指能夠互相連接或黏合的DNA區域。能夠互相連接或黏合的相容端元件對於本領域技術人員而言是顯而易見的，且其係指與另一股核酸序列呈現互補的末端元件，因而能夠與另一股進行鹼基配對。在許多非限制性實施例中，相容端元件包括含有相容突出端的限制位點、Gibson組裝序列、或由包括重組酶、大範圍核酸酶、TAL效應物/鋅指核酸酶、反式切割核酶/DNA酶或整合酶的任何DNA組裝方法所得的功能性元件。

【0059】 在某些實施例中，Gibson組裝法係被用於產生相容突出端。Gibson組裝法係指一種等溫DNA末端連接技術，其中在一個單一反應中可將多數DNA片段接合。此方法進一步描述於Gibson et al. (2009) Nature Methods 6:343-5，且其係以引用方式納入本文中。

【0060】 在其他實施例中，限制酶剪切係用於產生相容端，如圖1所描繪。使用這個方法，兩個獨特的限制酶可產生相容突出端。當這些突出端連結時，產生一個無法再被任何酵素識別的缺痕。應理解的是，可使用任何能產生相容突出端的限制酶。在某些非限制性實施例中，可使用標準生物性部件，例如BIOBRICKS ® (The BioBricks Foundation)或 BglBricks (Anderson et al. (2010) Journal of Biological Engineering 4:1)，以及關聯於這些標準生物性部件的酵素。使用這些標準生物性部件，例如BIOBRICKS ®或 BglBricks，對於本領域技術人員而言是常規的。應理解的是，雖然可使

用傳統的限制酶(例如：第一型、第二型或第三型限制酶)，亦可使用其他可剪切DNA的分子。舉例而言，標靶核酶可用於剪切特定的標靶位點。亦可使用大範圍核酸酶以使在插入的DNA元件處產生界面的可能性降到最低。亦可使用TALE或ZF核酸酶於靶向長鏈DNA位點，使得在插入的DNA元件內產生內部剪切的可能性降到最低。再者，可使用TOPO®選殖法達成限制酶剪切及連接。

【0061】 在某些實施例中，該第一相容端元件係經由限制酶BamHI的辨識及剪切而產生，而第二相容端元件則經由限制酶BglII的辨識及剪切而產生。在某些實施例中，該第三相容端元件係經由限制酶MfeI的辨識及剪切而產生，而第四相容端元件則經由限制酶RcoRI的辨識及剪切而產生。

【0062】 如本文所使用的，基因構築體的「分隔位點」係指可使該構築體線性化的區域。應理解的是，該分隔位點係位於該構築體核酸內經由接切而線性化的一個位點，其可容許額外的基因元件得以插入。在某些實施例中，該分隔位點係一個限制酶識別位點。舉例而言，圖1中，該分隔位點係由該第一及第四相容端元件所形成，分別以BamHI和EcoRI限制位點表示。使用對應的限制酶(BamHI和EcoRI)剪切該構築體而使其線性化後，可容許額外基因構築體得以插入。在某些實施例中，該分隔位點係由一個限制位點所形成。在某些實施例中，該分隔位點係由多於一個的限制位點所形成。

【0063】 本發明的多個方面關於用於製造一種包含本文所述基因構築體的組合性載體的方法。如圖1步驟3所描繪的，該方法涉及提供一種含有一個第一基因構築體的載體，該基因構築體包含一個簡稱為「20 bp引導

序列」的CRISPR引導序列，且其側邊有一個以「BglII」表示的第二相容端元件、以及一個針對一個第一限制酶且簡稱為「BbsI」的第一識別位點、一個簡稱為「條碼」的條碼元件，且其側邊有一個以「MfeI」表示的第三相容端元件，以及一個針對一個第二限制酶的第二識別位點。在某些實施例中，如圖1步驟2所示，該載體可經由使一個含有相容端第一基因構築體黏合及連接於一個經剪切的載體而產生。在某些實施例中，該第一基因構築體係經由如核苷酸陣列合成法所合成。可於該第一識別位點上剪切該第一基因構築體，進而產生一個第五相容端元件，而在該第二識別位點上剪切，則產生一個第六相容端元件。提供一個支架元件，其包含一個支架序列和一個以「BamHI」和「EcoRI」表示的分隔位點，其側邊有一個能夠黏合於該經剪切載體上的第五相容端元件的第七相容端元件，以及一個能夠黏合於該經剪切載體上的第六相容端元件的第八相容端元件。該支架元件係藉由相容端元件而黏合於該經剪切的載體上的第一基因構築體。黏合後，該支架序列係被整合在該CRISPR引導序列和該條碼序列之間，而該分隔位點則位於該支架序列和該條碼序列之間。

【0064】 應理解的是，可使用各種不同的酵素組合物進行該第一及第二識別位點的剪切。在某些實施例中，該兩個識別位點係位於該CRISPR引導序列外部，且以同樣的限制酶辨識該條碼元件，而產生含有支架元件的相容端。在其他實施例中，該兩個限制位點係位於該CRISPR引導序列外部，且以兩個不同的限制酶辨識該條碼元件，每個皆可產生含有支架元件的相容端。

【0065】 本發明進一步方面關於組合性構築體和用於生產組合性構

築體的方法。如本文所述，「組合性構築體」係指一個含有複數個DNA元件的基因構築體。如本文所述，複數個DNA元件係指多於一個DNA元件，每個該DNA元件包含一個CRISPR引導序列及一個支架序列。如圖1步驟5所示，產生一個組合性構築體涉及將包含一個關聯於本發明的第一基因構築體的載體線性化，其係藉由在該基因構築體內的分隔位點上剪切該載體而達成。一個關聯於本發明的第二基因構築體可被插入於該經剪切的載體，並黏合及連接於該載體。如本文所使用的「插入物」係指一個預定被插入於該經剪切載體中的基因構築體。在某些實施例中，該插入物係純化自一個載體，例如藉由PCR或限制酶剪切而得。該插入物可藉由位於該插入物內的終端相容端元件與該線性化載體的相容端元件黏合而被連接於該經剪切的載體 。

【0066】 圖1步驟5的(n)-次引導RNA核酸庫描繪一個組合後的組合性構築體，其含有複數個DNA元件以及複數個對應的條碼元件。在圖1步驟5中描繪的非限制性實例中，該基因構築體含有四個不同的DNA元件，和四個對應的條碼元件。組合性構築體進一步包含一個分隔位點，位於複數個條碼元件與複數個DNA元件之間。

【0067】 本文所述用於產生組合性構築體的方法可為迭代的。舉例而言，圖1中所描繪的組合性載體可以在其分隔位點上再次被剪切，而一或多個進一步的插入物可被連接至該組合性構築體上，同時維持一個分隔位點，以利進一步的插入。顯然，在整個迭代的過程中，當該基因構築體內的DNA元件數目持續增加時，關聯於每個DNA元件的獨特條碼如同其關聯的DNA元件一樣被維持在相同的基因構築體內。在某些實施例中，該組合

程序係重複至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20次、或多於20次。在某些實施例中，該程序係重複n次，其中n可為 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或大於20的數字。

【0068】 應理解的是，組合性構築體可包含任何數目的DNA元件和其關聯的條碼元件。在某些實施例中，一個組合性構築體含有2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或超過100個 DNA元件和其關聯的條碼元件。

【0069】 本發明其他方面關於包含關聯於一個單一條碼元件的多於一個CRISPR引導序列的基因構築體及載體。圖2顯示許多關聯於本發明的基因構築體非限制實例的示意圖。圖2A及2B的步驟1顯示一個基因構築體，其含有以「20 bp 引導序列 A」、「20 bp 引導序列 B」及「20 bp 引導序列 C」表示的三個CRISPR引導序列、一個以「條碼」表示的條碼元件、以及一個位於每個CRISPR引導序列之間，以及每個條碼元件與該CRISPR引導序列之間最靠近該條碼元件處的識別位點。在某些實施例中，該條碼元件可位於該CRISPR引導序列的下游，如圖2A所示。在其他實施例中，該條碼元件可位於該CRISPR引導序列的上游，如圖2B所示。在某些實施例中，位於

該CRISPR引導序列之間，以及在該條碼元件與該CRISPR引導序列之間最靠近該條碼元件處的識別位點為針對不同限制酶的每個不同識別位點。在某些實施例中，包含該至少兩個CRISPR引導序列、條碼元件以及識別位點的該基因構築體係以任何本技術所習知的方法合成，例如寡核苷酸陣列合成法。

【0070】 亦在本發明範疇中的為包含複數個DNA元件和一個條碼元件的基因構築體。在圖2步驟3中，一個基因構築體包含三個DNA元件，且其每個包含一個CRISPR引導序列和一個支架序列、一個條碼序列、以及一個位於每個DNA元件上游的啟動子序列。在某些實施例中，該條碼元件可位於該CRISPR引導序列元件的下游，如圖2A所示。在其他實施例中，該條碼元件可位於該CRISPR引導序列的上游，如圖2B所示。

【0071】 本發明多個方面關於生產一種包含本文所述基因構築體的組合性載體的方法。在某些實施例中，該方法涉及提供一種含有複數個CRISPR引導序列及一個位於該複數個CRISPR引導序列下游的條碼元件的載體。如圖2A及2B的步驟1所示，三個CRISPR引導序列以「20 bp 引導序列 A」、「20 bp 引導序列 B」及「20 bp 引導序列C」表示，而條碼元件則以「條碼」表示。該載體亦含有複數個針對複數個限制酶的限制位點。在圖2A步驟1中，每個限制位點係位於一個CRISPR引導序列的下游，且以「RE1」、「RE2」及「RE3」表示。在圖2B步驟1中，每個識別位點係位於一個CRISPR引導序列的上游，且以「RE1」、「RE2」及「RE3」表示。在某些實施例中，該載體亦含有一個位於至少一個該CRISPR引導序列上游的啟動子序列。可以任何本領域所習知的方法產生位於至少一個該CRISPR引導

序列下游的相容端元件。在某些實施例中，係在該載體內至少一個該識別位點上以一個限制酶進行剪切，而形成一個第一相容端元件及一個第二相容端元件。提供一個支架元件，其包含一個支架序列、視情況可選擇的一個啟動子序列、以及能夠分別黏合於該經剪切載體的第一相容端元件和第二相容端元件的一個第三相容端元件及第四相容端元件。該支架元件係藉由在該支架元件內的終端相容端元件與在該經剪切載體內的相容端元件黏合而黏合於該經剪切的載體上。本文所述的方法可為迭代的，進而產生一個組合性載體，其包含複數個CRISPR引導序列及支架序列，以及一個條碼元件。

【0072】 應理解的是，組合性載體可含有關聯於一個條碼元件的任何數量的DNA元件。在某些實施例中，一個組合性構築體含有2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或超過100個 DNA元件及一個條碼元件。關聯於一個條碼元件的DNA元件數量可取決於含有該CRISPR引導序列、條碼及識別位點且可被合成出的基因構築體長度。

【0073】 在本文所述的任何構築體或載體中，一或多個RNA結構域可被插入一或多個CRISPR引導序列中。在某些實施例中，該CRISPR 引導序列係融合於一或多個 RNA結構域。在某些實施例中，該RNA係一個非編

碼RNA或其片段。在這樣的實施例中，該RNA結構域可靶向一個DNA基因位點。此構築體或載體可用於呈現CRISPR。

【0074】 本發明進一步方面關於用於鑑定在一個基因構築體或載體內的一或多個DNA元件的方法。在組合過程中，一個關聯於一個特定的DNA元件的獨特條碼如同該特定DNA元件一樣，係保留在該基因構築體內。因此，鑑定一個條碼元件或複數個條碼元件可用於鑑定在相同基因構築體內的該被關聯的DNA元件或複數個DNA元件。在某些實施例中，一個條碼元件及/或一個DNA元件的序列係經由定序或微陣列分析而決定。應理解的是，任何訂定DNA序列的方法與鑑定一或多個條碼元件及對應的DNA元件的方法是相容的。顯著地，在一個組合性載體中，例如圖1步驟5所描繪的，該複數個條碼元件係在彼此鄰近處，因而可藉由例如DNA定序的方法，快速地鑑定多重的條碼元件，並因此同時鑑定多重DNA元件。

【0075】 本發明進一步方面關於包含二或多個本文所述基因構築體的核酸庫，其為與大量平行組合性基因學相容的方法。如本文所使用的，基因構築體庫係指二或多個基因構築體的集合。在某些實施例中，一個基因構築體的核酸庫係由在質體上的每個獨特DNA元件而產生。此質體庫可匯集形成一個載體庫。可形成一個插入物庫，例如，藉由在載體庫上進行PCR。在第一個組合過程中，所有載體皆可與所有的插入物配對，而產生一個配對組合的全組合性組。在這些配對庫與插入物庫間的進一步反應可形成由單一載體庫產生的三次、四次或多於四次的核酸庫。組合性構築體庫可用於篩選表現該組合性構築體庫的宿主細胞。在某些具體實施例中，該組合性構築體庫含有DNA元件DNA元件組合物，期含有靶向表觀遺傳基因

的CRISPR引導序列，例如表1中所列的CRISPR引導序列實例。

【0076】 應理解的是，由於組合步驟係在體外進行，這個技術可擴及任何可接收DNA的宿主細胞或有機體。在某些實施例中，該宿主細胞為細菌細胞。在某些實施例中，該有機體示細菌，而該構築體係由質體或噬菌體所攜帶。在某些實施例中，該宿主細胞為酵母菌細胞。在其他實施例中，該有機體是酵母菌，而該構築體係由質體或穿梭載體所攜帶。在其他實施例中，該宿主細胞是哺乳動物細胞，例如人類細胞。在此實施例中，本文所述的該基因構築體可由質體所攜帶，或由病毒所輸送，例如慢病毒或腺病毒。

【0077】 本文所述的基因構築體及載體關於包括一個CRISPR引導序列及支架序列的一個CRISPR系統元件的表現。表現CRISPR系統的該宿主細胞可表現一或多種額外的CRISPR元件，例如一個核酸內切酶。在某些實施例中，該宿主細胞亦表現一個核酸內切酶，例如一個Cas核酸內切酶。在某些實施例中，該Cas核酸內酶為Cas1、Cas2或Cas9核酸內切酶。在某些實施例中，表現一個Cas9核酸內切酶的該宿主細胞係來自化膿性鏈球菌、金黃色葡萄球菌、腦膜炎雙球菌、嗜熱鏈球菌、或齒垢密螺旋體。在某些實施例中，編碼了Cas9核酸內切酶的該核酸序列可為最適於在宿主細胞或有機體中表現的密碼子。在某些實施例中，該核酸內切酶是Cas9的同源酵素或直系酵素。

【0078】 在某些實施例中，該編碼Cas9核酸內切酶的核苷酸序列進一步被修飾以改變蛋白質的活性。在某些實施例中，該Cas9核酸內切酶是無催化活性的Cas9。舉例而言，dCas9含有具催化活性殘基(D10和H840)的突

變體，而不具有核酸酶活性。另可選擇或另外的，該Cas9核酸內切酶可融合於其他蛋白質或其蛋白質上。在某些實施例中，dCas9係融合於阻抑劑結構域上，例如KRAB結構域。在某些實施例中，這樣的dCas9融合蛋白係與本文所述的構築體一起使用，而用於多重基因抑制中(例如：CRISPR干擾(CRISPRi))。在某些實施例中，dCas9係融合於激活子結構域上，例如VP64或VPR。在某些實施例中，這樣的dCas9融合蛋白係與本文所述的構築體一起使用，而用於多重基因激活中(例如：CRISPR激活(CRISPRa))。在某些實施例中，dCas9係融合於表觀遺傳調控結構域上，例如組蛋白脫甲基結構域或組蛋白乙醯轉移酶結構域。在某些實施例中，dCas9係融合於LSD1或p300，或其一部分。在某些實施例中，dCas9的融合係用於基於CRISPR的表觀遺傳調控中。在某些實施例中，dCas9或Cas9係融合於Fok1核酸酶結構域上。在某些實施例中，融合於Fok1核酸酶結構域上的dCas9或Cas9係用於多重基因剪輯。在某些實施例中，dCas9或Cas9係融合於螢光蛋白上(例如：GFP、RFP、mCherry等)。在某些實施例中，融合於螢光蛋白上的Cas9/dCas9蛋白質係用於多重標記及/或觀察基因組位點。

【0079】 另可選擇的或另外的，該核酸內切酶是Cpf1核酸酶。在某些實施例中，該宿主細胞表現一個來自普雷沃菌屬(*Provetella* spp.)或法朗西斯菌屬(*Francisella* spp.)的Cpf1核酸酶。在某些實施例中，編碼了Cpf1核酸酶的核苷酸序列可為最適於在宿主細胞或有機體中表現的密碼子。

【0080】 本發明涵蓋任何可引入DNA的細胞類型，包括原核及真核細胞。在某些實施例中，該細胞為細菌細胞，例如埃希氏菌屬(*Escherichia* spp.)、鏈黴菌屬(*Streptomyces* spp.)、發酵單胞菌屬(*Zymomonas spp.)*、醋酸

桿菌屬(*Acetobacter* spp.)、檸檬酸桿菌屬(*Citrobacter* spp.)、集胞藻屬(*Synechocystis* spp.)、根瘤菌屬(*Rhizobium* spp.)、梭狀芽孢桿菌屬(*Clostridium* spp.)、棒狀桿菌屬(*Corynebacterium* spp.)、鏈球菌屬(*Streptococcus* spp.)、黃單胞菌屬(*Xanthomonas* spp.)、乳酸桿菌屬(*Lactobacillus* spp.)、乳酸球菌屬(*Lactococcus* spp.)、芽孢桿菌屬(*Bacillus* spp.)、產鹼桿菌屬(*Alcaligenes* spp.)、假性單胞菌屬(*Pseudomonas* spp.)、產氣單胞菌屬(*Aeromonas* spp.)、固氮菌屬(*Azotobacter* spp.)、叢毛單胞菌屬(*Comamonas* spp.)、分枝桿菌屬(*Mycobacterium* spp.)、紅球菌屬(*Rhodococcus* spp.)、葡萄糖酸桿菌屬(*Gluconobacter* spp.)、羅爾斯頓菌屬(*Ralstonia* spp.)、嗜酸菌屬(*Acidithiobacillus* spp.)、小月菌屬(*Microlunatus* spp.)、地桿菌屬(*Geobacillus* spp.)、土芽孢桿菌屬(*Geobacillus* spp.)、節桿菌屬(*Arthrobacter* spp.)、黃桿菌屬(*Flavobacterium* spp.)、沙雷氏菌屬(*Serratia* spp.)、糖多孢菌屬(*Saccharopolyspora* spp.)、棲熱菌屬(*Thermus* spp.)、寡養單胞菌屬(*Stenotrophomonas* spp.)、色桿菌屬(色桿菌屬)、中華根瘤菌屬(*Sinorhizobium* spp.)、糖多孢菌屬(*Saccharopolyspora* spp.)、土壤桿菌屬(*Agrobacterium* spp.)、以及泛菌屬(*Pantoea* spp.)。該細菌細胞可以是革蘭氏陰性菌，例如大腸桿菌(*Escherichia coli*；*E. coli*)，或革蘭氏陽性菌，例如芽孢桿菌。

【0081】　在其他實施例中，該細胞是一種真菌細胞，例如一種酵母菌細胞，如酵母菌屬(*Saccharomyces* spp.)、裂殖酵母菌屬(*Schizosaccharomyces* spp.)、畢赤氏酵母菌屬(*Pichia* spp.)、紅發夫酵母菌屬(*Phaffia* spp.)、克魯維酵母菌屬(*Kluyveromyces* spp.)、念珠菌屬(*Candida* spp.)、籃狀菌屬(*Talaromyces* spp.)、酒香酵母菌屬(*Brettanomyces* spp.)、管囊酵母菌屬

(*Pachysolen* spp.)、德巴利酵母菌屬(*Debaryomyces* spp.)、耶氏酵母菌屬(*Yarrowia* spp.)，以及工業用多倍體酵母菌株。較佳地，該酵母菌株為釀酒酵母菌株。其他真菌實例包括麴菌屬(*Aspergillus* spp.)、青黴菌屬(*Penicillium* spp.)、鐮刀菌屬(*Fusarium* spp.)、根黴菌屬(*Rhizopus* spp.)、枝頂孢黴菌屬(*Acremonium* spp.)、鏈孢黴菌屬(*Neurospora* spp.)、糞殼菌屬(*Sordaria* spp.)、稻熱病菌屬(*Magnaporthe* spp.)、壺菌異水黴菌屬(*Allomyces* spp.)、黑粉菌屬(*Ustilago* spp.)、葡萄孢菌屬(*Botrytis* spp.)、以及木黴屬(*Trichoderma* spp.)。

【0082】 在其他實施例中，該細胞是藻類細胞、植物細胞、昆蟲細胞、齧齒類動物細胞或哺乳動物細胞，包括齧齒類動物細胞或人類細胞(例如，人類胚胎腎細胞)(例如，HEK293T細胞)，人類真皮纖維母細胞、人類癌細胞，例如OVCAR8細胞或OVCAR8-ADR細胞。在某些實施例中，該細胞是人類癌細胞，例如人類卵巢癌細胞。

【0083】 本文亦提供包含靶向表觀遺傳基因的抑制劑的組成物。如本文所述，術語「表觀遺傳基因」係指任何會影響細胞中其他分子或程序的表觀遺傳調控狀態的基因。在某些實施例中，該表觀遺傳基因編碼一個核酸，例如RNA(例如，微型RNA)，其可影響表觀遺傳調控。一般而言，表觀遺傳學係指一個分子或程序的任何改變不會涉及細胞的基因組DNA的突變(Jaenisch and Bird *Nat. Gene.* (2003) 33: 245-254)。表觀遺傳調控涉及細胞中DNA介導的程序，例如經由包括DNA甲基化、組蛋白修飾、核小體重組、以及RNA介導的靶向作用等機制而進行轉錄、DNA修復、以及複製(Dawson and Kouzarides *Cell* (2012) 150(1): 12-27)。表觀遺傳基因的非限制性實例包括：DNMT1、DNMT3A、DNMT3B、DNMT3L、MBD1、MBD2、CREBBP、

EP300、HDAC1、HDAC2、SIRT1、CARM1、EZH1、EZH2、MLL、MLL2、NSD1、PRMT1、PRMT2、PRMT3、PRMT5、PRMT6、PRMT7、SETD2、KDM1A、KDM1B、KDM2A、KDM2B、KDM3A、KDM3B、KDM4A、KDM4B、KDM4C、KDM5A、KDM5B、KDM5C、KDM5D、KDM6A、KDM6B、PHF2、PHF8、BMI1、BRD1、BRD3、BRD4、ING1、ING2、ING3、ING4、以及ING5。

【0084】 如本文所使用的術語「抑制劑」係指任何可減少或防止表觀遺傳基因的表現，或者減少或防止由表觀遺傳基因編碼出的蛋白質活性的分子。在某些實施例中，相較於在不存有抑制劑組合物的情況下進行表觀遺傳基因的表現而言，二或多個表觀遺傳基因的抑制劑組合物可減少至少10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、或至少65%的表觀遺傳基因表現量。在某些實施例中，相較於在不存有抑制劑組合物的情況下由表觀遺傳基因所編碼的蛋白質活性而言，二或多個表觀遺傳基因的抑制劑組合物可減少至少 10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、或至少65% 由表觀遺傳基因編碼的蛋白質活性。

【0085】 在某些實施例中，表觀遺傳基因的抑制劑組合物包含至少2、3、4、5、6、7、8、9、或至少10個或多個表觀遺傳基因抑制劑。在某些實施例中，表觀遺傳基因的抑制劑組合物包含兩個表觀遺傳基因抑制劑。在某些實施例中，該抑制劑組合物抑制2、3、4、5、6、7、8、9、10或更多個表觀遺傳基因。在某些實施例中，該表觀遺傳基因的抑制劑組合物包含兩個靶向且抑制兩個表觀遺傳基因的抑制劑。

【0086】 在某些實施例中，至少一個該抑制劑組合物中的抑制劑為一種蛋白質，其可直接或間接減少或防止表觀遺傳基因的表現，或者減少或防止表觀遺傳基因所編碼的蛋白質的活性。舉例而言，該蛋白質可以是一個可減少或防止該表觀遺傳基因的表現的阻抑劑，或者一個由表觀遺傳基因所編碼的蛋白質的變構抑制劑。在某些實施例中，二或多個抑制劑為靶向二或多個表觀遺傳基因的蛋白質。在某些實施例中，該二或多個抑制劑為靶向表2中所述的任何表觀遺傳基因組合物的蛋白質。

【0087】 在某些實施例中，該抑制劑組合物的至少一個抑制劑為一種核酸，其可減少或防止表觀遺傳基因的表現，或者減少或防止由該表觀遺傳基因所編碼的蛋白質的活性。在某些實施例中，該核酸為一個CRISPR引導序列，連同一個支架序列，吸引一個核酸內切酶至該表觀遺傳基因。在某些實施例中，該二或多個抑制劑為靶向二或多個表觀遺傳基因的CRISPR引導序列。在某些實施例中，該二或多個抑制劑為靶向表2所述的任何表觀遺傳基因組合物的CRISPR引導序列。在某些實施例中，該二或多個抑制劑係選自靶向表1所述表觀遺傳基因的CRISPR引導序列的實例中的CRISPR引導序列。在某些實施例中，該表觀遺傳基因的組合物包含BRD4和KDM4C。在某些實施例中，表觀遺傳基因的組合物包含BRD4和KDM6B。

【0088】 在某些實施例中，該核酸為一個shRNA，其係藉由細胞的RNA干擾(RNAi)路徑 的進行而使標靶基因的表現沉默(例如，減少mRNA含量及/或蛋白質的產生)。在某些實施例中，二或多個抑制劑為靶向二或多個表觀遺傳基因的shRNA。在某些實施例中，該二或多個抑制劑為靶向表2所述的任何表觀遺傳基因組合物的shRNA。在某些實施例中，該二或多個

抑制劑係選自表4所提供的靶向表觀遺傳基因的shRNA實例的shRNA。在某些實施例中，該表觀遺傳基因的組合物包含BRD4和KDM4C。在某些實施例中，表觀遺傳基因的組合物包含BRD4和KDM6B。

【0089】 在某些實施例中，該抑制劑組合物的至少一個抑制劑為一種小分子，其可減少或防止表觀遺傳基因的表現，或者減少或防止由該表觀遺傳基因所編碼的蛋白質的活性。在某些實施例中，該二或多個抑制劑為靶向二或多個表觀遺傳基因的小分子。在某些實施例中，該二或多個抑制劑為靶向表2所述的任何表觀遺傳基因組合物的小分子。在某些實施例中，該表觀遺傳基因的組合物包含BRD4和KDM4C。在某些實施例中，表觀遺傳基因的組合物包含BRD4和KDM6B。

【0090】 任何可減少或防止BRD4的表現，或者減少或防止由BRD4所編碼的蛋白質活性的小分子可相容於本文所述的組成物和方法。BRD4抑制劑的實例包括，但不限定於JQ1 ((6S)-4-(4-氯苯基)-2,3,9-三甲基-6H-噻吩並[3,2-f][1,2,4]三唑並[4,3-a][1,4]二氮雜-6-乙酸1,1-二甲基乙基酯)、MS417 ([(6S)-4-(4-氯苯基)-2,3,9-三甲基-6*H*-噻吩並[3,2-*f*][1,2,4]三唑並[4,3-*a*][1,4]二氮雜-6-基]乙酸甲酯)、或RVX-208 (2-[4-(2-羥基乙氧基)-3,5-二甲基苯基]-5,7-二甲氧基-4(3H)- 喹唑啉酮)。在某些實施例中，該BRD4的抑制劑為JQ1。額外的BRD4抑制劑對於本領域技術人員而言是顯而易見的，且可在，例如，PCT公開案號WO 2014/154760 A1以及Vidler et al. *J. Med. Chem.* (2013) 56: 8073-8088中發現。

【0091】 任何可減少或防止KDM4C (亦指稱為JMJD2)的表現，或者減少或防止由KDM4C所編碼的蛋白質活性的小分子可相容於本文所述的

組成物和方法。在某些實施例中，該KDM4C抑制劑為SD70 (N-(呋喃-2-基(8-羥基喹啉-7-基)甲基)異丁醯胺)或咖啡酸。額外的KDM4C抑制劑對於本領域技術人員而言是顯而易見的，且可在，例如，Leurs et al. *Bioorg. & Med. Chem. Lett.* (2012) 22(12): 5811-5813以及Hamada et al. *Bioorg. & Med. Chem. Lett.* (2009) 19: 2852-2855)中發現。

【0092】 任何可減少或防止KDM6B (亦指稱為JMJD3)的表現，或者減少或防止由KDM6B所編碼的蛋白質活性的小分子可相容於本文所述的組成物和方法。KDM6B抑制劑的實例包括，但不限定於GSK-J4 (3-((6-(4,5-二氫-1H-苯並[d]氮雜-3(2H)-基)-2-(吡啶-2-基)嘧啶-4-基)氨基)丙酸乙酯，單鹽酸鹽)、GSK-J1 (N-[2-(2-吡啶基)-6-(1,2,4,5-四氫-3H-3-苯並氮雜-3-基)-4-嘧啶基]-β-丙氨酸)、以及IOX1 (8-羥基-5-喹啉羧酸；8-羥基-5-喹啉羧酸)。在某些實施例中，該KDM6B的抑制劑為GSK-J4。額外的KDM6B抑制劑對於本領域技術人員而言是顯而易見的。

【0093】 二或多個抑制劑的組合物可包含二或多個蛋白質抑制劑、二或多個CRISPR引導序列、二或多個shRNA、或者二或多個小分子抑制劑。在某些實施例中，該二或多個抑制劑為不同種類的抑制劑(例如：蛋白質、核酸、小分子)。在某些實施例中，該組合物包含一個蛋白質抑制劑及一或多個額外的抑制劑(例如：CRISPR引導序列、shRNA、及/或小分子抑制劑)。在其他實施例中，該組合物包含一個CRISPR引導序列及一或多個額外的抑制劑(例如：蛋白質、shRNA、及/或小分子抑制劑)。在其他實施例中，該組合物包含一個shRNA及一或多個額外的抑制劑(例如：蛋白質、CRISPR引導序列、及/或小分子抑制劑)。在其他實施例中，該組合物包含一個小分

子抑制劑及一或多個額外的抑制劑(例如：蛋白質、shRNA、及/或shRNA)。

【0094】 本文所述的方法和組成物可用於減少細胞增生，例如癌細胞或其他需要使其增生減少的細胞。在某些實施例中，使細胞與表觀遺傳基因(例如：靶向表2所述的表觀遺傳基因抑制劑的組合物)的二或多個抑制劑的組合物接觸，其可部分或完全地減少該細胞的增生。在某些實施例中，相較於不與該抑制劑組合物接觸的細胞而言，使細胞與表觀遺傳基因的二或多個抑制劑的組合物接觸可部分或完全地減少該細胞的增生。在某些實施例中，相較於不與該抑制劑組合物接觸的細胞而言，使細胞與表觀遺傳基因的二或多個抑制劑的組合物接觸可使細胞增生減少至少10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、或至少65%。在某些實施例中，在某些實施例中，相較於與該抑制劑組合物接觸的非癌細胞而言，使癌細胞與表觀遺傳基因的二或多個抑制劑的組合物接觸可部分或完全地減少該癌細胞的增生。可藉由本領域所習知的方法評估及定量細胞增生，例如使用細胞活性檢測法、MTT檢測法、或BrdU細胞增生檢測法。

【0095】 本發明其他方面關於用於治療受試者的癌症的方法及組成物。癌症是一種特徵為在細胞增生及其他惡性細胞特性上無法控制或控制異常的疾病。本文所使用的術語「癌症」係指習知的任何類型的癌症，包括但不限定於，乳腺癌、膽道癌、膀胱癌、腦癌、子宮頸癌、絨毛膜癌、結腸癌、子宮內膜癌、食道癌、胃癌、血液系統腫瘤、 T細胞急性淋巴細胞白血病/淋巴瘤、毛細胞白血病、慢性骨髓性白血病、多發性骨髓瘤、 AIDS相關白血病及成人T細胞白血病/淋巴瘤、上皮內腫瘤、肝癌、肺癌、淋巴瘤、神經母細胞瘤、口腔癌、卵巢癌、胰腺癌、前列腺癌、直腸癌、肉瘤、皮

膚癌、睾丸癌、甲狀腺癌、以及腎癌。癌細胞可以是活體內癌細胞(即：在有機體中)、離體(即：從有機體上取出並在體外維持)、或體外。

【0096】 該方法涉及以一個有效量將表觀遺傳基因的二或多個抑制劑的組合物給藥予受試者。在某些實施例中，該受試者是具有、疑有、或者有發展出癌症的風險的受試者。在某些實施例中，該受試者為一個哺乳動物受試者，包括但不限定於，狗、貓、馬、牛、豬、綿羊、山羊、雞、嚙齒動物、或靈長類動物。在某些實施例中，該受試者為人類受試者，例如一個病患。該人類受試者可為兒童或成人受試者。受試者是否被認為處於癌症的「風險」，可以由熟練的醫師來確認。

【0097】 如本文所使用的「治療」包括改善、治療、防止其敗壞、減緩進展速率、或預防該病症再發生(即：避免復發)。一個有效量的組成物係指一個可產生治療效果的劑量的組成物。例如，在治療一個受試者的癌症的方法中，一個有效量的靶向表觀遺傳基因的抑制劑組合物為任何劑量，其可提供抗癌效果，例如減少或防止癌細胞增生或對於癌細胞而言具有細胞毒性。在某些實施例中，當伴隨或連同一或多個額外的靶向表觀遺傳基因的抑制劑進行一個抑制劑的給藥時，相較於在不存有一或多個靶向表觀遺傳基因的額外抑制劑的情況下給藥一個有效劑量的抑制劑而言，該靶向表觀遺傳基因的抑制劑的有效量將減少。在某些實施例中，當在不存有一或多個靶向表觀遺傳基因的額外抑制劑的情況下進行給藥時，該靶向一個表觀遺傳基因的抑制劑不會減少或防止癌細胞的增生。

【0098】 可使用本領域所習知的任何方法將靶向表觀遺傳基因的抑制劑或靶向表觀遺傳基因的抑制劑組合物(例如，表2中所述的表觀遺傳基因

組合物)給藥予受試者。在某些實施例中，該抑製劑係由局部、腸內或腸胃外給藥途徑進行給藥。在某些實施例中， 該抑制劑係經由靜脈內、肌肉內、或皮下注射方式給藥。

【0099】 可以任何本領域所習知的方法將任何本文所述的表觀遺傳基因抑制劑給藥予受試者，或輸送至或與細胞接觸。在某些實施例中，可藉由奈米顆粒、細胞穿膜胜肽、聚合物、脂質體、或重組表現載體將表觀遺傳基因的抑制劑輸送至細胞中。在其他實施例中，表觀遺傳基因的抑制劑係結合於一或多個奈米顆粒、細胞穿膜胜肽、及/或聚合物上。在其他實施例中，表觀遺傳基因的抑制劑被包含於脂質體中。

【0100】 亦提供用於鑑定可減少或防止細胞或細胞群增生的表觀遺傳基因的抑制劑組合物的方法。如圖4A所描繪，該方法涉及使兩群細胞與一個含有靶向表觀遺傳基因的CRISPR引導序列及支架序列(例如：條碼化的CRISPR核酸庫)以及一個Cas9核酸內切酶的組合性核酸庫接觸。該兩群細胞被培養一段不同的時間。舉例而言，一群細胞可培養15天，另一群細胞培養20天。以，例如定序方法，鑑定該二或多個CRISPR引導序列及支架序列的組合物。舉例而言，可藉由定序一個對於CRISPR引導序列而言為一種獨特辨識物的條碼序列，而鑑定該CRISPR引導序列及支架序列。在培養一段較長時間的群細胞中每個CRISPR引導序列及支架序列的組合物的豐度與在培養一段較短時間的群細胞中每個CRISPR引導序列及支架序列的組合物的豐度進行比較。相較於培養一段較短時間的細胞群中的CRISPR引導序列豐度而言，使細胞增生減少的CRISPR引導序列及支架序列組合物在培養一段較長時間的細胞群中的豐度較低。這樣的組合物被鑑定為可減少細胞

增生的表觀遺傳基因抑制劑的組合物。

【0101】 提供其他用於在抑制作用產生而減少或防止細胞群增生的情況下，鑑定表觀遺傳基因組合物的方法。如圖4A所描繪的，該方法涉及使兩群細胞與一個含有靶向表觀遺傳基因的CRISPR引導序列及支架序列(例如：條碼化的CRISPR核酸庫)以及一個Cas9核酸內切酶的組合性核酸庫接觸。該兩群細胞被培養一段不同的時間。舉例而言，一群細胞可培養15天，另一群細胞培養20天。以，例如定序方法，鑑定該二或多個CRISPR引導序列及支架序列的組合物。舉例而言，可藉由定序一個對於CRISPR引導序列而言為一種獨特辨識物的條碼序列，而鑑定該CRISPR引導序列及支架序列。在培養一段較長時間的群細胞中每個CRISPR引導序列及支架序列的組合物的豐度與在培養一段較短時間的群細胞中每個CRISPR引導序列及支架序列的組合物的豐度進行比較。使細胞增生減少的CRISPR引導序列及支架序列組合物在培養一段較長時間的細胞群中的豐度較低。這樣的組合物被鑑定為可減少或防止細胞增生的被抑制劑靶向的表觀遺傳基因組合物。

【0102】 在某些實施例中，靶向關聯於本發明的表觀遺傳基因的一或多個基因或抑制劑係表現於一個重組表現載體中。如本文所使用的「載體」可以是可藉由限制酶作用及連接作用(例如，使用CombiGEM方法)，或藉由重組而在不同基因環境下進行輸送，或用於在宿主(例如，癌細胞)中表現而插入所需的一個序列或多個序列的數種核酸中的任何一種。典型而言，載體是由DNA構成，雖然亦有RNA載體。載體包括，但不限定於：質體、fosmid、噬菌粒、病毒基因組以及人工染色體。在某些實施例中，該載體為

慢病毒載體。在某些實施例中，二或多個靶向表觀遺傳基因的基因或抑制劑可在相同的重組表現載體中表現。在某些實施例中，二或多個靶向表觀遺傳基因的基因或抑制劑可在二或多個重組表現載體中表現。

【0103】 選殖載體為一種可自主複製或嵌入於宿主細胞基因組中的載體，其進一步的特徵在於具有一或多個核酸內切酶的限制位點，使得該載體可由可決定的方式進行剪切，並且可與一個所需的DNA序列進行連接，或者具有重組位點，其可容許具有相容端的插入物得以嵌入其中，使得新的重組載體可保有其在宿主細胞中複製的能力。在質體的例子中，所需序列可隨著在宿主細胞，例如宿主細菌中該質體的拷貝數目的增加而發生數次的複製，或者在該宿主進行有絲分裂而複製前，僅僅發生一次複製。在噬菌體的例子中，複製可主動地發生在裂解階段中或被動地發生在溶原性階段中。

【0104】 表現載體是一種可藉由限制及連接或重組反應將所需DNA序列插入其中的載體，使得其可連接於基因調控序列，且可能可以RNA轉錄體的方式表現。載體可進一步含有一或多個適用於鑑定被轉化或尚未被轉化或者經由載體轉染的細胞的標誌序列。標誌序列包括，例如，編碼了可增加或減少針對抗生素或其他化合物的抗性或敏感性的蛋白質基因、編碼了可藉由本領域所習知的標準檢測法偵測出其活性的酵素基因(例如：β半乳糖苷酶、螢光素酶或鹼性磷酸酶)、以及明顯可影響被轉化或轉染細胞、宿主、殖株或斑塊(例如：綠螢光蛋白、紅螢光蛋白)的表型的基因。較佳的載體為能夠自主複製並表現出存在於DNA片段中且可操作性連接的結構基因產物。

【0105】 如本文所使用的編碼序列及調控序列係指「可操作性地」連接，當其以此方式共價連接時，其係將編碼序列的表現或轉錄置於調控序列的影響或控制下。如果期望中，編碼序列可轉譯成一個功能性蛋白質，則兩個DNA序列可被認為是可操作性地連接，其係假使在5'端調控序列上誘導一個啟動子而造成該編碼序列的轉錄，且假使在該兩個DNA序列間的鏈結不會(1)造成引入一個架構轉移突變；(2)干擾啟動子區域引導該編碼序列進行轉錄的能力；或(3)干擾對應的RNA轉錄體轉譯成蛋白質的能力。因此，假使一個啟動子區域能夠影響DNA序列的轉錄，使得所得的轉錄體可被轉譯成所需蛋白質或聚胜肽，則該啟動子區域係可操作性地連接於該編碼區域。

【0106】 當核酸表現於細胞中時，可使用各種不同的轉錄控制序列(例如，啟動子/增強子序列)引導其表現。啟動子可為天然啟動子，亦即，該基因的啟動子係位於其內源性的上下游，其可提供正常的基因調控作用。在某些實施例中，該啟動子可以是組成型的，亦即，該啟動子係被上調，而可持續地轉錄其關聯基因。亦可使用各種條件式啟動子，例如在存有或缺乏某一分子的情況下而被調控的啟動子。在某些實施例中，該啟動子為第二型RNA聚合酶啟動子，例如哺乳動物第二型RNA聚合酶啟動子。在某些實施例中，該啟動子為人類泛素C啟動子(UBCp)。在某些實施例中，該啟動子為病毒啟動子。在某些實施例中，該啟動子為人類巨細胞病毒啟動子(CMVp)。在某些實施例中，該啟動子為第三型RNA聚合酶啟動子。第三型RNA聚合酶啟動子的實例包括，但不限定於H1啟動子、U6啟動子、小鼠U6啟動子、豬U6啟動子。在某些實施例中， 該啟動子為U6啟動子(U6p)。

