

公 告 本

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：96147782

※ 申請日期：96.12.13

※IPC 分類：~~C07K~~

一、發明名稱：(中文/英文)

最佳化之非標準鋅指蛋白

OPTIMIZED NON-CANONICAL ZINC FINGER PROTEINS

C12N 15/66 (2006.01)
C12N 15/82 (2006.01)
C12N 15/29 (2006.01)

二、申請人：(共 2 人)

姓名或名稱：(中文/英文)

1. 美商道禮責任有限公司

DOW AGROSCIENCES LLC

2. 美商聖加莫生物科技公司

SANGAMO BIOSCIENCES, INC.

代表人：(中文/英文)

1. 安德歷亞 T 伯魯奇

BORUCKI, ANDREA T.

2. 大衛 市川

ICHIKAWA, DAVID

住居所或營業所地址：(中文/英文)

1. 美國印第安那州印第安那普利市瓊斯維爾路9330號

9330 ZIONSVILLE ROAD, INDIANAPOLIS, IN 46268-1054, U.S.A.

2. 美國加州里奇蒙市運河大道501號尖端禮齊盟科技中心

POINT RICHMOND TECH CENTER, 501 CANAL BLVD., SUITE

A100, RICHMOND, CA 94804, U.S.A.

國 籍：(中文/英文)

1. 美國 U.S.A.

2. 美國 U.S.A.

三、發明人：(共 11 人)

姓 名：(中文/英文)

1. 蔡其華
CAI, QIHUA C.
2. 喬瑟夫 F 派特李諾
PETOLINO, JOSEPH F.
3. 羅比 J 蓋瑞森
GARRISON, ROBBI J.
4. 瓊 C 米切爾
MITCHELL, JON C.
5. 維普拉 K 舒克拉
SHUKLA, VIPULA K.
6. 麗莎 W 貝克
BAKER, LISA W.
7. 雷恩 C 布魯
BLUE, RYAN C.
8. 妮可 L 阿諾
ARNOLD, NICOLE L.
9. 莎拉 E 華登
WORDEN, SARAH E.
10. 傑佛瑞 米勒
MILLER, JEFFREY
11. 法佑德 爾諾夫
URNOV, FYODOR

國 籍：(中文/英文)

1. 中華人民共和國 P.R.C.
2. 美國 U.S.A.
3. 美國 U.S.A.
4. 美國 U.S.A.
5. 美國 U.S.A.
6. 美國 U.S.A.
7. 美國 U.S.A.
8. 美國 U.S.A.
9. 美國 U.S.A.
10. 美國 U.S.A.
11. 俄羅斯 RUSSIA

四、聲明事項：

☐ 主張專利法第二十二條第二項☐第一款或☐第二款規定之事實，其事實發生日期為： 年 月 日。

☒ 申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

☒ 有主張專利法第二十七條第一項國際優先權：

1. 美國；2006年12月14日；60/874,911

2. 美國；2007年05月30日；60/932,497

☐ 無主張專利法第二十七條第一項國際優先權：

1.

2.

☐ 主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

☒ 主張專利法第三十條生物材料：

☐ 須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

☒ 不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

九、發明說明：

【發明所屬之技術領域】

本揭示案係在基因組工程化、基因靶向、靶向染色體整合、蛋白質表現及表觀基因組編輯之領域內。

本申請案主張2006年12月14日申請之美國臨時申請案第60/874,911號及2007年5月30日申請之美國臨時申請案第60/932,497號之權利，該等案件之全部揭示內容以引用的方式併入本文中。

【先前技術】

蛋白質與DNA、RNA、蛋白質及其他分子之序列特異性結合牽涉於諸如轉錄、複製、染色質結構、重組、DNA修復、RNA加工及轉譯之許多細胞過程中。參與蛋白質-DNA、蛋白質-RNA及蛋白質-蛋白質相互作用之細胞結合蛋白之結合特異性促進發育、分化及內穩定。

鋅指蛋白(ZFP)為可以序列特異性方式與DNA結合之蛋白質。鋅指首先在來自非洲爪蟾蜍(*Xenopus laevis*)之卵母細胞之轉錄因子TFIIIA中鑑別出。該類ZFP之單鋅指域為約30個胺基酸長，且若干結構研究已經證明，其含有 β 轉折(含有2個保守半胱胺酸殘基)及 α 螺旋(含有2個保守組胺酸殘基)，其藉由2個半胱胺酸及2個組胺酸，經由鋅原子之配位而保持一特定構形。該類ZFP亦稱為C2H2 ZFP。亦已提示另一類ZFP。參見(例如)，Jiang等人(1996) J. Biol. Chem. 271:10723-10730關於Cys-Cys-His-Cys(C3H)ZFP之討論。迄今為止，已在數千個已知或推定轉錄因子中鑑別出超過10,000個鋅指序列。鋅指域不僅牽涉於DNA識別

中，而且亦牽涉於RNA結合及蛋白質-蛋白質結合中。當前估計，該類分子將構成所有人類基因之約2%。

大多數鋅指蛋白在各指域中具有四面體配位單一鋅原子之保守半胱氨酸及組氨酸殘基。詳言之，大多數ZFP特徵在於具有通用序列：-Cys-(X)₂₋₄-Cys-(X)₁₂-His-(X)₃₋₅-His-(SEQ ID NO:1)之指組分，其中X表示任何胺基酸(C₂H₂ZFP)。該最廣泛表現種類之鋅配位序列含有具有特定間隔之2個半胱氨酸及2個組氨酸。各指之摺疊結構含有反向平行β轉折、指尖區域及短的兩性α螺旋。金屬配位配位體結合於鋅離子且在zif268型鋅指之狀況下，短的兩性α螺旋結合於DNA之大溝中。另外，鋅指之結構藉由某些保守疏水性胺基酸殘基(例如，直接在第一個保守Cys前方之殘基及在指之螺旋區段之位置+4處的殘基)及藉由經由保守半胱氨酸及組氨酸殘基之鋅配位得以穩定。

已描述在產生直接鹼基接觸、'支撐'或'扶持'直接鄰近於鹼基接觸位置之殘基之位置中，及能夠接觸DNA之磷酸酯骨架之位置中具有變異的標準(C₂H₂)鋅指蛋白。參見(例如)，美國專利第6,007,988；6,013,453；6,140,081；6,866,997；6,746,838；6,140,081；6,610,512；7,101,972；6,453,242；6,785,613；7,013,219號；PCT WO 98/53059；Choo等人(2000) Curr. Opin. Struct. Biol. 10:411-416；Segal等人(2000) Curr. Opin. Chem. Biol. 4:34-39。

另外，亦已描述含有具有經修飾鋅配位殘基之鋅指之鋅指蛋白(參見(例如)，美國專利申請第20030108880、

20060246567及20060246588號；其揭示內容以引用之方式併入本文中)。然而，雖然含有該等非標準鋅指之鋅指蛋白保留基因轉錄調節功能，但是其充當鋅指核酸酶(ZFN)之能力在一些狀況下，相對於專有地由標準C₂H₂鋅指組成之鋅指蛋白而言減小。

因此，尤其在建構鋅指核酸酶中，仍對含有具有最佳化非標準鋅配位區域之鋅指之其他工程化鋅指結合蛋白質存在需要。

【發明內容】

本揭示案提供在至少一個鋅配位殘基中具有變異之鋅指DNA結合域。詳言之，本文中描述CCHC鋅指。該等CCHC鋅指可進一步在鋅配位殘基附近，例如在環繞鋅指之最靠C末端之鋅配位殘基的殘基中，包含額外變異(取代、插入及/或缺失)。亦描述包含一或多個該等CCHC鋅指之鋅指多肽及融合蛋白，編碼該等鋅指及融合蛋白之聚核苷酸，及使用該等鋅指多肽及/或融合蛋白之方法。

因此，本揭示案涵蓋(但不限於)以下編號實施例：

1.一種包含非標準(非C₂H₂)鋅指之鋅指蛋白，其中該非標準鋅指具有與DNA結合有關之螺旋部分，且其中該螺旋部分之鋅配位區域包含胺基酸序列HX₁X₂RCX_L(SEQ ID NO:2)；且其中該鋅指蛋白係經工程化以結合於標靶序列。