【0107】 進行基因表現所需的調控序列的精確特質在物種或細胞種類間有所不同，但大致上應包括，有必要的，分別參與轉錄以及轉譯的起始的5’端非轉錄以及5’端非轉譯序列，例如TATA盒、帽序列、CAAT序列等。尤其是，這類5’端非轉錄的調控序列將包括一個啟動子區域，其包括一個用於可操作性地連接基因的轉錄控制的啟動子序列。調控序列亦可包括有需要的增強子序列，或上游激活子序列。本發明的載體可視情況選擇包括5’端前導或訊息序列。合適載體的選擇及設計是在本領域技術人員的能力及判斷內。

【0108】 含有所有用於基因表現的必需元件的表現載體可由商業上獲得，且為本領域技術人員所習知。參見，例如，Sambrook et al., Molecular Cloning: A Laboratory Manual, Fourth Edition, Cold Spring Harbor Laboratory Press, 2012。以基因工程方式將異源DNA(RNA)引入細胞中。異源性DNA(RNA)係被置於可操作的轉錄控制元件下，以允許在宿主細胞中表現該異源性DNA。

【0109】 可使用本領域所習知的方法及技術將關聯於本發明的一個核酸分子引入一個細胞或多個細胞中。舉例而言，可藉由例如轉化的標準程序將核酸分子引入，包括化學轉化及電穿孔、病毒輔導、粒子轟擊等。某些實施例中，可使用慢病毒而達成病毒輔導。表現核酸分子亦可藉由將該核酸分子嵌入基因組中而達成。

【0110】 本發明在其應用上並不限定於由下述說明的或附圖中所示的結構細節及部件的佈置。本發明能夠進行其它實施例，且可以各種方式被實踐或被執行。同樣地，本文所使用的措辭及術語是為了描述的目的，

而不應被認為是限制性。使用「包括」、「包含」或「具有」、「含有」、「涉及」及其變體，是指包括其後列出的項目及其等同物以及額外的項目。

【0111】 本發明透過下面的實例進一步說明，其不應該被解釋為進一步限制。在本申請案中所引用的所有參考資料(包括參考文獻、授權的專利、公開的專利申請、以及共同未決的專利申請）的全部內容均於此明確的引入作為參考，特別是對於本文引用的教導。

【0112】 實例

【0113】 實例1

【0114】 如圖1所示，可使用本文所述的方法製作一種高複合度的條碼化CRISPR分子的組合性核酸庫。步驟1中，使用陣列式寡核苷酸合成法產生一種大量的CRISPR引導RNA(gRNA)核酸庫，其包括針對每個CRISPR引導序列的正向及反向寡核苷酸(例如：Oligo F-A和Oligo R-A；Oligo F-B和Oligo R-B)。被合成的寡核苷酸的長度與終端產物庫的複合度無關。步驟2中，進行針對gRNA的寡核苷酸對的分離及黏合。針對每個gRNA的寡核苷酸含有20個鹼基對的CRISPR引導序列、兩個BbsI位點、以及一個條碼元件。寡核苷酸亦可含有用於使該寡核苷酸連接於儲存載體的5’及3’端單股突出區域。步驟3中，將含有BbsI及MfeI突出端的被黏合的寡核苷酸對匯集在一起，而用於單鍋連接反應，以將gRNA庫插入至被BbsI及MfeI剪切的AWp28儲存載體上。於是產生一個AWp28儲存載體庫，每個皆含有一個CRISPR引導序列、兩個BbsI位點，以及一個條碼元件

【0115】 步驟4中，接著使用BbsI使匯集的儲存載體庫進行單鍋剪切反應，而將該載體上介於該CRISPR引導序列和條碼元件之間打開，以利在

該CRISPR引導序列和該條碼元件之間插入gRNA支架元件。由於gRNA支架元件的插入物含有一個在該支架元件3'端由限制酶識別位點BamHI和EcoRI所形成的分隔位點，可針對該gRNA儲存載體庫進行迭代的選殖步驟，以接續地產生更多複合度的含有多重gRNA表現區域的n次條碼化核酸庫，其可使用Cas9核酸酶或效應物而介導組合性的基因敲除、激活或抑制，如步驟5所示。簡言之，使用與位於分隔位點的該限制酶識別位點(例如：BamHI和EcoRI)互補的限制酶剪切該條碼化引導RNA庫。剪切作用使其得以插入編碼一個CRISPR引導序列、支架序列、分隔位點及條碼元件的額外片段。因為相同組的限制酶可用於多個回合的選殖步驟，以建立一個n次核酸庫，此策略的一個優點為(n+1)次核酸庫不需要增加所需限制酶組的數目。

【0116】 為了建立一個(n)-次的(m)gRNA核酸庫，其會進行(n+2)回合的選殖步驟。本策略的額外優點是可使用相同的CRISPR引導RNA庫以產生更高複合度核酸庫。

【0117】 實例2：

【0118】 如圖2A所示，可使用本文所述的方法製作一個條碼化CRISPR分子的複合組合性核酸庫。步驟1中，使多重CRISPR引導序列與一個單一的條碼元件在單一核苷酸上合成，以產生一個大量的組合性gRNA核酸庫。在每個該CRISPR引導序列之後存有限制酶識別位點。藉由不同的限制酶位點將引導序列和條碼連接在一起。圖2A的基因構築體和載體顯示出一種含有三個CRISPR引導序列及一個位於該CRISPR引導序列下游的條碼元件的示例性核苷酸。

【0119】 步驟2中，以單鍋組裝法將匯集合成的核苷酸連接於標的載

體上。標的載體可含有一個啟動子，以起始至少一個CRISPR引導序列的表現。如步驟3所示，依序在每個CRISPR引導序列之後的限制酶識別位點上以不同限制酶剪切載體，而得以插入支架元件，而在某些例子中，插入一個啟動子元件以起始下游CRISPR引導序列的表現。該方法因而可產生一個編碼了多重CRISPR引導序列及支架序列且帶有一個單一條碼元件的條碼化組合性引導RNA庫，而藉由使用Cas9核酸酶或效應物以進行組合性的基因敲除、激活或抑制。

【0120】 圖2B描繪產生一個條碼化CRISPR分子的複合組合性核酸庫的另可選擇策略。步驟1中，步驟1中，使多重CRISPR引導序列與一個單一的條碼元件在單一核苷酸上合成，以產生一個大量的組合性gRNA核酸庫。在每個該CRISPR引導序列以及一個在該第一CRISPR引導序列上游的限制位點之後存有限制酶識別位點，其用於插入一個啟動子元件。藉由不同的限制酶位點將引導序列和條碼連接在一起。圖2B的基因構築體和載體顯示出一種含有三個CRISPR引導序列及一個位於該CRISPR引導序列上游的條碼元件的示例性核苷酸。

【0121】 步驟2中，以單鍋組裝法將匯集合成的核苷酸連接於標的載體上。標的載體可含有一個啟動子，以起始至少一個CRISPR引導序列的表現。如步驟3所示，依序在每個CRISPR引導序列之後的限制酶識別位點上以不同限制酶剪切載體，而得以插入支架元件，而在某些例子中，插入一個啟動子元件以起始下游CRISPR引導序列的表現。該方法因而可產生一個編碼了多重CRISPR引導序列及支架序列且帶有一個單一條碼元件的條碼化組合性引導RNA庫，而藉由使用Cas9核酸酶或效應物以進行組合性的基

因敲除、激活或抑制。

【0122】 使用圖2A及2B中所描繪的策略，可藉由(n+1)回合的選殖步驟建立一個(n)次的(m)gRNA核酸庫。該核酸庫的複合度取決於步驟1中合成的核苷酸長度。因為針對每個CRISPR引導序列以及啟動子元件而言，用於插入支架序列的限制酶識別位點各自不同，增加核苷酸的複合度(即，CRISPR引導序列的數目)亦可增加剪切步驟所需的限制酶數目。

【0123】 實例3：

【0124】 使用根據CombiGEM的DNA組裝方法以有效且可擴增地組裝條碼化組合性gRNA核酸庫。可藉由慢病毒將該核酸庫輸送至人類細胞中，藉此製造出基因上高度多樣性的帶有獨特gRNA組合物的細胞群，因而可藉由在匯集的檢測法中進行條碼定序而追蹤。如此的策略，稱為CombiGEM-CRISPR，其係使用簡易的單鍋選殖步驟，而得以進行可擴增式高度組合性gRNA核酸庫的組裝，因而針對組合性基因功能性的系統性分析而言，可以簡易化且可加速其進行。

【0125】 為了創建初始的條碼化sgRNA核酸庫，首先合成編碼了條碼化gRNA標靶序列庫的寡核苷酸對陣列，並在等比例下進行黏合、匯集，而用於選殖儲存載體中下游的U6啟動子(圖1)。接著，在單鍋連接反應中，將gRNA的支架序列插入該儲存載體庫中。將CombiGEM方法用於高度組合性gRNA核酸庫的可擴增式組裝(圖1)。在條碼化的sgRNA構築體中，將BamHI和EcoRI位點置於gRNA序列與其條碼之間，而BglII和MfeI則位於其末端。策略性地放置這些限制酶位點，可使條碼在酵素剪切後從其gRNA序列中產生間隔化，以及在與插入物連接後代表其相對應的gRNA的條碼形成串聯結

構。為了建構一次核酸庫，以BglII和MfeI限制酶剪切該儲存載體，並使其連接於被BamHI和EcoRI剪切的慢病毒標的載體上與其相容的DNA末端，而製備該條碼化sgRNA表現單元的匯集插入物。然後，一次核酸庫可做為標的載體，而用於下一回合進行條碼化sgRNA表現單元的匯集式插入，而產生二次核酸庫，其中代表著每個sgRNA的條碼係位於每個慢病毒構築體的一端。可迭代地重複此程序，而產生高次的條碼化組合性gRNA核酸庫。可藉由偵測串聯式的條碼而追蹤組合性gRNA的識別性，對於每個組合物而言，該條碼是獨特的(例如，藉由高通量定序)。

【0126】 為了驗證慢病毒組合性gRNA表現系統的功能性，建構靶向編碼了綠螢光蛋白(GFP)和紅螢光蛋白(RFP)的序列的gRNA組合物(表1)。使用流式細胞儀(圖6A和6B)及螢光顯微鏡(圖6E)測定該組合性基因的擾動表型。使用攜帶雙重RFP及GFP報導基因連同條碼化組合性gRNA表現單元的慢病毒感染穩定表現了經人類密碼子最適化的Cas9核酸酶(*OVCAR8-ADR-Cas9*)的人類卵巢癌細胞(OVCAR8-ADR) (Honma, et al. *Nat. Med.* (2008) 14: 939-948)(圖6A)。可預見的是，具活性的gRNA將會靶向編碼了GFP及RFP的序列，而產生插入缺失序列以敲除GFP及RFP的表現。可觀察到GFP及RFP螢光強度被有效地抑制，主因是相較於<0.7%的載體對照組而言，在感染後4及8天的兩組中，GFP及RFP雙重陰性的細胞群為觀察到的主要帶有Cas9核酸酶和靶向RFP及GFP兩者的gRNA表現單元的主要細胞群(總細胞群中佔~83至97%)(圖6C及6D)。並無法從表現有靶向GFP及/或RFP但沒有Cas9核酸酶的gRNA的對照組細胞株中觀察到這樣的抑制現象(圖6B)。基因擾動的特異性已被確認，因為僅攜帶有靶向GFP的sgRNA的細胞

呈現出GFP訊號的消失，而非RFP訊號。相似地，含有靶向RFP的sgRNA的細胞呈現出RFP表現量的降低，但在GFP表現量上並無影響(圖6E)。這些結果證明，慢病毒載體編碼組合性gRNA構築體而可抑制多重基因表現的能力係同時存在於單一人類細胞中。

【0127】 多樣化的表觀遺傳修飾作用往往共同作用於調節基因表現的模式(Wang, et al. *Nat. Genet.* (2008) 40:897-903)，且組合性表觀遺傳的調控是新興且可望作為癌症治療的有效策略(Dawson, et al. *Cell* (2012) 150: 12-27; Juergens, et al. *Cancer Discov.* (2011) 1: 598-607)。使用如本文所述的CombiGEM-CRISPR方法及組成物，有系統地驗證表觀遺傳基因擾動對於抗癌表型的組合性效果。建構一個含有153個靶向一組50個表觀遺傳基因(每個基因有3個sgRNA) 的條碼化sgRNA和3個根據GeCKOv2核酸庫的對照組sgRNA的核酸庫(Shalem, et al. *Science* (2014) 343:84-87)(表1)。

【0128】 表1：sgRNA標靶序列

sgRNA ID	**sgRNA 標靶序列**
GFP-sg1	GGGCGAGGAGCTGTTCACCG (SEQ ID NO: 7)
RFP-sg1	CACCCAGACCATGAAGATCA (SEQ ID NO: 8)
RFP-sg2	CCACTTCAAGTGCACATCCG (SEQ ID NO: 9)
對照組-sg1	ATCGTTTCCGCTTAACGGCG (SEQ ID NO: 10)
對照組-sg2	AAACGGTACGACAGCGTGTG (SEQ ID NO: 11)
對照組-sg3	CCATCACCGATCGTGAGCCT (SEQ ID NO: 12)
DNMT1-sg1	CTAGACGTCCATTCACTTCC (SEQ ID NO: 13)
DNMT1-sg2	TTTCCAAACCTCGCACGCCC (SEQ ID NO: 14)
DNMT1-sg3	ACGTAAAGAAGAATTATCCG (SEQ ID NO: 15)
DNMT3A-sg1	CCGCCCCACCTTCCGTGCCG (SEQ ID NO: 16)
DNMT3A-sg2	TGGCGCTCCTCCTTGCCACG (SEQ ID NO: 17)
DNMT3A-sg3	CCGCTCCGCAGCAGAGCTGC (SEQ ID NO: 18)
DNMT3B-sg1	AGAGTCGCGAGCTTGATCTT (SEQ ID NO: 19)
DNMT3B-sg2	ATCCGCACCCCGGAGATCAG (SEQ ID NO: 20)
DNMT3B-sg3	GAAGACTCGATCCTCGTCAA (SEQ ID NO: 21)
DNMT3L-sg1	AGGGATCTGCGCCCCATGTA (SEQ ID NO: 22)

DNMT3L-sg2	ACTCACCTTCTATATTTCGC (SEQ ID NO: 23)
DNMT3L-sg3	CACCAAAATCACGTCCATGC (SEQ ID NO: 24)
MBD1-sg1	TCACCCGTAGGCAACGTCGC (SEQ ID NO: 25)
MBD1-sg2	GTGTCCAGCGACGTTGCCTA (SEQ ID NO: 26)
MBD1-sg3	ACGTTGTGCAAAGACTGTCG (SEQ ID NO: 27)
MBD2-sg1	CGGCGACTCCGCCATAGAGC (SEQ ID NO: 28)
MBD2-sg2	GGAGCCGGTCCCTTTCCCGT (SEQ ID NO: 29)
MBD2-sg3	AGTCTTGAAAGCGCATGCCA (SEQ ID NO: 30)
CREBBP-sg1	AGCGGCTCTAGTATCAACCC (SEQ ID NO: 31)
CREBBP-sg2	GAATCACATGACGCATTGTC (SEQ ID NO: 32)
CREBBP-sg3	CCCGCAAATGACTGGTCACG (SEQ ID NO: 33)
EP300-sg1	CTGTCAGAATTGCTGCGATC (SEQ ID NO: 34)
EP300-sg2	CTTGGCAAGACTTGCCTGAC (SEQ ID NO: 35)
EP300-sg3	TAGTTCCCCTAACCTCAATA (SEQ ID NO: 36)
HDAC1-sg1	ACACCATTCGTAACGTTGCC (SEQ ID NO: 37)
HDAC1-sg2	TCACTCGAGATGCGCTTGTC (SEQ ID NO: 38)
HDAC1-sg3	AGAATGCTGCCGCACGCACC (SEQ ID NO: 39)
HDAC2-sg1	TCCGTAATGTTGCTCGATGT (SEQ ID NO: 40)
HDAC2-sg2	TCCAACATCGAGCAACATTA (SEQ ID NO: 41)
HDAC2-sg3	TACAACAGATCGTGTAATGA (SEQ ID NO: 42)
SIRT1-sg1	GTTGACTGTGAAGCTGTACG (SEQ ID NO: 43)
SIRT1-sg2	AACAGGTTGCGGGAATCCAA (SEQ ID NO: 44)
SIRT1-sg3	TACCCAGAACATAGACACGC (SEQ ID NO: 45)
CARM1-sg1	CTCGCCGTTCGCGTCGCCGA (SEQ ID NO: 46)
CARM1-sg2	CCCGTACTCACGGCTGTAGA (SEQ ID NO: 47)
CARM1-sg3	GGGCCACGTACCGTTGGGTG (SEQ ID NO: 48)
EZH1-sg1	ACAGGCTTCATTGACTGAAC (SEQ ID NO: 49)
EZH1-sg2	AGCTGATCAATAACTATGAT (SEQ ID NO: 50)
EZH1-sg3	CCTCATCTGAGTACTGATTC (SEQ ID NO: 51)
EZH2-sg1	ACACGCTTCCGCCAACAAAC (SEQ ID NO: 52)
EZH2-sg2	TGCGACTGAGACAGCTCAAG (SEQ ID NO: 53)
EZH2-sg3	AAAACTTCATCTCCCATATA (SEQ ID NO: 54)
MLL-sg1	GTACAAATTGTACGACGGAG (SEQ ID NO: 55)
MLL-sg2	GACCCCTCGGCGGTTTATAG (SEQ ID NO: 56)
MLL-sg3	TATATTGCGACCACCAAACT (SEQ ID NO: 57)
MLL2-sg1	CAGAGAGCACAACGCCGCAC (SEQ ID NO: 58)
MLL2-sg2	GGAACCGCTGGCAGTCGCGC (SEQ ID NO: 59)
MLL2-sg3	CTCCCGCTGCCCGTGTAGAC (SEQ ID NO: 60)
NSD1-sg1	CTGGCTCGAGATTTAGCGCA (SEQ ID NO: 61)
NSD1-sg2	AATCTGTTCATGCGCTTACG (SEQ ID NO: 62)
NSD1-sg3	GATTCCAGTACCAGTACATT (SEQ ID NO: 63)
PRMT1-sg1	CTCACCGTGGTCTAACTTGT (SEQ ID NO: 64)

PRMT1-sg2	GGATGTCATGTCCTCAGCGT (SEQ ID NO: 65)
PRMT1-sg3	TTTGACTCCTACGCACACTT (SEQ ID NO: 66)
PRMT2-sg1	CGTGGATGAGTACGACCCCG (SEQ ID NO: 67)
PRMT2-sg2	TCTTCTGTGCACACTATGCG (SEQ ID NO: 68)
PRMT2-sg3	CTGTCCCAGAAGTGAATCGC (SEQ ID NO: 69)
PRMT3-sg1	GCCATGTGCTCGTTAGCGTC (SEQ ID NO: 70)
PRMT3-sg2	GCCTGACGCTAACGAGCACA (SEQ ID NO: 71)
PRMT3-sg3	GAATTCATGTACTCAACTGT (SEQ ID NO: 72)
PRMT5-sg1	CGGAATGCGGGGTCCGAACT (SEQ ID NO: 73)
PRMT5-sg2	CAGCATACAGCTTTATCCGC (SEQ ID NO: 74)
PRMT5-sg3	ATGAACTCCCTCTTGAAACG (SEQ ID NO: 75)
PRMT6-sg1	ATTGTCCGGCGAGGACGTGC (SEQ ID NO: 76)
PRMT6-sg2	CTTCGCCACGCGCTGTCTCA (SEQ ID NO: 77)
PRMT6-sg3	GACGGTACTGGACGTGGGCG (SEQ ID NO: 78)
PRMT7-sg1	CAATCCGACCACGGGGTCTG (SEQ ID NO: 79)
PRMT7-sg2	GAGGTTCAAACCGCCTGCTA (SEQ ID NO: 80)
PRMT7-sg3	TAAAGTCGGCTGGTGACACC (SEQ ID NO: 81)
SETD2-sg1	AGTTCTTCTCGGTGTCCAAA (SEQ ID NO: 82)
SETD2-sg2	GACTATCAGTTCCAGAGATA (SEQ ID NO: 83)
SETD2-sg3	AACTTACGAAGGAAGGTCTT (SEQ ID NO: 84)
KDM1A-sg1	TTACCTTCGCCCGCTTGCGC (SEQ ID NO: 85)
KDM1A-sg2	CCGGCCCTACTGTCGTGCCT (SEQ ID NO: 86)
KDM1A-sg3	AGAGCCGACTTCCTCATGAC (SEQ ID NO: 87)
KDM1B-sg1	CATACCGCATCGATAAGTCT (SEQ ID NO: 88)
KDM1B-sg2	ATAGCCAAGACTTATCGATG (SEQ ID NO: 89)
KDM1B-sg3	GAACATACCTTCTGTAGTAA (SEQ ID NO: 90)
KDM2A-sg1	ACGCTACTATGAGACCCCAG (SEQ ID NO: 91)
KDM2A-sg2	TATGGCAGGGAGTCGTCGCA (SEQ ID NO: 92)
KDM2A-sg3	GTAACGAATCCTTTCTTCTT (SEQ ID NO: 93)
KDM2B-sg1	CCTCGTTCTCGTCGTATCGC (SEQ ID NO: 94)
KDM2B-sg2	GCGTTACTACGAGACGCCCG (SEQ ID NO: 95)
KDM2B-sg3	CTTGGTCAAGCGTCCGACTG (SEQ ID NO: 96)
KDM3A-sg1	TAAATGCCGAGAGTGTCGCT (SEQ ID NO: 97)
KDM3A-sg2	GTCTGTCAAAACCGACTTCC (SEQ ID NO: 98)
KDM3A-sg3	GATACTGCTTGGCTGTACTG (SEQ ID NO: 99)
KDM3B-sg1	TCTTGTATGGGCGCCCCGTG (SEQ ID NO: 100)
KDM3B-sg2	GCCTTGACTGTTACCGGCTC (SEQ ID NO: 101)
KDM3B-sg3	TCCTGAGCCGGTAACAGTCA (SEQ ID NO: 102)
KDM4A-sg1	ACTCCGCACAGTTAAAACCA (SEQ ID NO: 103)
KDM4A-sg2	GCGGAACTCTCGAACAGTCA (SEQ ID NO: 104)
KDM4A-sg3	TTCCACTCACTTATCGCTAT (SEQ ID NO: 105)
KDM4B-sg1	CCCCGCGTACTTCTCGCTGT (SEQ ID NO: 106)

KDM4B-sg2	GTATGATGACATCGACGACG (SEQ ID NO: 107)
KDM4B-sg3	TCACCAGGTACTGTACCCCG (SEQ ID NO: 108)
KDM4C-sg1	CCTTTGCAAGACCCGCACGA (SEQ ID NO: 109)
KDM4C-sg2	AGTAGGCTTCGTGTGATCAA (SEQ ID NO: 110)
KDM4C-sg3	GTCTAAAGGAGCCCATCGTG (SEQ ID NO: 111)
KDM5A-sg1	CATGAACCCCAACGTGCTAA (SEQ ID NO: 112)
KDM5A-sg2	CTGGGATTCAAATAACTCGG (SEQ ID NO: 113)
KDM5A-sg3	TCTCTGGTATGAAAGTGCCG (SEQ ID NO: 114)
KDM5B-sg1	GTCCGCGAACTCTTCCCAGC (SEQ ID NO: 115)
KDM5B-sg2	TCGAAGACCGGGCACTCGGG (SEQ ID NO: 116)
KDM5B-sg3	GGACTTATTTCAGCTTAATA (SEQ ID NO: 117)
KDM5C-sg1	CTTACCGCCATGACACACTT (SEQ ID NO: 118)
KDM5C-sg2	GATAAACAATGCGTTCGTAG (SEQ ID NO: 119)
KDM5C-sg3	GGGCTACCCGAGCCCACCGA (SEQ ID NO: 120)
KDM5D-sg1	GATTTACTCCTCGCGTCCAA (SEQ ID NO: 121)
KDM5D-sg2	AAAGACTTACCGCGGGTGGG (SEQ ID NO: 122)
KDM5D-sg3	TAAGGCCCGACATGGAACCG (SEQ ID NO: 123)
KDM6A-sg1	ACTGTAAACTGTAGTACCTC (SEQ ID NO: 124)
KDM6A-sg2	CAGCATTATCTGCATACCAG (SEQ ID NO: 125)
KDM6A-sg3	AGACTATGAGTCTAGTTTAA (SEQ ID NO: 126)
KDM6B-sg1	TACCACAGCGCCCTTCGATA (SEQ ID NO: 127)
KDM6B-sg2	ATCCCCCTCCTCGTAGCGCA (SEQ ID NO: 128)
KDM6B-sg3	CAAAGGCTTCCCGTGCAGCG (SEQ ID NO: 129)
PHF2-sg1	TTCTGCACGGGCTTGACGTC (SEQ ID NO: 130)
PHF2-sg2	GACGTCAAGCCCGTGCAGAA (SEQ ID NO: 131)
PHF2-sg3	CAGTGACGTCGAGAACTACG (SEQ ID NO: 132)
PHF8-sg1	TCTGACGAACGTAGGGCTCC (SEQ ID NO: 133)
PHF8-sg2	GGCTTAGTGAAAAAACGCCG (SEQ ID NO: 134)
PHF8-sg3	CCTCGCCATCATTCACTGTG (SEQ ID NO: 135)
BMI1-sg1	AACGTGTATTGTTCGTTACC (SEQ ID NO: 136)
BMI1-sg2	TCCTACCTTATATTCAGTAG (SEQ ID NO: 137)
BMI1-sg3	AAAGGTTTACCATCAGCAGA (SEQ ID NO: 138)
BRD1-sg1	CACCGTGTTCTATAGAGCCG (SEQ ID NO: 139)
BRD1-sg2	CGGCGCGAGGTGGACAGCAT (SEQ ID NO: 140)
BRD1-sg3	CGACTCACCGGCTGCGATCC (SEQ ID NO: 141)
BRD3-sg1	CGACGTGACGTTTGCAGTGA (SEQ ID NO: 142)
BRD3-sg2	CAAAGGTCGGAAGCCGGCTG (SEQ ID NO: 143)
BRD3-sg3	CATCACTGCAAACGTCACGT (SEQ ID NO: 144)
BRD4-sg1	ACTAGCATGTCTGCGGAGAG (SEQ ID NO: 145)
BRD4-sg2	TCTAGTCCATCCCCCATTAC (SEQ ID NO: 146)
BRD4-sg3	GGGAACAATAAAGAAGCGCT (SEQ ID NO: 147)
ING1-sg1	GAGATCGACGCGAAATACCA (SEQ ID NO: 148)

ING1-sg2	TATAAATCCGCGCCCGAAAG (SEQ ID NO: 149)
ING1-sg3	CCCATACCAGTTATTGCGCT (SEQ ID NO: 150)
ING2-sg1	GCAGCAGCAACTGTACTCGT (SEQ ID NO: 151)
ING2-sg2	GCAGCGACTCCACGCACTCA (SEQ ID NO: 152)
ING2-sg3	GATCTTCAAGAAGACCCCGC (SEQ ID NO: 153)
ING3-sg1	TCTCGCGCATTTCCGTGAAG (SEQ ID NO: 154)
ING3-sg2	CTTCACGGAAATGCGCGAGA (SEQ ID NO: 155)
ING3-sg3	TCGATACTGCATTTGTAATC (SEQ ID NO: 156)
ING4-sg1	GCTGCTCGTGCTCGTTCCAA (SEQ ID NO: 157)
ING4-sg2	CCTAGAAGGCCGGACTCAAA (SEQ ID NO: 158)
ING4-sg3	GGCACTACTCATATACTCAG (SEQ ID NO: 159)
ING5-sg1	GATCTGCTTCAAAGCGCGCC (SEQ ID NO: 160)
ING5-sg2	CTTCCAGCTGATGCGAGAGC (SEQ ID NO: 161)
ING5-sg3	GAAGTTCCTCTGAAGTTCGC (SEQ ID NO: 162)

【0129】 使用qRT-PCR在 OVCAR8-ADR細胞中驗證這些50個表觀遺傳基因。使用CombiGEM 方法產生一個二次 (153 x 153 sgRNAs = 總共23,409種組合物)匯集式條碼化gRNA核酸庫。製造慢病毒庫以將該庫輸送至OVCAR8-ADR-Cas9細胞中。分離由匯集細胞群所得的基因組DNA並藉由聚合酶鏈鎖反應(PCR)而進行無偏差條碼擴增。使用Illumina HiSeq 定序法定量儲存於大腸桿菌以及被感染的人類細胞庫中的個別條碼化組合物(圖3A-3D)。在質體及被感染細胞庫內達成二次核酸庫的近全覆蓋，而每個樣本中具有介於23至34百萬個讀值(圖3B)，並且可觀察到條碼化gRNA組合物表示值的分佈是相對平均的(圖 3A和3B)。其次，在質體及被感染細胞庫之間的條碼呈現值具有高度關聯性(圖3C)，且在被感染細胞庫生物性複製體中所呈現的條碼具有高度再現性(圖3D)。因此，可使用CombiGEM-CRISPR以有效地組裝並輸送條碼化組合性gRNA核酸庫至人類細胞中。

【0130】 為了確認gRNA對於剪輯OVCAR8-ADR-Cas9細胞中內源性基因的功能，進行Surveyor檢測法以估算在8個從核酸庫中隨機挑選並由gRNA靶向的位點的剪切效率。在感染後12天，可觀察到插入缺失序列產生

的效率介於1.9%至26.2%之間(圖7A-7C)。可偵測到感染後12天在所有gRNA靶向的位點上DNA錯誤配對的剪切(圖 7A和8A)。在本文雙重gRNA系統內的多重位點上可測定出同時剪切效率，且在表現了個別gRNA或雙重gRNA的細胞中可觀察到相當的剪切量(圖7A和8A)。亦可偵測到在個別gRNA及雙重gRNA表現細胞中靶向蛋白質含量的缺少(圖8B)。這些結果說明多重系統並不會妨礙gRNA的活性。為了區別單一細胞與全部感染細胞群之間由雙重gRNA所定向的雙重剪切效果，先分離以雙重gRNA感染的單一細胞所得的殖株，並使用Sanger定序法偵測在兩個靶向的基因組位點上有插入、缺失或突變現象的細胞(圖9A及9B)。

【0131】 藉由在標靶基因組位點上進行深度定序以估算插入缺失序列的產生效率。在不同的gRNA間，可觀察到插入缺失序列產生的速率(即，14至93%；圖16A)以及在架構轉移突變上(即，所有插入缺失序列中有52至95%；圖5B)有大的變異。此外，在先前研究中經A375黑色素瘤細胞驗證的gRNA(Shalem et al. 2014 *Science* 343:84-7)在OVCAR8-ADR-Cas9細胞中展現出活性降低的現象(例如，對於NF1-sg4及MED12-sg1 sgRNAs而言)，以及產生差異化插入缺失序列的傾向(例如，對於NF1-sg1 sgRNA而言)(圖16C)。這種差異可能部分因為在不同種類細胞中，標靶位點上染色質可及性的變異(Wu et al. (2014) Nat. *Biotechnol.* 32: 670-6)以及DNA斷裂修復的機制(Ghezraoui et al (2014) *Mol Cell* 55:829-42)可能有所不同所致。在gRNA設計最適化的持續努力，包括改善對靶剪切率(Donesch et al (2014) *Nat Biotechnol* 32:1262-7)及使脫靶剪切現象最小化，應可創建出更有效的gRNA組，進而可改善其在廣範圍細胞種類中進行大規模基因擾動篩選的應用

性。插入缺失序列的產生可進一步藉由在多重系統中的gRNA進行評估。深度定序分析偵測出與在sgRNA或雙重gRNA系統下表現出的相同gRNA而言，可比較的插入缺失序列產生頻率及傾向(圖16D)。為了在單細胞中區分出由雙重gRNA導引的雙重剪切事件與分佈於整個細胞群的剪切事件，從感染了雙重gRNA構築體的單細胞群中分離出殖株。偵測到在兩個標靶基因組位點上有基因插入、缺失或突變的細胞(圖9A-9C；表6)。本文結果說明CombiGEM組合性gRNA核酸庫可用於在OVCAR8-ADR-Cas9細胞中產生雙重基因突變體。

【0132】 首先以OVCAR8-ADR-Cas9細胞進行匯集式組合性基因篩選以鑑定可調控癌細胞增生的gRNA組合物。根據各種不同的參數，建構一個數學模型而比對出在一個群體中各個核酸庫成員豐度的相對改變量(參見以下方法；圖17A和17B)。使用含有異質次群體且攜帶不同gRNA組合物的群體進行模擬。特別地，在模擬開始時，先定義整個群體的特定百分比為攜帶抗增生性(f_s)及前期增生性(f_f)gRNA組合物的次群體。在模擬開始時，於每個次群體中，藉由CRISPR-Cas9系統使一個分率(p)的細胞產生突變，而產生一個經修飾的倍增時間($T_{doubling,m}$)。此模型說明了在整個細胞群體中，在模擬的條件下(即，f_s及 f_f = 2、5、或10%；p = 0.2、0.4、0.6、0.8、或1.0；$T_{doubling,m}$= 36、48、或60小時)，帶有抗增生性gRNA組的條碼化細胞呈現值可衰減約23至97%(圖17B)。一般而言，增加突變效率、增加抗增生性細胞的倍增時間、減少前期增生性細胞的倍增時間、以及在群體中增加前期增生性組合物的百分比(圖17C)，預期可導致在整個群體中對於抗增生性條碼而言有較大的條碼量衰減。

【0133】 在試驗篩選中，為了鑑定可調控癌細胞增生的gRNA組合物，以二次組合性gRNA核酸庫感染OVCAR8-ADR-Cas9細胞群，並培養15及20天，之後，從該細胞中分離出基因組DNA以進行嵌入條碼的無偏差擴增以及定量(圖4A、10A和10B)。比較培養20天及15天的組別間的條碼豐度(針對每百萬個讀值進行標準化)可產生$\log_2$(條碼計數比值)數值(圖4A、10A和10B)。預期中，可抑制細胞增生的引導RNA組合物會產生負數的$\log_2$比值，而具有生長優勢的匯合細胞被預期具有正數的$\log_2$比值。為了降低變異性，在15天組別中，將少於~100個絕對讀值的組合物濾除，並將針對每個gRNA對(即，*sgRNA-A* + *sgRNA-B*以及*sgRNA-B* + *sgRNA-A*)的兩種潛在排列的$\log_2$比值取平均(圖11)。針對兩個生物性複製體測定針對每個gRNA組合物的$\log_2$比值，並排序(圖5B和12A)。在15天及20天組別中，大多數的gRNA組合物，包括作為內部對照組且來自GeCKOv2庫(Shalem et al (2014) *Science* 343: 84-7)而在人類基因組中不含對靶位點的三個對照組gRNA，在條碼呈現值上不會展現出顯著的變化。據顯示，在兩個生物性複製體中(Q-值<0.01，表2和圖12B)，有61個gRNA組合物表現出相當的抗增生性效果($\log_2$ 比值 < -0.90)，因而產生有潛力的基因組，並可進一步探討其抑制癌細胞生長的能力。

【0134】 表2：依照匯集式篩選而可抑制OVCAR8-ADR細胞增生的二次sgRNA篩中物

sgRNA-A	sgRNA-B	$\log_2$ 比值 – 20天/15天 複製體 1	$\log_2$ 比值 – 20天/15天 複製體 2	Z-分數 複製體 1	Z-分數 複製體 2	Q 值
BRD4_sg3	MLL_sg3	-2.61	-3.32	-8.45	-10.72	7.01E-33

BMI1_sg2	HDAC2_sg3	-3.91	-1.93	-12.62	-6.29	3.41E-32
BMI1_sg2	KDM1B_sg3	-1.34	-3.71	-4.39	-11.99	1.20E-23
ING3_sg3	BMI1_sg1	-3.68	-1.11	-11.87	-3.64	4.57E-21
BRD4_sg3	KDM6A_sg2	-2.30	-2.20	-7.45	-7.15	1.52E-18
BRD4_sg3	PHF2_sg2	-1.57	-2.88	-5.12	-9.31	4.05E-18
BMI1_sg2	PRMT6_sg1	-2.80	-1.59	-9.05	-5.19	1.19E-17
ING3_sg2	KDM5A_sg2	-2.28	-2.04	-7.40	-6.65	3.48E-17
KDM5B_sg3	MLL_sg3	-3.17	-1.05	-10.23	-3.47	2.61E-16
KDM6A_sg3	KDM6A_sg2	-1.95	-2.26	-6.35	-7.33	2.61E-16
BRD4_sg3	KDM4C_sg1	-1.48	-2.28	-4.84	-7.40	7.88E-13
KDM6B_sg1	KDM3A_sg2	-0.98	-2.70	-3.23	-8.74	3.00E-12
ING3_sg3	KDM6A_sg2	-1.46	-2.10	-4.79	-6.81	2.02E-11
PRMT5_sg3	PRMT5_sg3	-1.54	-1.92	-5.04	-6.26	9.13E-11
BRD4_sg3	BRD4_sg2	-2.45	-1.00	-7.93	-3.30	1.24E-10
BMI1_sg2	NSD1_sg3	-2.19	-1.20	-7.13	-3.94	2.68E-10
BMI1_sg2	MBD2_sg3	-1.44	-1.93	-4.70	-6.27	3.94E-10
BMI1_sg2	KDM1A_sg1	-1.62	-1.40	-5.29	-4.59	5.85E-08
KDM3A_sg2	PRMT5_sg3	-1.28	-1.72	-4.21	-5.62	6.62E-08
BRD4_sg3	EP300_sg3	-1.28	-1.71	-4.22	-5.58	7.28E-08
BMI1_sg2	KDM3A_sg1	-1.47	-1.48	-4.79	-4.85	1.32E-07
BRD4_sg3	KDM6B_sg1	-1.20	-1.64	-3.94	-5.35	5.58E-07
PHF8_sg2	KDM1A_sg1	-1.11	-1.69	-3.65	-5.52	8.77E-07
KDM6A_sg3	HDAC2_sg3	-1.03	-1.77	-3.40	-5.77	8.77E-07
BRD4_sg3	KDM6B_sg2	-0.92	-1.86	-3.05	-6.05	1.13E-06
PRMT5_sg3	HDAC1_sg1	-1.26	-1.43	-4.15	-4.68	3.14E-06
PHF8_sg2	PRMT5_sg3	-1.60	-1.08	-5.23	-3.57	3.50E-06
BRD4_sg3	PHF2_sg1	-1.51	-1.13	-4.94	-3.71	5.35E-06
BRD4_sg3	EZH2_sg3	-1.02	-1.57	-3.38	-5.12	8.80E-06
PRMT5_sg3	EZH1_sg1	-0.96	-1.62	-3.19	-5.28	9.53E-06
KDM6A_sg2	KDM5A_sg2	-1.48	-1.04	-4.86	-3.44	1.72E-05
BRD4_sg3	KDM2B_sg2	-1.00	-1.49	-3.30	-4.88	2.59E-05
KDM6B_sg3	PRMT5_sg3	-1.03	-1.43	-3.41	-4.67	3.60E-05
BRD4_sg1	KDM6B_sg1	-1.30	-1.15	-4.27	-3.79	3.74E-05
KDM1A_sg3	PRMT5_sg3	-1.52	-0.91	-4.96	-3.01	4.81E-05
PRMT5_sg3	EP300_sg2	-1.41	-1.00	-4.61	-3.31	5.44E-05
BRD4_sg3	PRMT5_sg2	-1.46	-0.91	-4.79	-3.02	7.59E-05
KDM5D_sg2	MLL_sg3	-1.01	-1.34	-3.35	-4.40	9.28E-05
BRD4_sg3	PRMT5_sg3	-1.03	-1.33	-3.39	-4.35	9.28E-05
PHF2_sg1	PRMT5_sg3	-1.19	-1.13	-3.92	-3.71	1.32E-04
PRMT5_sg2	DNMT1_sg1	-1.28	-1.02	-4.21	-3.37	1.54E-04
KDM6A_sg2	KDM5C_sg1	-1.28	-1.01	-4.19	-3.34	1.73E-04
BMI1_sg2	KDM6A_sg2	-0.96	-1.30	-3.17	-4.25	2.47E-04

KDM1A_sg1	PRMT5_sg3	-0.95	-1.29	-3.15	-4.24	2.68E-04
BRD4_sg2	KDM2B_sg1	-1.06	-1.17	-3.51	-3.85	2.82E-04
KDM4A_sg2	PRMT5_sg3	-1.07	-1.16	-3.52	-3.81	3.18E-04
PRMT6_sg2	PRMT5_sg2	-1.02	-1.13	-3.37	-3.74	6.20E-04
KDM2B_sg3	PRMT5_sg3	-1.13	-1.01	-3.73	-3.34	6.95E-04
KDM6B_sg1	MLL_sg3	-0.92	-1.19	-3.05	-3.90	9.39E-04
BRD4_sg3	BMI1_sg2	-0.94	-1.14	-3.11	-3.76	1.10E-03
KDM5A_sg3	NSD1_sg3	-0.91	-1.14	-3.03	-3.75	1.37E-03
BRD4_sg3	KDM3A_sg2	-0.92	-1.13	-3.04	-3.74	1.37E-03
KDM6B_sg3	PRMT7_sg3	-1.11	-0.90	-3.65	-3.00	1.85E-03
BRD3_sg3	KDM4C_sg1	-0.99	-1.01	-3.27	-3.33	2.13E-03
KDM4C_sg1	PRMT5_sg3	-0.93	-1.06	-3.08	-3.51	2.15E-03
KDM3B_sg1	PRMT5_sg3	-1.00	-0.99	-3.31	-3.28	2.15E-03
PRMT5_sg2	MBD1_sg1	-0.94	-1.04	-3.11	-3.42	2.47E-03
EP300_sg3	MBD1_sg3	-0.92	-1.03	-3.05	-3.41	2.93E-03
ING3_sg1	BRD4_sg3	-0.99	-0.96	-3.28	-3.17	2.93E-03
KDM1A_sg1	HDAC2_sg3	-0.94	-0.95	-3.12	-3.14	4.71E-03
PRMT5_sg2	CARM1_sg1	-0.95	-0.91	-3.14	-3.02	5.78E-03