2.如實施例1之鋅指蛋白，其中X₁為A且X₂為Q。

3.如實施例1之鋅指蛋白，其中X₁為K且X₂為E。

4.如實施例1之鋅指蛋白，其中X₁為T且X₂為R。

- 5.如實施例1之鋅指蛋白，其中 X_L 為G。
- 6.一種包含兩個或兩個以上鋅指之鋅指蛋白，其中至少一個鋅指包含序列 $\text{Cys}-(X^A)_{2-4}-\text{Cys}-(X^B)_{12}-\text{His}-(X^C)_{3-5}-\text{Cys}-(X^D)_{1-10}$ (SEQ ID NO:3)，其中 X^A 、 X^B 、 X^C 及 X^D 可為任何胺基酸。
- 7.如實施例1至6中任一實施例之鋅指蛋白，其包含表1、2、3或4之任一表中所示之序列中之任一者。
- 8.如實施例6或7中任一實施例之鋅指蛋白，其中 X^D 包含序列QLV或QKP。
- 9.如實施例8之鋅指蛋白，其中序列QLV或QKP為鋅指之3個C末端胺基酸殘基。
- 10.如實施例6至9中任一實施例之鋅指蛋白，其中 X^D 包含1、2或3個Gly(G)殘基。
- 11.一種包含複數個鋅指之鋅指蛋白，其中至少一個鋅指包含根據實施例1至10中任一實施例之CCHC鋅指。
- 12.如實施例11之鋅指蛋白，其中該鋅指蛋白包含3、4、5或6個鋅指。
- 13.如實施例11或12之鋅指蛋白，其中鋅指2包含CCHC鋅指。
- 14.如實施例11至13中任一實施例之鋅指蛋白，其中C末端鋅指包含CCHC指。
- 15.如實施例11至14中任一實施例之鋅指蛋白，其中至少2個鋅指包含CCHC鋅指。
- 16.如實施例11至15中任一實施例之鋅指蛋白，其中該鋅指蛋白包含表8中所示序列中之任一者，且其係經工程

化以結合於IPP2-K基因中之標靶序列。

17.一種包含實施例1至16中任一實施例之鋅指蛋白及一或多個功能域融合蛋白。

18.一種融合蛋白，其包含：

(a)裂解半域、

(b)實施例1至16中任一實施例之鋅指蛋白，及

(c)插入裂解半域與鋅指蛋白之間的ZC連接子。

19.如實施例18之融合蛋白，其中該ZC連接子之長度為5個胺基酸。

20.如實施例19之融合蛋白，其中該ZC連接子之胺基酸序列為GLRGS(SEQ ID NO:4)。

21.如實施例18之融合蛋白，其中該ZC連接子之長度為6個胺基酸。

22.如實施例21之融合蛋白，其中該ZC連接子之胺基酸序列為GGLRGS(SEQ ID NO:5)。

23.一種編碼根據實施例1至16中任一實施例之鋅指蛋白或根據實施例17至22中任一實施例之融合蛋白的聚核苷酸。

24.一種靶向裂解植物細胞中之細胞染色質之方法，該方法包含在細胞中表現一對根據實施例18至22中任一實施例之融合蛋白；其中：

(a)該等融合蛋白之標靶序列彼此在10個核苷酸內；且

(b)該等融合蛋白二聚化並裂解位於標靶序列之間的DNA。

25.一種在宿主植物細胞中靶向基因重組之方法，該方法包含：

(a)在該宿主細胞中表現一對根據實施例18至22中任一實施例之融合蛋白，其中該等融合蛋白之標靶序列存在於所選的宿主標靶基因座中；及

(b)鑑別在宿主標靶基因座中具有序列變異之重組宿主細胞。

26.如實施例24或25中任一者之方法，其中序列變異為選自由遺傳物質缺失、遺傳物質插入、遺傳物質取代及其任何組合組成之群的突變。

27.如實施例24至26中任一實施例之方法，其進一步包含將外源性聚核苷酸引入宿主細胞中。

28.如實施例27之方法，其中該外源性聚核苷酸包含與宿主標靶基因座同源之序列。

29.如實施例24至28中任一實施例之方法，其中該植物係選自由單子葉植物、雙子葉植物、裸子植物及真核藻類組成之群。

30.如實施例29之方法，其中該植物係選自由玉米、水稻、小麥、馬鈴薯、大豆、番茄、煙草、芸苔科之成員及芥菜屬(*Arabidopsis*)組成之群。

31.如實施例24至29中任一實施例之方法，其中該植物為樹木。

32.如實施例24至31中任一實施例之方法，其中該標靶序列係在IPP2K基因中。

33.一種降低種子中植酸含量之方法，該方法包含使根據實施例32之IPP2-K基因失活或變異。

34.一種使磷在種子中代謝可利用性更高之方法，該方法包含使根據實施例32之IPP2-K基因失活或變異。

35.一種包含根據實施例1至16中任一實施例之鋅指蛋白、根據實施例17至22中任一實施例之融合蛋白或根據實施例23之聚核苷酸的植物細胞。

36.如實施例35之植物細胞，其中該細胞為種子。

37.如實施例36之植物細胞，其中該種子為玉米種子。

38.如實施例35至37中任一實施例之植物細胞，其中該IPP2-K係部分或完全失活。

39.如實施例38之植物細胞，其中該種子中之植酸含量降低。

40.如實施例35至39之植物細胞，其中該細胞中磷之代謝可利用含量增加。

【實施方式】

本文中揭示包含含有具有Cys-Cys-His-Cys格式之非標準鋅指之鋅指結合多肽(ZFP)的組合物。由於鋅配位為鋅指提供主要摺疊能量，故鋅配位殘基之調整提供改良指穩定性及結構之現成方式，該指穩定性及結構影響鋅指蛋白之多種重要功能性特徵，包括(例如)細胞半衰期、與其它細胞因子之相互作用、DNA結合特異性及親和力及功能域之相對定向。