【0135】 藉由評估該gRNA對，在個別的(非匯集的)細胞生長測定法中，使用由慢病毒所輸送且與其對應的gRNA對，針對抑制OVCAR8-ADR-Cas9細胞增生的能力(即，在5天內抑制 ~33%)而驗證本篩選過程中所得到的篩中物(圖4C)。在匯集式篩選與個別驗證法中所收集的數據之間，具有高度的一致性(圖13)。綜上所述，本文所述的方法提供一種試驗流程，用於系統性篩選出能在卵巢癌細胞中展現出抗增生性效果的條碼化組合性gRNA。

【0136】 相較於與對照組gRNA合併使用而言，當其與其他靶向表觀遺傳基因的gRNA合併使用時，許多靶向表觀遺傳基因的gRNA表現出較強的抗增生效果(圖 4B及12A)。

【0137】 藉由深度定序評估gRNA的脫靶活性，其顯示出，藉由針對該兩個gRNA的CRISPR設計與CCTop工具而進行電腦運算所預測出的所有

外顯子脫靶基因組位點具有低的插入缺失序列產生率(即，0.15至0.38%)(圖19)。綜上所述，建立了一個試驗程序，並驗證其可用於系統性篩選出能在卵巢癌細胞中展現出抗增生性效果的條碼化組合性gRNA。

【0138】 以驗證檢測法確認同時靶向KDM4C和BRD4的gRNA對(圖5A和8)以及shRNA對(圖 5B和14)，其協同地導致癌細胞生長效果的衰減。再者，以小分子KDM4C 抑制劑SD70 (Jin, et al. *PNAS* (2014) 111:9235-9240)以及小分子BRD4抑制劑JQ1 (Asangani, et al. *Nature* (2014) 510:278-282)共同處理，可協同地抑制OVCAR8-ADR細胞的增生(圖5C)。相似地，如同以KDM6B/6A抑制劑SK-J4 (Kruidenier, et al. *Nature* (2012) 488: 404-408)以及JQ1 (FIG. 5D)共同處理一般，同時靶向KDM6B和BRD4的gRNA對(圖 5A和8)以及shRNA對(圖 5B和14)顯示出其協同性。藉由Bliss independence (Bliss *Ann. Appl. Biol.* (1939) 6: 585-615)以及Highest Single Agent (Borisy, et al. *PNAS* (2003) 100: 7977-7982)模型，可確認在這些成對的小分子藥物組合物兩者之間的協同性(圖 5C和 5D)。

【0139】 本文所述的方法可用於鑑定可抑制癌細胞增生的新穎表觀遺傳標靶基因對，且具有發展出協同性藥物治療的潛力。本方法亦可增加根據CRISPR-Cas9系統而在高通量下進行系統性多重基因擾動篩選的實用性。

【0140】 這些方法亦有助於鑑定生物探究中的新領域，例如探討以觀察到的表型為基礎的機制。舉例而言，在以編碼了靶向KDM4C和BRD4，或者KDM6B和BRD4兩者的gRNA的慢病毒感染的細胞群中，評估其基因表現型態(圖21A)。顯著擾動的基因係與涉及癌症相關路徑的基因組有關聯

性，包括TNFα/NFκB訊息路徑、p53路徑、以及凋亡(圖21B)。此外，表觀遺傳基因擾動的組合性效果是複雜的，且可因不同的細胞種類而有變異(圖22A和22B)。

【0141】 方法

【0142】 **載體建構**

【0143】 使用標準的分子選殖技術，包括限制酶剪切、連接、PCR及Gibson組裝法進行載體的建構(表3)。一般的寡核苷酸係購自Integrated DNA Technologies。將該載體構築物轉染至DH5α大腸桿菌株中，使用50 μg/ml羧苄青黴素(Teknova)分離攜帶該構築體的殖株。使用Plasmid Mini或Midi Kits (Qiagen)萃取並純化DNA。以Genewiz's DNA 定序服務驗證載體構築物的序列。

【0144】 表 3：構築體

構築體 ID	設計
pAWp28	pBT264-U6p-{2xBbsI}-sgRNA scaffold-{MfeI}
pAWp28-1	pBT264-U6p-GFP-sg1
pAWp28-2	pBT264-U6p-RFP-sg1
pAWp28-3	pBT264-U6p-RFP-sg2
pAWp28-4	pBT264-U6p-KDM4C-sg1
pAWp28-5	pBT264-U6p-PHF2-sg1
pAWp28-6	pBT264-U6p-KDM6B-sg2
pAWp28-7	pBT264-U6p-PHF2-sg2
pAWp28-8	pBT264-U6p-DNMT1-sg1
pAWp28-9	pBT264-U6p-DNMT3B-sg1
pAWp28-10	pBT264-U6p-PRMT2-sg3
pAWp28-11	pBT264-U6p-HDAC2-sg1
pAWp28-12	pBT264-U6p-ING4-sg1
pAWp28-13	pBT264-U6p-KDM1B-sg3
pAWp28-14	pBT264-U6p-KDM2A-sg3
pAWp28-15	pBT264-U6p-PRMT6-sg1
pAWp28-16	pBT264-U6p-BMI1-sg2
pAWp28-17	pBT264-U6p-PHF8-sg2

pAWp9	pFUGW-UBCp-*RFP*-CMVp-*GFP*-{BamHI+EcoRI}
pAWp9-1	pFUGW-UBCp-*RFP*-CMVp-*GFP*-U6p-GFP-sg1
pAWp9-2	pFUGW-UBCp-*RFP*-CMVp-*GFP*-U6p-RFP-sg1
pAWp9-3	pFUGW-UBCp-*RFP*-CMVp-*GFP*-U6p-RFP-sg2
pAWp9-4	pFUGW-UBCp-*RFP*-CMVp-*GFP*-U6p-RFP-sg1-U6p-GFP-sg1
pAWp9-5	pFUGW-UBCp-*RFP*-CMVp-*GFP*-U6p-RFP-sg2-U6p-GFP-sg1
pAWp11	pFUGW-CMVp
pAWp12	pFUGW-CMVp-*GFP*
pAWp12-1	pFUGW-CMVp-*GFP*-[U6p-BRD4-sg3]-[U6p-PHF2-sg1]
pAWp12-2	pFUGW-CMVp-*GFP*-[U6p-BRD4-sg3]-[U6p-KDM6B-sg2]
pAWp12-3	pFUGW-CMVp-*GFP*-[U6p-BRD4-sg3]-[U6p-KDM4C-sg1]
pAWp12-4	pFUGW-CMVp-*GFP*-[U6p-BRD4-sg3]-[U6p-PHF2-sg2]
pAWp12-5	pFUGW-CMVp-*GFP*-[U6p-BRD4-sg3]
pAWp12-6	pFUGW-CMVp-*GFP*-[U6p-KDM4C-sg1]
pAWp12-7	pFUGW-CMVp-*GFP*-[U6p-KDM6B-sg2]
pAWp12-8	pFUGW-CMVp-*GFP*-[U6p-DNMT1-sg1]
pAWp12-9	pFUGW-CMVp-*GFP*-[U6p-DNMT3B-sg1]
pAWp12-10	pFUGW-CMVp-*GFP*-[U6p-PRMT2-sg3]
pAWp12-11	pFUGW-CMVp-*GFP*-[U6p-HDAC2-sg1]
pAWp12-12	pFUGW-CMVp-*GFP*-[U6p-ING4-sg1]
pAWp12-13	pFUGW-CMVp-*GFP*-[U6p-KDM1B-sg3]
pAWp12-14	pFUGW-CMVp-*GFP*-[U6p-KDM2A-sg3]
pAWp12-15	pFUGW-CMVp-*GFP*-[U6p-PRMT6-sg1]
pAWp12-16	pFUGW-CMVp-*GFP*-[U6p-BMI1-sg2]-[U6p-PHF8-sg2]
pAWp21	pLKO.1-Control sh
pAWp21-1	pLKO.1-KDM4C-sh1
pAWp21-2	pLKO.1-KDM4C-sh2
pAWp21-3	pLKO.1-KDM6B-sh
pAWp21-4	pLKO.1-BRD4-sh1
pAWp21-5	pLKO.1-BRD4-sh2
pAWp30	pFUGW-EFSp-*Cas9-P2A-Zeo*

【0145】 為了產生一個編碼shRNA且靶向特定基因的慢病毒，進行攜帶有正義及反義序列的寡核苷酸對的合成、黏合，並藉由連接反應將其選殖於經AgeI及EcoRI剪切的pLKO.1載體29(Addgene質體 #10879)中。設計shRNA的正義及反義序列，並依據siRNA Selection Program (sirna.wi.mit.edu/)進行設計及建構(表4)。

【0146】 表4：於個別驗證檢測中所使用的shRNA反義序列

shRNA ID	shRNA 反義序列
Control-sh	CGAGGGCGACTTAACCTTAGG (SEQ ID NO: 163)
KDM4C-sh1	AAATCTTCGTAATCCAAGTAT (SEQ ID NO: 164)
KDM4C-sh2	GTAATACCGGGTGTTCCGATG (SEQ ID NO: 165)
KDM6B-sh	ATTAATCCACACGAGGTCTCC (SEQ ID NO: 166)
BRD4-sh1	TATAGTAATCAGGGAGGTTCA (SEQ ID NO: 167)
BRD4-sh2	TTTAGACTTGATTGTGCTCAT (SEQ ID NO: 168)

【0147】 為了產生編碼了Cas9蛋白質及使用博萊黴素(Zeocin)抗性作為篩選標記的pAWp30慢病毒表現載體，使用Phusion DNA聚合酶(New England Biolabs)進行PCR而從Addgene質體#49535中擴增EFS啟動子和Cas9序列，而博萊黴素序列則由Addgene質體#25736進行擴增。使用Gibson Assembly Master Mix (New England Biolabs)將PCR產物選殖入pAWp11慢病毒載體骨架中。

【0148】 為了建構一個含有由U6啟動子(U6p)啟動且靶向特定基因的sgRNA的表現，進行與20 bp的sgRNA標靶序列配對的寡核苷酸對的合成、黏合，並使用T4連接酶(New England Biolabs)將其選殖至經BbsI剪切的pAWp28載體中。為了建構一個用於由U6p所起始的單一或組合性sgRNA的表現的慢病毒載體，以BglII和MfeI酵素(Thermo Scientific)剪切儲存載體而製備U6p-sgRNA表現區域，並分別將其插入pAWp12載體骨架或單一sgRNA表現載體中，其係藉由BamHI和EcoRI酵素(Thermo Scientific)剪切該載體後進行相容黏合端的連接反應而得。為了同時表現帶有雙重RFP及GFP螢光蛋白報導基因的sgRNA，使用與上述相同的策略，將經U6p起始的sgRAN表現區域插入pAWp9中，以取代pAWp12及慢病毒骨架。藉由將獨特的BamHI和EcoRI位點引入載體中而得以將U6p-sgRNA表現區域插入，進而修飾pAWp7骨架而獲得pAWp9載體。

【0149】 **條碼化組合性sgRNA核酸庫的組裝**

【0150】 合成一個攜帶條碼化sgRNA序列且具有153個寡核苷酸對(Oligo F-(x)及Oligo R-(x)，其中x = 1 to 153)的陣列，並將其黏合以產生一雙股插入物，其攜帶有20 bp sgRNA標靶序列、兩個BbsI限制酶位點、獨特於每個sgRNA但互相間有至少兩個鹼基不同、且其末端為5'突出端的8 bp條碼的雙股插入物。為了產生匯集的儲存載體庫，將該153個黏合的插入物等比例混合，並藉由單鍋連接反應，利用其相容端將其選殖至pAWp28儲存載體(以BbsI和MfeI進行剪切)中。為了建構條碼化sgRNA庫，使用經由BbsI剪切的匯集式儲存載體，以及一個藉由合成並黏合一個寡核苷酸對S1和S2而製備出的一個含有sgRNA支架序列、BamHI和EcoRI限制酶位點、及其末端為5'突出端的插入物，進行另一個單鍋連接反應。該匯集式儲存載體及該條碼化sgRNA核酸庫兩者皆於Endura勝任細胞(Lucigen)製備，並以Plasmid Midi kit (Qiagen)進行純化。

【0151】 藉由與上述產生單一及組合性sgRNA構築體的相同策略建構攜帶單一或組合性gRNA的匯集式慢病毒載體庫，不同的是，以匯集的而非個別的插入物及載體進行組裝。簡言之，藉由單鍋反應而使用BglII和MfeI剪切匯集的儲存載體庫而產生匯集式U6p-sgRNA插入物。以BamHI和EcoRI剪切標的慢病毒載體(pAWp12)。藉由相容端(即，BamHI + BglII & EcoRI + MfeI)而連接經剪切的插入物及載體，以在慢病毒載體中創建匯集式一次sgRNA核酸庫(153個sgRNA)。再以BamHI和EcoRI剪切該一次sgRNA載體庫，並以相同的U6p-sgRNA插入物庫組裝二次sgRNA核酸庫(153 x 153 sgRNAs = 總共23,409個組合物)。在匯集式組裝步驟後，該sgRNA被定位於

該載體構築體的一端，而其相對應的條碼則在串聯於另一端。以XL10-Gold超勝任細胞(Agilent Technologies)製備慢病毒sgRNA核酸庫，並以Plasmid Midi kit (Qiagen)進行純化。

【0152】 **細胞培養**

【0153】 由ATC取得HEK293T細胞，並在37℃含5% CO_2情況下，將其培養於補充10%熱失活胎牛血清與1X抗生素-抗真菌劑(Life Technologies)的DMEM中。OVCAR8-ADR細胞係由T. Ochiya所贈(Japanese National Cancer Center Research Institute, Japan)。認證OVCAR8-ADR細胞的身份(Genetica DNA Laboratories)。以慢病毒pAWp30載體感染OVCAR8-ADR細胞，且在存有200 μg/ml博萊黴素(Life Technologies)的情況下進行篩選3週，而產生可穩定表現Cas9蛋白質(OVCAR-ADR-Cas9)的OVCAR8-ADR細胞。在37℃含5% CO_2情況下，將OVCAR8-ADR及OVCAR8-ADR-Cas9細胞培養於補充了10%熱失活胎牛血清與1X抗生素-抗真菌劑的RPMI中。針對藥物處理，在進行細胞活性測試之前，使用SD70 (Xcessbio #M60194)、GSK-J4 (Cayman Chemical #12073)、及/或(+)-JQ1 (Cayman Chemical #11187)，並以指定劑量處理OVCAR8-ADR細胞。

【0154】 **慢病毒生產及輔導作用**

【0155】 在6孔培養盤中生產慢病毒，並將其裝載於HEK283T細胞中。進行轉染前，將HEK293T細胞維持在~70%匯合度。將FuGENE HD轉染試劑(Promega)與0.5 μg的慢病毒載體、1 μg的pCMV-dR8.2-dvpr載體、及0.5 μg的pCMV-VSV-G載體混合於100 μl的OptiMEM培養基(Life Technologies)中，並在添加於細胞培養前，使其在室溫下置放15分鐘。隔天

更換培養液。在轉染後48小時及96小時收集含有新產生病毒的上清液，並經由0.45 μm 聚醚碸膜(Pall)進行過濾。在存有8 μg/ml聚凝胺(Sigma)的情況下，使用500 μl經過濾的病毒上清液感染250,000個細胞過夜，而以個別載體構築體進行輔導作用。對於在篩選中所使用的匯集慢病毒庫的生產而言，使用相同的試驗程序，放大慢病毒生產及輔導程序。使用Amicon Ultra Centrifugal Filter Unit (Millipore)收集經過濾的病毒上清液。在存有8 μg/ml聚凝胺(Sigma)的情況下，以0.3至0.5的感染多重性進行細胞的感染，藉此而確保單拷貝數可嵌入大部分細胞中，其相當於30-40%的感染效率。使用於篩選的細胞總數較核酸庫大小多約300倍，藉此維持核酸庫的覆蓋率，並減少任何因慢病毒隨機嵌入基因組中而導致假性效應。於感染後隔天更換細胞培養基，並在進行試驗前培養一段指定的時間。

【0156】 用於條碼定序的樣本製備

為了由培養的細胞中製備用於條碼定序的樣本，使用 DNeasy Blood & Tissue Kit (Qiagen)，並根據製造商的操作程序萃取並製備基因組 DNA。針對條碼化 sgRNA 質體庫，使用 Plasmid Midi Kit (Qiagen)萃取轉染至大腸桿菌的質體 DNA。使用 Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies)測定 DNA 的濃度。

【0157】 使用Kapa Hotstart Ready Mix (Kapa Biosystems)，從質體/基因組DNA樣本中以PCR放大含有代表著匯集載體及被感染細胞庫中各種組合物的獨特條碼的一段~360 bp的片段。對於質體DNA而言，於25-μl的PCR反應中添加1 ng的DNA模板。對於基因組DNA而言，於50-μl的PCR反應中

添加800 ng的DNA，並且，針對每個基因組DNA進行總共64個PCR反應，以確保將被擴增的細胞基因組數目較該核酸庫大小大於100倍。此外，使PCR參數最適化，以確保PCR擴增的步驟可維持在指數期，以防止PCR偏差。在進行多重定序的PCR過程中，加入Illumina錨定序列與一個8個鹼基對的索引條碼。擴增條碼序列所使用的引子對序列為：

5’-AATGATACGGCGACCACCGAGATCTACACGGATCCGCAACGGAATTC-3’ (SEQ ID NO: 1)和

5’CAAGCAGAAGACGGCATACGAGATNNNNNNNNGGTTGCGTCAGCAAACACAG-3’ (SEQ ID NO: 2)，其中NNNNNNNN代表一個指定給每個試驗樣本的特定索引條碼。

【0158】 進行1.5%瓊脂糖膠體電泳並根據片段大小純化含有條碼序列的PCR產物，進一步以QIAquick Gel Extraction Kit (Qiagen)進行萃取。使用KAPA SYBR Fast qPCR Master Mix (Kapa Biosystems)以及Illumina Library Quantification Kit (Kapa Biosystems)進行定量PCR，以測定PCR產物的濃度。用於定量PCR的正向及反向引子分別為：

5’-AATGATACGGCGACCACCGA-3’ (SEQ ID NO: 3)以及

5’-CAAGCAGAAGACGGCATACGA-3’ (SEQ ID NO: 4)。之後，將來自不同樣本的PCR產物以一所需比例匯集，而用於多重樣本定序，並至於Illumina HiSeq系統中以CombiGEM條碼引子

(5’-CCACCGAGATCTACACGGATCCGCAACGGAATTC-3’ (SEQ ID NO: 5)) 以及索引條碼引子

(5’-GTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACC-3’ (SEQ ID NO:

6))進行分析。

【0159】 **條碼序列數據分析**

【0160】 從Illumina定序資料中處理每個sgRNA組合物的條碼讀值。將代表每種組合物的條碼讀值以每個由索引條碼分類的樣本的每百萬個讀值進行標準化。如同細胞增生的測量，計算20天對比於15天的群組中標準化條碼讀值的條碼計數比例，以倍數變化表示。前期增生性與抗增生性表型分別具有>1及<1的標準化條碼讀值的倍數變化，而表型沒有改變的倍數變化=1。在15天的組別中，將具有少於~100個的條碼絕對讀值濾除以改善數據的可信度。取得每個相同的sgRNA組合物的不同可能等級的倍數變化的平均值，並可觀察到在倍數變化尚的高度一致性(即，針對超過82%及95%的組合物，其變異係數(CV)分別為 <0.2及 <0.4，(圖11)。經計算的倍數變化經過Log轉化後以Log2比值表示。在由相同慢病毒庫進行獨立感染的兩個生物性複製體中進行篩選。在所有試驗的條件下以Log2比值排序組合物。前幾名的篩中物組(空心圓)定義為在兩個生物性複製體中，與僅攜帶對照組sgRNA(空心三角形)的sgRNA組合物的$\log_2$比值的平均值差距至少3個標準差的組別(圖 4B、12A和12B)。

【0161】 **細胞活性測試**

【0162】 進行MTT呈色法以評估細胞活性。 針對每個96孔盤的細胞培養，加入100 μl的MTT (3-(4,5-二甲基吡啶-2-基)-2,5-二苯基溴化)溶液(Sigma)。在37°C，含5% CO_2的條件下培養細胞3小時。存活的細胞將可溶性的MTT鹽類轉化成不可溶的藍色三苯基甲臜晶體。在37°C下將三苯基甲臜晶體溶解於100 μl的溶解緩衝液中。使用Synergy H1 Microplate Reader

(BioTek)測量在光密度 (OD) 570nm及650nm (參考值)的吸光讀值。

【0163】 **藥物協同作用定量**

【0164】 一般常用Bliss independence (BI)模型(Bliss *Ann. Appl.Biol.* (1939) 26: 585-615)以及Highest Single Agent (HSA)模型(Borisy, et al. *PNAS* (2003) 100: 7977-7982)驗證藥物組合之間的協同作用。

【0165】 根據BI模型，預測效果(E_{Exp})如下：

$$E_{Exp} = E_A + E_B - (E_A \times E_B),$$

【0166】 其中E_A是在單獨使用某個濃度的藥物A時所觀察到的生長抑制效果，而E_B則是在單獨使用某個濃度的藥物B時所觀察到的生長抑制效果。E_{Obs}是針對藥物組合物(A+B)所觀察到的生長抑制效果，其分別具有與E_A和E_B情況下相同的濃度。每個效果係以介於0與1之間的分級抑制效果表示。當$E_{Obs} - E_{Exp} > 0$時，兩種藥物被視為具有協同的交互作用。

【0167】 HAS模型與BI模型相似，不同點在於，根據HAS模型，E_{Exp}等於由組合物的單一藥劑(E_A或E_B)在與藥物組合物(A + B)中濃度相同的情況下所產生的較大生長抑制效果。

【0168】 假使$E_{Obs} - E_{Exp} > 0.1$ (即，在圖 5C與5D中，由預測的BI和HAS模型中有 > 10%的過量抑制現象)，則兩個藥物被視為在兩個模型中至少兩個不同濃度的組合物下具有協同作用，以增加我們標準的嚴格性。

【0169】 **流式細胞儀**

【0170】 在感染後4天及8天，收集細胞。以補充2%胎牛血清的1x PBS清洗並重新懸浮樣本。為了移除任何的細胞聚集，在裝載至LSRII Fortessa flow cytometer (Becton Dickinson)之前，使重新懸浮的細胞通過細胞過濾

器。每個樣本收集至少20,000個訊號。根據細胞樣本，選擇適當的雷射組及過濾器。使用前向散射及側向散射於鑑定適當的細胞群。使用製造商的內建軟體分析數據。

【0171】 **螢光顯微鏡**

【0172】 直接在倒置螢光顯微鏡(Zeiss)下觀察經慢病毒感染三天後的培養細胞。使用Zeiss內建軟體擷取影像。

【0173】 **免疫墨點分析**

【0174】 以補充了蛋白酶抑制劑的2X RIPA緩衝液使細胞裂解。使用杵式電動攪拌機(Agros)均質細胞裂解物30秒，然後在4°C 下以15,000 rpm離心15分鐘。使用BCA測定法(Thermo Scientific)定量上清液。在4-15%聚丙烯醯胺膠體(Bio-Rad)上進行膠體電泳前，於99 °C下使蛋白質變性5分鐘。在4°C下以80V的電壓操作2小時，將蛋白質轉移至硝化纖維素薄膜上。所使用的初級抗體：抗-BRD4 (1:2,000, Cell Signaling #13440)、抗-KDM4C (1:1,000, Abcam ab85454)、抗-KDM6B (1:1,000, Abcam ab85392)、以及抗-β-肌動蛋白 (1:4,000, Abcam ab6276)。所使用的二及抗體為：經HRP-鏈結的抗 –兔子IgG (1:2,000, Cell Signaling #7074)、以及經HRP-鏈結的抗-小鼠IgG (1:4,000, Cell Signaling #7076)。藉由SuperSignal West Pico化學發光受質 (Thermo Scientific) 進行薄膜的呈色反應，並使用 ChemiDoc Touch 影像系統(BioRad)擷取影像。

【0175】 **用於基因組修飾的Surveyor測定法及定序分析**

【0176】 進行Surveyor測定法以驗證DNA剪切效率。使用QuickExtract DNA萃取溶液(Epicentre)並依照製造商的操作方法從細胞培養

中萃取基因組DNA。使用Phusion DNA聚合酶及表5所列的引子進行PCR反應而產生攜帶標靶位點的擴增子。使約200 ng的PCR擴增子變性，自我黏合，並與1.5 μl的Surveyor核酸酶(Transgenomic)在42ºC培養30分鐘。然後在2%瓊脂糖膠體上分析樣本。使用ImageJ軟體定量DAN條帶的強度，並以下述公式估算插入缺失序列的發生性(Ran et al. *Nat. Protoc.* (2013) 8:2281-2308)：

插入缺失序列 (%) = 100 x (1- (1- f_{cut})的平方根),

【0177】 而 $f_{cut} = (b + c)/(a + b + c)$，其中$a$是未被剪切的PCR擴增子的條帶強度，而$b$ 和 c則為每個被剪切條帶的強度。針對標靶等位基因的預期未被剪切與被剪切的條帶表列於圖 7及8中。

【0178】 進行Sanger定序以分析因組合性sgRNA的表現而產生的基因組修飾作用。培養經由組合性sgRNA構築體感染的細胞12天，並藉由序列稀釋法，將其重新播種於96孔培養盤中形成單一細胞群。在培養5至12天後，使用QuickExtract DNA萃取溶液從分離的單細胞擴增殖株中萃取出基因組DNA，並藉由如上所述的PCR方法製備攜帶有標靶等位基因的擴增子。使用TA Cloning Kit (Life Technologies)並依照製造商的操作方式將PCR擴增子選殖至TOPO載體中，之後使用Sanger定序法鑑定核苷酸的突變、插入及缺失。

【0179】 表5：於Surveyor測定法及Sanger定序中所使用的PCR引子列表

標靶 sgRNA ID	**正向 (5' to 3')**	**反向 (5' to 3')**
BMI1-sg2	AGAAATTAAACGGCTACCCTCCA	GTTGGTACAAAGTGGTGAAGGC (SEQ ID

	(SEQ ID NO: 169)	NO: 183)
BRD4-sg2	TCCATAGTGTCTTGAGCACCAC (SEQ ID NO: 170)	ACGTGGCTTCATTGTACATCCT (SEQ ID NO: 184)
BRD4-sg3	CACTTGCTGATGCCAGTAGGAG (SEQ ID NO: 171)	AAGCACATGCTTCAGGCTAACA (SEQ ID NO: 185)
DNMT1-sg1	GTGAATAGCTTGGGAATGTGGG (SEQ ID NO: 172)	TCATCTGCTCTTACGCTTAGCC (SEQ ID NO: 186)
DNMT3B-sg1	GCCACACTCTACATGGGAGC (SEQ ID NO: 173)	CTCGGCAACCCTCCATACAT (SEQ ID NO: 187)
HDAC2-sg1	GACTTTTCCATCAGGGACACCT (SEQ ID NO: 174)	AACCATGCACAGAATCCAGATTTA (SEQ ID NO: 188)
ING4-sg1	GGTGGACAAACACATTCGGC (SEQ ID NO: 175)	AAGAGTTCTTGGCGCAGACA (SEQ ID NO: 189)
KDM1B-sg3	CCTATCATTGCCCCAAGGAGTC (SEQ ID NO: 176)	TCGTCCAAGTTACAGTCATCACA (SEQ ID NO: 190)
KDM2A-sg3	CTAGGCCTCCGACAGTTGTAAT (SEQ ID NO: 177)	TCCTCTGGTGCACAGAAAAGTC (SEQ ID NO: 191)
KDM4C-sg1	AGCCACCCTTGGTTGGTTTT (SEQ ID NO: 178)	TTCTCTCCAGACACTGCCCT (SEQ ID NO: 192)
KDM6B-sg2	GGTAAGGGAAACTCTGGGGC (SEQ ID NO: 179)	GTGCCCAGAACTACTGCCAT (SEQ ID NO: 193)
PHF8-sg2	CTCCCTCCCTTCCTAAGGCT (SEQ ID NO: 180)	GAGGTGAGTTCCAGCTTCCC (SEQ ID NO: 194)
PRMT2-sg3	ATTGCCTTAAGTCGACACCTGAT (SEQ ID NO: 181)	CACCTTACAGGCACTGCGTT (SEQ ID NO: 195)
PRMT6-sg1	GACTGTAGAGTTGCCGGAACAG (SEQ ID NO: 182)	CTCCCTCCCTAGAGGCTATGAG (SEQ ID NO: 196)

【0180】 表6：在攜帶有BMI1-sg2 and PHF8-sg2表現構築體的OVCAR8-ADR-Cas9單細胞群中標靶等位基因的序列。粗體字的核苷酸代表PAM序列；劃底線的核苷酸則為插入或突變的鹼基對。

單細胞	sgRNA ID	序列 (5' to 3')	SEQ ID NO.	插入缺失序列
1	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAG----------GTCTTGT GAACTTGGACATCACAAATAGGAC	231	-10bp
1	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAG----------GTCTTGT GAACTTGGACATCACAAATAGGAC	231	-10bp

1	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAC----------TTCAAAGGGGCATGATAC ACACAAGGGGAAACCAGTGAAGACC	232	-10bp
1	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAC----------TTCAAAGGGGCATGATAC ACACAAGGGGAAACCAGTGAAGACC	232	-10bp
2	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATA---------**GG**TCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	233	-9bp
2	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGT---**GG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	234	-3bp
2	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA-CGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	235	-1bp
2	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA-C----**TGG**ATCTTCAAAGGGGCATGAT ACACACAAGGGGAAACCAGTGAAGACC	236	-5bp
3	BMI1-sg 2	ACTCCAATAATAATTACAAACAAGGAATTTCAACA GTTTCCTACCTTATATTCTGTGATCTG**TGG**TCTGGT CTTGTGAACTTGGACATCACAAATAGGAC	237	2mut & +4bp
3	BMI1-sg 2	ACTCCAATAATAATTACAAACAAGGAATTTCAACA GTTTCCTACCTTATATTCTGTGATCTG**TGG**TCTGGT CTTGTGAACTTGGACATCACAAATAGGAC	237	2mut & +4bp
3	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	238	-3bp
3	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	238	-3bp
4	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
4	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
4	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
4	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT

5	BMI1-sg2	ATTTCAACAGTTTCCTACCTTATATACTATACTATAT ATATATATATATACTATATATATAGTGGTCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	241	+30bp
5	BMI1-sg2	ATTTCAACAGTTTCCTACCTTATATACTATACTATAT ATATATATATATACTATATATATAGTGGTCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	241	+30bp
5	PHF8-sg2	CTCTCATGTTTTTCTGGCTTAGTGAAAAAACTTCA CTAAGTTTTTACTTAGTGGATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	242	+16bp & 2mut
5	PHF8-sg2	CTCTCATGTTTTTCTGGCTTAGTGAAAAAACTTCA CTAAGTTTTTACTTAGTGGATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	242	+16bp & 2mut
6	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC-----TGGTCTGGTCT TGTGAACTTGGACATCACAAATAGGAC	243	-5bp
6	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTAC----------TTGTGAACTT GGACATCACAAATAGGAC	244	-24bp
6	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA----CATGGATCTTCAAAGGGGCATGA TACACACAAGGGGAAACCAGTGAAGACC	245	-4bp & 1mut
6	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACG---TGGATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
7	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCT---------AGTGGTCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	247	-10bp
7	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCT---------AGTGGTCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	247	-10bp
7	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CGTGGATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
7	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CGTGGATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
8	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCT---GTGGTCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	248	1mut & -3bp
8	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCT---GTGGTCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	248	1mut & -3bp

8	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
8	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
9	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
9	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC-----**TGG**TCTGGTCT TGTGAACTTGGACATCACAAATAGGAC	243	-5bp
9	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA----C<u>A</u>**TGG**ATCTTCAAAGGGGCATGA TACACACAAGGGGAAACCAGTGAAGACC	245	-4bp & 1mut
9	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
10	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
10	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
10	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAC<u>TAA</u>GTGGATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	249	3mut
10	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
11	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
11	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTAC----------TAG**TGG**TCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	250	-11bp
11	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
11	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT

12	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
12	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
12	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
12	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
13	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
13	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
13	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGA----------ATCTTCAAAGGGGCATGATACAC ACAAGGGGAAACCAGTGAAGACC	251	-13bp
13	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGA----------ATCTTCAAAGGGGCATGATACAC ACAAGGGGAAACCAGTGAAGACC	251	-13bp
14	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
14	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
14	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACG---**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
14	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACG---**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
15	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
15	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT

15	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
15	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
16	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
16	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC-----**TGG**TCTGGTCT TGTGAACTTGGACATCACAAATAGGAC	243	-5bp
16	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA----C<u>A</u>**TGG**ATCTTCAAAGGGGCATGA TACACACAAGGGGAAACCAGTGAAGACC	245	-4bp & 1mut
16	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
17	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	234	-3bp
17	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	234	-3bp
17	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
17	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
18	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
18	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
18	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
18	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT

19	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
19	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCAA----------G**TGG**TCTGGTCTTGT GAACTTGGACATCACAAATAGGAC	252	-14bp & 1mut
19	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
19	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA----C**A**TGGATCTTCAAAGGGGCATGA TACACACAAGGGGAAACCAGTGAAGACC	245	-4bp & 1mut
20	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
20	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
20	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAC---G**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	253	-3bp
20	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAC---G**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	253	-3bp
21	BMI1-sg2	ACAACTCCAATAATAATTACAAACAAGGAATTTCA ACAGTTTCCTACCTTATATTCAG**A**TAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	254	+1bp
21	BMI1-sg2	ACAACTCCAATAATAATTACAAACAAGGAATTTCA ACAGTTTCCTACCTTATATTCAG**A**TAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	254	+1bp
21	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
21	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACG---**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
22	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTA----------G**TGG**TCTGGTCTTGTG AACTTGGACATCACAAATAGGAC	255	-14bp
22	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC--TAG-**GG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	256	-3bp

22	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
22	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
23	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTACCTTATATTCAG----------TTGTGAACTTGGACATCACAAATAGGAC	257	-13bp
23	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTACCTTATA-----TAG**TGG**TCTGGTCTTGTGAACTTGGACATCACAAATAGGAC	258	-5bp
23	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
23	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
24	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTCTTGTGAACTTGGACATCACAAATAGGAC	239	-3bp
24	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTA----------G**TGG**TCTGGTCTTGTGAACTTGGACATCACAAATAGGAC	255	-14bp
24	PHF8-sg2	TCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAA<u>CATG</u>C<u>C</u>CCG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	259	+3bp & 2mut
24	PHF8-sg2	TCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAA<u>CATG</u>C<u>C</u>CCG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	259	+3bp & 2mut
25	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTCTTGTGAACTTGGACATCACAAATAGGAC	239	-3bp
25	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTCTTGTGAACTTGGACATCACAAATAGGAC	239	-3bp
25	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
25	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATGATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp

26	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
26	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGT---**GG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	234	-3bp
26	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
26	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
27	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTAT---------**TGG**TCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	260	-9bp
27	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTAT---------**TGG**TCTGGTCTTG TGAACTTGGACATCACAAATAGGAC	260	-9bp
27	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
27	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACG---**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
28	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACA----------GT**GG**TCTGGTCTTGTGAAC TTGGACATCACAAATAGGAC	261	-22bp
28	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACA----------GT**GG**TCTGGTCTTGTGAAC TTGGACATCACAAATAGGAC	261	-22bp
28	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
28	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
29	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
29	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC-----**TGG**TCTGGTCT TGTGAACTTGGACATCACAAATAGGAC	243	-5bp

29	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA----C**A**T**GG**ATCTTCAAAGGGGCATGA TACACACAAGGGGAAACCAGTGAAGACC	245	-4bp & 1mut
29	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
30	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
30	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
30	PHF8-sg 2	TCTCATGTTTTTCTGGCTTAGTGAAA<u>TCT</u>AA<u>G</u>C<u>TTA</u> <u>GTGAAATCTAA</u>GCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	262	+17bp & 1mut
30	PHF8-sg 2	TCTCATGTTTTTCTGGCTTAGTGAAA<u>TCT</u>AA<u>G</u>C<u>TTA</u> <u>GTGAAATCTAA</u>GCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	262	+17bp & 1mut
31	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC<u>TTATA</u>**TGG**TCTGGT CTTGTGAACTTGGACATCACAAATAGGAC	263	5mut
31	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTA----------G**TGG**TCTGGTCTTGTG AACTTGGACATCACAAATAGGAC	255	-14bp
31	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
31	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAA---CG**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
32	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
32	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATAT<u>A</u>---TAG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	264	-3bp & 1mut
32	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
32	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT

33	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAG----------AAC TTGGACATCACAAATAGGAC	265	-18bp
33	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAG----------AAC TTGGACATCACAAATAGGAC	265	-18bp
33	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
33	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
34	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
34	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
34	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
34	PHF8-sg2	GTTCTGACTCACAATCTCTCATGTTTTTCTGGCTTA GTGAAAAAAA<u>A</u>CGCCGTGGATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	266	+1bp
35	BMI1-sg2	TAATTACAAACAAGGAATTTCAACAGTTTCCTACC TTATATTCAG<u>TATAATATATTCATT</u>AG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	267	+14bp
35	BMI1-sg2	TAATTACAAACAAGGAATTTCAACAGTTTCCTACC TTATATTCAG<u>TATAATATATTCATT</u>AG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	267	+14bp
35	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTG----------**G**ATCTTCAAAGGGGCATGATACAC ACAAGGGGAAACCAGTGAAGACC	268	-13bp
35	PHF8-sg2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTG----------**G**ATCTTCAAAGGGGCATGATACAC ACAAGGGGAAACCAGTGAAGACC	268	-13bp
36	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC--TAG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	269	-2bp
36	BMI1-sg2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC--TAG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	269	-2bp

36	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAACG---**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	246	-3bp
36	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
37	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTCAACAG TTTCCTACCTTATATTATATAG**TGG**TCTGGTCTTGTGAAC TTGGACATCACAAATAGGAC	270	3mut
37	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCAGTAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	239	WT
37	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA-C----**TGG**ATCTTCAAAGGGGCATGAT ACACACAAGGGGAAACCAGTGAAGACC	236	-5bp
37	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAA-C----**TGG**ATCTTCAAAGGGGCATGAT ACACACAAGGGGAAACCAGTGAAGACC	236	-5bp
38	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	239	-3bp
38	BMI1-sg 2	TAATTACAAACAAGGAATTTCAACAGTTTCCTACC TTATATTCAGGTAGTGAATCTGAATAG**TGG**TCTGG TCTTGTGAACTTGGACATCACAAATAGGAC	271	+14bp
38	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAAC---G**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	253	-3bp
38	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAAAC---G**TGG**ATCTTCAAAGGGGCATG ATACACACAAGGGGAAACCAGTGAAGACC	253	-3bp
39	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTT----------CTGGTCTTGTGA ACTTGGACATCACAAATAGGAC	272	-15bp
39	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTCA---G**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	239	-3bp
39	PHF8-sg 2	ACAATCTCTCATGTTTTTCTGGCTTAGTGAATCTTC AAAGGGATCTTCAAAA**GG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	273	+11bp & 4mut
39	PHF8-sg 2	ACAATCTCTCATGTTTTTCTGGCTTAGTGAATCTTC AAAGGGATCTTCAAAA**GG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	273	+11bp & 4mut

40	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATT----------GTGAAC TTGGACATCACAAATAGGAC	274	-18bp
40	BMI1-sg 2	TACAACTCCAATAATAATTACAAACAAGGAATTTC AACAGTTTCCTACCTTATATTC---AG**TGG**TCTGGTC TTGTGAACTTGGACATCACAAATAGGAC	275	-3bp
40	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT
40	PHF8-sg 2	CGTTCTGACTCACAATCTCTCATGTTTTTCTGGCTT AGTGAAAAAACGCCG**TGG**ATCTTCAAAGGGGCAT GATACACACAAGGGGAAACCAGTGAAGACC	240	WT

【0181】 **RNA萃取及定量RT-PCR (qRT-PCR)**

【0182】 使用TRIzol Plus RNA Purification Kit (Life Technologies)並依照製造商的操作方法從細胞中萃取出RNA，並以PureLink on-column DNase kit (Life Technologies)處理。使用NanoDrop分光光度計測定RNA的品質及濃度。使用SuperScript III Reverse Transcriptase (Life Technologies)、Random Primer Mix (New England Biolabs)及RNAse OUT (Invitrogen)進行RNA樣本的反轉錄。為了評估基因表現量，LightCycler480 system (Roche)上使用SYBR FAST qPCR MasterMix (KAPA)進行定量PCR。使用LifeCyler480 SW 1.1內建軟體定量數據。使用PrimerBlast (NCBI)設計並評估PCR引子。引子序列表列於表7中。

【0183】 表 7：進行 gRT-PCR 所使用的 PCR 引子

標靶基因 ID	正向引子 (5' to 3')	反向引子 (5' to 3')
BRD4	GTTGATGTGATTGCCGGCTC (SEQ ID NO: 197)	TTAGGCAGGACCTGTTTCGG (SEQ ID NO: 200)
KDM4C	CGTACGGGTTCATGCAAGTT (SEQ ID NO: 198)	CGTTTGCTTAAGAGCACCTCC (SEQ ID NO: 201)
KDM6B	CCCCTCACCGCCTATCAGTA (SEQ ID NO: 199)	TCTTGAACAAGTCGGGGTCG (SEQ ID NO: 202)