已展示包含非標準鋅指之鋅指蛋白(諸如美國專利申請

案第 20030108880；20060246567；及 20060246588 號中所揭示者)結合 DNA 且改變轉錄。然而，當併入至鋅指核酸酶(ZFN，參見(例如)美國專利申請公開案第 2005/0064474 號)中時，該等先前所述之非標準鋅指蛋白可有時在裂解標靶 DNA 時顯示次最佳活性。

本文中描述包含一或多個 CCHC 鋅指之鋅指蛋白，其中鋅配位殘基周圍之 C 末端對之特定序列已改變。本文中亦描述融合蛋白，例如鋅指核酸酶(ZFN)，其包含該等最佳化非標準鋅指，其中 ZFN 以可與使用包含標準(CCHH)鋅指之 ZFN 所達到之裂解相比的速率裂解標靶 DNA。

如本文中所揭示之融合多肽可增強或抑制基因之轉錄，及/或裂解標靶序列。亦提供編碼最佳化非標準鋅指之聚核苷酸及編碼包含一或多個最佳化非標準鋅指之融合蛋白之聚核苷酸。另外提供醫藥組合物，其包含與醫藥學上可接受之載劑組合的治療有效量之本文所述之鋅指-核苷酸結合多肽中之任一者或其功能性片段；或治療有效量之編碼經修飾鋅指-核苷酸結合多肽中之任一者或其功能性片段的核苷酸序列。另外提供農業組合物，其包含與農業上可接受之載劑組合的農藝學有效量之本文所述之鋅指-核苷酸結合多肽中之任一者或其功能性片段；或農藝學有效量之編碼經修飾鋅指-核苷酸結合多肽中之任一者或其功能性片段的核苷酸序列。亦提供用於獲得結合於基因組序列之經修飾鋅指-核苷酸結合多肽之篩選方法。

基因組序列包括存在於以下各物中之彼等序列：染色

體、離合染色小體、細胞器基因組(例如線粒體、葉綠體)、人工染色體及諸如擴增序列、雙微染色體及內源性或感染細菌及病毒基因組之存在於細胞中之任何其它類型之核酸。基因組序列可為正常(亦即野生型)或突變體；突變序列可包含(例如)插入、缺失、取代、移位、重排及/或點突變。基因組序列亦可包含許多不同等位基因之一。

概要

除非另外指示，否則本文中揭示之方法以及組合物之製備及用途的實施使用分子生物學、生物化學、染色質結構及分析、計算化學、細胞培養、重組DNA及此項技術技藝內之相關領域中的習知技術。該等技術在文獻中充分闡釋。參見(例如)，Sambrook等人MOLECULAR CLONING: A LABORATORY MANUAL，第二版，Cold Spring Harbor Laboratory Press, 1989及第三版，2001；Ausubel等人，CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987及定期更新；系列METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION，第三版，Academic Press, San Diego, 1998；METHODS IN ENZYMOLOGY，第304卷，"Chromatin" (P.M. Wassarman及A. P. Wolffe編)，Academic Press, San Diego, 1999；及METHODS IN MOLECULAR BIOLOGY，第119卷，"Chromatin Protocols" (P.B. Becker編) Humana Press, Totowa, 1999。

定義

術語"核酸"、"聚核苷酸"及"寡核苷酸"可互換使用且係指呈線性或環形構形且呈單鏈或雙股形式之脫氧核糖核苷酸或核糖核苷酸聚合物。為本揭示案之目的起見，關於聚合物之長度，該等術語不欲解釋為具有限制性。該等術語可涵蓋天然核苷酸之已知類似物，以及在鹼基、糖及/或磷酸酯部分(例如硫代磷酸酯骨架)中經修飾之核苷酸。一般而言，特定核苷酸之類似物具有相同鹼基配對特異性；亦即，A之類似物將與T鹼基配對。

術語"多肽"、"肽"及"蛋白質"可互換使用以指胺基酸殘基之聚合物。該術語亦適用於一或多種胺基酸為相應天然存在胺基酸之化學類似物或經修飾衍生物之胺基酸聚合物。

"結合"係指大分子之間的(例如蛋白質與核酸之間)序列特異性、非共價相互作用。並非結合相互作用之所有組分需要具有序列特異性(例如，與DNA骨架中磷酸酯殘基之接觸)，只要相互作用總體上為序列特異性的即可。該相互作用特徵通常在於 10^{-6} M^{-1} 或更低之解離常數(K_d)。"親和力"係指結合強度：增加之結合親和力與較低之 K_d 相關。

"結合蛋白質"為能夠與另一分子非共價結合之蛋白質。結合蛋白質可結合於(例如)DNA分子(DNA結合蛋白質)、RNA分子(RNA結合蛋白質)及/或蛋白質分子(蛋白質結合蛋白質)。在蛋白質結合蛋白質之狀況下，其可與其自身

結合(以形成同源二聚體、同源三聚體等等)，及/或其可結合於不同蛋白質之一或多個分子。結合蛋白質可具有超過一種類型之結合活性。舉例而言，鋅指蛋白具有DNA結合、RNA結合及蛋白質結合活性。

"鋅指DNA結合蛋白質"(或結合域)為蛋白質或為較大蛋白質內部之域，其以序列特異性方式，經由一或多個鋅指來結合DNA，該等鋅指為結合域內之胺基酸序列區域，該結合域之結構經由鋅離子之配位來穩定。術語鋅指DNA結合蛋白質常常縮寫為鋅指蛋白或ZFP。

鋅指結合域可經"工程化"以結合於預定核苷酸序列。用於工程化鋅指蛋白之方法之非限制性實例為設計及選擇。經設計鋅指蛋白為不存在於自然界中之蛋白質，其設計/組成主要由合理標準產生。用於設計之合理標準包括應用取代規則及處理儲存現有ZFP設計及結合資料之資訊之資料庫中的資訊之電腦化算法。參見(例如)，美國專利6,140,081；6,453,242；6,534,261；及6,785,613；亦參見WO 98/53058；WO 98/53059；WO 98/53060；WO 02/016536及WO 03/016496；及美國專利6,746,838；6,866,997；及7,030,215。

"所選"鋅指蛋白為自然界中未發現之蛋白質，其製造主要由諸如噬菌體顯示、相互作用捕獲或雜交選擇之經驗性方法產生。參見(例如)，US 5,789,538；US 5,925,523；US 6,007,988；US 6,013,453；US 6,200,759；US 6,733,970；US RE39,229；及WO 95/19431；WO 96/06166；WO

98/53057 ; WO 98/54311 ; WO 00/27878 ; WO 01/60970
WO 01/88197及WO 02/099084。

"非標準"鋅指蛋白為包含非標準(非C₂H₂)鋅指之蛋白質。因此與天然存在C₂H₂鋅指蛋白相比，非標準鋅指包含至少一個胺基酸之取代、添加及/或缺失。非標準鋅指之非限制性實例包括包含Cys-Cys-His-Cys(例如C₃H)之鋅配位殘基(自胺基至羧基)之彼等鋅指。

"同源序列"係指第一序列與第二序列共同具有一定程度之序列一致性，且其序列可與第二序列之序列相同。"同源、非相同序列"係指第一序列與第二序列共同具有一定程度之序列一致性，但其序列不同於第二序列之序列。舉例而言，構成突變基因之野生型序列之聚核苷酸與突變基因之序列同源且不相同。在某些實施例中，2個序列之間的同源程度足以允許在其間利用正常細胞機制進行同源重組。2個同源非相同序列可具有任何長度且其非同源程度可低至單個核苷酸(例如，對藉由靶向同源重組來校正基因組點突變而言)或高至10個或10個以上千鹼基(kilobase)(例如，對在染色體中之預定位點處插入基因而言)。2個包含同源非相同序列之聚核苷酸不需要具有相同長度。舉例而言，可使用具有20與10,000個之間的核苷酸或核苷酸對之外源性聚核苷酸(亦即供體聚核苷酸)。