【0184】 表 8：用於偵測插入缺失序列的深度定序法所使用的 PCR 引子

標靶 gRNA ID	標靶位點的類型	20 bp gRNA 標靶序列 (5' to 3')	SEQ ID NO.	正向引子 (5' to 3')	SEQ ID NO.	反向引子 (5' to 3')	SEQ ID NO.
NF1-sg1	對靶	GTTGTGCTCAGTACTGACTT	276	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GGTATCTGTGGTTGATGCAGTTTTCC	304	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-AGTAGTGAGGCCGCTTATAACC	332
NF1-sg4	對靶	TTTCAGCTTCCAATAAAAAC	277	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GGTATCTGTGGTTGATGCAGTTTTCC	304	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-AGTAGTGAGGCCGCTTATAACC	332
NF2-sg2	對靶	ATTCCACGGGAAGGAGATCT	278	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GCACAGAGCTGCTGCTTGGAGTG	305	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-TAACAAGGAGATGCCCTGGCTGG	333
MED12-sg1	對靶	AGGATTGAAGCTGACGTTCT	279	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CGCTTTCCTGCCTCAGGATGAAC	306	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-AGGTCATGAAGGCAAACTCAGCC	334
PHF8-sg2	對靶	GGCTTAGTGAAAAACGCCG	280	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-TTGGAAGAGAAGGATCTGCTGAGGC	307	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CACCTGTCAAAAGTCCTACTCCGG	335
BMI-sg2	對靶	TCCTACCTTATATTCAGTAG	281	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GCTACCCTCCACAAAGCACACAC	308	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CCTGGAGACCAGCAAGTATTGTCC	336
KDM4C-sg1	對靶	CCTTTGCAAGACCCGCACGA	109	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CCTTCAGAAACAATGTCCCAAATCG	309	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-GTCCTCTGAACCCCAGCTGTAAG	337
KDM4C-sg1	脫靶	GCTTTGCCCGAACCGC	282	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-AGCCTT	310	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CAAA	338

		ACGA		TCTGAGAGCGGG CTAG		CAGAGGCCAAAG GGTGTCCC	
KDM4 C-sg1	脫靶	CCTAG GCCAG ACCTG CACGA	283	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-GCCTCC TCTCATCCTCTCG CTTC	311	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-ACAG GAGGTCGTGGTG CAGTTCTC	339
KDM4 C-sg1	脫靶	GCTCT GGAAG ACCCG CACCA	284	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-GATTGG CTCCAAGCGGCCA TCAAAC	312	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-TTGTG TGAGGAACGTTG ACGCTACC	340
KDM4 C-sg1	脫靶	CCTTAT CAAGA CCCAC ACCA	285	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-TTGAAC TCAAGGCTCAGC CAACAGGC	313	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-GAGC GTAGGTCCTCTGC ATGGAG	341
KDM6 B-sg2	對靶	ATCCCC CTCCTC GTAGC GCA	286	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-CAACTC AGGCTGGATGCAT CGG	314	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-CCACA GAATGACAGGAA CCCATGG	342
KDM6 B-sg2	脫靶	CTGCTC CTCCTC GTAGC GCT	287	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-TTGGTG GCCGCTGAGTGT GTGTAC	315	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-ACTGA GCAGAGCCTAGG AGGCAG	343
KDM6 B-sg2	脫靶	TGCGC CCTCCT CCTAG CGCA	288	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-CAGCAT GTTGACATAGCGG C	316	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-TGTTG CCAGATCCAGAG GCGTC	344
KDM6 B-sg2	脫靶	CTCCTC CTCCG CGTAG CGCT	289	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-TGAGA GGAGATGAGTCG GGGTC	317	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-AACTG GCCCGAGTAGTCG GAGCAG	345
KDM6 B-sg2	脫靶	CTGCC CCTCCT GGTAG CGCC	290	ACACTCTTTCCCT ACACGACGCTCTT CCGATCT-GACGG GTCAAAGCCTCA GGAGAG	318	GTGACTGGAGTTC AGACGTGTGCTCT TCCGATCT-TCCTC AGAGTGTGTGGA AGTGCTGG	346
KDM6	脫靶	AACCA	291	ACACTCTTTCCCT	319	GTGACTGGAGTTC	347

B-sg2		GCTCCTCGTAGCTCA		ACACGACGCTCTTCCGATCT-TTAGCTGCCCAGCTCACAGCTACC		AGACGTGTGCTCTTCCGATCT-AAGAGCTCCTAGGGGAGGATCAG	
KDM6B-sg2	脫靶	ACCGCCCTCCTCCTAGCTCA	292	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GAGCCCCAAGAGCGAGACAA	320	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-TGGCAGGAGCACAGCCTAAGGA	348
KDM6B-sg2	脫靶	AGCCCGCTCCTCGTGGGGCA	293	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GGCGCTCAGAAGGCTGTGCAG	321	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-ACACCCGCCTCGGAGATCAACAC	349
BRD4-sg3	對靶	GGGAACAATAAAGAAGCGCT	294	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CCTAGGTGACACTGGACTTTTGC	322	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CCACCCCTACATCTCACCTTGTTG	350
BRD4-sg3	脫靶	TGGAAAAACAAAGAAGAGCT	295	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-AGGTTCACCTCAGGCTGCTCAGAAG	323	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-GTGAGGTTTCCACGTGCCAGC	351
BRD4-sg3	脫靶	GGGAAGTATAAGGAAGAGCT	296	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CGTCTCTCTCCATGTGAGCTTGTG	324	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CCAACAATTCCAGGTATGAAACTCCC	352
BRD4-sg3	脫靶	GTGAGCAATAAAGCAGCCCT	297	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GGAAAGATCATCTGATCAGGCCCATC	325	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CTCCCCACTTGTAGGTTCCTAATCC	353
BRD4-sg2	對靶	TCTAGTCCATCCCCCATTAC	298	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CTTGTTAGGGTTGGAGGTCTCTGG	326	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-GTAGAGTGCCTGGTGAAGAATGTG	354
BRD4-sg2	脫靶	AATATTCCATTCCCCATTAC	299	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-AAGGCCCGTAAAGGGCA	327	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-TCCAGACTGTTGTTCAGT	355

				AGTTCAG		CCTGT	
BRD4-sg2	脫靶	TGTTGTCCATACCTCATTAC	300	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-GGTGACAGGAAGCTGTCGGAACAT	328	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CTCTGGATTTGCCCACACCTAGTC	356
BRD4-sg2	脫靶	TCTAGGTCATGCACCATTAC	301	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CTCACTGTGATCTGACACCAAACAC	329	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-GCATGCTTGCTTTCTGAAGGTGGC	357
BRD4-sg2	脫靶	CCCATTCCTTCCCCCATTAC	302	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CTAGTTGCCTTCATGCCTTACAGAC	330	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-TTCCAAGCAAGTGAGCTTCAGCACC	358
BRD4-sg2	脫靶	TCCACACCCTCCCCCATTAC	303	ACACTCTTTCCCTACACGACGCTCTTCCGATCT-CTGCTCCCACTCCAGACTACCC	331	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-CCACCCATGACACAGGAGGG	359

【0185】 表 9：用於偵測插入缺失序列的全基因組擴增法所使用的PCR 引子

標靶 sgRNA ID	正向引子 (5' to 3')	反向引子 (5' to 3')
BMI1-sg2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTACCCTCCACAAAGCACACAC (SEQ ID NO: 360)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTGGAGACCAGCAAGTATTGTCC (SEQ ID NO: 362)
PHF8-sg2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTGGAAGAGAAGGATCTGCTGAGGC (SEQ ID NO: 361)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCACCTGTCAAAAGTCCTACTCCGG (SEQ ID NO: 363)

【0186】 **RNA-Seq及數據分析**

【0187】 使用TRIzol Plus RNA純化套組(Life Technologies)，根據製造商的操作說明從細胞中萃取出RNA，並以PureLink on-column DNase kit (Life Technologies)處理。使用NanoDrop分光光度計測定RNA的品質及濃度。使用SuperScript III Reverse Transcriptase (Life Technologies)、Random

Primer Mix (New England Biolabs)及RNAse OUT (Invitrogen)進行RNA樣本的反轉錄。使用Illumina Library Prep Kit製備定序核酸庫，並依照製造商的操作建議由1 μg的總RNA輸入量開始進行。在PCR擴增後，在2%瓊脂糖膠體(E-Gel EX, Invitrogen)上依照大小篩選出大小為300 +/- 25 bp的核酸庫，接著在Illumina HiSeq 2000儀器上進行單末端定序。在兩個生物性複製體中進行RNA-Seq試驗。

【0188】 使用TopHat2 (Kim, et al. *Genome Biol* (2013) 14:R36)以及Bowtie (Langmead, et al. *Genome Biol* (2009) 10:R25)而由人類轉錄組(RefSeq, hg19)進行cDNA片段的原始單末端讀值的校對。使用Cuffdiff2 (Trapnell, et al. *Nat Biotechnol* (2013) 31:46-53)，且藉由偏差修正功能，遮蔽對應到粒線體及核糖體RNA轉錄體的讀值，而獲得樣本間具有表現差異性的基因。當在至少一個測試條件下，基因符合最小值為每百萬個讀值每千個鹼基中有0.1個片段(FPKM)，則稱為差異性表現的基因，絕對的$\log_2$倍數變化為至少0.5，而在進行多重假定校正(Q-值)後P值小於0.05。使用MSigDB資料庫 (broadinstitute.org/gsea/index.jsp) (Subramanian, et al. *Proc Natl Acad Sci USA* (2005) 102:15545-50)進行基因組富集性分析。

【0189】 **在混合細胞群中細胞增生的數學模型**

【0190】 為了估算不同參數如何影響本文中匯集式篩選的結果，本文模擬在攜帶有gRNA組合物且展現出不同生長速率的混合細胞群中的細胞增生效果。在給定時間(t)下，細胞群中的總細胞計數值(C_N)是以含有不同組合物(C_i；1至 N，其中N是總組合物數目)的個別細胞計數值的總和表示，而使$C_N(t) = \sum_{i=1}^{N} C_i(t)$。

【0191】 對於每個個別的gRNA組合物而言，細胞生長係以方程式(1)表示。根據指數細胞生長模型，帶有個別gRNA組合物的細胞包括兩個細胞群：一個是因CRISPR-Cas9系統所造成的基因破壞而具有修飾的生長速率(k_m)，而另一個(未被修飾的細胞)是具有野生型生長速率(k_{wt})。前者細胞群定義為一個分率p的細胞，其被CRISPR-Cas9系統的剪切效率所限制。簡言之，假設p在整個檢測時程中為常數。

$$C_i(t) = p\,C_i(t) + (1-p)C_i(t) = pC_0e^{k_mt} + (1-p)C_0e^{k_{wt}t} \quad \text{(方程式1)}$$

其中p代表具有修飾生長速率的突變細胞分率，而C_0代表帶有相同條碼化gRNA組合物的初始細胞數目。細胞生長速率(k)是由方程式(2)中的細胞倍增時間($T_{doubling}$)進行估算。由試驗測得野生型OVCAR8-ADR-Cas9細胞的倍增時間為~24小時(數據未呈現)。

$$k = \ln 2 \,/\, T_{doubling} \quad \text{(方程式2)}$$

【0192】 為了簡化模型，將整體細胞群分隔成如方程式(3)所述的3個具有不同生長表型的次群體。

$$C_N(t) = \sum_{i=1}^{N} C_i(t) \approx f_{wt}N \times C_0e^{k_{wt}t} + f_sN \times C_{i,slow}(t) + f_fN \times C_{i,fast}(t)$$

(方程式3)

其中$C_{i,slow}(t)$及$C_{i,fast}(t)$分別代表含有抗增生性gRNA及前期增生性gRNA的細胞的平均生長圖譜，其係由方程式(1)所決定。試驗開始時，整個細胞群中野生型或含有抗增生性gRNA及前期增生性gRNA的百分比分別以f_{wt}、f_s以及f_f表示。

【0193】 根據這個混合細胞生長模型，本文模擬了在整體細胞群中含

有前期增生性gRNA及抗增生性gRNA呈現值的相對頻率(*R. F.*)。相對頻率定義為在給定時間下，相對於初始時間點的條碼豐度($\mathbf{i.e.}, \boldsymbol{R.F.} = \frac{\boldsymbol{F(t)}}{\boldsymbol{F(t=0)}}, \boldsymbol{where\ F} = \boldsymbol{C_i(t)/C_N(t)}$)。在庫中的組合物總數，N，以及初始細胞數，$\boldsymbol{C_0}$，並不會影響相對頻率的結果。以定義的參數進行模擬後，可在細胞群中分別觀察到前期增生性gRNA及抗增生性gRNA的富集與缺乏現象(圖17A)。可觀察到，富集與缺乏的程度隨著定義為具有抗增生性($\boldsymbol{f_s}$)及前期增生性($\boldsymbol{f_f}$)反應的初始gRNA組合物的不同百分比(即，2、5、或10%)而變化(圖17A)。本文進一步藉由調控經修飾細胞的倍增時間($\boldsymbol{T_{doubling,m}}$)，以及針對抗增生性gRNA且具有經修飾生長速率的細胞分率($\boldsymbol{p}$)，而評估抗增生性gRNA呈現值的相對頻率。假定$\boldsymbol{p}$在整個試驗時程中維持常數，在圖17B-17C中所顯示的參數範圍下，整個細胞群中抗增生性殖株的呈現值可能衰減~23%至97%。這個模型藉由將細胞群區分成具有平均生長速率的次細胞群，且不計算細胞間潛在的交互作用，而代表著一個簡易版的細胞生長動力學。根據本文的模型，可藉由增加具有修飾生長速率的細胞分率而提升的gRNA效率，以及增加檢測時間，而提高基因篩選的靈敏度。

【0194】 至此描述了本發明的至少一個實施例的許多方面，但應理解各種變體、修飾及改進將對於本領域技術人員而言是顯而易見的。這樣的變體、修飾及改進旨在作為本公開的一部分，且旨在本發明的精神和範圍內。因此，前面的描述及附圖僅僅是作為示例性。

【0195】

References

1. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343,** 84–7 (2014).

2. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343,** 80–4 (2014).

3. Zhou, Y. *et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509,** 487–491 (2014).

4. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32,** 267–73 (2014).

5. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661 (2014).

6. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* doi:10.1038/nature14136 (2014).

7. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–23 (2013).

8. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339,** 823–6 (2013).

9. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154,** 442–451 (2013).

10. Cheng, A. A., Ding, H. & Lu, T. K. Enhanced killing of antibiotic-resistant bacteria enabled by massively parallel combinatorial genetics. *Proc. Natl. Acad. Sci.* **111,** 12462–7 (2014).

11. Honma, K. *et al.* RPN2 gene confers docetaxel resistance in breast cancer. *Nat. Med.* **14,** 939–948 (2008).

12. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40,** 897–903 (2008).

13. Dawson, M. A. & Kouzarides, T. Cancer epigenetics: From mechanism to therapy. *Cell* **150,** 12–27 (2012).

14. Juergens, R. A. *et al.* Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer Discov.* **1,** 598–607 (2011).

15. Jones, P. A. & Baylin, S. B. The Epigenomics of Cancer. *Cell* **128,** 683–692 (2007).

16. Yoo, C. B. & Jones, P. A. Epigenetic therapy of cancer: past, present and future. *Nat. Rev. Drug Discov.* **5,** 37–50 (2006).

17. Jin, C. *et al.* Chem-seq permits identification of genomic targets of drugs against androgen receptor regulation selected by functional phenotypic screens. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 9235–40 (2014).

18. Asangani, I. a *et al.* Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature* **510,** 278–82 (2014).

19. Kruidenier, L. *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488,** 404–408 (2012).

20. Bliss, C. I. THE TOXICITY OF POISONS APPLIED JOINTLY1. *Ann. Appl. Biol.* **26,** 585–615 (1939).

21. Borisy, A. A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7977–7982 (2003).

22. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31,** 839–43 (2013).

23. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32,** 677–683 (2014).

24. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* **32,** 670–676 (2014).

25. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32,** 1262–1267 (2014).

26. Essletzbichler, P. *et al.* Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res.* **24,** 2059–2065 (2014).

27. Blasco, R. B. *et al.* Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Rep.* **9,** 1219–1227 (2014).

28. Choi, P. S. & Meyerson, M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* **5,** 3728 (2014).

29. Moffat, J. *et al.* A Lentiviral RNAi Library for Human and Mouse Genes Applied to an Arrayed Viral High-Content Screen. *Cell* **124,** 1283–1298 (2006).

30. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–308 (2013).

31. Kim D, *et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.

32. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.

33. Trapnell C, *et al.* (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31:46–53.

34. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545–50.

35. Stemmer M (2015) CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* 10(4):e0124633.

【0196】 **等同物**

【0197】 雖然本文中已描述及說明許多發明性實施例，本領域技術人員將可立即地預想各式其他方法及/或結構而執行其功能及/或獲得其結果及/或如本文所述的一或多種優點，且每個此類變體及/或修飾被認為是在本文所述發明性實施例的範圍中。更概括而言，本領域技術人員可立即理解的是，本文所述的所有參數、尺寸、材料以及配置意指為示例性，而實際的參數、尺寸、材料以及配置將依照本發明教導所使用的特定應用或多項

應用而定。本領域技術人員將理解或能夠使用不超越常規試驗的方式而確認許多等同於本文所述的具體發明性實施例。因此，應理解的是，前述實施例僅是以示例性呈現，且其係在所附的申請專利範圍以及其同等物的範圍內，而發明性實施例可以不同於具體描述以及申請專利範圍宣稱的情況下實施。本發明的發明性實施例是針對本文所述的每個單獨的特徵、系統、製品、材料、套件、及/或方法。此外，假使此類特徵、系統、製品、材料、套件、及/或方法不相互矛盾時，任何二或多個此類特徵、系統、製品、材料、套件、及/或方法的組合係包含於本發明的發明性範圍中。

【0198】 如本文所定義及使用的所有定義，應理解為掌控了字典的定義、藉由引用而納入的文件定義、及/或被定義的術語的普通意義。

【0199】 除非明確指出其相反性，在本文的說明書及申請專利範圍中所使用的不定冠詞「一」以及「一個」應被理解為係指「至少一個」。

【0200】 如本文的說明書及申請專利範圍中所使用的短語「及/或」應被理解為係指這樣連結的元件中「任一個或兩者」，亦即，在某些情況下結合性地存在，而在其他情況下分離地性的存在的元件。以「及/或」表列的複數個元件應以相同方式被解釋，亦即，「一或多個」這樣連結的元件。無論關連於或非關聯於該等被具體辨認的元件而言，視情況可選擇呈現有別於由「及/或」短語所具體辨認的元件的其他元件。因此，作為一個非限制性實例，參照「A及/或B」，當與開放式詞語，例如「包含」，一起使用時，在一個實施例中，僅意指A(視情況可包括除B以外的元件)；而在另一個實施例中，僅意指B(視情況可包括除A以外的元件)；在又另一個實施例中，意指A及B兩者(視情況可包括其他元件)等。

【0201】 如本文的說明書及申請專利範圍中所使用的「或」應被理解為與如上所述的「及/或」具有相同的含義。舉例而言，當在列表中分隔項目時，「或」或者「及/或」應被解釋為包括性，亦即包括至少一個，但亦包括多於一個的數個或表列的元件，且視情況可選擇的額外未被表列的項目。只有明確被指出為相反性的術語，例如「只有一個」或「確僅一個」，或者，當使用於申請專利範圍中，「由.....組成」，將意指包括數個或表列元件中確僅一個元件。一般而言，如本文所使用的術語「或」，當置於排他性術語之前時，例如「任一」、「其中之一」、「只有一個」或「確僅一個」時應僅被理解為表示排他性(亦即，「一個或另一個但不是兩者」)。「基本上由.....組成」，當用於申請專利範圍時，如同在專利法領域中所使用，應具有其普通含義。

【0202】 如本文的說明書及申請專利範圍中所使用的短語「至少一個」，在參照一或多個元件的列表時，應被理解為係指選自在該列表中元件的任何一或多個元件的至少一個元件，但不必要包括具體表列於該元件列表中的各個及每個元件中的至少一個，並且不排除該元件列表中元件的任意組合。無論關連於或非關聯於該等被具體辨認的元件而言，此定義尚允許有別於在術語「至少一個」所指稱的該元件列表中被具體識別的元件而視情況可選擇呈現的元件。因此，作為一個非限制性實例，「A及B中至少一個」(或，等同地，「A或B中至少一個」，或者等同地，「A及/或B中至少一個」)，可在一個實施例中，意指至少一個，視情況可選擇包括多於一個A，且不存有B(且視情況可選擇包括除了B以外的元件)；而在另一個實施例中，意指至少一個，視情況可選擇包括多於一個B，且不存有A(且視情況可選擇包括

除了A以外的元件)；在又另一個實施例中，意指至少一個，視情況可選擇包括多於一個A，以及至少一個，視情況可選擇包括多於一個B(而視情況可選擇包括其他元件)。

【0203】 亦應當理解，除非明確指出相反性，在包括多於一個步驟或動作的任何本文所要求保護的方法中，該方法的步驟或動作的次序，並不一定限定於該方法的步驟或動作所敘述的次序。

【0204】 本文所述的所有參考文獻、專利和專利申請案都透過引用相對於每個被引用的主題而納入本文中，其在某些情況下，可以包括本文件的全部內容。

【0205】 在申請專利範圍中，以及在上述說明書中，所有過渡短語，例如「包含」、「包括」、「攜帶」、「具有」、「含有」、「涉及」、「持有」、「組成」等應被理解為是開放式的，亦即，意指包括但不限定於。僅過渡短語，「由.....組成」以及「基本上由.....組成」應分別為封閉性或半封閉性過渡短語，如美國專利局手冊中的專利審查程序，第2111.03節所述。

【符號說明】

無

【生物材料寄存】無

【序列表】(請換頁單獨記載)

序列表

<110> 麻省理工學院

<120> 用於常間回文重複序列叢集之大量平行組合性基因學

<130>

<140> 104135960
<141> 2015-10-30

<150> US 62/166,302
<151> 2015-05-26

<150> US 62/073,126
<151> 2014-10-31

<160> 387

<170> PatentIn version 3.5

<210> 1
<211> 47
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 1
aatgatacgg cgaccaccga gatctacacg gatccgcaac ggaattc 47

<210> 2
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<220>
<221> misc_feature
<222> (25)..(32)
<223> n is a, c, g, or t

<400> 2
caagcagaag acggcatacg agatnnnnnn nnggttgcgt cagcaaacac ag 52

<210> 3
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 3
aatgatacgg cgaccaccga 20

<210> 4
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 4
caagcagaag acggcatacg a 21

<210> 5
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 5
ccaccgagat ctacacggat ccgcaacgga attc 34

<210> 6
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 6
gtggcgtggt gtgcactgtg tttgctgacg caacc 35

<210> 7
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 7
gggcgaggag ctgttcaccg 20

<210> 8
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 8
cacccagacc atgaagatca 20

<210> 9
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 9
ccacttcaag tgcacatccg 20

<210> 10
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 10
atcgtttccg cttaacggcg 20

<210> 11
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 11
aaacggtacg acagcgtgtg 20

<210> 12
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 12
ccatcaccga tcgtgagcct 20

<210> 13
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 13
ctagacgtcc attcacttcc 20

<210> 14
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 14
tttccaaacc tcgcacgccc 20

<210> 15
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 15
acgtaaagaa gaattatccg 20

<210> 16
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 16
ccgccccacc ttccgtgccg 20

<210> 17
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 17
tggcgctcct ccttgccacg 20

<210> 18

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 18
ccgctccgca gcagagctgc 20

<210> 19
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 19
agagtcgcga gcttgatctt 20

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

<220>
<223> 合成性聚核苷酸

<400> 20
atccgcaccc cggagatcag 20

<210> 21
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 21
gaagactcga tcctcgtcaa 20

<210> 22
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 22
agggatctgc gccccatgta 20

<210> 23
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 23
actcaccttc tatatttcgc 20

<210> 24
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 24
caccaaaatc acgtccatgc 20

<210> 25
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 25
tcacccgtag gcaacgtcgc 20

<210> 26
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 26
gtgtccagcg acgttgccta 20

<210> 27
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 27

acgttgtgca aagactgtcg 20

<210> 28
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 28
cggcgactcc gccatagagc 20

<210> 29
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 29
ggagccggtc cctttcccgt 20

<210> 30
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 30
agtcttgaaa gcgcatgcca 20

<210> 31
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 31
agcggctcta gtatcaaccc 20

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

<220>
<223> 合成性聚核苷酸

<400> 32
gaatcacatg acgcattgtc 20

<210> 33
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 33
cccgcaaatg actggtcacg 20

<210> 34
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 34
ctgtcagaat tgctgcgatc 20

<210> 35
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 35
cttggcaaga cttgcctgac 20

<210> 36
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 36
tagttcccct aacctcaata 20

<210> 37
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 37
acaccattcg taacgttgcc 20

<210> 38
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 38
tcactcgaga tgcgcttgtc 20

<210> 39
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 39
agaatgctgc cgcacgcacc 20

<210> 40
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 40
tccgtaatgt tgctcgatgt 20

<210> 41
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 41
tccaacatcg agcaacatta 20

<210> 42
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 42
tacaacagat cgtgtaatga 20

<210> 43
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 43
gttgactgtg aagctgtacg 20

<210> 44
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 44
aacaggttgc gggaatccaa 20

<210> 45
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 45
tacccagaac atagacacgc 20

<210> 46
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 46
ctcgccgttc gcgtcgccga 20

<210> 47

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 47
cccgtactca cggctgtaga 20

<210> 48
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 48
gggccacgta ccgttgggtg 20

<210> 49
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 49
acaggcttca ttgactgaac 20

<210> 50
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 50
agctgatcaa taactatgat 20

<210> 51
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 51
cctcatctga gtactgattc 20

<210> 52
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 52
acacgcttcc gccaacaaac 20

<210> 53
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 53
tgcgactgag acagctcaag 20

<210> 54
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 54
aaaacttcat ctcccatata 20

<210> 55
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 55
gtacaaattg tacgacggag 20

<210> 56
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 56

gacccctcgg cggtttatag 20

<210> 57
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 57
tatattgcga ccaccaaact 20

<210> 58
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 58
cagagagcac aacgccgcac 20

<210> 59
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 59
ggaaccgctg gcagtcgcgc 20

<210> 60
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 60
ctcccgctgc ccgtgtagac 20

<210> 61
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 61
ctggctcgag atttagcgca 20

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

<220>
<223> 合成性聚核苷酸

<400> 62
aatctgttca tgcgcttacg 20

<210> 63
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 63
gattccagta ccagtacatt 20

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

<220>
<223> 合成性聚核苷酸

<400> 64
ctcaccgtgg tctaacttgt 20

<210> 65
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 65
ggatgtcatg tcctcagcgt 20

<210> 66
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 66
tttgactcct acgcacactt 20

<210> 67
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 67
cgtggatgag tacgaccccg 20

<210> 68
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 68
tcttctgtgc acactatgcg 20

<210> 69
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 69
ctgtcccaga agtgaatcgc 20

<210> 70
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 70
gccatgtgct cgttagcgtc 20

<210> 71
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 71
gcctgacgct aacgagcaca 20

<210> 72
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 72
gaattcatgt actcaactgt 20

<210> 73
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 73
cggaatgcgg ggtccgaact 20

<210> 74
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 74
cagcatacag ctttatccgc 20

<210> 75
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 75
atgaactccc tcttgaaacg 20

<210> 76

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 76
attgtccggc gaggacgtgc 20

<210> 77
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 77
cttcgccacg cgctgtctca 20

<210> 78
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 78
gacggtactg gacgtgggcg 20

<210> 79
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 79
caatccgacc acggggtctg 20

<210> 80
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 80
gaggttcaaa ccgcctgcta 20

<210> 81
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 81
taaagtcggc tggtgacacc 20

<210> 82
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 82
agttcttctc ggtgtccaaa 20

<210> 83
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 83
gactatcagt tccagagata 20

<210> 84
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 84
aacttacgaa ggaaggtctt 20

<210> 85
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 85

ttaccttcgc ccgcttgcgc 20

<210> 86
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 86
ccggccctac tgtcgtgcct 20

<210> 87
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 87
agagccgact tcctcatgac 20

<210> 88
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 88
ataccgcat cgataagtct 20

<210> 89
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 89
atagccaaga cttatcgatg 20

<210> 90
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 90
gaacatacct tctgtagtaa 20

<210> 91
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 91
acgctactat gagaccccag 20

<210> 92
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 92
tatggcaggg agtcgtcgca 20

<210> 93
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 93
gtaacgaatc ctttcttctt 20

<210> 94
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 94
cctcgttctc gtcgtatcgc 20

<210> 95
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 95
gcgttactac gagacgcccg 20

<210> 96
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 96
cttggtcaag cgtccgactg 20

<210> 97
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 97
taaatgccga gagtgtcgct 20

<210> 98
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 98
gtctgtcaaa accgacttcc 20

<210> 99
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 99
gatactgctt ggctgtactg 20

<210> 100
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 100
tcttgtatgg gcgccccgtg 20

<210> 101
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 101
gccttgactg ttaccggctc 20

<210> 102
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 102
tcctgagccg gtaacagtca 20

<210> 103
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 103
actccgcaca gttaaaacca 20

<210> 104
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 104
gcggaactct cgaacagtca 20

<210> 105

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 105
ttccactcac ttatcgctat 20

<210> 106
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 106
ccccgcgtac ttctcgctgt 20

<210> 107
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 107
gtatgatgac atcgacgacg 20

<210> 108
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 108
tcaccaggta ctgtaccccg 20

<210> 109
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 109
cctttgcaag acccgcacga 20

<210> 110
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 110
agtaggcttc gtgtgatcaa 20

<210> 111
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 111
gtctaaagga gcccatcgtg 20

<210> 112
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 112
catgaacccc aacgtgctaa 20

<210> 113
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 113
ctgggattca ataactcgg 20

<210> 114
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 114

tctctggtat gaaagtgccg 20

<210> 115
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 115
gtccgcgaac tcttcccagc 20

<210> 116
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 116
tcgaagaccg ggcactcggg 20

<210> 117
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 117
ggacttattt cagcttaata 20

<210> 118
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 118
cttaccgcca tgacacactt 20

<210> 119
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 119
gataaacaat gcgttcgtag 20

<210> 120
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 120
gggctacccg agcccaccga 20

<210> 121
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 121
gatttactcc tcgcgtccaa 20

<210> 122
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 122
aaagacttac cgcgggtggg 20

<210> 123
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 123
taaggcccga catggaaccg 20

<210> 124
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 124
actgtaaact gtagtacctc 20

<210> 125
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 125
cagcattatc tgcataccag 20

<210> 126
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 126
agactatgag tctagtttaa 20

<210> 127
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 127
taccacagcg cccttcgata 20

<210> 128
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 128
atccccctcc tcgtagcgca 20

<210> 129
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 129
caaaggcttc ccgtgcagcg 20

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

<220>
<223> 合成性聚核苷酸

<400> 130
ttctgcacgg gcttgacgtc 20

<210> 131
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 131
gacgtcaagc ccgtgcagaa 20

<210> 132
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 132
cagtgacgtc gagaactacg 20

<210> 133
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 133
tctgacgaac gtagggctcc 20

<210> 134

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 134
ggcttagtga aaaaacgccg 20

<210> 135
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 135
cctcgccatc attcactgtg 20

<210> 136
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 136
aacgtgtatt gttcgttacc 20

<210> 137
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 137
tcctacctta tattcagtag 20

<210> 138
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 138
aaaggtttac catcagcaga 20

<210> 139
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 139
caccgtgttc tatagagccg 20

<210> 140
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 140
cggcgcgagg tggacagcat 20

<210> 141
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 141
cgactcaccg gctgcgatcc 20

<210> 142
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 142
cgacgtgacg tttgcagtga 20

<210> 143
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 143

caaaggtcgg aagccggctg 20

<210> 144
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 144
catcactgca aacgtcacgt 20

<210> 145
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 145
actagcatgt ctgcggagag 20

<210> 146
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 146
tctagtccat cccccattac 20

<210> 147
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 147
gggaacaata aagaagcgct 20

<210> 148
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 148
gagatcgacg cgaaatacca 20

<210> 149
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 149
tataaatccg cgcccgaaag 20

<210> 150
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 150
cccataccag ttattgcgct 20

<210> 151
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 151
gcagcagcaa ctgtactcgt 20

<210> 152
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 152
gcagcgactc cacgcactca 20

<210> 153
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 153
gatcttcaag aagaccccgc 20

<210> 154
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 154
tctcgcgcat ttccgtgaag 20

<210> 155
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 155
cttcacggaa atgcgcgaga 20

<210> 156
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 156
tcgatactgc atttgtaatc 20

<210> 157
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 157
gctgctcgtg ctcgttccaa 20

<210> 158
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 158
cctagaaggc cggactcaaa 20

<210> 159
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 159
ggcactactc atatactcag 20

<210> 160
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 160
gatctgcttc aaagcgcgcc 20

<210> 161
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 161
cttccagctg atgcgagagc 20

<210> 162
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 162
gaagttcctc tgaagttcgc 20

<210> 163

<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 163
cgagggcgac ttaaccttag g 21

<210> 164
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 164
aaatcttcgt aatccaagta t 21

<210> 165
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 165
gtaataccgg gtgttccgat g 21

<210> 166
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 166
attaatccac acgaggtctc c 21

<210> 167
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 167
tatagtaatc agggaggttc a 21

<210> 168
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 168
tttagacttg attgtgctca t 21

<210> 169
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 169
agaaattaaa cggctaccct cca 23

<210> 170
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 170
tccatagtgt cttgagcacc ac 22

<210> 171
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 171
cacttgctga tgccagtagg ag 22

<210> 172
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 172

gtgaatagct tgggaatgtg gg 22

<210> 173
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 173
gccacactct acatgggagc 20

<210> 174
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 174
gacttttcca tcagggacac ct 22

<210> 175
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 175
ggtggacaaa cacattcggc 20

<210> 176
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 176
cctatcattg ccccaaggag tc 22

<210> 177
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 177
ctaggcctcc gacagttgta at 22

<210> 178
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 178
agccaccctt ggttggtttt 20

<210> 179
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 179
ggtaagggaa actctggggc 20

<210> 180
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 180
ctccctccct tcctaaggct 20

<210> 181
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 181
attgccttaa gtcgacacct gat 23

<210> 182
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 182
gactgtagag ttgccggaac ag 22

<210> 183
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 183
gttggtacaa agtggtgaag gc 22

<210> 184
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 184
acgtggcttc attgtacatc ct 22

<210> 185
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 185
aagcacatgc ttcaggctaa ca 22

<210> 186
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 186
tcatctgctc ttacgcttag cc 22

<210> 187
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 187
ctcggcaacc ctccatacat 20

<210> 188
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 188
aaccatgcac agaatccaga ttta 24

<210> 189
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 189
aagagttctt ggcgcagaca 20

<210> 190
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 190
tcgtccaagt tacagtcatc aca 23

<210> 191
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 191
tcctctggtg cacagaaaag tc 22

<210> 192

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 192
ttctctccag acactgccct 20

<210> 193
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 193
gtgcccagaa ctactgccat 20

<210> 194
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 194
gaggtgagtt ccagcttccc 20

<210> 195
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 195
caccttacag gcactgcgtt 20

<210> 196
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 196
ctccctccct agaggctatg ag 22

<210> 197
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 197
gttgatgtga ttgccggctc 20

<210> 198
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 198
cgtacgggtt catgcaagtt 20

<210> 199
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 199
cccctcaccg cctatcagta 20

<210> 200
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 200
ttaggcagga cctgtttcgg 20

<210> 201
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 201

cgtttgctta agagcacctc c 21

<210> 202
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 202
tcttgaacaa gtcggggtcg 20

<210> 203
<211> 61
<212> DNA
<213> Homo sapiens

<400> 203
tttcaacagt ttcctacctt atatcagtag tggtctggtc ttgtgaactt ggacatcaca 60

a 61

<210> 204
<211> 62
<212> DNA
<213> Homo sapiens

<400> 204
tgactcacaa tctctcatgt ttttctggct tagtgaaaaa acgccgtgga tcttcaaagg 60

gg 62

<210> 205
<211> 44
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 205
tttcaacagt ttcctacctt atattgtgaa cttggacatc acaa 44

<210> 206
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 206

tgactcacaa tctctcatgt ttttctggct tagtgaaaaa acggccgtgg atcttcaaag 60

ggg 63

<210> 207
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 207
tgactcacaa tctctcatgt ttttctggct tagtgaaaaa acgtggatct tcaaagggg 59

<210> 208
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 208
tttcaacagt ttcctacctt atattcagtg gtctggtctt gtgaacttgg acatcacaa 59

<210> 209
<211> 13
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 209
ttcaacatca caa 13

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

<220>
<223> 合成性聚核苷酸

<400> 210
tgactcacaa tctctcatgt ttttctggct tagtgaaaaa acgtggatct tcaaagggg 59

<210> 211
<211> 60
<212> DNA
<213> Homo sapiens

<400> 211
tagacatttg ggaagtttct agtccatccc ccattactgg cagatttctc aatctcgtcc 60

<210> 212
<211> 59
<212> DNA
<213> Homo sapiens

<400> 212
gaaggataat cacctttgca agacccgcac gatgggctcc tttagactcc atgtatgca 59

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

<220>
<223> 合成性聚核苷酸

<400> 213
gaaggataat cacctttgca agacccacga tgggctcctt tagactccat gtatgca 57

<210> 214
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 214
gaaggataat cacctttgca agacccgcga tgggctcctt tagactccat gtatgca 57

<210> 215
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 215
tagacatttg ggaagtttct agtccatcct ggcagatttc tcaatctcgt cc 52

<210> 216
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 216

gaaggataat cacctttgca agacccgccg atgggctcct ttagactcca tgtatgca 58

<210> 217
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 217
gaaggataat cacctttgca aatccatgta tgca 34

<210> 218
<211> 50
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 218
gaaggataat cacctttgca agatgggctc ctttagactc catgtatgca 50

<210> 219
<211> 49
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 219
tttcaacagt ttcctacctt atattcagtt gtgaacttgg acatcacaa 49

<210> 220
<211> 65
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 220
tgactcacaa tctctcatgt ttttctggct tagtgaaaaa acaaaaccgt ggatcttcaa 60

agggg 65

<210> 221
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 221
tttcaacagt ttcctacctt attagtggtc tggtcttgtg aacttggaca tcacaa 56

<210> 222
<211> 52
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 222
tttcaacagt ttcctacctt atattctggt cttgtgaact tggacatcac aa 52

<210> 223
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 223
tttcaacagt ttcctacctt atattctagt ggtctggtct tgtgaacttg gacatcacaa 60

<210> 224
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 224
tttcaacagt ttcctacctt ttagtggtct ggtcttgtga acttggacat cacaa 55

<210> 225
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 225
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcagg 60

tcttgtgaac ttggacatca caaataggac 90

<210> 226

<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 226
cgttctgact cacaatctct catgttttc tggcttagtg aaaaaacttc aaaggggcat 60

gatacacaca aggggaaacc agtgaagacc 90

<210> 227
<211> 91
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 227
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tataggtctg 60

gtcttgtgaa cttggacatc acaaatagga c 91

<210> 228
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 228
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcagt 60

ggtctggtct tgtgaacttg gacatcacaa ataggac 97

<210> 229
<211> 99
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 229
cgttctgact cacaatctct catgttttc tggcttagtg aaaaacgccg tggatcttca 60

aaggggcatg atacacacaa ggggaaacca gtgaagacc 99

<210> 230
<211> 95
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 230
cgttctgact cacaatctct catgttttc tggcttagtg aaaaactgga tcttcaaagg 60

ggcatgatac acacaagggg aaaccagtga agacc 95

<210> 231
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 231
actccaataa taattacaaa caaggaattt caacagtttc ctaccttata ttctgtgatc 60

tgtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 232
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 232
cgttctgact cacaatctct catgttttc tggcttagtg aaaaaacgtg gatcttcaaa 60

ggggcatgat acacacaagg ggaaaccagt gaagacc 97

<210> 233
<211> 100
<212> DNA
<213> Homo sapiens

<400> 233
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcagt 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 234
<211> 100
<212> DNA
<213> Homo sapiens

<400> 234
cgttctgact cacaatctct catgttttc tggcttagtg aaaaaacgcc gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 235
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 235
atttcaacag tttcctacct tatatactat actatatata tatatatata ctatatatat 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 236
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 236
ctctcatgtt tttctggctt agtgaaaaaa cttcactaag tttttactta gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 237
<211> 95
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 237
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattctgg 60

tctggtcttg tgaacttgga catcacaaat aggac 95

<210> 238
<211> 76
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 238
tacaactcca ataataatta caaacaagga atttcaacag tttcctactt gtgaacttgg 60

acatcacaaa taggac 76

<210> 239
<211> 96
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 239
cgttctgact cacaatctct catgttttc tggcttagtg aaaaacatgg atcttcaaag 60
gggcatgata cacacaaggg gaaaccagtg aagacc 96

<210> 240
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 240
cgttctgact cacaatctct catgttttc tggcttagtg aaaaaacgtg gatcttcaaa 60
ggggcatgat acacacaagg ggaaaccagt gaagacc 97

<210> 241
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 241
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct agtggtctgg 60
tcttgtgaac ttggacatca caaataggac 90

<210> 242
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 242
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattctgt 60
ggtctggtct tgtgaacttg gacatcacaa ataggac 97