用於測定核酸及胺基酸序列一致性之技術為此項技術中所已知。通常，該等技術包括測定基因之mRNA之核苷酸序列及/或測定其所編碼之胺基酸序列，且將該等序列與

第二核苷酸或胺基酸序列比較。基因組序列亦可以該方式測定及比較。一般而言，一致性分別指2個聚核苷酸或多肽序列之確切核苷酸與核苷酸或胺基酸與胺基酸對應性。兩個或兩個以上序列(聚核苷酸或胺基酸)可藉由測定其一致性百分比來比較。2個序列(無論為核酸還是胺基酸序列)之一致性百分比為2個經比對序列之間之確切匹配數除以較短序列之長度且乘以100。核酸序列之近似比對由Smith及Waterman, **Advances in Applied Mathematics** 2:482-489 (1981)之局部同源算法來提供。該算法可藉由使用由Dayhoff, **Atlas of Protein Sequences and Structure**, M.O. Dayhoff編，5增刊3:353-358, National Biomedical Research Foundation, Washington, D.C., USA開發且由Gribskov, **Nucl. Acids Res.** 14(6):6745-6763 (1986)正規化之計分矩陣應用於胺基酸序列。測定序列一致性百分比之該算法的例示性實施例由Genetics Computer Group(Madison, WI)於"BestFit"實用申請案中提供。該方法之預設參數描述於Wisconsin Sequence Analysis Package Program Manual，第8版(1995)(可獲自Genetics Computer Group, Madison, WI)中。在本揭示案之上下文中建立一致性百分比之例示性方法欲使用由University of Edinburgh授權，由John F. Collins及Shane S. Sturrok開發，且由IntelliGenetics, Inc. (Mountain View, CA)銷售之MPSRCH程式包。自此套程式包，可使用Smith-Waterman算法，其中對於計分表使用預設參數(例如，空隙開口罰分為12，空隙擴展罰分為1且空

隙為6)。自所產生之資料，"匹配"值反映序列一致性。用於計算序列之間的一致性百分比或相似性之其它適合程式通常為此項技術中所已知，例如，另一比對程式為連同預設參數一起使用之BLAST。舉例而言，BLASTN及BLASTP可使用以下預設參數來使用：遺傳密碼=標準；過濾器=無；鏈=兩個；截止=60；預期=10；Matrix=BLOSUM62；說明=50個序列；分類依據(sort by)=高分(HIGH SCORE)；資料庫=非冗餘、GenBank+EMBL+DDBJ+PDB+GenBank CDS轉譯+Swiss蛋白質+Spupdate+PIR。該等程式之細節可見於網際網路。關於本文中所述之序列，序列一致性之所要程度之範圍為約35%至100%及其間任何整數值。通常，序列之間的一致性百分比為至少35%-40%；40%-45%；45%-50%；50%-60%；60%-70%；70-75%，較佳80-82%，更佳85-90%，甚至更佳92%，仍更佳95%，且最佳98%之序列一致性。

或者，聚核苷酸之間的序列相似性程度可藉由在允許在同源區域之間形成穩定雙股體之條件下使聚核苷酸雜交，接著用單鏈特異性核酸酶消化，且測定所消化片段之大小來測定。如使用上述方法所測定，2個核酸或2個多肽序列當該等序列在所確定之分子長度範圍內，顯示至少約70%-75%，較佳80%-82%，更佳85%-90%，甚至更佳92%，仍更佳95%，且最佳98%之序列一致性時，大體上彼此同源。如本文中所使用，大體上同源亦係指與指定DNA或多肽序列展示完全一致性之序列。大體上同源之DNA序列可

以(例如)如對彼特定系統所確定之嚴格條件下之南方雜交實驗(Southern hybridization experiment)鑑別。確定適當雜交條件係在熟習此項技術者之技術範圍內。參見(例如)，Sambrook等人，同上；**Nucleic Acid Hybridization: A Practical Approach**，編者 B.D. Hames 及 S.J. Higgins, (1985) Oxford；Washington, DC；IRL Press)。

2個核酸片段之選擇性雜交可如下測定。2個核酸分子之間的序列一致性程度影響該等分子之間雜交事件之效率及強度。部分一致之核酸序列將至少部分抑制完全一致序列與標靶分子之雜交。對完全一致序列之雜交的抑制可使用在此項技術中熟知之雜交檢定(例如，南方(DNA)墨點(Southern (DNA) blot)、北方(RNA)墨點(Northern (RNA) blot)、溶液雜交或其類似方法，參見 Sambrook，等人，*Molecular Cloning: A Laboratory Manual*，第二版，(1989) Cold Spring Harbor, N.Y.)來評估。該等檢定可使用不同程度之選擇性，例如使用自低嚴格性變化至高嚴格性之條件來進行。若使用具有低嚴格性之條件，則非特異性結合之缺少可使用甚至缺乏部分程度之序列一致性的第二探針(例如與標靶分子具有小於約30%序列一致性之探針)來評估，以使得在不存在非特異性結合事件之情況下，第二探針將不與標靶雜交。