<210> 243

<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 243
cgttctgact cacaatctct catgttttc tggcttagtg aaaaaactaa gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 244
<211> 89
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 244
tacaactcca ataataatta caaacaagga atttcaacag tttcctacta gtggtctggt 60

cttgtgaact tggacatcac aaataggac 89

<210> 245
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 245
cgttctgact cacaatctct catgttttc tggcttagtg aatcttcaaa ggggcatgat 60

acacacaagg ggaaaccagt gaagacc 87

<210> 246
<211> 86
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 246
tacaactcca ataataatta caaacaagga atttcaacag tttccaagtg gtctggtctt 60

gtgaacttgg acatcacaaa taggac 86

<210> 247
<211> 97
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 247
cgttctgact cacaatctct catgtttttc tggcttagtg aaaaaacgtg gatcttcaaa 60

ggggcatgat acacacaagg ggaaaccagt gaagacc 97

<210> 248
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 248
acaactccaa taataattac aaacaaggaa tttcaacagt ttcctacctt atattcagat 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 249
<211> 86
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 249
tacaactcca ataataatta caaacaagga atttcaacag tttcctagtg gtctggtctt 60

gtgaacttgg acatcacaaa taggac 86

<210> 250
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 250
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattctag 60

ggtctggtct tgtgaacttg gacatcacaa ataggac 97

<210> 251
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 251
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcagt 60

tgtgaacttg gacatcacaa ataggac 87

<210> 252
<211> 95
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 252
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatatagtgg 60

tctggtcttg tgaacttgga catcacaaat aggac 95

<210> 253
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 253
tctgactcac aatctctcat gtttttctgg cttagtgaaa aacatgcccc gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 254
<211> 91
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 254
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tattggtctg 60

gtcttgtgaa cttggacatc acaaatagga c 91

<210> 255
<211> 78
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 255
tacaactcca ataataatta caaacaagga atttcaacag tggtctggtc ttgtgaactt 60

ggacatcaca aataggac 78

<210> 256
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 256
tctcatgttt ttctggctta gtgaaatcta agcttagtga aatctaagcc gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 257
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 257
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattctta 60

tatggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 258
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 258
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatatatagt 60

ggtctggtct tgtgaacttg gacatcacaa ataggac 97

<210> 259
<211> 82
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 259

tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcaga 60

acttggacat cacaaatagg ac 82

<210> 260
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 260
gttctgactc acaatctctc atgtttttct ggcttagtga aaaaaacgcc gtggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 261
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 261
taattacaaa caaggaattt caacagtttc ctaccttata ttcagtataa tatattcatt 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 262
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 262
cgttctgact cacaatctct catgtttttc tggcttagtg gatcttcaaa ggggcatgat 60

acacacaagg ggaaaccagt gaagacc 87

<210> 263
<211> 98
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 263
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattctag 60

tggtctggtc ttgtgaactt ggacatcaca aataggac 98

<210> 264
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 264
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattatat 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 265
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 265
taattacaaa caaggaattt caacagtttc ctaccttata ttcaggtagt gaatctgaat 60

agtggtctgg tcttgtgaac ttggacatca caaataggac 100

<210> 266
<211> 85
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 266
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tctggtcttg 60

tgaacttgga catcacaaat aggac 85

<210> 267
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 267
acaatctctc atgtttttct ggcttagtga atcttcaaag ggatcttcaa aaggatcttc 60

aaaggggcat gatacacaca aggggaaacc agtgaagacc 100

<210> 268
<211> 82
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 268
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattgtga 60
acttggacat cacaaatagg ac 82

<210> 269
<211> 97
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 269
tacaactcca ataataatta caaacaagga atttcaacag tttcctacct tatattcagt 60
ggtctggtct tgtgaacttg gacatcacaa ataggac 97

<210> 270
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 270
gttgtgctca gtactgactt 20

<210> 271
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 271
tttcagcttc caataaaaac 20

<210> 272
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 272
attccacggg aaggagatct 20

<210> 273
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 273
aggattgaag ctgacgttct 20

<210> 274
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 274
ggcttagtga aaaaacgccg 20

<210> 275
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 275
tcctaccttа tattcagtag 20

<210> 276
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 276
gctttgcccg aaccgcacga 20

<210> 277
<211> 20
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 277
cctaggccag acctgcacga 20

<210> 278
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 278
gctctggaag acccgcacca 20

<210> 279
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 279
ccttatcaag acccacacca 20

<210> 280
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 280
atccccctcc tcgtagcgca 20

<210> 281
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 281
ctgctcctcc tcgtagcgct 20

<210> 282

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 282
tgcgccctcc tcctagcgca 20

<210> 283
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 283
ctcctcctcc gcgtagcgct 20

<210> 284
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 284
ctgcccctcc tggtagcgcc 20

<210> 285
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 285
aaccagctcc tcgtagctca 20

<210> 286
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 286
accgccctcc tcctagctca 20

<210> 287
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 287
agcccgctcc tcgtggggca 20

<210> 288
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 288
gggaacaata aagaagcgct 20

<210> 289
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 289
tggaaaaaca aagaagagct 20

<210> 290
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 290
gggaagtata aggaagagct 20

<210> 291
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 291

gtgagcaata aagcagccct 20

<210> 292
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 292
tctagtccat cccccattac 20

<210> 293
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 293
aatattccat tccccattac 20

<210> 294
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 294
tgttgtccat acctcattac 20

<210> 295
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 295
tctaggtcat gcaccattac 20

<210> 296
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 296
cccattcctt cccccattac 20

<210> 297
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 297
tccacaccct cccccattac 20

<210> 298
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 298
acactctttc cctacacgac gctcttccga tctggtatct gtggttgatg cagttttcc 59

<210> 299
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 299
acactctttc cctacacgac gctcttccga tctgcacaga gctgctgctt ggagtg 56

<210> 300
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 300
acactctttc cctacacgac gctcttccga tctcgctttc ctgcctcagg atgaac 56

<210> 301
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 301
acactctttc cctacacgac gctcttccga tctttggaag agaaggatct gctgaggc 58

<210> 302
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 302
acactctttc cctacacgac gctcttccga tctgctaccc tccacaaagc acacac 56

<210> 303
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 303
acactctttc cctacacgac gctcttccga tctccttcag aaacaatgtc ccaaatcg 58

<210> 304
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 304
acactctttc cctacacgac gctcttccga tctagccttt ctgagagcgg gctag 55

<210> 305
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 305
acactctttc cctacacgac gctcttccga tctgcctcct ctcatcctct cgcttc 56

<210> 306
<211> 58
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 306
acactctttc cctacacgac gctcttccga tctgattggc tccaagcggc catcaaac 58

<210> 307
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 307
acactctttc cctacacgac gctcttccga tctttgaact caaggctcag ccaacaggc 59

<210> 308
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 308
acactctttc cctacacgac gctcttccga tctcaactca ggctggatgc atcgg 55

<210> 309
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 309
acactctttc cctacacgac gctcttccga tctttggtgg ccgctgagtg tgtgtac 57

<210> 310
<211> 53
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 310
acactctttc cctacacgac gctcttccga tctcagcatg ttgacatagc ggc 53

<210> 311

<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 311
acactctttc cctacacgac gctcttccga tcttgagagg agatgagtcg gggtc 55

<210> 312
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 312
acactctttc cctacacgac gctcttccga tctgacgggt caaagcctca ggagag 56

<210> 313
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 313
acactctttc cctacacgac gctcttccga tctttagctg cccagctcac agctacc 57

<210> 314
<211> 53
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 314
acactctttc cctacacgac gctcttccga tctgagcccc aagagcgaga caa 53

<210> 315
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 315
acactctttc cctacacgac gctcttccga tctggcgctc agaaggctgt gcag 54

<210> 316
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 316
acactctttc cctacacgac gctcttccga tctcctaggt gacactggac ttttgc 56

<210> 317
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 317
acactctttc cctacacgac gctcttccga tctaggttca cctcaggctg ctcagaag 58

<210> 318
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 318
acactctttc cctacacgac gctcttccga tctcgtctct ctccatgtga gcttgtg 57

<210> 319
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 319
acactctttc cctacacgac gctcttccga tctggaaaga tcatctgatc aggcccatc 59

<210> 320
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 320

acactctttc cctacacgac gctcttccga tctcttgtta gggttggagg tctctgg 57

<210> 321
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 321
acactctttc cctacacgac gctcttccga tctaaggccc gtaaagggca agttcag 57

<210> 322
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 322
acactctttc cctacacgac gctcttccga tctggtgaca ggaagctgtc ggaacat 57

<210> 323
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 323
acactctttc cctacacgac gctcttccga tctctcactg tgatctgaca ccaaacac 58

<210> 324
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 324
acactctttc cctacacgac gctcttccga tctctagttg ccttcatgcc ttacagac 58

<210> 325
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 325
acactctttc cctacacgac gctcttccga tctctgctcc cactccagac taccc 55

<210> 326
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 326
gtgactggag ttcagacgtg tgctcttccg atctagtagt gaggccgctt ataacc 56

<210> 327
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 327
gtgactggag ttcagacgtg tgctcttccg atcttaacaa ggagatgccc tggctgg 57

<210> 328
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 328
gtgactggag ttcagacgtg tgctcttccg atctaggtca tgaaggcaaa ctcagcc 57

<210> 329
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 329
gtgactggag ttcagacgtg tgctcttccg atctcacctg tcaaaagtcc tactccgg 58

<210> 330
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 330
gtgactggag ttcagacgtg tgctcttccg atctcctgga gaccagcaag tattgtcc 58

<210> 331
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 331
gtgactggag ttcagacgtg tgctcttccg atctgtcctc tgaaccccag ctgtaag 57

<210> 332
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 332
gtgactggag ttcagacgtg tgctcttccg atctcaaaca gaggccaaag ggtgtccc 58

<210> 333
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 333
gtgactggag ttcagacgtg tgctcttccg atctacagga ggtcgtggtg cagttctc 58

<210> 334
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 334
gtgactggag ttcagacgtg tgctcttccg atctttgtgt gaggaacgtt gacgctacc 59

<210> 335
<211> 57
<212> DNA

<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 335
gtgactggag ttcagacgtg tgctcttccg atctgagcgt aggtcctctg catggag 57

<210> 336
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 336
gtgactggag ttcagacgtg tgctcttccg atctccacag aatgacagga acccatgg 58

<210> 337
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 337
gtgactggag ttcagacgtg tgctcttccg atctactgag cagagcctag gaggcag 57

<210> 338
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 338
gtgactggag ttcagacgtg tgctcttccg atcttgttgc cagatccaga ggcgtc 56

<210> 339
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 339
gtgactggag ttcagacgtg tgctcttccg atctaactgg cccgagtagt cggagcag 58

<210> 340

<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 340
gtgactggag ttcagacgtg tgctcttccg atcttcctca gagtgtgtgg aagtgctgg 59

<210> 341
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 341
gtgactggag ttcagacgtg tgctcttccg atctaagagc tcctagggga ggatcag 57

<210> 342
<211> 56
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 342
gtgactggag ttcagacgtg tgctcttccg atcttggcag gagcacagcc taagga 56

<210> 343
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 343
gtgactggag ttcagacgtg tgctcttccg atctacaccc gcctcggaga tcaacac 57

<210> 344
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 344
gtgactggag ttcagacgtg tgctcttccg atctccaccc ctacatctca ccttgttg 58

<210> 345
<211> 55
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 345
gtgactggag ttcagacgtg tgctcttccg atctgtgagg tttccacgtg ccagc 55

<210> 346
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 346
gtgactggag ttcagacgtg tgctcttccg atctccaaca attccaggta tgaaactccc 60

<210> 347
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 347
gtgactggag ttcagacgtg tgctcttccg atctctcccc acttgtaggt tcctaatcc 59

<210> 348
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 348
gtgactggag ttcagacgtg tgctcttccg atctgtagag tgcctggtga agaatgtg 58

<210> 349
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 349

gtgactggag ttcagacgtg tgctcttccg atcttccaga ctgttgttca gtcctgt 57

<210> 350
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 350
gtgactggag ttcagacgtg tgctcttccg atctctctgg atttgcccac acctagtc 58

<210> 351
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 351
gtgactggag ttcagacgtg tgctcttccg atctgcatgc ttgctttctg aaggtggc 58

<210> 352
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 352
gtgactggag ttcagacgtg tgctcttccg atctttccaa gcaagtgagc ttcagcacc 59

<210> 353
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 353
gtgactggag ttcagacgtg tgctcttccg atctccaccc atgacacagg aggg 54

<210> 354
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 354
gtctcgtggg ctcggagatg tgtataagag acaggctacc ctccacaaag cacacac 57

<210> 355
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 355
gtctcgtggg ctcggagatg tgtataagag acagttggaa gagaaggatc tgctgaggc 59

<210> 356
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 356
tcgtcggcag cgtcagatgt gtataagaga cagcctggag accagcaagt attgtcc 57

<210> 357
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 357
tcgtcggcag cgtcagatgt gtataagaga cagcacctgt caaaagtcct actccgg 57

<210> 358
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<220>
<221> misc_feature
<222> (6)..(24)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (51)..(58)

<223> n is a, c, g, or t

<400> 358
caccgnnnnn nnnnnnnnnn nnnngtttgg gtcttcgaga agacctattc nnnnnnnnc 59

<210> 359
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<220>
<221> misc_feature
<222> (6)..(13)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (40)..(58)
<223> n is a, c, g, or t

<400> 359
aattgnnnnn nnngaatagg tcttctcgaa gacccaaacn nnnnnnnnnn nnnnnnnnc 59

<210> 360
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<220>
<221> misc_feature
<222> (6)..(24)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (51)..(58)
<223> n is a, c, g, or t

<400> 360
caccgnnnnn nnnnnnnnnn nnnngtttgg gtcttcgaga agacctattc nnnnnnnnca 60

att 63

<210> 361
<211> 63
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<220>
<221> misc_feature
<222> (6)..(13)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (40)..(58)
<223> n is a, c, g, or t

<400> 361
aattgnnnnn nnngaatagg tcttctcgaa gacccaaacn nnnnnnnnnn nnnnnnncg 60

gtg 63

<210> 362
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 362
cctttgcaag acccgcacga tgg 23

<210> 363
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 363
gctttgcccg aaccgcacga gag 23

<210> 364
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 364
cctaggccag acctgcacga tgg 23

<210> 365

<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 365
gctctggaag acccgcacca ggg 23

<210> 366
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 366
ccttatcaag acccacacca gag 23

<210> 367
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 367
atcccctcc tcgtagcgca tgg 23

<210> 368
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 368
tgcgccctcc tcctagcgca tgg 23

<210> 369
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 369
ctgcccctcc tggtagcgcc tgg 23

<210> 370
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 370
ctgctcctcc tcgtagcgct ggg 23

<210> 371
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 371
aaccagctcc tcgtagctca ggg 23

<210> 372
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 372
agcccgctcc tcgtggggca cgg 23

<210> 373
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 373
accgccctcc tcctagctca ggg 23

<210> 374
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 374

ctcctcctcc gcgtagcgct tgg 23

<210> 375
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 375
ctcctcctcc tcctaccgca agg 23

<210> 376
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 376
attcctctcc tggtaccgca agg 23

<210> 377
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 377
ctcaccctcc tcctagcaca tag 23

<210> 378
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 378
gggaacaata aagaagcgct tgg 23

<210> 379
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 379
tggaaaaaca aagaagagct tgg 23

<210> 380
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 380
gggaagtata aggaagagct cag 23

<210> 381
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 381
gtgagcaata aagcagccct aag 23

<210> 382
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 382
tctagtccat cccccattac tgg 23

<210> 383
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 383
aatattccat tccccattac tgg 23

<210> 384
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 384
tgttgtccat acctcattac tgg 23

<210> 385
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 385
tctaggtcat gcaccattac tgg 23

<210> 386
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 386
tccacaccct cccccattac tag 23

<210> 387
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成性聚核苷酸

<400> 387
cccattcctt cccccattac cag 23

【發明申請專利範圍】

【請求項1】　一種基因構築體，其包含

第一 DNA 元件，其包含

CRISPR 引導序列以及

支架序列，其係將一種核酸內切酶吸引至標靶核酸；

在該第一 DNA 元件兩側的第一相容端元件及第二相容端元件，其中該第一及第二相容端元件能夠互相黏合；

條碼元件；

在該條碼元件兩側的第三相容端元件及第四相容端元件，其中該第三及第四相容端元件能夠互相黏合，但不能黏合於該第一或第二相容端元件；以及

位於該第四相容端元件及該第一相容端元件之間的分隔位點，其中該 DNA 元件、第一相容端元件、以及第二相容端元件係位於該分隔位點的一側，而該條碼元件、該第三相容端元件、以及該第四相容端元件係位於該分隔位點的另一側。

【請求項2】　如請求項 1 之基因構築體，其進一步包含位於該第一 DNA 元件上游的啟動子元件。

【請求項3】　一種載體，其包含根據請求項 1 或 2 之基因構築體。

【請求項4】　一種基因構築體，其包含：

複數個 DNA 元件，其中該複數個 DNA 元件中的各個 DNA 元件包含 CRISPR 引導序列及支架序列，其係將一種核酸內切酶吸引至標靶核酸；

在該複數個 DNA 元件兩側的第一相容端元件和第二相容端元件，其中該第一及第二相容端元件能夠互相黏合；

複數個條碼元件；

在該複數個條碼元件兩側的第三相容端元件和第四相容端元件，其中該第三和第四相容元件能夠互相黏合，但不會與該第一或第二相容端元件黏合；以及

位於該複數個 DNA 元件及該複數個條碼元件之間的分隔位點。

【請求項5】　一種載體，其包含：

根據請求項 4 之基因構築體，及

位於各個該 CRISPR 引導序列上游的啟動子序列。

【請求項6】　一種產生組合性載體之方法，其包含：

(a) 提供一種載體，其含有第一基因構築體，其包含：

CRISPR 引導序列；

第二相容端元件及在該 CRISPR 引導序列側邊且針對第一限制酶的第一識別位點；

條碼元件；以及

第三相容端及在該條碼元件側邊且針對第二限制酶的第二識別位點；

(b) 在該第一識別位點上剪切該第一基因構築體，而產生第五相容端元件，且在該第二識別位點上剪切該載體而產生第六相容端元件；

(c) 提供一種支架元件，其包含：

支架序列，其係將一種核酸內切酶吸引至標靶核酸；

包含第一相容端元件及第四相容端元件的分隔位點；以及

在該支架元件兩側的第七相容端元件及第八相容端元件，其中該第七相容端元件能夠黏合於該第五相容端元件，而該第八相容端元件則能夠黏合於該第六相容端元件；

(d) 將該支架元件黏合於該經剪切的第一基因構築體，其中該黏合係發生在該載體和該支架元件內能夠互相黏合的相容端元件處，且其中，在黏合後，該支架元件係被整合於該 CRISPR 引導序列和條碼元件之間，且其中，該分隔位點係位於該支架序列和該條碼元件之間，進而創建出組合性載體。

【請求項7】　如請求項 6 之方法，其進一步包含：

(a) 提供根據請求項 6 之組合性載體；

(b) 在該支架元件內的分隔位點上剪切該載體，而產生第一相容端元件及第四相容端元件；

(c) 提供第二基因構築體，其包含

CRISPR 引導序列；

支架序列，其係將一種核酸內切酶吸引至標靶核酸；

條碼元件；以及

在該第二基因構築體兩側的第二相容端元件和第三相容端元件，其中該第二基因構築體的第二相容端元件能夠與該載體的第一相容端

元件黏合，而該第二基因構築體的第三相容端元件則能夠與該載體的第四相容端元件黏合；

(d)使該第二基因構築體黏合於該經剪切的載體上，其中該黏合係發生在該第二基因構築體與載體內且能夠互相黏合的相容端元件處，且其中，在黏合後，該第二基因構築體係被整合於該載體，進而創建出包含串聯的條碼元件和串聯的 CRISPR 引導及支架序列的組合性載體。

【請求項8】　如請求項 7 之方法，其中該組合性載體進一步包含位於該 CRISPR 引導序列上游的啟動子元件。

【請求項9】　如請求項 7 之方法，其中該方法係反覆進行的。

【請求項10】　如請求項 8 之方法，其中該方法係反覆進行的。

【請求項11】　如請求項 6 至 10 項中任一項之方法，其中該第一識別位點及該第二識別位點具有相同的識別位點序列，且該第一限制酶及該第二限制酶為相同的限制酶。

【請求項12】　一種基因構築體，其包含

至少兩個 CRISPR 引導序列；

支架序列，其係將一種核酸內切酶吸引至標靶核酸；

條碼序列；以及

介於各個 CRISPR 引導序列之間，且介於該條碼元件和該 CRISPR 引導序列之間最靠近該條碼元件的限制酶識別位點。

【請求項13】　一種基因構築體，其包含

複數個 DNA 元件，各個包含

CRISPR 引導序列，及

支架序列，其係將一種核酸內切酶吸引至標靶核酸；；

條碼元件；以及

位於該複數個 DNA 元件的各 DNA 元件上游的啟動子序列。

【請求項14】　如請求項 12 或 13 之基因構築體，其中該條碼序列係位於該基因構築體的 5’端。

【請求項15】　如請求項 12 或 13 之基因構築體，其中該條碼序列係位於該基因構築體的 3’端。

【請求項16】　一種載體，其包含根據請求項 12 至 15 之任何基因構築體。

【請求項17】　一種產生一個組合性載體的方法，其包含

(a) 提供一種載體，其包含：

複數個 CRISPR 引導序列；

條碼元件，其中該條碼元件係位於該複數個 CRISPR 引導序列的下游；

以及

複數個針對複數個限制酶的識別位點，其中各個該複數個識別位點係位於該複數個 CRISPR 引導序列中之一個的下游；

(b) 以至少一個的該複數個限制酶在至少一個的該複數個識別位點上剪切該載體，而產生第一相容端元件及第二相容端元件；

(c) 提供第一支架元件，其包含：

支架序列，其係將一種核酸內切酶吸引至標靶核酸；以及

在該第一支架元件兩側的第三相容端元件及第四相容端元件，其中該第三相容端元件能夠黏合於該經剪切載體的第一相容端元件，且該第四相容端元件能夠黏合於該經剪切載體的第二相容端元件；

(d) 使該第一支架元件黏合於該經剪切的載體，其中該黏合係發生在該第一支架元件與經剪切載體內的相容端元件處，且其中，在黏合後，該第一支架元件係被整合於該複數個 CRISPR 引導序列中之一個的下游，因而產生組合性載體。

【請求項18】　如請求項 17 之方法，其中在(a)中，該載體進一步包含位於該複數個 CRISPR 引導序列中至少一個的上游的啟動子序列；及/或在(c)中，該第一支架元件進一步包含啟動子序列。

【請求項19】　如請求項 17 或 18 之方法，其中該方法係反覆進行的。

【請求項20】　一種產生組合性載體的方法，其包含

(a) 提供一個載體，其包含：

複數個 CRISPR 引導序列，

條碼元件，其中該條碼元件係位於該複數個 CRISPR 引導序列的上游；；以及

複數個針對複數個限制酶的識別位點，其中各個該複數個識別位點係位於該複數個 CRISPR 引導序列中之一個的上游；

(b) 以至少一個的該複數個限制酶在至少一個的該複數個識別位點上剪切該載體，而產生第一相容端元件及第二相容端元件；

(c) 提供第一支架元件，其包含：

支架序列，其係將一種核酸內切酶吸引至標靶核酸，

啟動子序列，以及

在該第一支架元件兩側的第三相容端元件及第四相容端元件，其中該第三相容端元件能夠黏合於該經剪切載體的第一相容端元件，且該第四相容端元件能夠黏合於該經剪切載體的第二相容端元件；

(d) 使該第一支架元件黏合於該經剪切的載體，其中該黏合係發生在該第一支架元件與經剪切載體內的相容端元件處，且其中，在黏合後，該第一支架元件係被整合於該複數個 CRISPR 引導序列中之一個的上游，因而產生組合性載體。

【請求項21】 如請求項 20 之方法，其中在(a)中，該載體進一步包含位於該複數個 CRISPR 引導序列中至少一個的上游的啟動子序列。

【請求項22】 如請求項 20 或 21 之方法，其中該方法係反覆進行的。

圖式

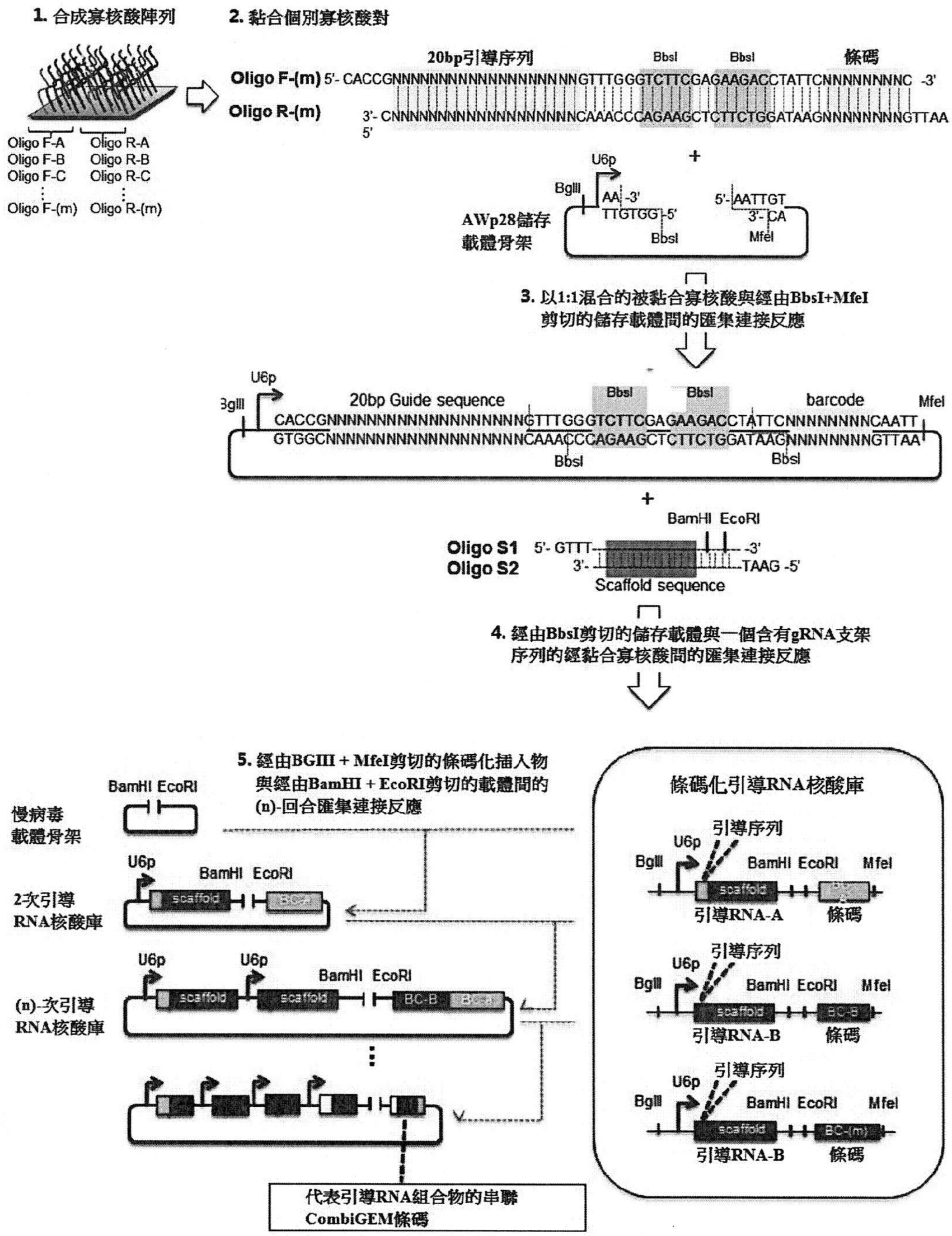

1. 合成寡核酸陣列
2. 黏合個別寡核酸對
20bp引導序列
BbsI
BbsI
條碼
Oligo F-(m) 5'- CACCGNNNNNNNNNNNNNNNNNNNGTTTGGGTCTTCGAGAAGACCTATTCNNNNNNNNC -3'
Oligo R-(m)
3'- CNNNNNNNNNNNNNNNNNNCAAACCCAGAAGCTCTTCTGGATAAGNNNNNNNNGTTAA
5'
Oligo F-A
Oligo F-B
Oligo F-C
Oligo F-(m)
Oligo R-A
Oligo R-B
Oligo R-C
Oligo R-(m)
U6p
BglII
AA -3'
TTGTGG -5'
BbsI
5'- AATTGT
3'- CA
MfeI
AWp28儲存
載體骨架
3. 以1:1混合的被黏合寡核酸與經由BbsI+MfeI
剪切的儲存載體間的匯集連接反應
U6p
BglII
20bp Guide sequence
BbsI
BbsI
barcode
MfeI
CACCGNNNNNNNNNNNNNNNNNNNGTTTGGGTCTTCGAGAAGACCTATTCNNNNNNNNCAATT
GTGGCNNNNNNNNNNNNNNNNNNNCAAACCCAGAAGCTCTTCTGGATAAGNNNNNNNNGTTAA
BbsI
BbsI
BamHI
EcoRI
Oligo S1
Oligo S2
5'- GTTT
-3'
3'-
TAAG -5'
Scaffold sequence
4. 經由BbsI剪切的儲存載體與一個含有gRNA支架
序列的經黏合寡核酸間的匯集連接反應
5. 經由BGIII + MfeI剪切的條碼化插入物
與經由BamHI + EcoRI剪切的載體間的
(n)-回合匯集連接反應
BamHI EcoRI
慢病毒
載體骨架
U6p
BamHI EcoRI
scaffold
2次引導
RNA核酸庫
U6p
U6p
BamHI EcoRI
scaffold
scaffold
BC-B
(n)-次引導
RNA核酸庫
代表引導RNA組合物的串聯
CombiGEM條碼
條碼化引導RNA核酸庫
引導序列
U6p
BglII
BamHI EcoRI
MfeI
scaffold
引導RNA-A
條碼
引導序列
U6p
BglII
BamHI EcoRI
MfeI
scaffold
BC-B
引導RNA-B
條碼
引導序列
U6p
BglII
BamHI EcoRI
MfeI
scaffold
BC-(m)
引導RNA-B
條碼

圖 1

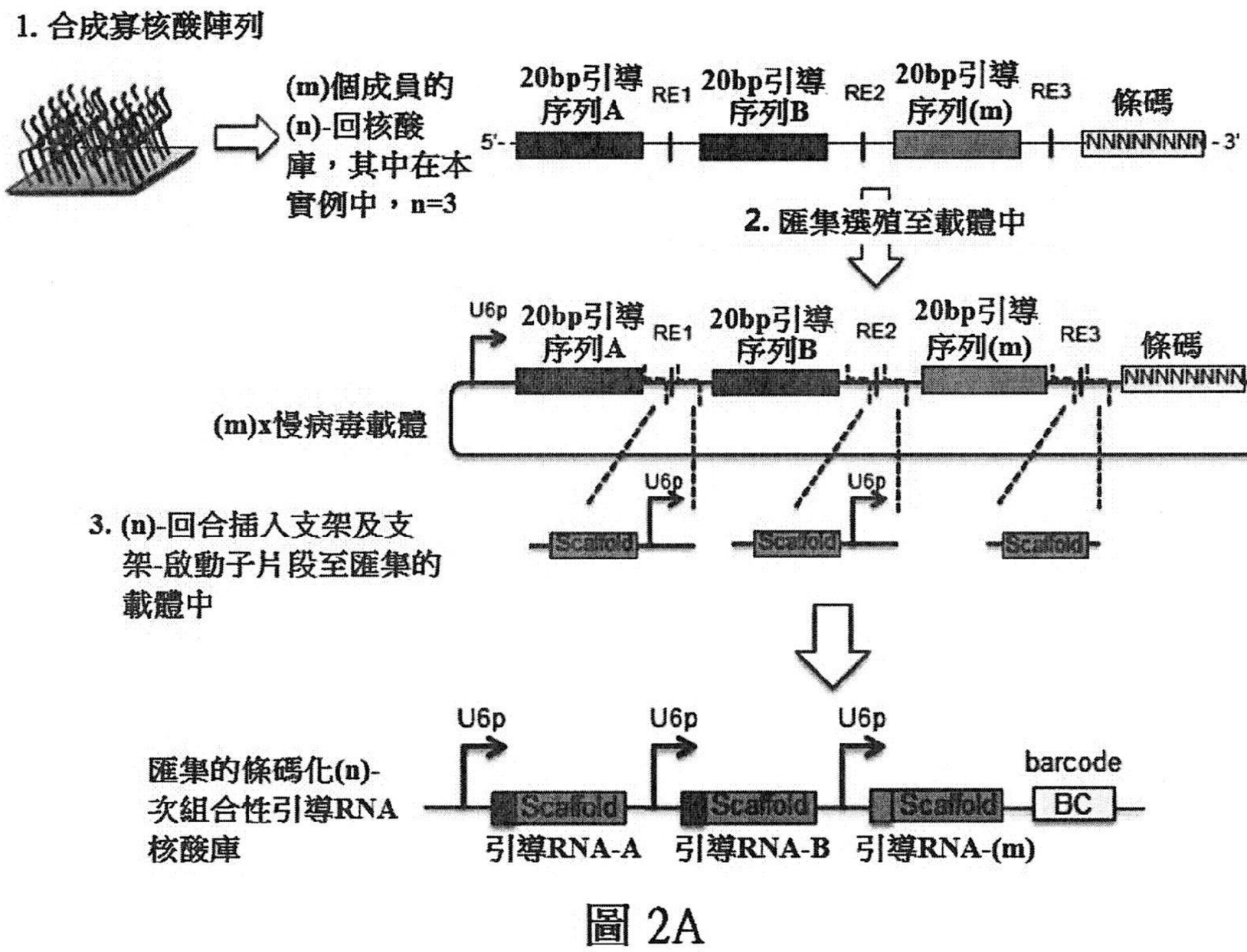

1. 合成寡核酸陣列
(m)個成員的(n)-回核酸庫，其中在本實例中，n=3
20bp引導序列A
RE1
20bp引導序列B
RE2
20bp引導序列(m)
RE3
條碼
5'-
NNNNNNNN
-3'
2. 匯集選殖至載體中
U6p
(m)x慢病毒載體
Scaffold
3. (n)-回合插入支架及支架-啟動子片段至匯集的載體中
匯集的條碼化(n)-次組合性引導RNA核酸庫
barcode
BC
引導RNA-A
引導RNA-B
引導RNA-(m)

圖 2A

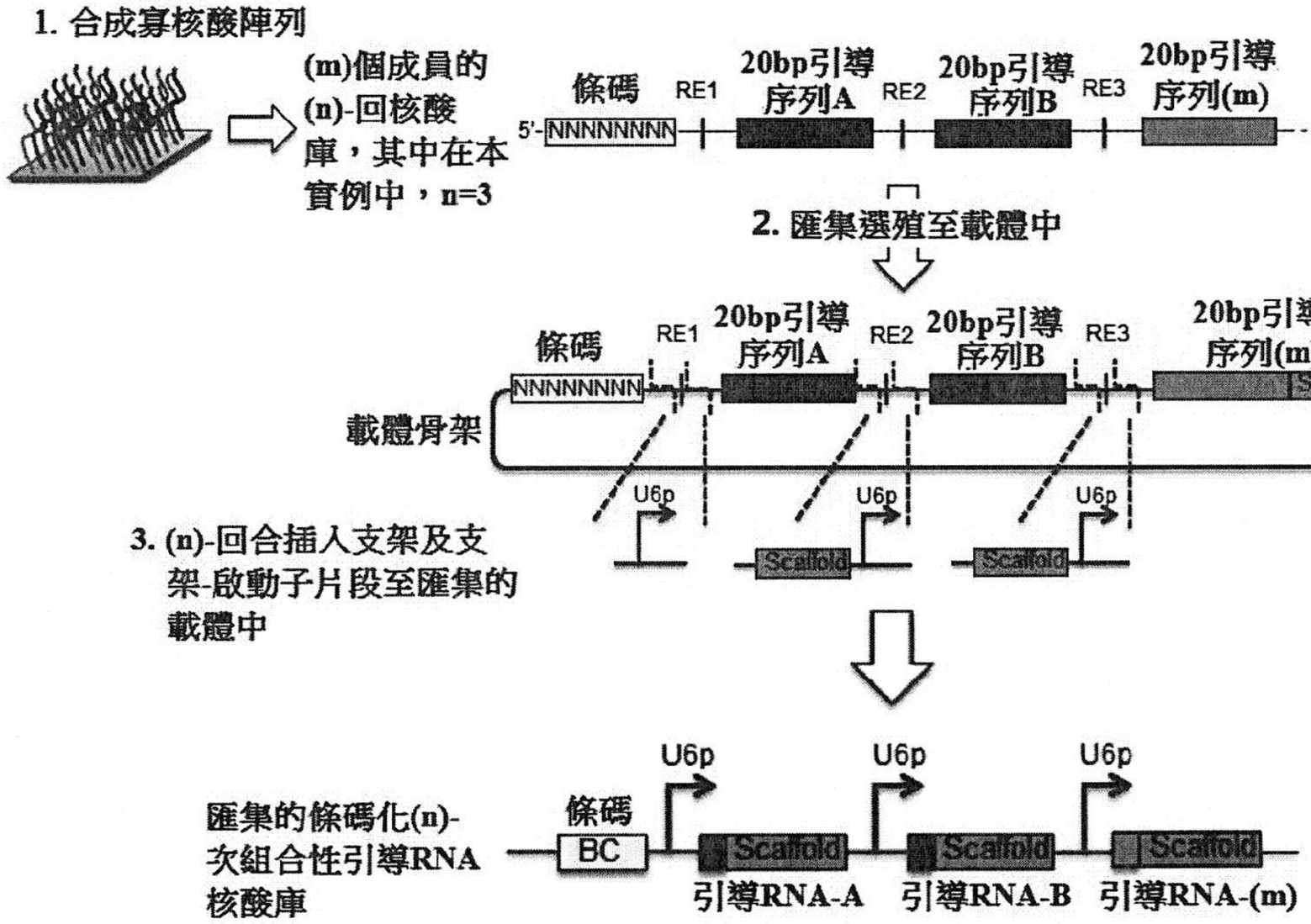

1. 合成寡核酸陣列
(m)個成員的(n)-回核酸庫，其中在本實例中，n=3
條碼
RE1
20bp引導序列A
RE2
20bp引導序列B
RE3
20bp引導序列(m)
5'-
NNNNNNNN
-3'
2. 匯集選殖至載體中
載體骨架
Scaffold
U6p
3. (n)-回合插入支架及支架-啟動子片段至匯集的載體中
匯集的條碼化(n)-次組合性引導RNA核酸庫
BC
引導RNA-A
引導RNA-B
引導RNA-(m)

圖 2B

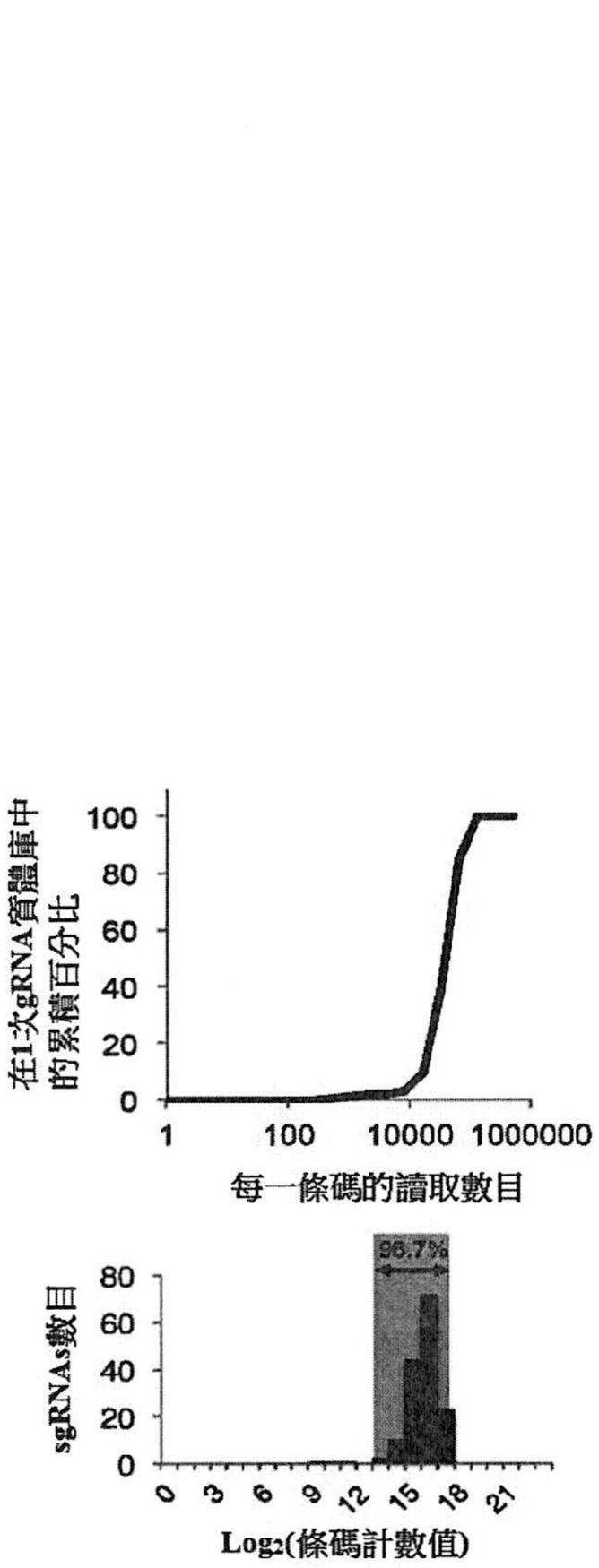


圖 3A

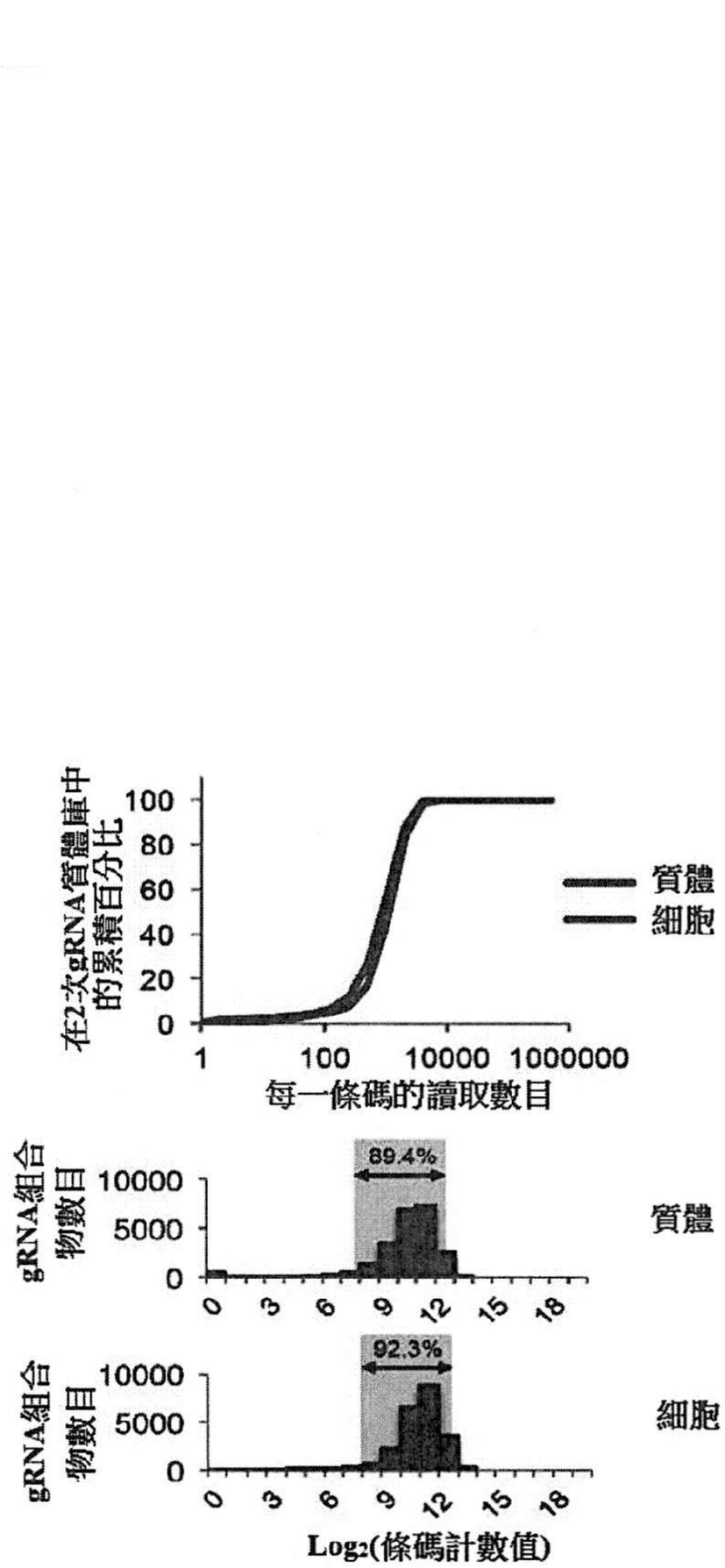


圖 3B

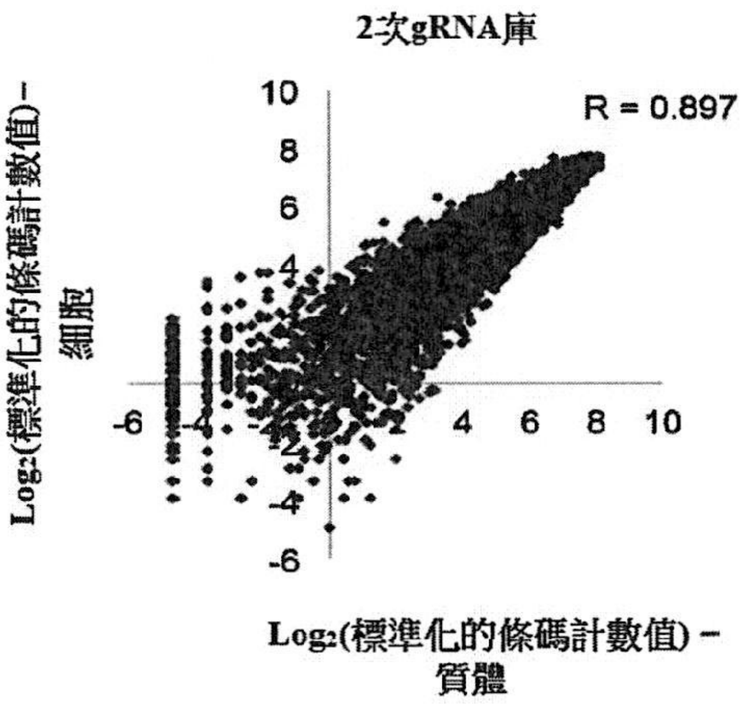


圖 3C

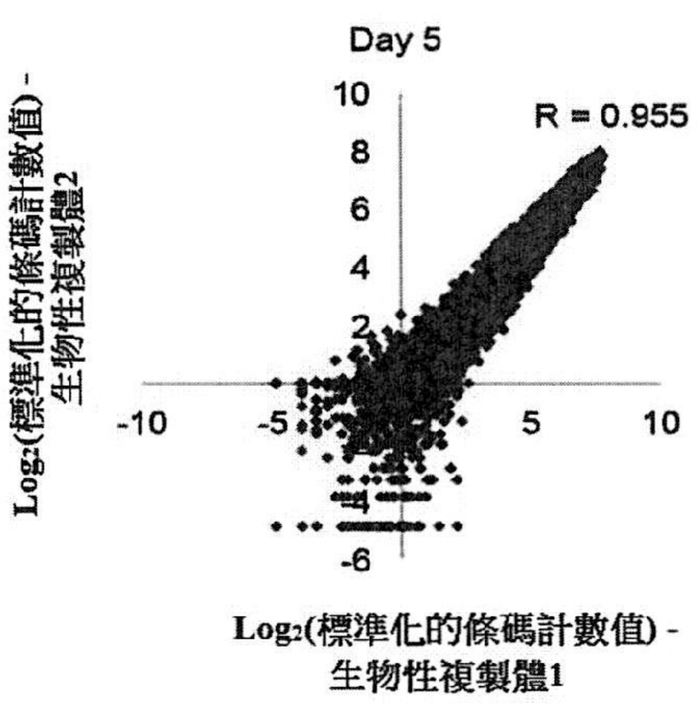


圖 3D

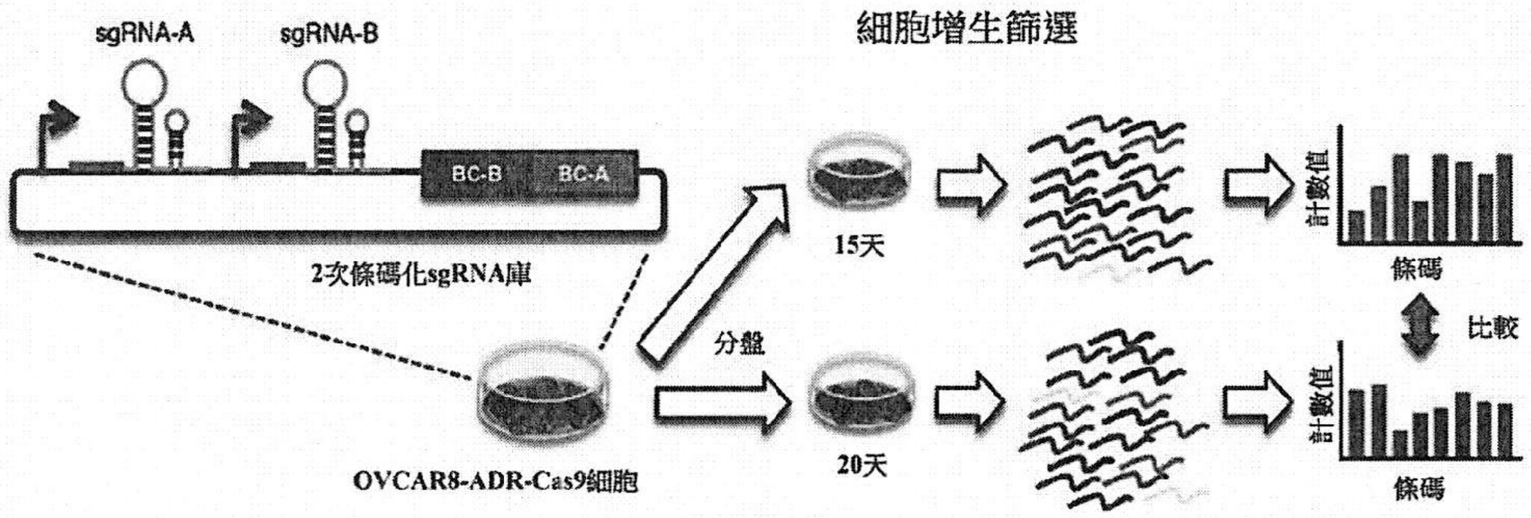
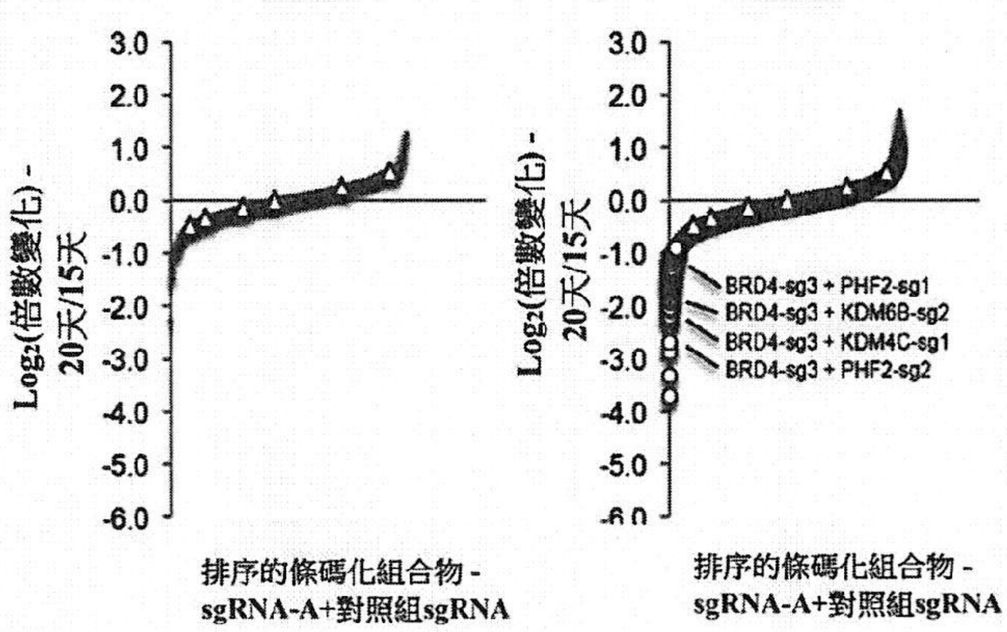

sgRNA-A
sgRNA-B
BC-B
BC-A
2次條碼化sgRNA庫
OVCAR8-ADR-Cas9細胞
分盤
細胞增生篩選
15天
20天
計數值
條碼
比較

圖 4A


複製體1
Log2(倍數變化) - 20天/15天
排序的條碼化組合物 - sgRNA-A+對照組sgRNA

複製體1
Log2(倍數變化) - 20天/15天
BRD4-sg3 + PHF2-sg1
BRD4-sg3 + KDM6B-sg2
BRD4-sg3 + KDM4C-sg1
BRD4-sg3 + PHF2-sg2
排序的條碼化組合物 - sgRNA-A+對照組sgRNA

圖 4B

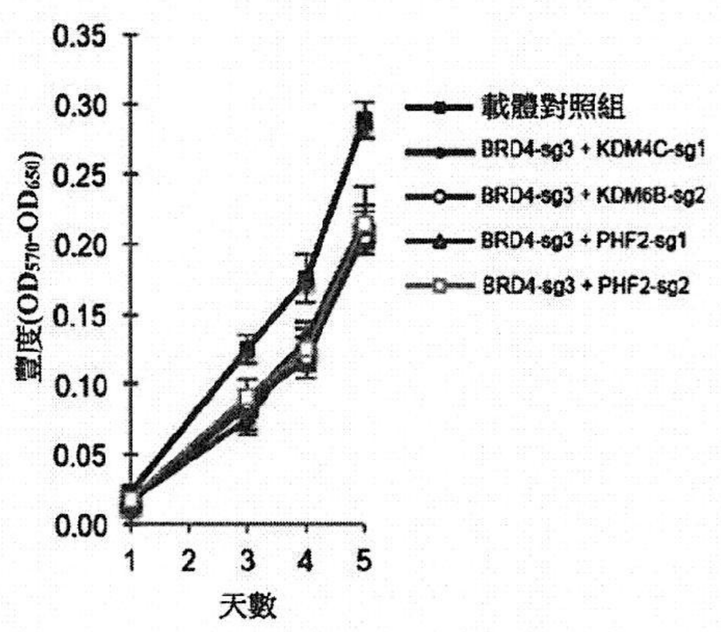

豐度(OD570-OD650)
天數
載體對照組
BRD4-sg3 + KDM4C-sg1
BRD4-sg3 + KDM6B-sg2
BRD4-sg3 + PHF2-sg1
BRD4-sg3 + PHF2-sg2

圖 4C

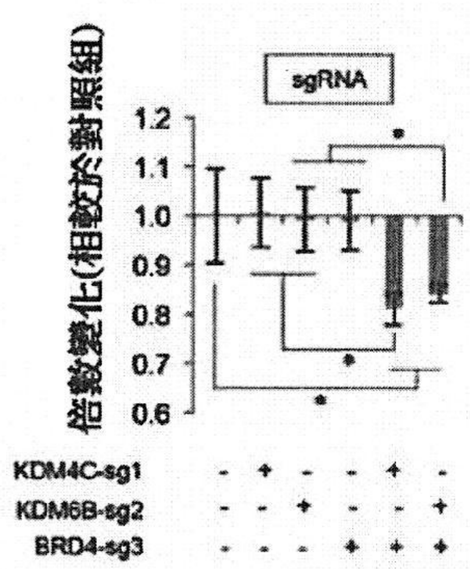


圖 5A

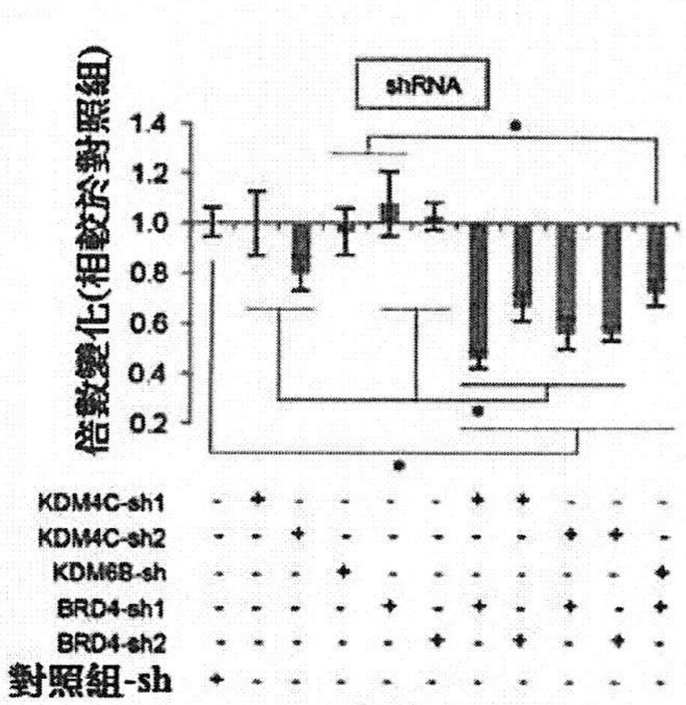


圖 5B

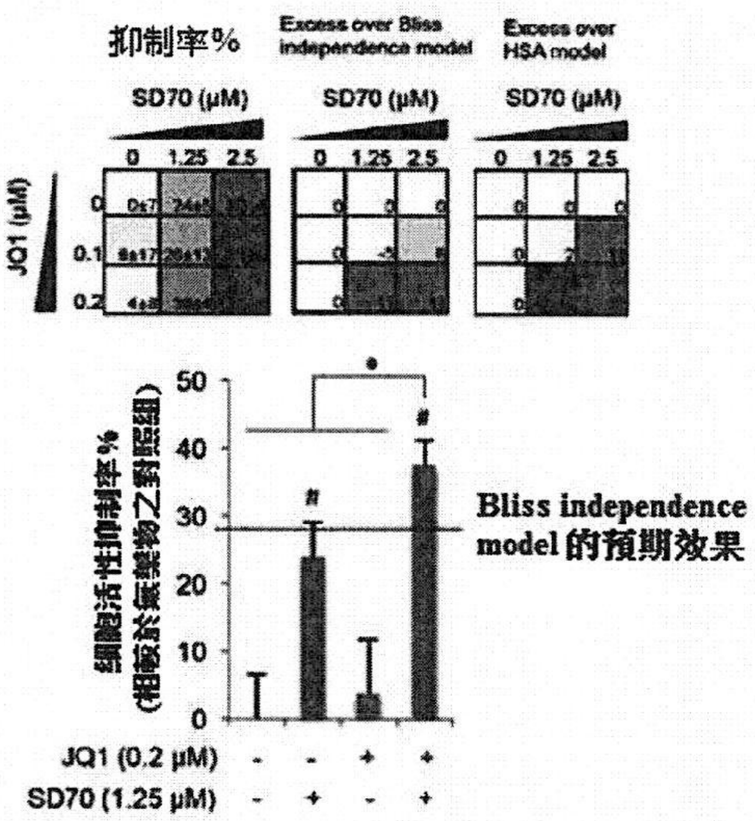


圖 5C

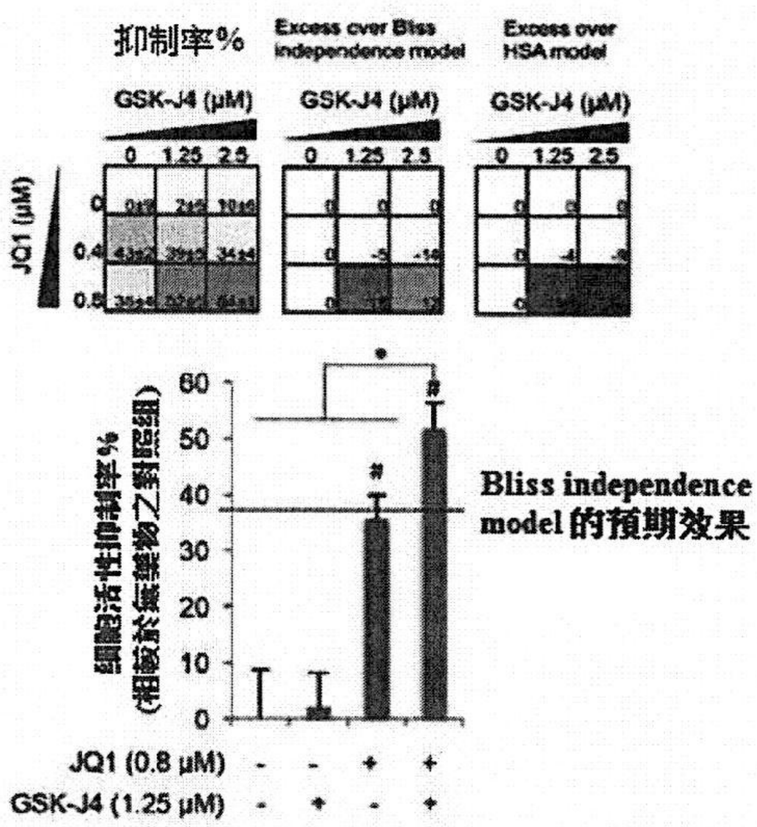


圖 5D

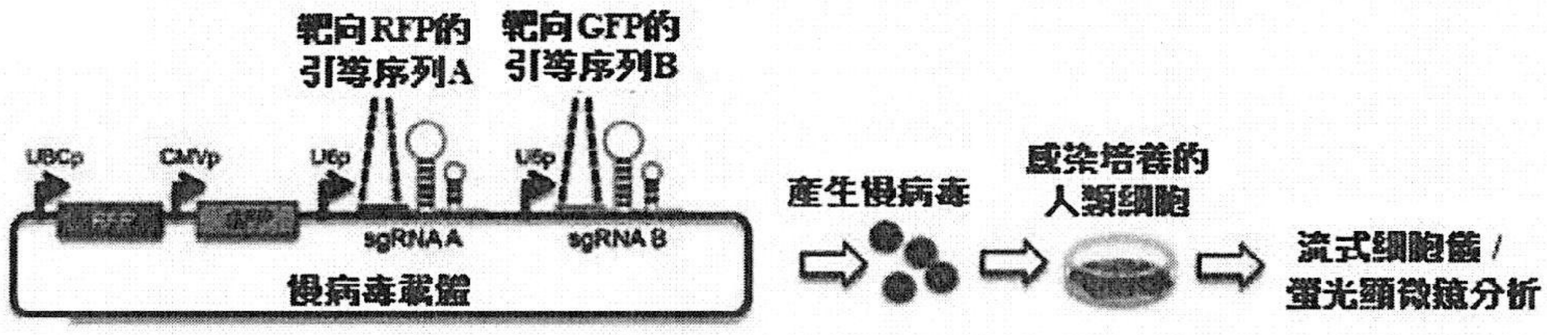


圖 6A

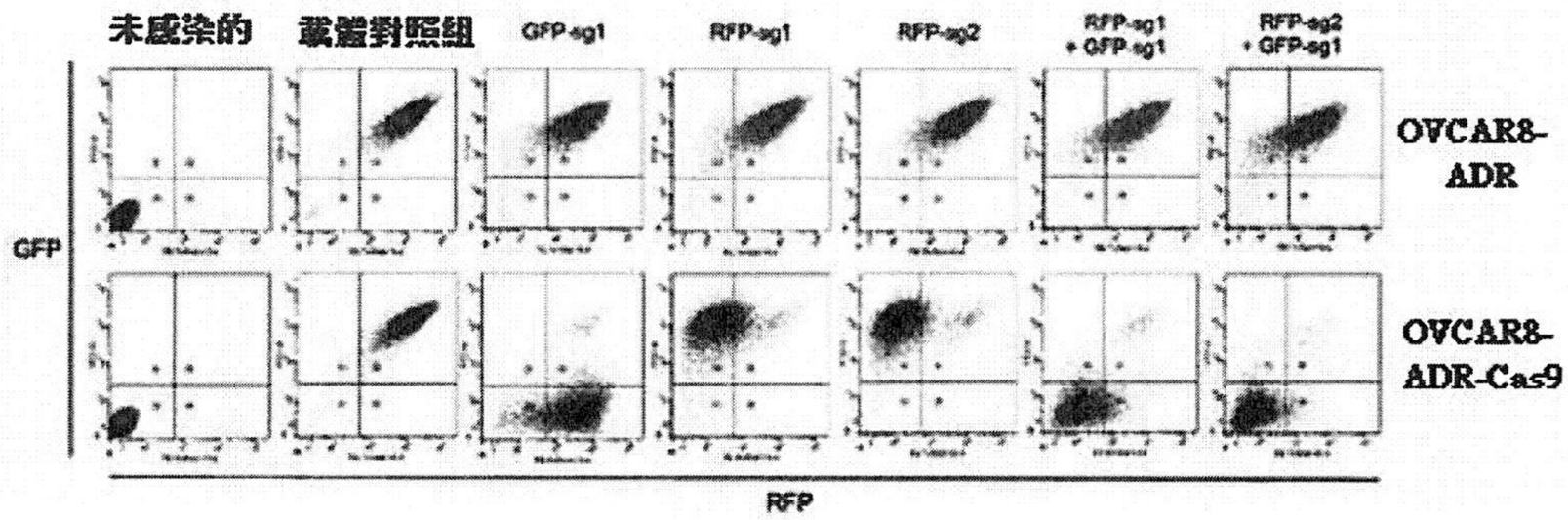


圖 6B

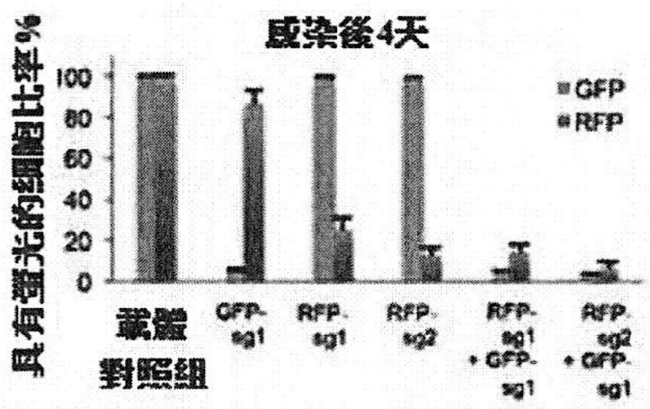


圖 6C

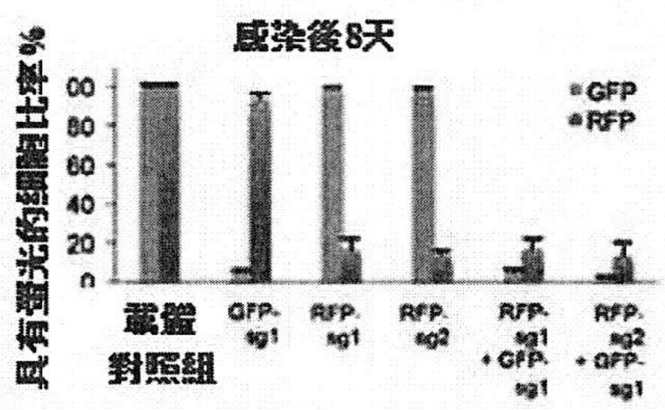


圖 6D

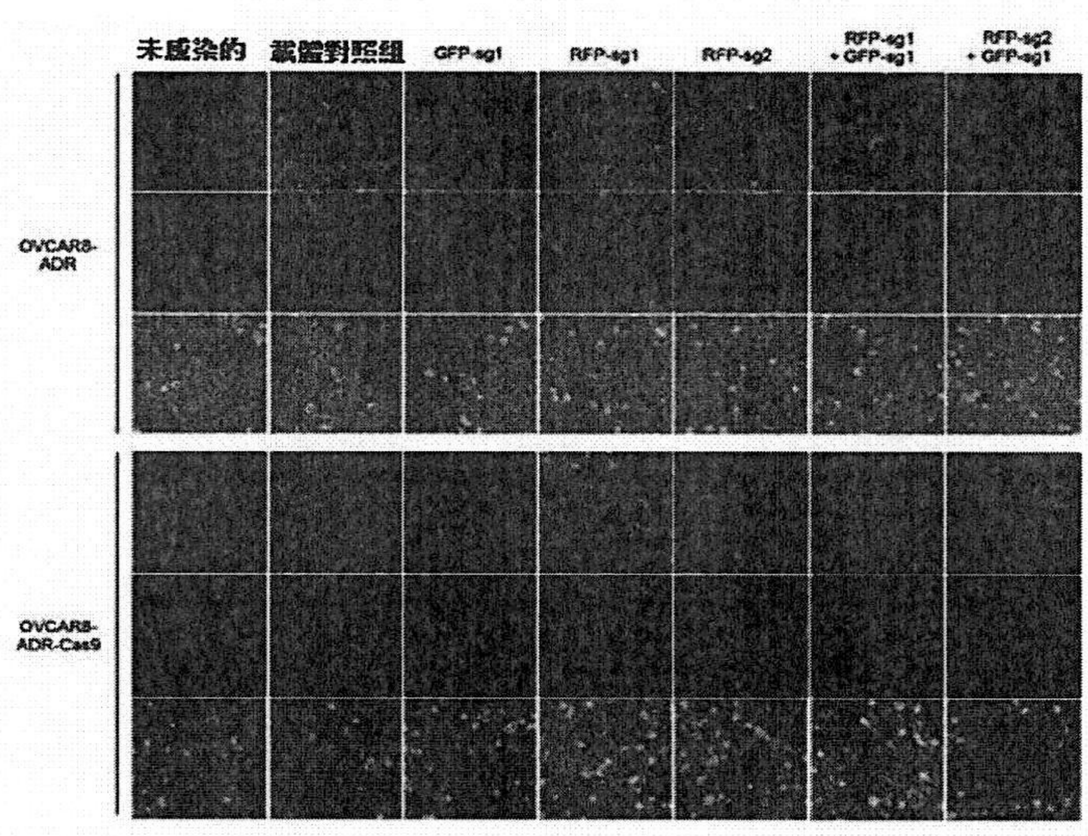

未感染的
載體對照組
GFP-sg1
RFP-sg1
RFP-sg2
RFP-sg1 + GFP-sg1
RFP-sg2 + GFP-sg1
OVCAR8-ADR
OVCAR8-ADR-Cas9

圖 6E

sgRNA	未被剪切的條帶	被剪切的條帶
DNMT1-sg1	600	450/150
DNMT3B-sg1	544	400/150
PRMT2-sg3	559	370/170
HDAC2-sg1	871	500/300
ING4-sg1	502	300/200
KDM1B-sg3	499	350/150
KDM2A-sg3	455	300/150
PRMT6-sg1	849	550/300
KDM4C-sg1	419	284/155
KDM6B-sg2	489	273/216
BRD4-sg3	541	341/200

圖 7A

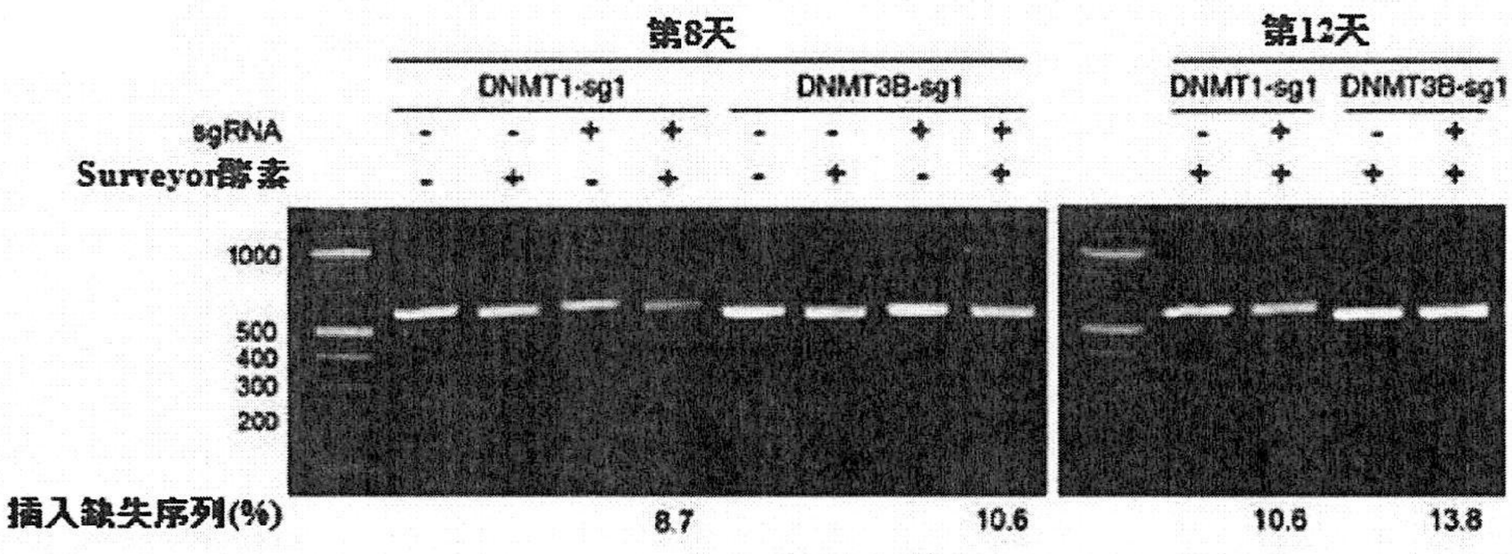


圖 7B

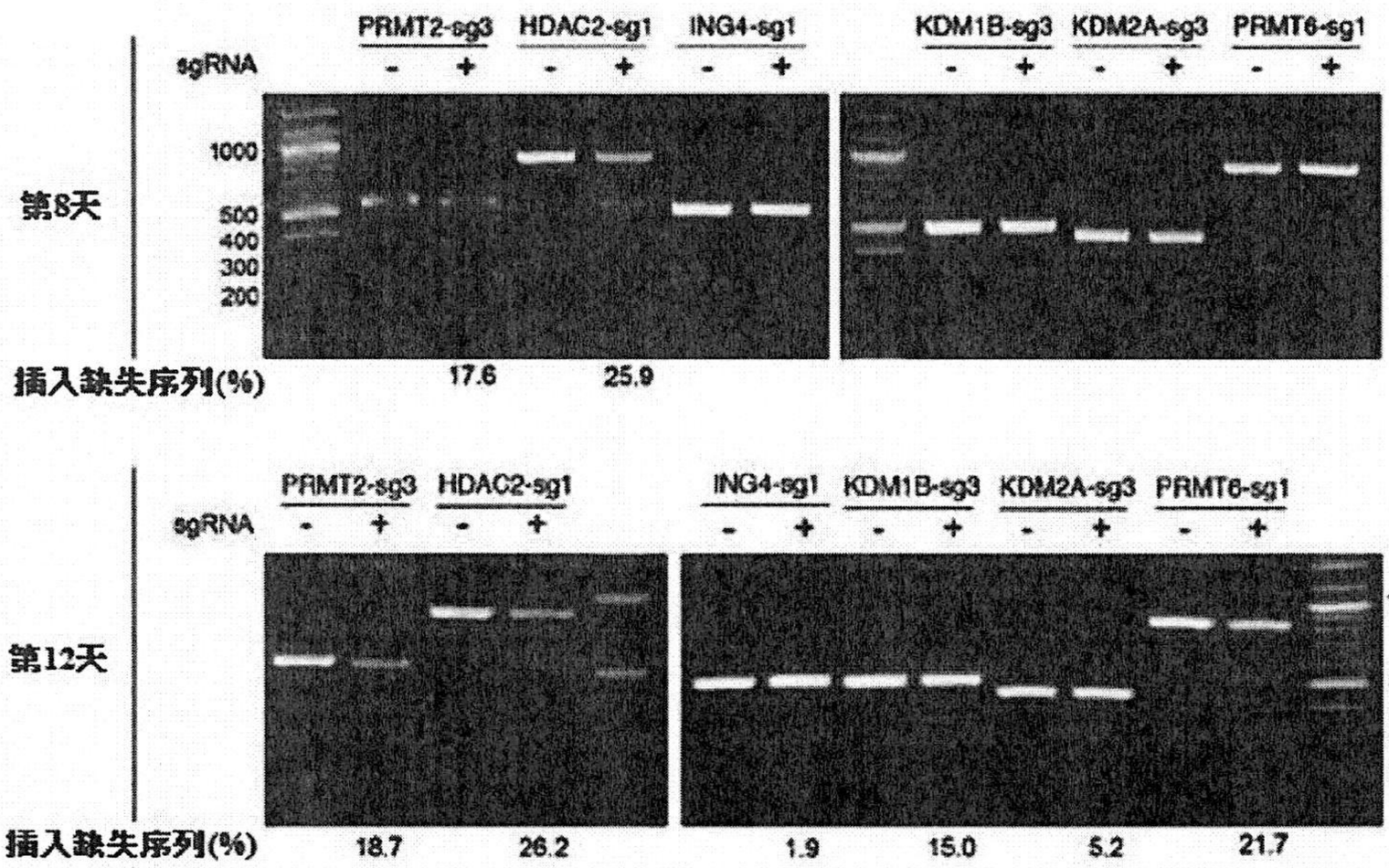


圖 7C

sgRNA	未被剪切的條帶	被剪切的條帶
KDM4C-sg1	419	264/155
KDM6B-sg2	489	273/216
BRD4-sg3	541	341/200

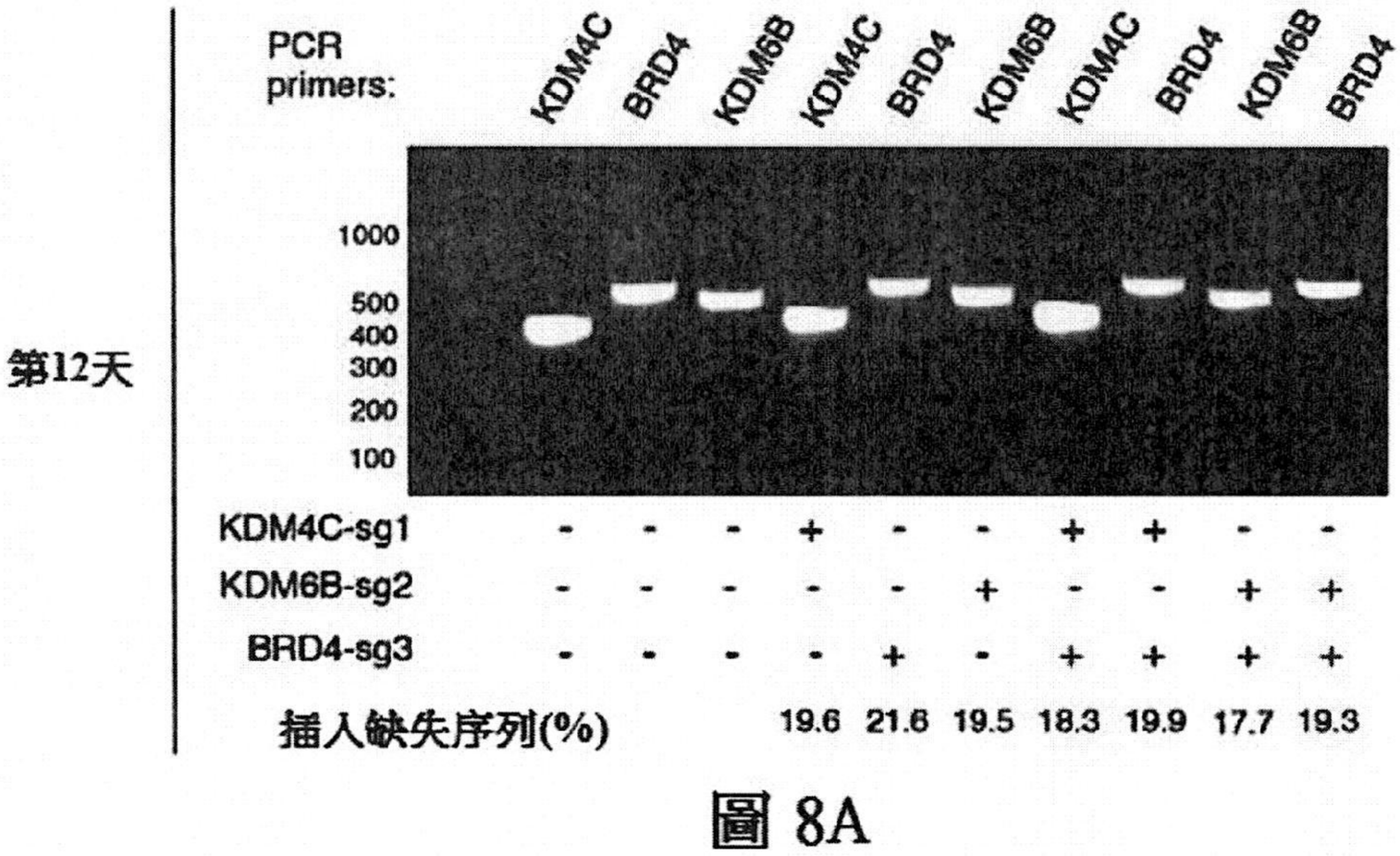


圖 8A

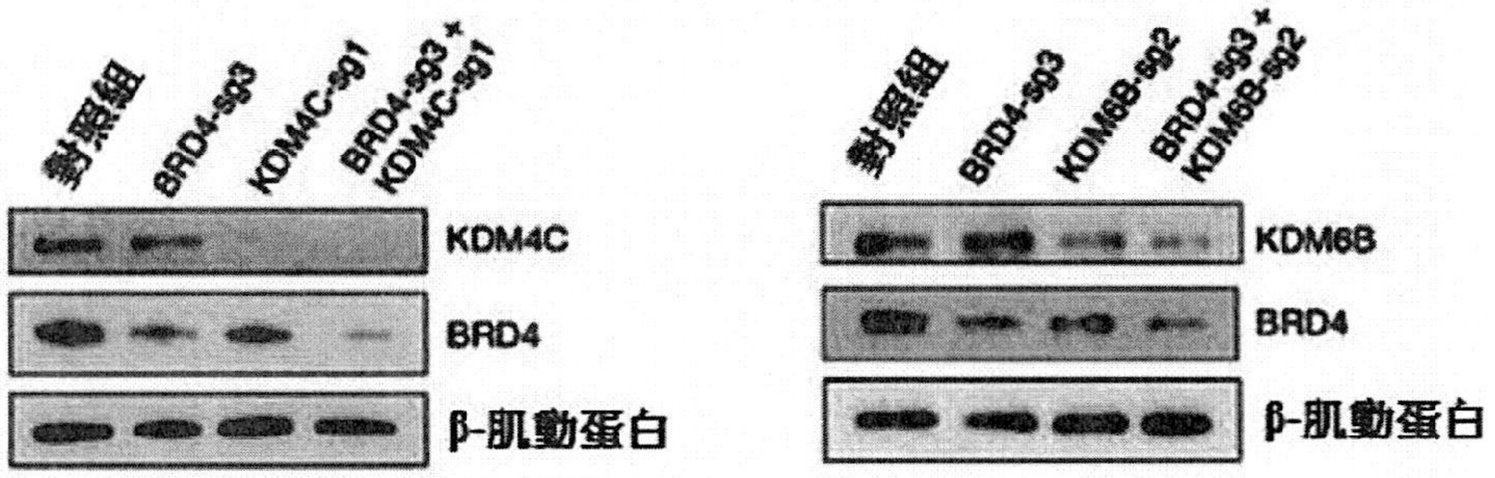


圖 8B

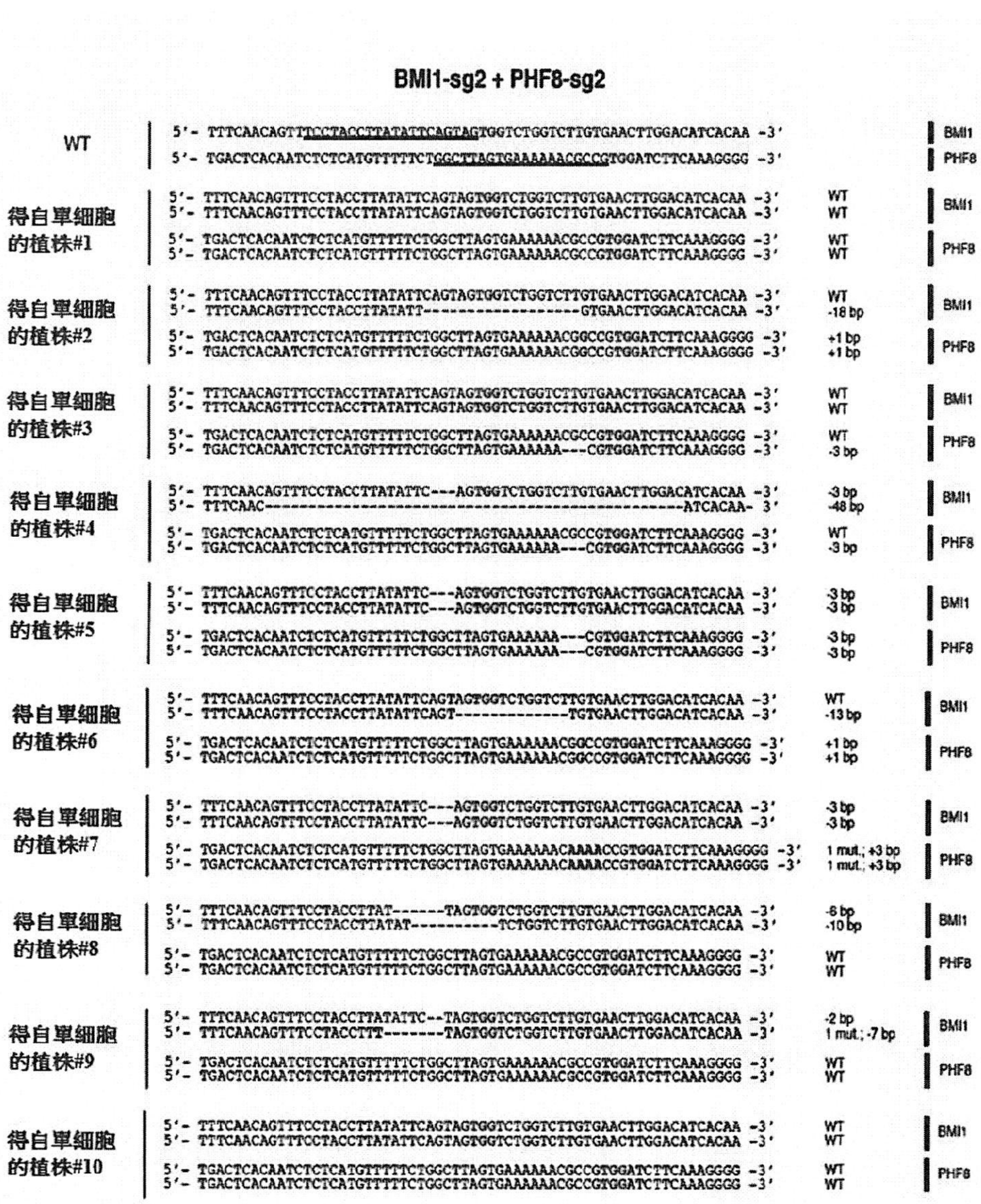

BMI1-sg2 + PHF8-sg2
WT
5'- TTTCAACAGTTTCCTACCTTATATTCAGTAGTGGTCTGGTCTTGTGAACTTGGACATCACAA -3'
5'- TGACTCACAATCTCTCATGTTTTTCTGGCTTAGTGAAAAAACGCCGTGGATCTTCAAAGGGG -3'
BMI1
PHF8
得自單細胞的植株#1
得自單細胞的植株#2
得自單細胞的植株#3
得自單細胞的植株#4
得自單細胞的植株#5
得自單細胞的植株#6
得自單細胞的植株#7
得自單細胞的植株#8
得自單細胞的植株#9
得自單細胞的植株#10
WT
-18 bp
+1 bp
-3 bp
-48 bp
-13 bp
1 mut.; +3 bp
-6 bp
-10 bp
-2 bp
1 mut.; -7 bp