當利用基於雜交之偵測系統時，選擇與參考核酸序列互補之核酸探針，且隨後藉由選擇探針及參考序列彼此選擇性雜交或結合以形成雙股體分子之適當條件來選擇核酸探

針。能夠在適度嚴格雜交條件下與參考序列選擇性雜交之核酸分子通常在允許偵測長度至少為約10-14個核苷酸與所選核酸探針之序列具有至少約70%序列一致性的標靶核酸序列之條件下雜交。嚴格雜交條件通常允許偵測長度至少為約10-14個核苷酸與所選核酸探針之序列具有大於約90-95%序列一致性的標靶核酸序列。適用於其中探針及參考序列具有特定程度之序列一致性之探針/參考序列雜交的雜交條件，可如此項技術中所已知來確定(參見(例如)，**Nucleic Acid Hybridization: A Practical Approach**，編者 B.D. Hames 及 S.J. Higgins, (1985) Oxford ; Washington, DC ; IRL Press)。

用於雜交之條件對熟習此項技術者而言為熟知的。雜交嚴格性係指雜交條件不利於形成含有錯配核苷酸雜交體之程度，其中較高嚴格性與錯配雜交體之較低耐受性相關。影響雜交嚴格性之因素對熟習此項技術者而言為熟知的且包括(但不限於)溫度、pH值、離子強度及諸如甲醯胺及二甲亞砷之有機溶劑之濃度。如熟習此項技術者所已知，雜交嚴格性因更高溫度、更低離子強度及更低溶劑濃度而增加。

關於用於雜交之嚴格條件，在此項技術中熟知，可使用許多等效條件藉由改變(例如)以下因素來建立特定嚴格性：序列之長度及性質、各種序列之鹼基組成、鹽及其他雜交溶液組分之濃度、雜交溶液中存在或不存在阻斷劑(例如硫酸葡聚糖及聚乙二醇)、雜交反應溫度及時間參

數，以及變化之洗滌條件。雜交條件之特定設置之選擇係按照此項技術中之標準方法來選擇(參見(例如)，Sambrook，等人，**Molecular Cloning: A Laboratory Manual**，第二版，(1989) Cold Spring Harbor, N.Y.)。

"重組"係指交換2個聚核苷酸之間的遺傳資訊之方法。對該揭示案之目的而言，"同源重組(HR)"係指(例如)在細胞中修復雙股斷裂期間發生的該交換之指定形式。該過程需要核苷酸序列同源性，使用"供體"分子以作為修復"標靶"分子(亦即，經歷雙股斷裂之分子)之模板，且不同地稱為"非交換基因轉變"或"短管基因轉變"，此係因為其使得遺傳資訊自供體轉移至標靶。不希望受任何特定理論約束，該轉移可包括在斷裂標靶與供體之間形成之異源雙股DNA的錯配校正；及/或"合成依賴性鏈黏接"，其中供體用以再合成將成為標靶之部分之遺傳資訊；及/或相關方法。該指定HR常常導致標靶分子之序列變異，致使供體聚核苷酸之部分或所有序列併入標靶聚核苷酸中。

"裂解"係指DNA分子之共價骨架斷裂。裂解可藉由包括(但不限於)磷酸二酯鍵之酶促或化學水解之多種方法來引發。單鏈裂解及雙股裂解為可能的，且雙股裂解可作為2個不同單鏈裂解事件之結果發生。DNA裂解可造成鈍末端或交錯末端之產生。在某些實施例中，對於靶向雙股DNA裂解使用融合多肽。

"裂解域"包含一或多個具有DNA裂解催化活性之多肽序列。裂解域可包含於單個多肽鏈中或裂解活性可由2個(或

2個以上)多肽之結合而產生。

"裂解半域"為與第二多肽(相同或不同)結合形成具有裂解活性(例如雙股裂解活性)之複合物的多肽序列。

術語"裂解域"及"裂解半域"包括野生型域及保留多聚化(例如二聚化)以形成功能裂解域之能力的裂解域或裂解半域之部分或突變體。

"染色質"為包含細胞基因組之核蛋白質結構。細胞染色質包含核酸，主要為DNA，及蛋白質，包括組蛋白及非組蛋白染色體蛋白質。大多數真核細胞染色質以核小體形式存在，其中核小體核心包含約150個與包含各具有組蛋白H2A、H2B、H3及H4中之2者之八聚體結合的DNA鹼基對；且連接子DNA(具有視生物體而定之可變長度)在核小體核心之間延伸。組蛋白H1之分子通常與連接子DNA結合。就本揭示案之目的而言，術語"染色質"意欲涵蓋所有類型之原核及真核之細胞核蛋白質。細胞染色質包括染色體染色質及游離型染色質。