圖 9A

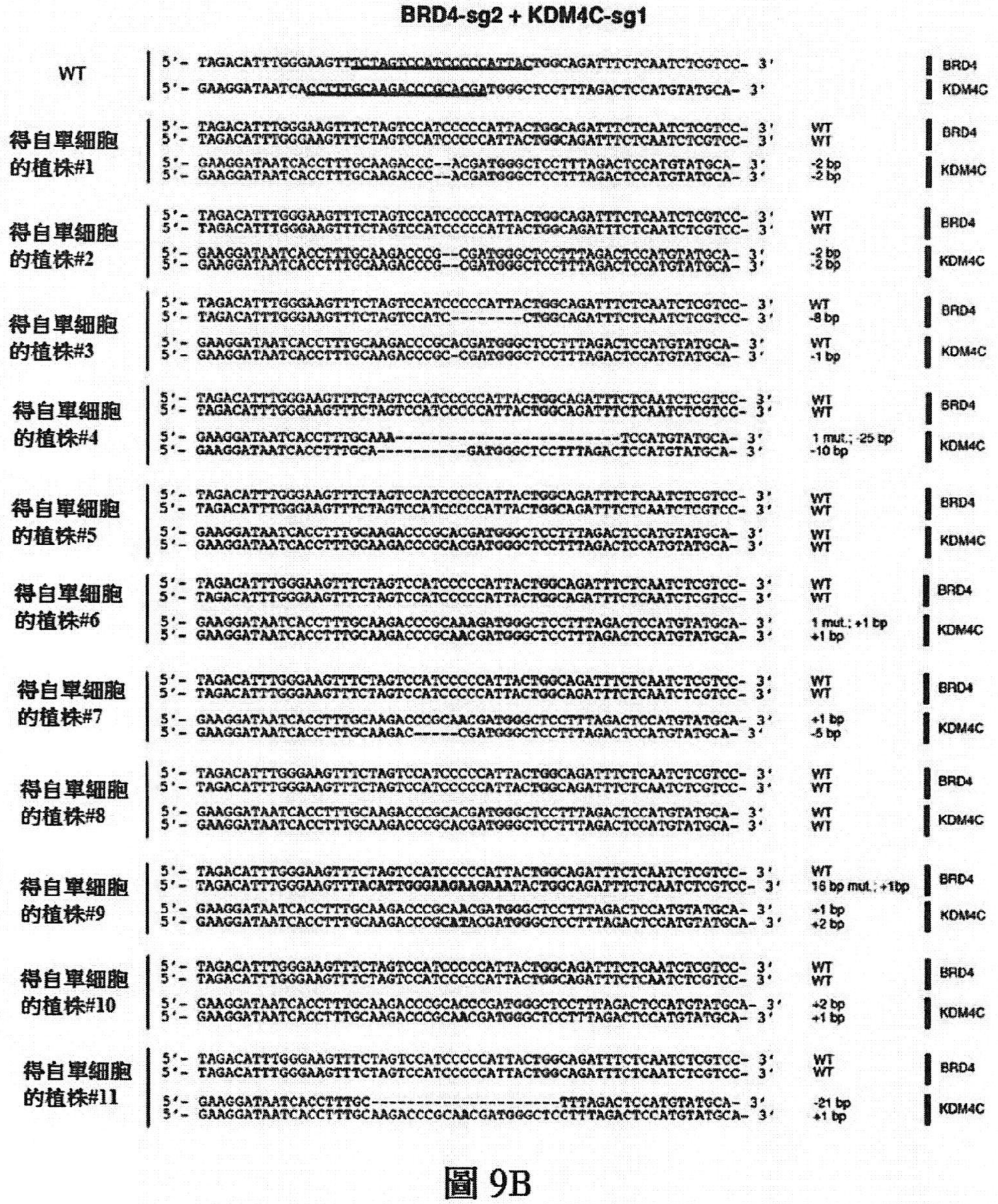

BRD4-sg2 + KDM4C-sg1
WT
5'- TAGACATTTGGGAAGTTTCTAGTCCATCCCCCATTACTGGCAGATTTCTCAATCTCGTCC- 3'
5'- GAAGGATAATCACCTTTGCAAGACCCGCACGATGGGCTCCTTTAGACTCCATGTATGCA- 3'
BRD4
KDM4C
得自單細胞的植株#1
得自單細胞的植株#2
得自單細胞的植株#3
得自單細胞的植株#4
得自單細胞的植株#5
得自單細胞的植株#6
得自單細胞的植株#7
得自單細胞的植株#8
得自單細胞的植株#9
得自單細胞的植株#10
得自單細胞的植株#11
WT
-2 bp
-8 bp
-1 bp
1 mut.; -25 bp
-10 bp
1 mut.; +1 bp
+1 bp
-5 bp
16 bp mut.; +1bp
+2 bp
-21 bp

圖 9B

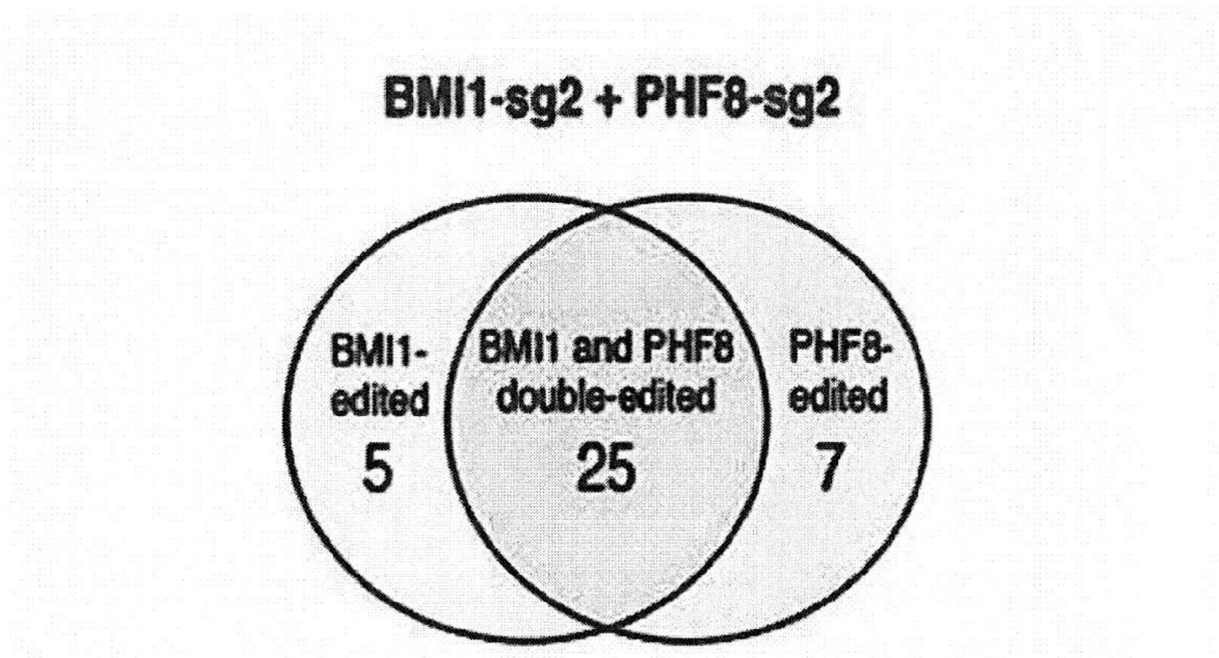

BMI1-sg2 + PHF8-sg2
BMI1-edited
5
BMI1 and PHF8 double-edited
25
PHF8-edited
7

圖 9C

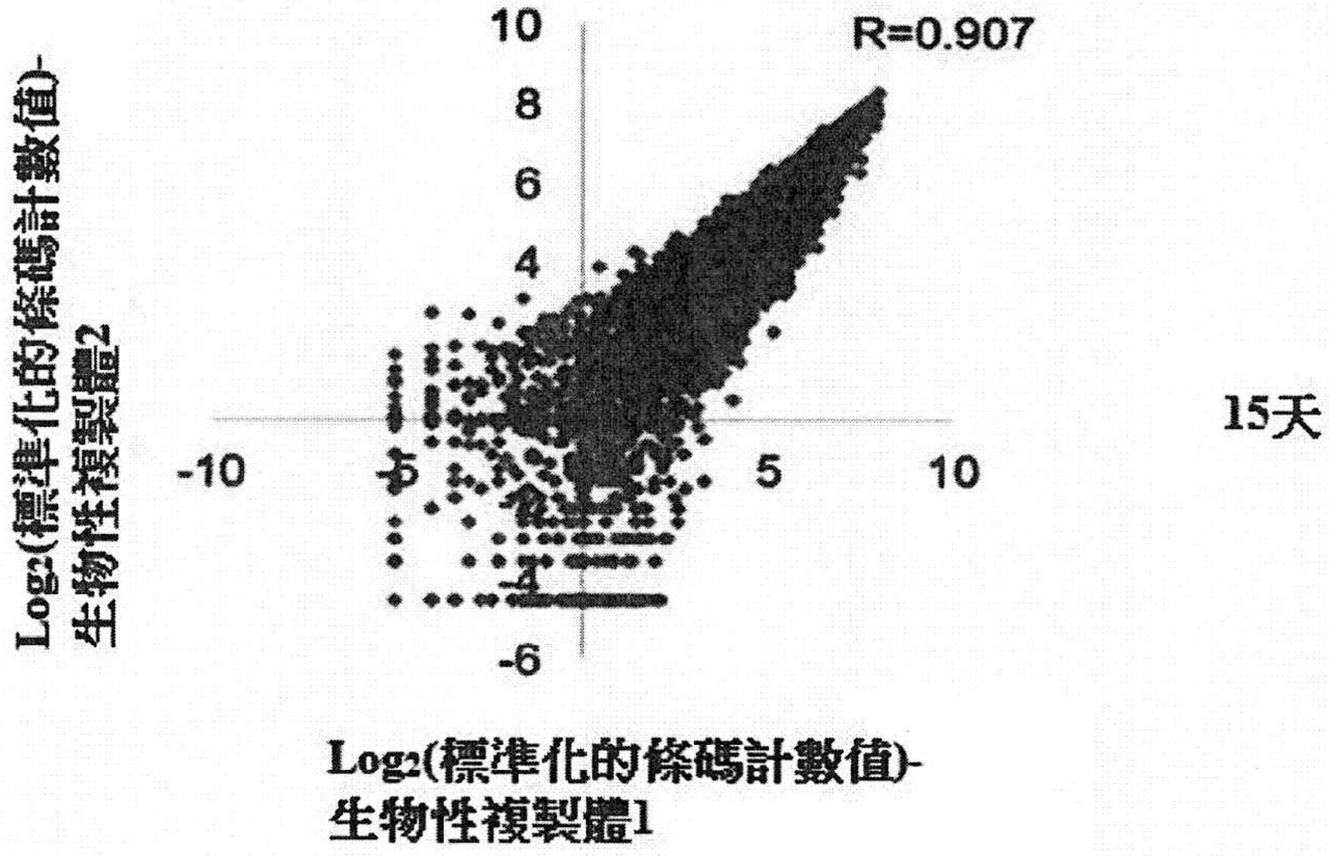

R=0.907
10
8
6
4
-10
5
10
-6
Log2(標準化的條碼計數值)-生物性複製體2
Log2(標準化的條碼計數值)-生物性複製體1
15天

圖 10A

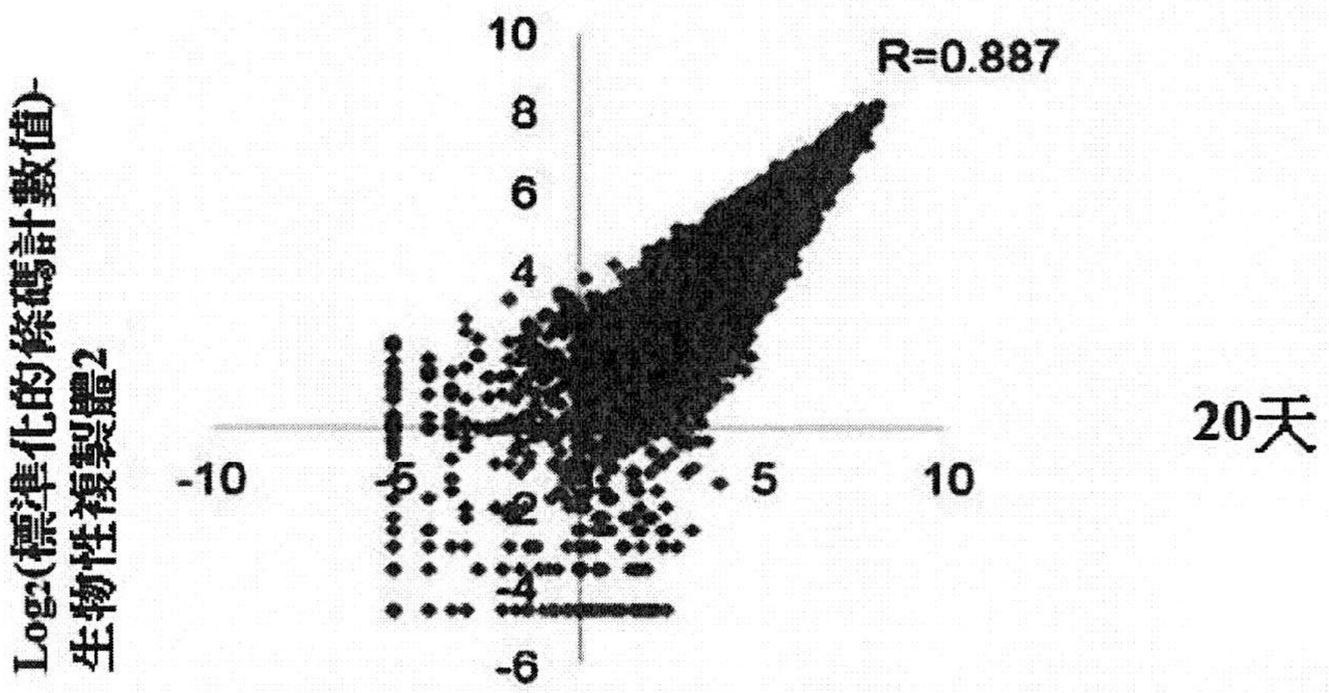

R=0.887
10
8
6
4
-10
5
10
-6
Log2(標準化的條碼計數值)-生物性複製體2
Log2(標準化的條碼計數值)-生物性複製體1
20天

圖 10B

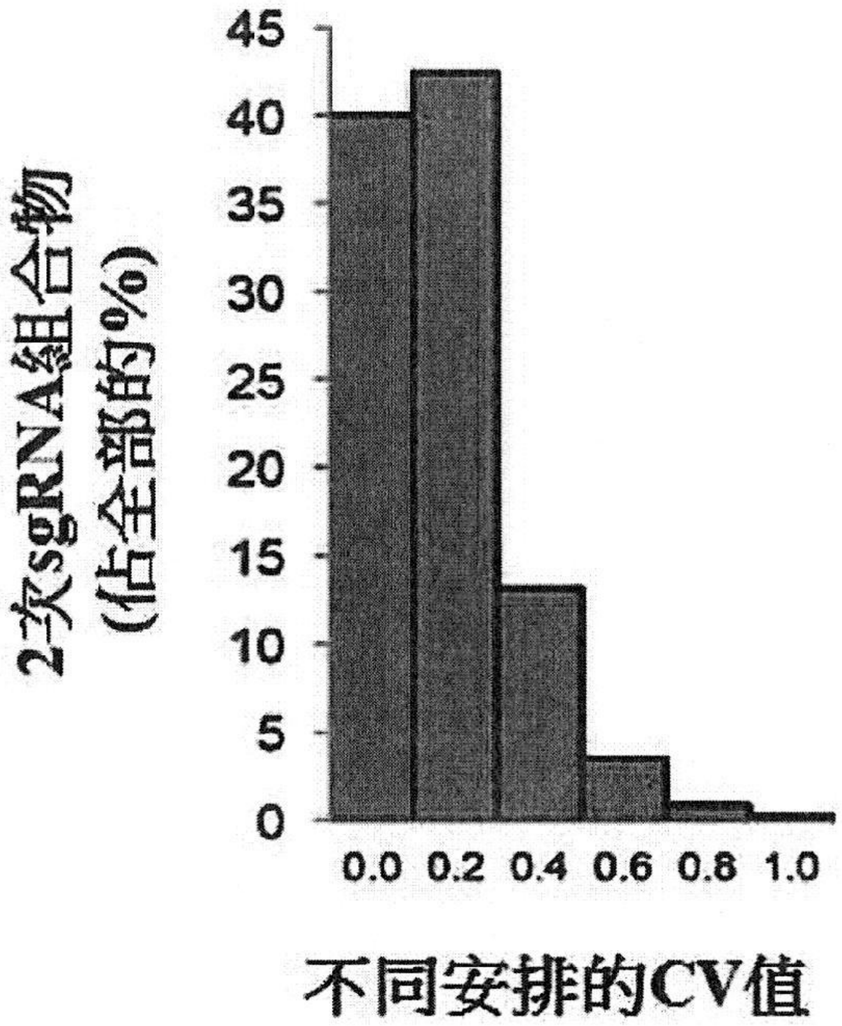

2次sgRNA組合物
(佔全部的%)
45
40
35
30
25
20
15
10
5
0
0.0 0.2 0.4 0.6 0.8 1.0
不同安排的CV值

圖 11

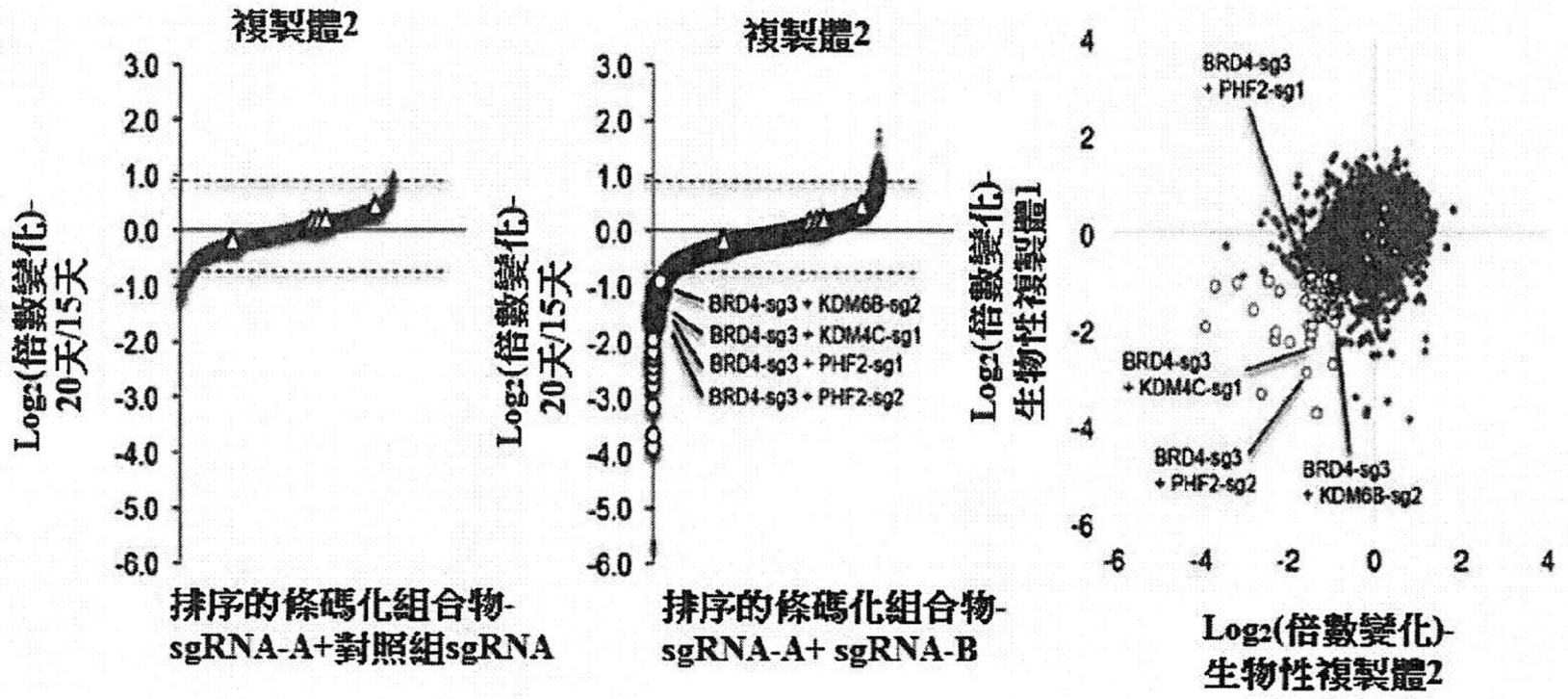


圖 12A

圖 12B

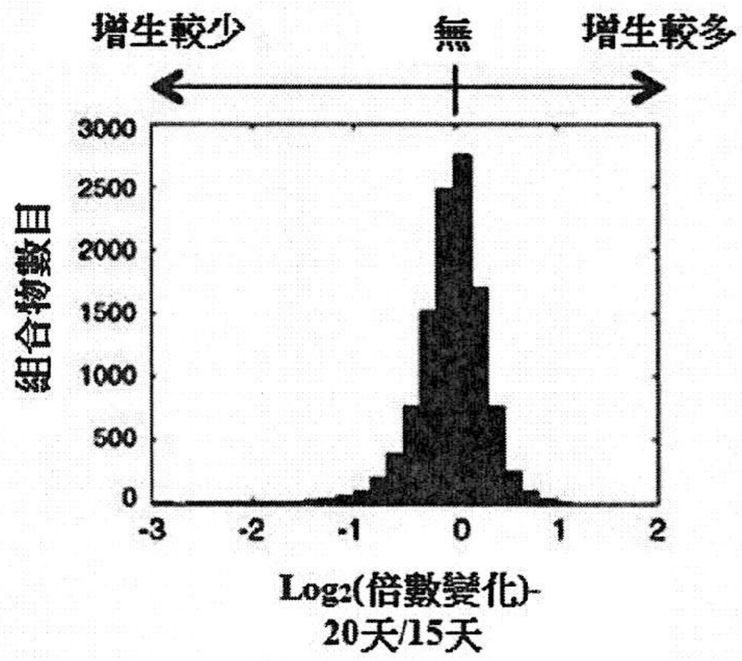


圖 12C

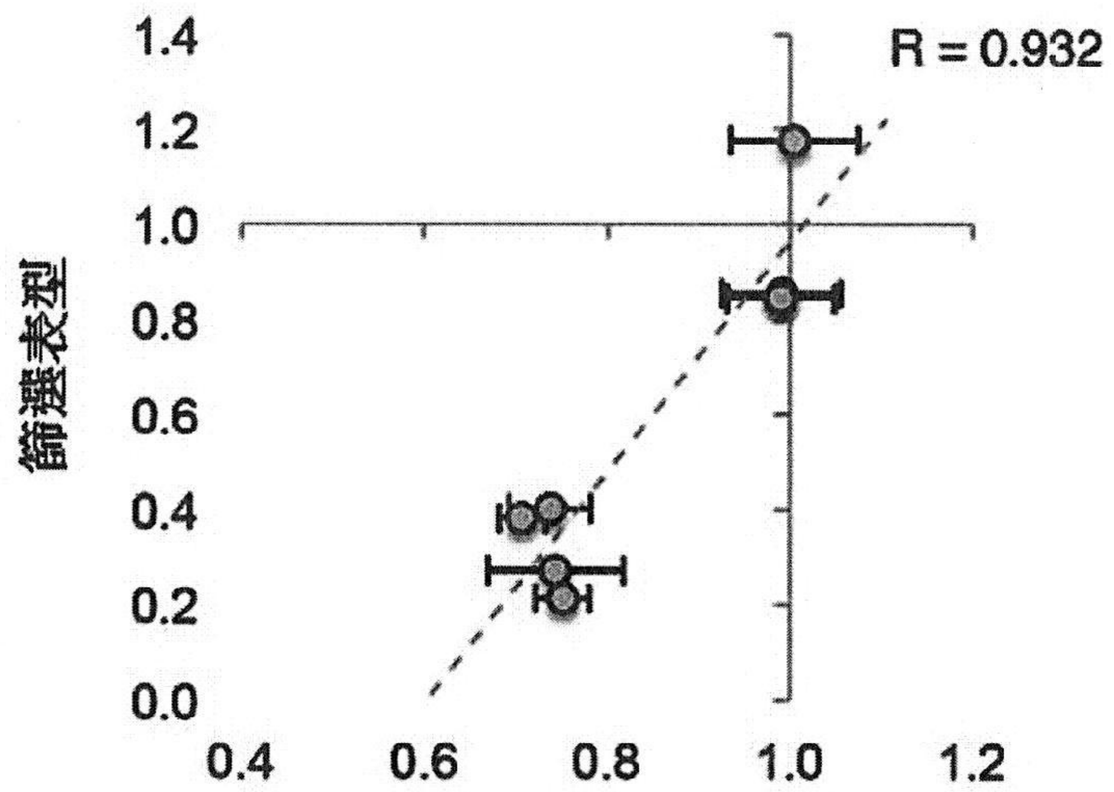


圖 13

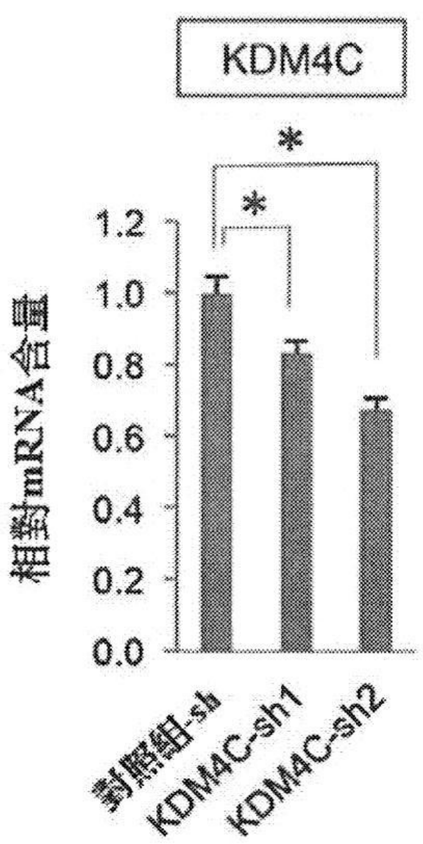


圖 14A

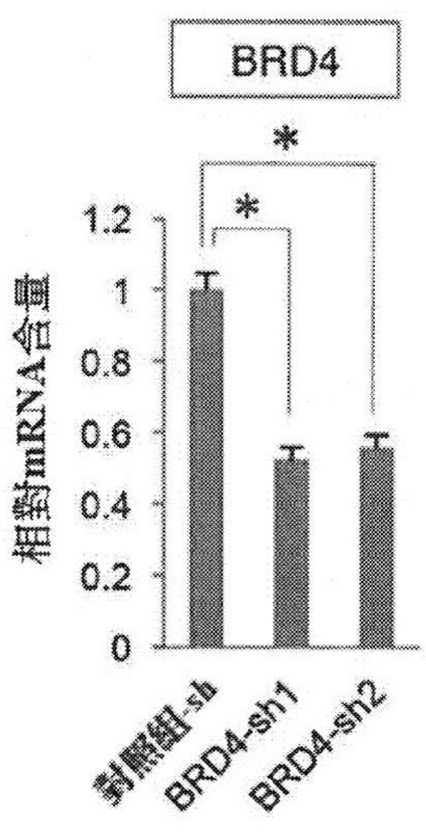


圖 14B

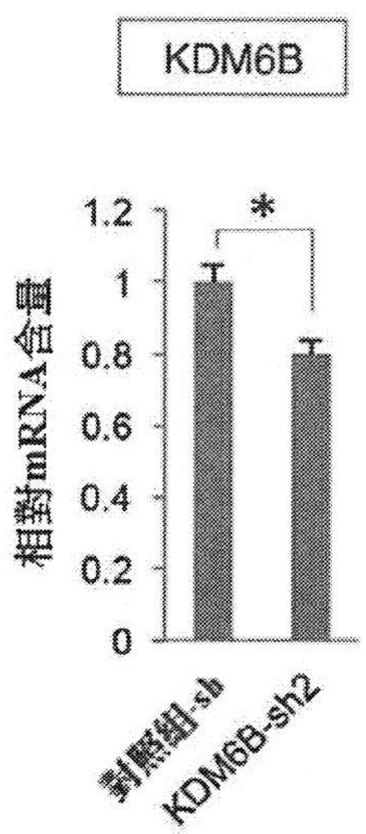
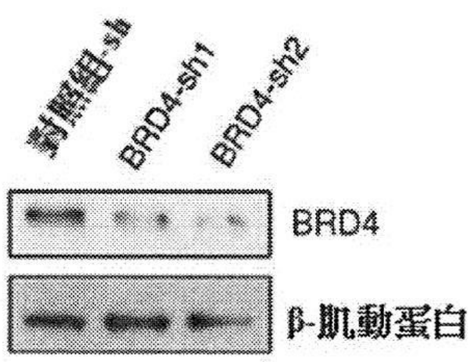
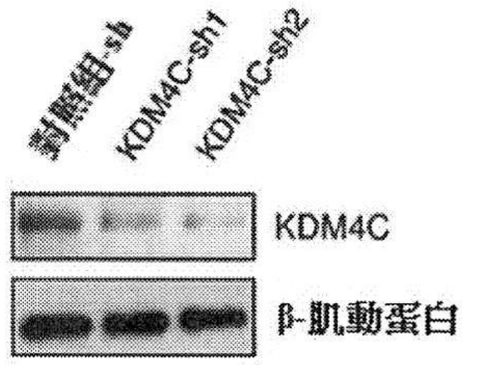
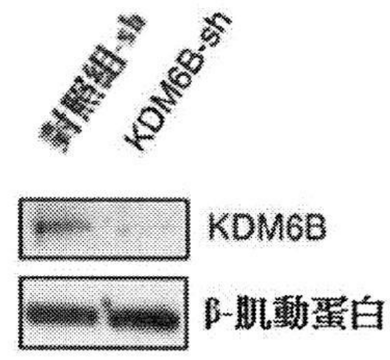


圖 14C







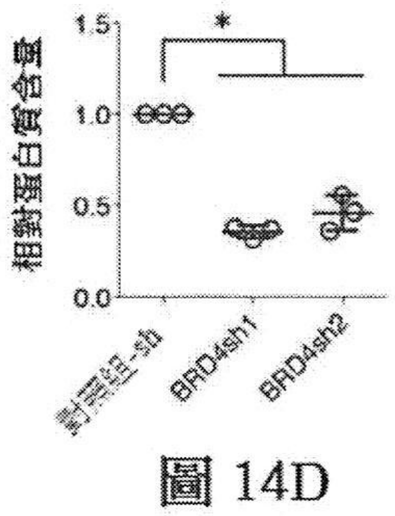


圖 14D

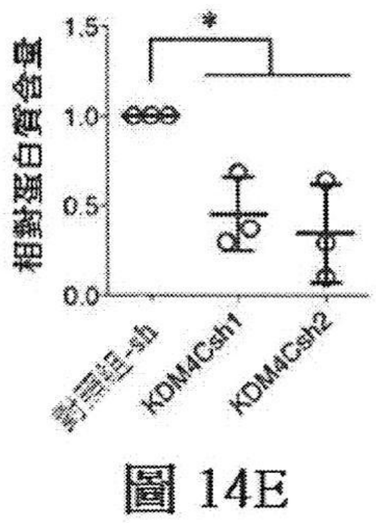


圖 14E

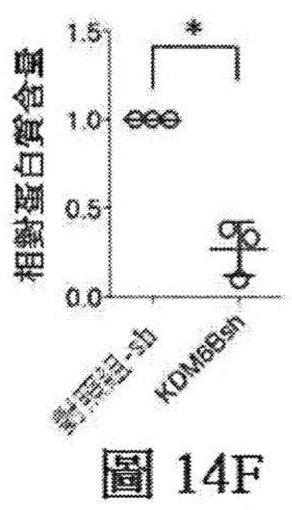


圖 14F

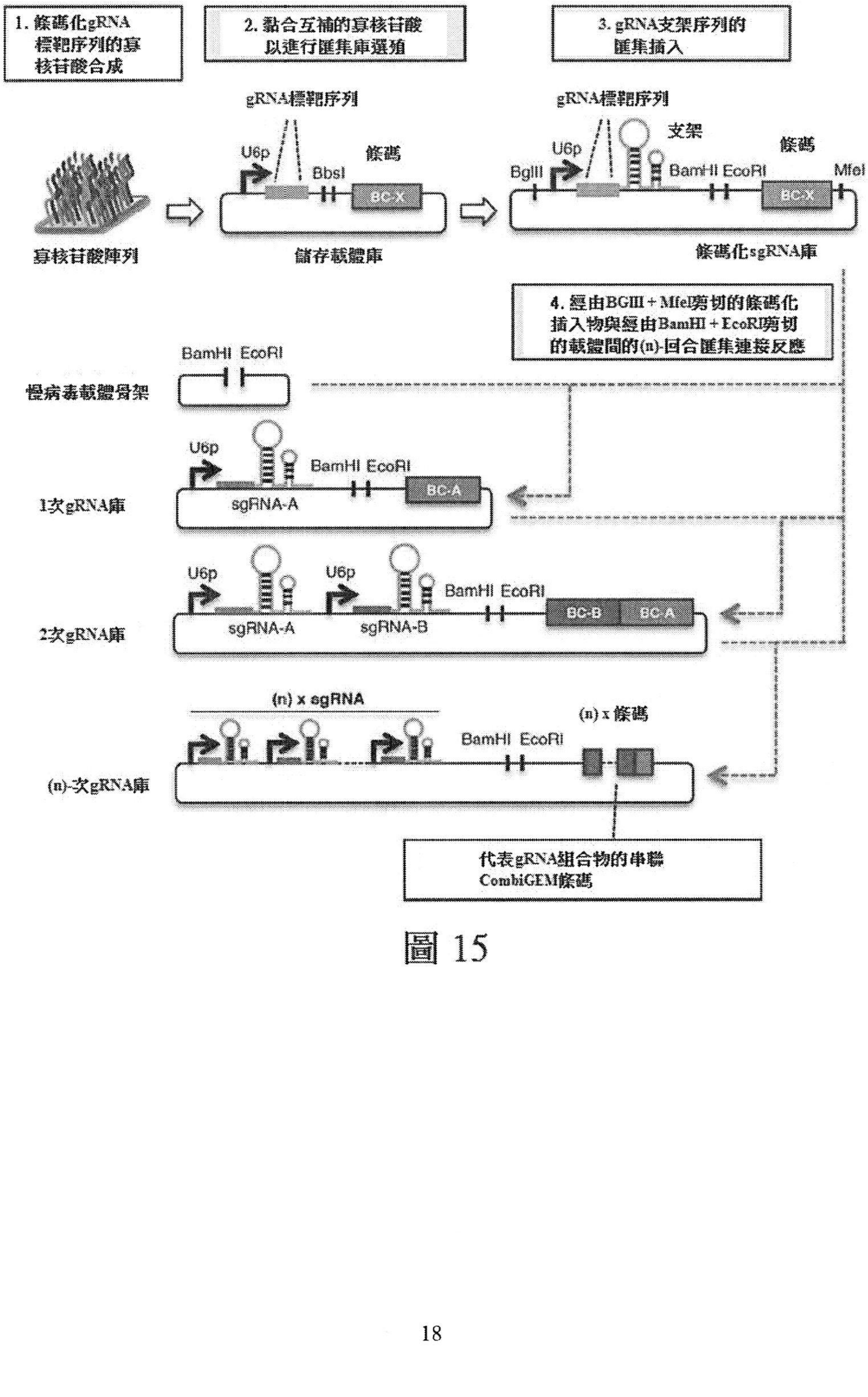

1. 條碼化gRNA標靶序列的寡核苷酸合成
2. 黏合互補的寡核苷酸以進行匯集庫選殖
3. gRNA支架序列的匯集插入
gRNA標靶序列
U6p
BbsI
條碼
BC-X
寡核苷酸陣列
儲存載體庫
gRNA標靶序列
支架
U6p
BglII
BamHI EcoRI
條碼
MfeI
BC-X
條碼化sgRNA庫
4. 經由BGIII + MfeI剪切的條碼化插入物與經由BamHI + EcoRI剪切的載體間的(n)-回合匯集連接反應
BamHI EcoRI
慢病毒載體骨架
U6p
BamHI EcoRI
BC-A
sgRNA-A
1次gRNA庫
U6p
U6p
BamHI EcoRI
BC-B
BC-A
sgRNA-A
sgRNA-B
2次gRNA庫
(n) x sgRNA
(n) x 條碼
BamHI EcoRI
(n)-次gRNA庫
代表gRNA組合物的串聯
CombiGEM條碼

圖 15

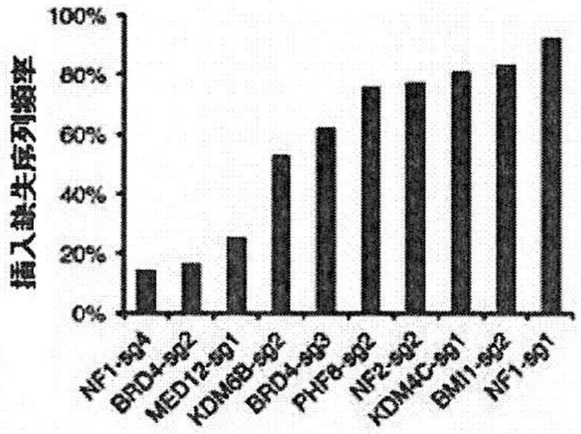


圖 16A

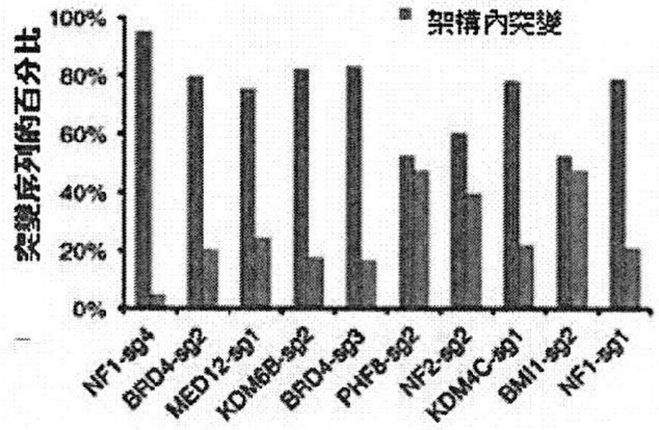


圖 16B

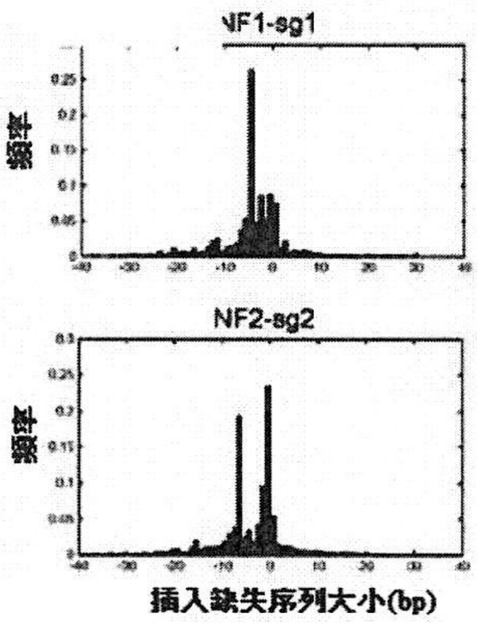


圖 16C

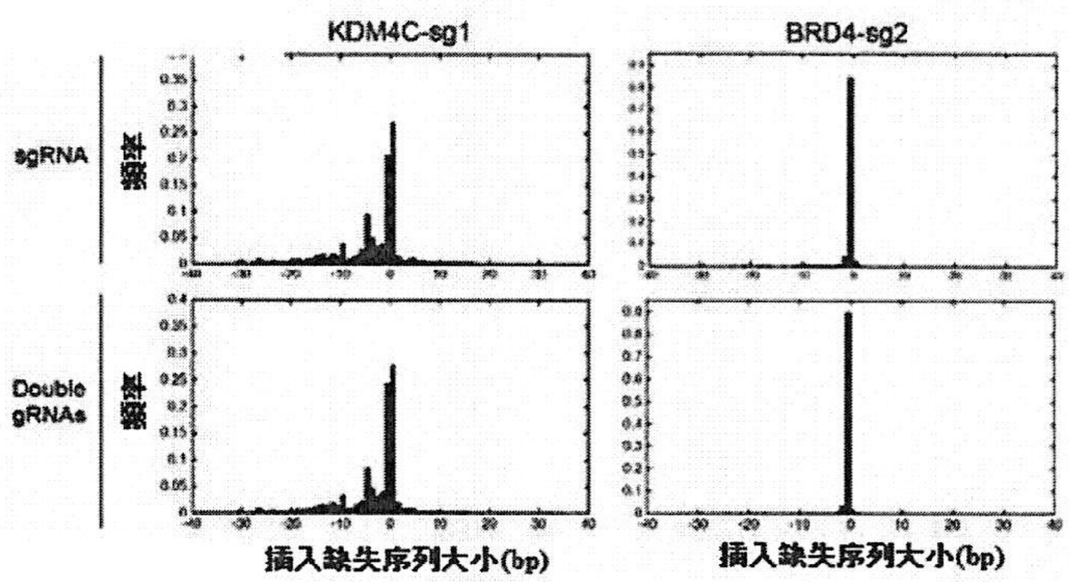
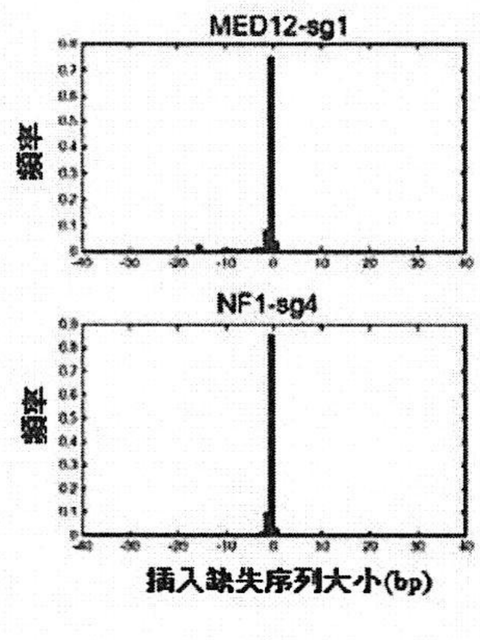


圖 16D



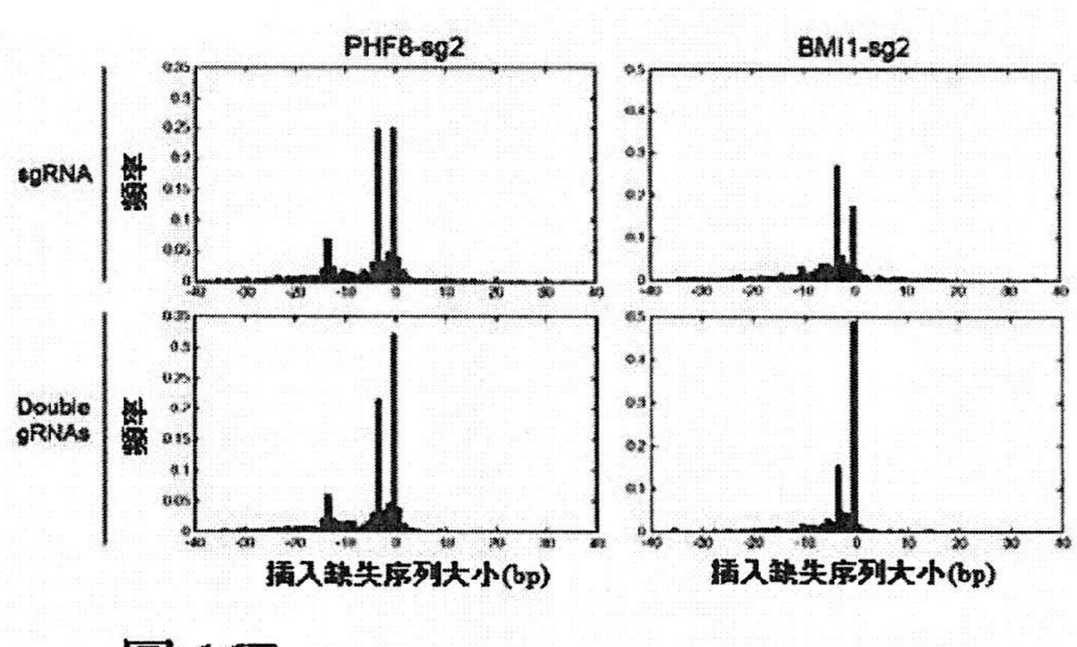


圖 16E

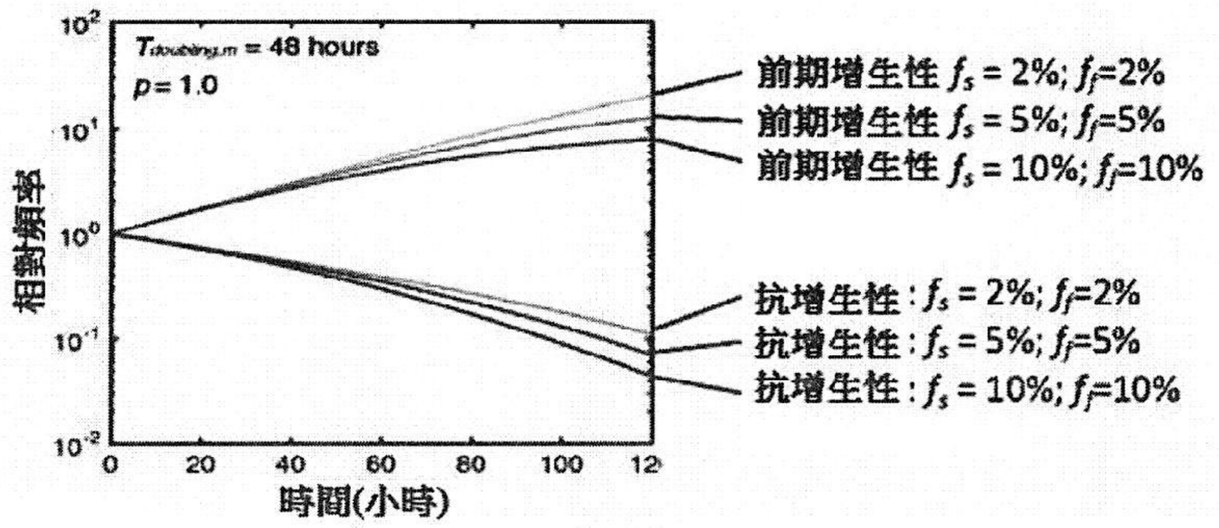


圖 17A

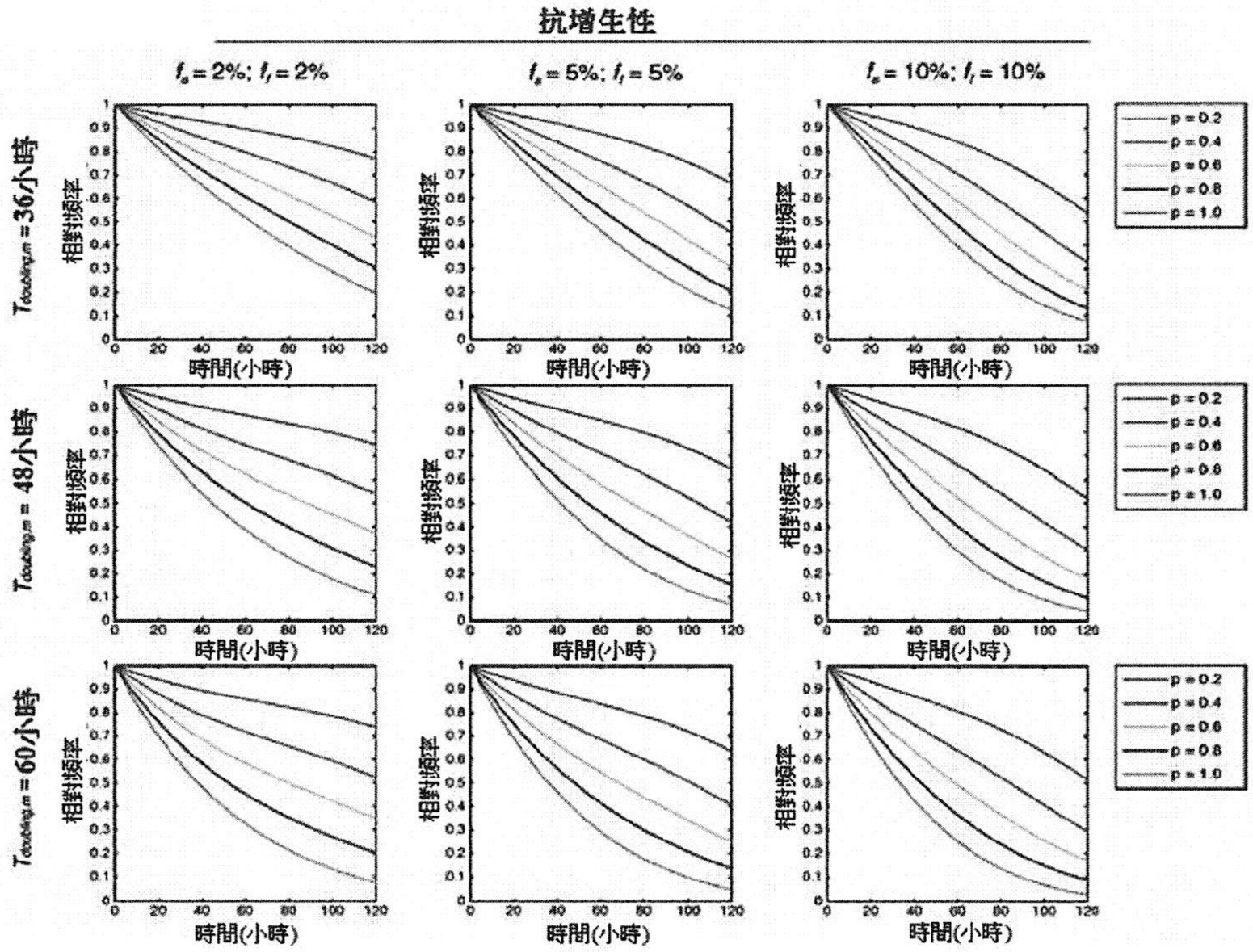


圖 17B

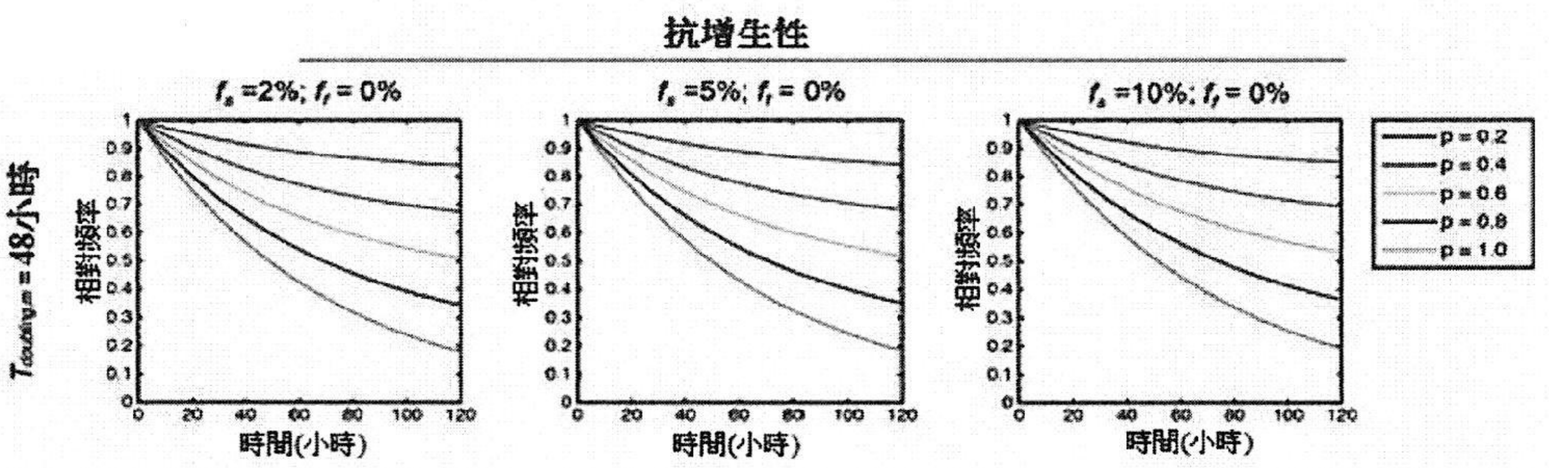


圖 17C

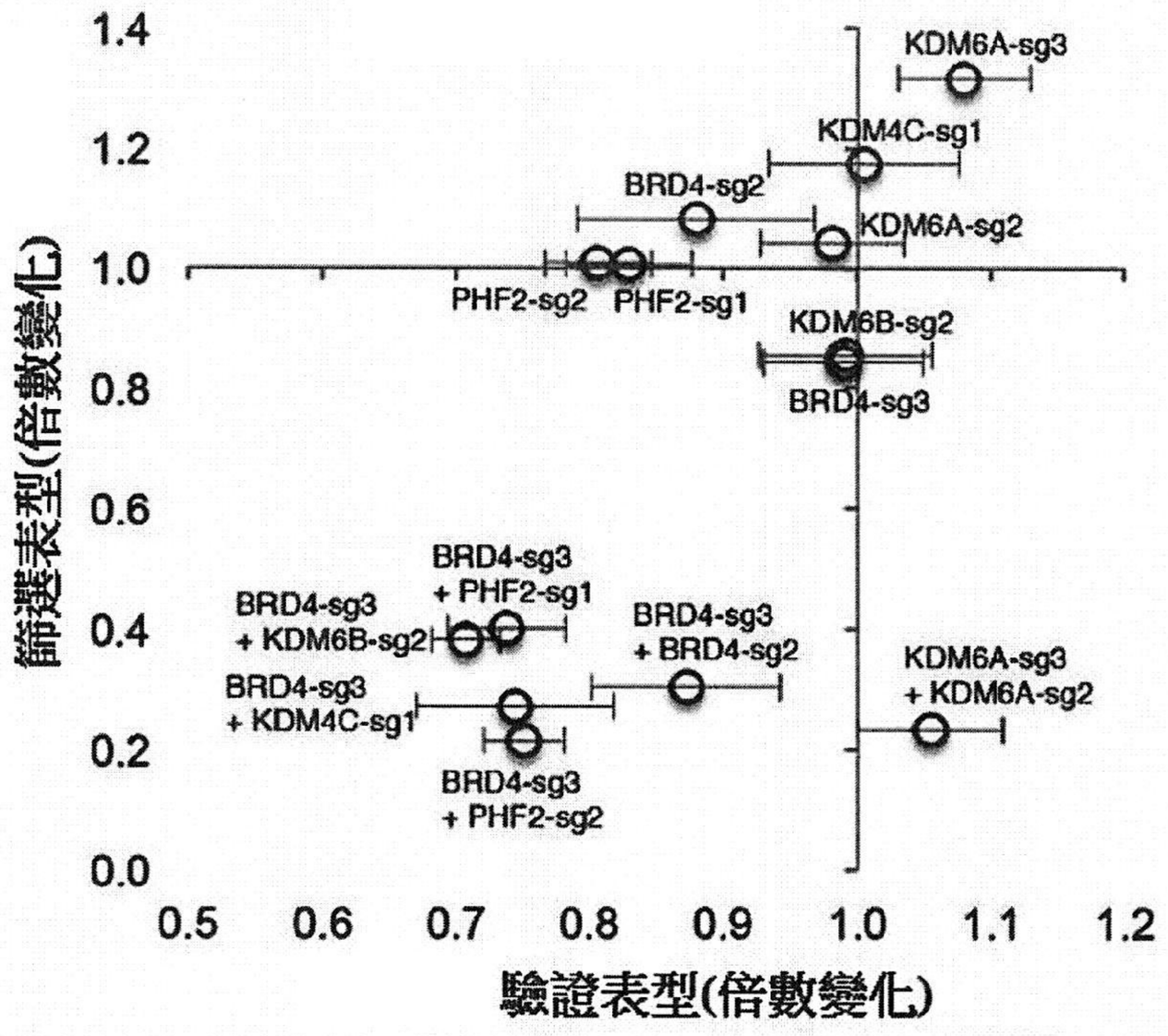

1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
篩選表型(倍數變化)
KDM6A-sg3
KDM4C-sg1
BRD4-sg2
KDM6A-sg2
PHF2-sg2
PHF2-sg1
KDM6B-sg2
BRD4-sg3
BRD4-sg3 + PHF2-sg1
BRD4-sg3 + KDM6B-sg2
BRD4-sg3 + BRD4-sg2
KDM6A-sg3 + KDM6A-sg2
BRD4-sg3 + KDM4C-sg1
BRD4-sg3 + PHF2-sg2
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
驗證表型(倍數變化)

圖 18

	引導序列	PAM		
KDM4C-sg1	**CCTTTGCAAGACCCGCACGA**	TGG	80.85%	
	GCTTTGCCCGAACCGCACGA	GAG	0.29%	
	CCTAGGCCAGACCTGCACGA	TGG	0.23%	
	GCTCTGGAAGACCCGCACCA	GGG	0.21%	外顯子
	CCTTATCAAGACCCACACCA	GAG	0.15%	
KDM6B-sg2	**ATCCCCCTCCTCGTAGCGCA**	**TGG**	52.74%	
	TGCGCCCTCCTCCTAGCGCA	TGG	3.74%	
	CTGCCCCTCCTGGTAGCGCC	TGG	1.01%	
	CTGCTCCTCCTCGTAGCGCT	GGG	0.43%	
	AACCAGCTCCTCGTAGCTCA	GGG	0.32%	
	AGCCCGCTCCTCGTGGGGCA	CGG	0.19%	外顯子
	ACCGCCCTCCTCCTAGCTCA	GGG	0.17%	
	CTCCTCCTCCGCGTAGCGCT	TGG	0.14%	
	CTCCTCCTCCTCCTACCGCA	AGG	n.d.	
	ATTCCTCTCCTGGTACCGCA	AGG	n.d.	
	CTCACCCTCCTCCTAGCACA	TAG	n.d.	
BRD4-sg3	**GGGAACAATAAAGAAGCGCT**	**TGG**	61.97%	
	TGGAAAAACAAAGAAGAGCT	TGG	0.38%	
	GGGAAGTATAAGGAAGAGCT	CAG	0.26%	外顯子
	GTGAGCAATAAAGCAGCCCT	AAG	0.18%	
BRD4-sg2	**TCTAGTCCATCCCCCATTAC**	**TGG**	16.65%	
	AATATTCCATTCCCCATTAC	TGG	0.33%	
	TGTTGTCCATACCTCATTAC	TGG	0.19%	外顯子
	TCTAGGTCATGCACCATTAC	TGG	0.19%	
	TCCACACCCTCCCCCATTAC	TAG	0.25%	內含子
	CCCATTCCTTCCCCCATTAC	CAG	0.21%	

圖 19

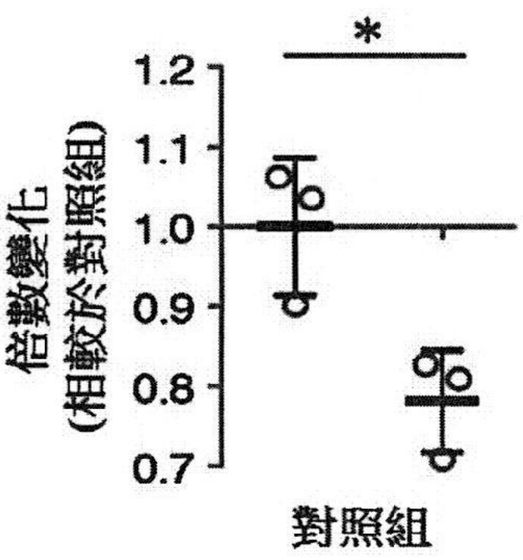


圖 20A

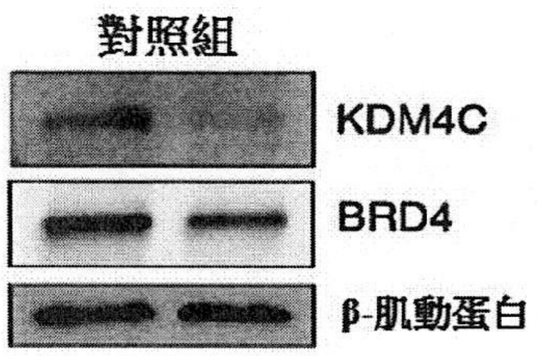


圖 20B

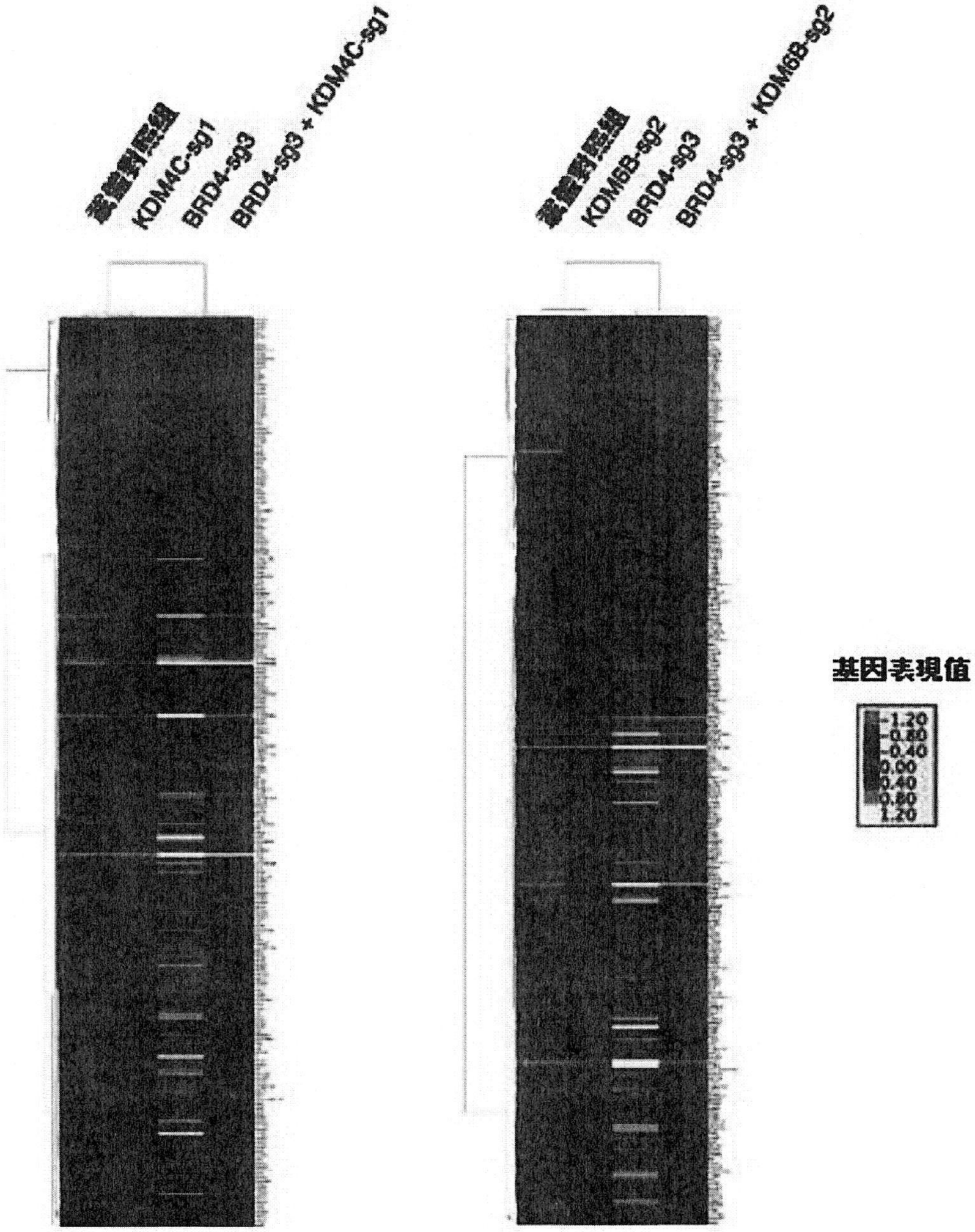

表皮對照組
KDM4C-sg1
BRD4-sg3
BRD4-sg3 + KDM4C-sg1
表皮對照組
KDM6B-sg2
BRD4-sg3
BRD4-sg3 + KDM6B-sg2
基因表現值
-1.20
-0.80
-0.40
0.00
0.40
0.80
1.20

圖 21A

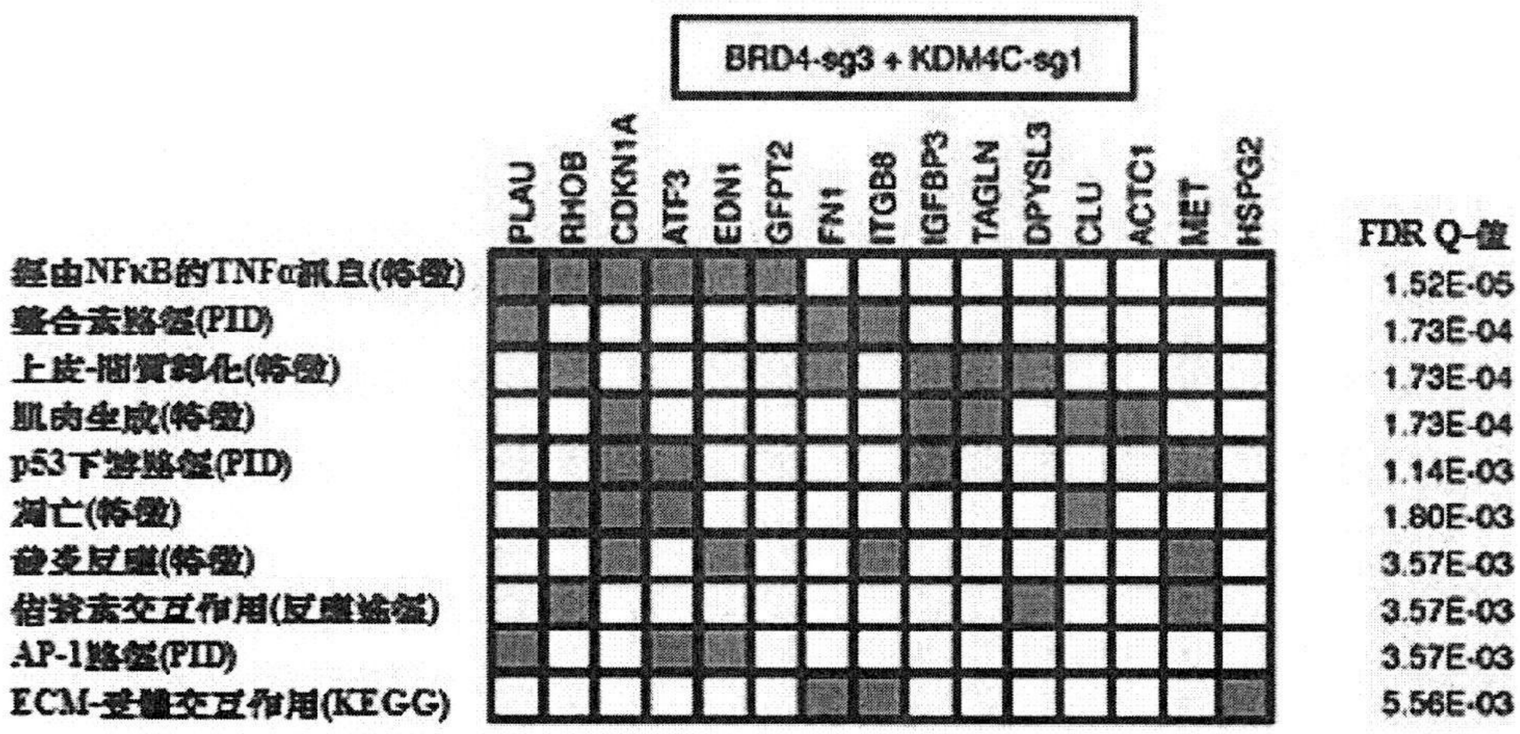

BRD4-sg3 + KDM4C-sg1
PLAU
RHOB
CDKN1A
ATF3
EDN1
GFPT2
FN1
ITGB8
IGFBP3
TAGLN
DPYSL3
CLU
ACTC1
MET
HSPG2
FDR Q-值
經由NFκB的TNFα訊息(特徵) 1.52E-05
整合素路徑(PID) 1.73E-04
上皮-間質轉化(特徵) 1.73E-04
肌肉生成(特徵) 1.73E-04
p53下游路徑(PID) 1.14E-03
凋亡(特徵) 1.80E-03
發炎反應(特徵) 3.57E-03
細胞素交互作用(反應途徑) 3.57E-03
AP-1路徑(PID) 3.57E-03
ECM-受體交互作用(KEGG) 5.56E-03
BRD4-sg3 + KDM6B-sg2
ATF3
FOS
EGR1
PLAU
EDN1
RHOB
CDKN1A
SAT1
PTGS2
GFPT2
CLU
NPTXR
ITGB8
FDR Q-值
經由NFκB的TNFα訊息(特徵) 2.67E-12
AP-1路徑(PID) 3.51E-06
凋亡(特徵) 1.54E-04
NFAT/TF路徑(PID) 3.40E-03
UV反應提高(特徵) 3.40E-03
ATF2路徑(PID) 4.76E-03
p53路徑(特徵) 6.13E-03
轉皮質激素受體訊息路徑(PID) 8.52E-03
被Smad2/3訊息路徑(PID) 8.52E-03
整合素路徑(PID) 1.45E-02

圖 21B

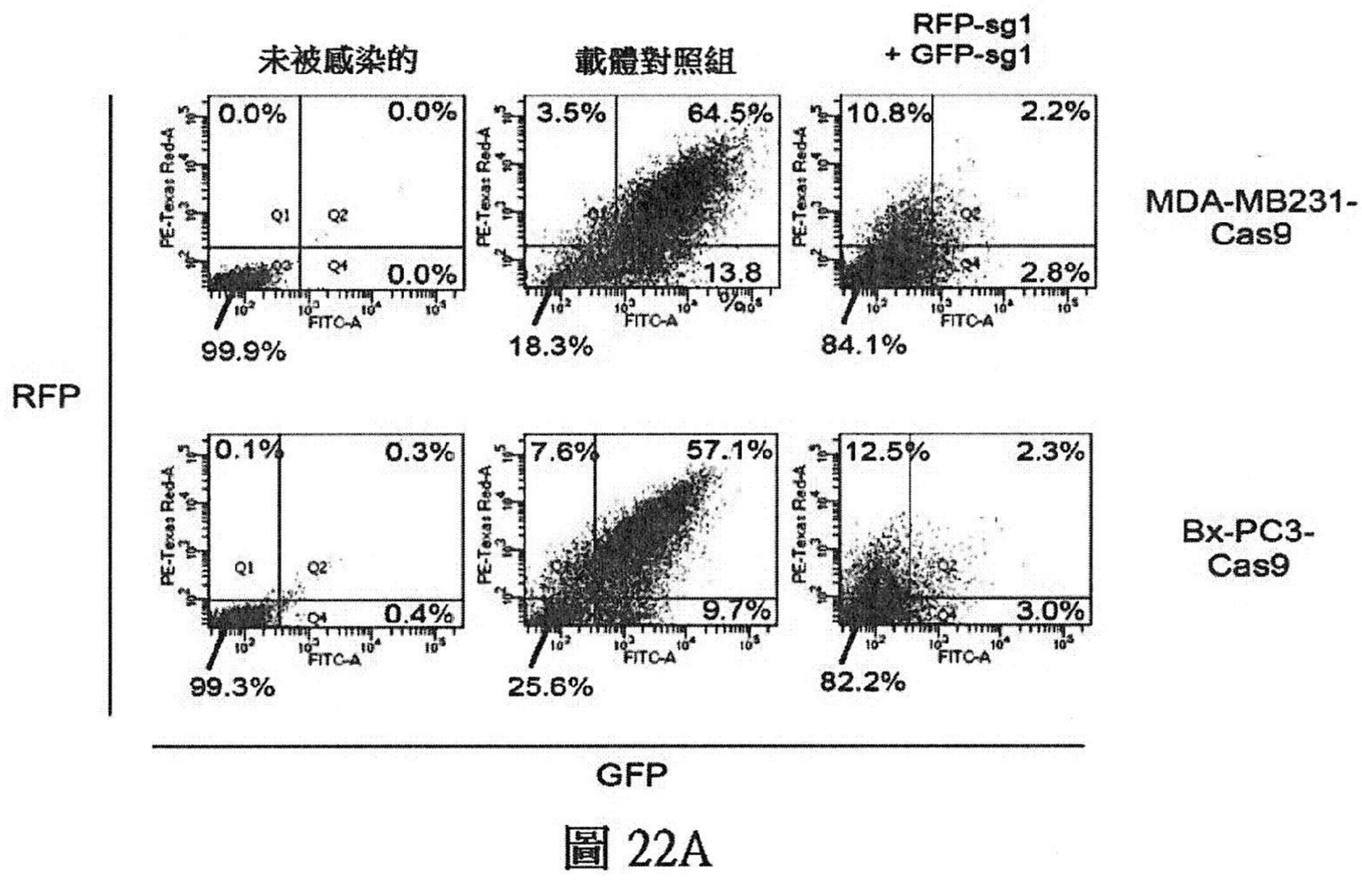

未被感染的
載體對照組
RFP-sg1
+ GFP-sg1
MDA-MB231-
Cas9
Bx-PC3-
Cas9
RFP
GFP
PE-Texas Red-A
FITC-A
0.0%
0.0%
0.0%
99.9%
3.5%
64.5%
13.8%
18.3%
10.8%
2.2%
2.8%
84.1%
0.1%
0.3%
0.4%
99.3%
7.6%
57.1%
9.7%
25.6%
12.5%
2.3%
3.0%
82.2%

圖 22A

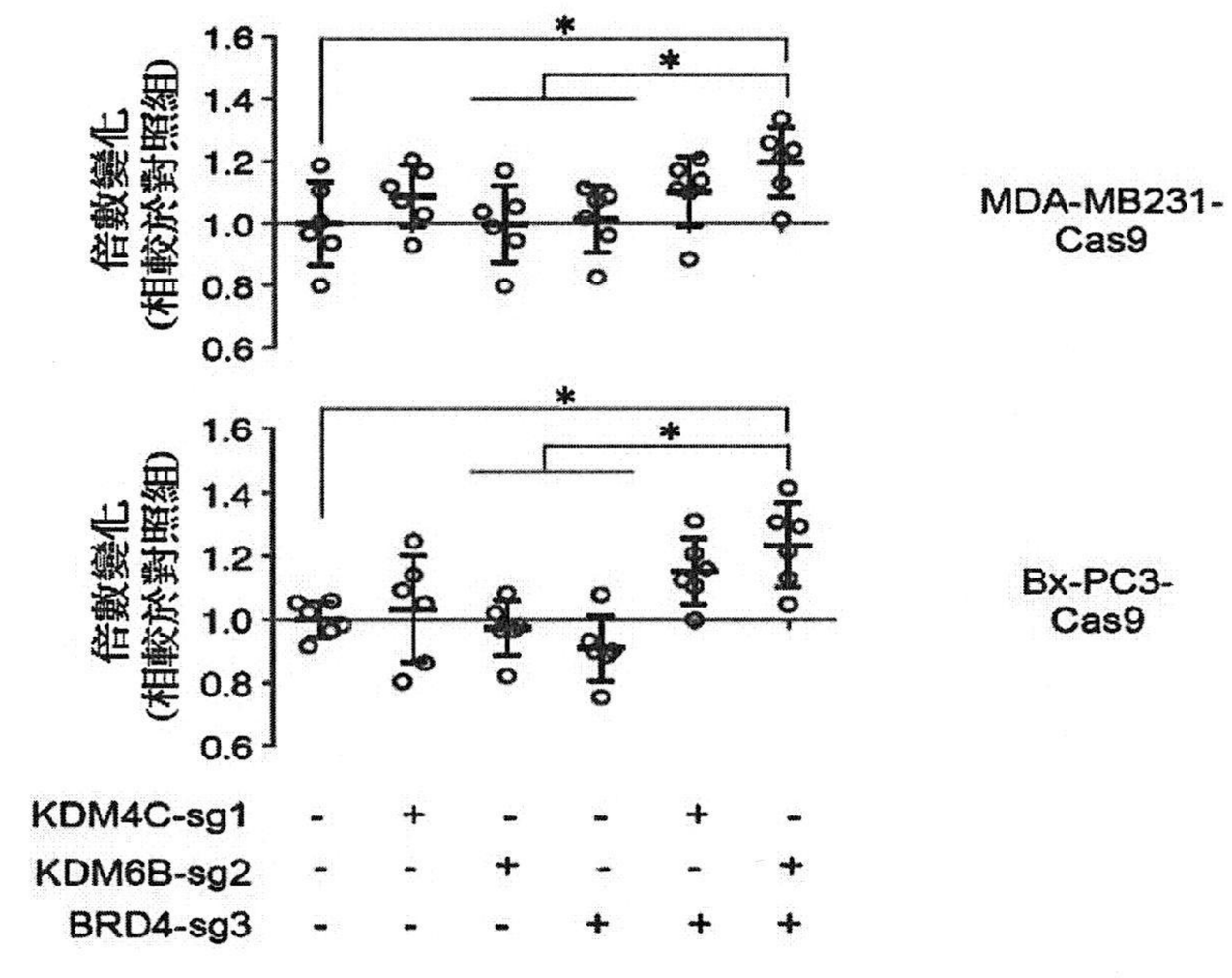

倍數變化
(相較於對照組)
1.6
1.4
1.2
1.0
0.8
0.6
*
*
MDA-MB231-
Cas9
Bx-PC3-
Cas9
KDM4C-sg1 - + - - + -
KDM6B-sg2 - - + - - +
BRD4-sg3 - - - + + +

圖 22B