"染色體"為包含細胞之全部或部分基因組之染色質複合物。細胞之基因組常常由其核型表徵，該核型為所有構成細胞基因組之染色體的集合。細胞基因組可包含一或多個染色體。

"離合染色小體"為複製核酸、核蛋白質複合物或包含不為細胞染色體核型之部分的核酸之其它結構。離合染色小體之實例包括質體及某些病毒基因組。

"可達區域"為細胞染色質中存在於核酸中之標靶位點可

由識別該標靶位點之外源性分子結合的位點。不希望受任何特定理論約束，據信可達區域為未包裝於核小體結構中之區域。可達區域之不同結構可常常因其對化學及酶探針(例如核酸酶)之敏感性而得以偵測。

"標靶位點"或"標靶序列"為界定結合分子所欲結合之核酸之部分的核酸序列，其限制條件為存在足以結合之充分條件。舉例而言，序列5'-GAATTC-3'為Eco RI限制性核酸內切酶之標靶位點。

"外源性"分子為通常不存在於細胞中，但可由一或多種遺傳、生物化學或其它方法引入細胞中之分子。"在細胞中正常存在"係相對於細胞之特定發育階段及環境條件而言來判定。因此，舉例而言，僅在肌肉之胚胎發育期間存在之分子相對於成熟肌細胞而言為外源性分子。類似地，由熱休克誘導之分子相對於非熱休克細胞而言為外源性分子。外源性分子可包含(例如)功能異常內源分子之功能型式或正常作用內源分子之功能異常型式。

外源性分子可尤其為諸如由組合化學方法產生之小分子，或大分子，諸如蛋白質、核酸、碳水化合物、脂質、醣蛋白、脂蛋白、多醣、上述分子之任何經修飾衍生物，或包含上述分子中之一或多者之任何複合物。核酸包括DNA及RNA，可為單鏈或雙股；可為線性、分枝或環形；且可具有任何長度。核酸包括能夠形成雙股體之彼等核酸，以及三鏈體形成核酸。參見(例如)，美國專利第5,176,996及5,422,251號。蛋白質包括(但不限於)，DNA結

合蛋白質、轉錄因子、染色質重塑因子、甲基化DNA結合蛋白質、聚合酶、甲基酶、去甲基酶、乙醯基酶、脫乙醯基酶、激酶、磷酸酶、整合酶、重組酶、連接酶、拓撲異構酶、旋轉酶及解螺旋酶。

外源性分子可為與內源分子相同類型之分子，例如外源性蛋白質或核酸。舉例而言，外源性核酸可包含感染病毒基因組、根癌農桿菌T鏈(*Agrobacterium tumefaciens* T-strand)、引入細胞中之質體或離合染色小體，或通常不存在於細胞中之染色體。然而，外源性核酸或聚核苷酸可含有與內源序列同源或一致之序列。相對於特定內源基因組區域，"外源性序列"係指在彼區域不存在之核苷酸序列。該外源性序列可存在於另一內源染色體位置或其可根本不存在於基因組中。因此，外源性聚核苷酸可含有外源性及內源序列：例如，側接有與基因組區域同源之序列之轉殖基因。該等外源性核酸係用於如下文所述之靶向整合及靶向重組之方法中。用於將外源性分子引入細胞中之方法對熟習此項技術者而言為已知的且包括(但不限於)脂質介導轉移(亦即，包括中性及陽離子性脂質之脂質體)、電穿孔、直接注射、細胞融合、粒子轟擊、磷酸鈣共沈澱、DEAE-葡聚糖介導轉移及病毒載體介導轉移。

相反，"內源分子"為通常存在於特定環境條件下處於特定發育階段之特定細胞中之分子。舉例而言，內源核酸可包含染色體，線粒體、葉綠體或其它細胞器之基因組，或天然存在之游離型核酸。其他內源分子可包括蛋白質，例

如轉錄因子及酶。

"融合"分子為兩個或兩個以上子單元分子(例如)共價連接之分子。子單元分子可為同一化學類型之分子，或可為不同化學類型之分子。第一類型融合分子之實例包括(但不限於)融合蛋白(例如 ZFP DNA結合域與裂解域之間的融合體)及融合核酸(例如，編碼上文所述之融合蛋白之核酸)。第二類型融合分子之實例包括(但不限於)三鏈體形成核酸與多肽之間的融合體，及小溝結合子與核酸之間的融合。

細胞中融合蛋白之表現可由將融合蛋白傳遞至細胞或將編碼融合蛋白之聚核苷酸傳遞至細胞來產生，其中轉錄聚核苷酸，且轉譯轉錄物以產生融合蛋白。反式剪接、多肽裂解及多肽連接亦可牽涉於細胞中蛋白質之表現中。將聚核苷酸及多肽傳遞至細胞中之方法在本揭示案中其它地方呈示。

對本揭示案之目的而言，"基因"包括編碼基因產物(參看下文)之DNA區域，以及調節基因產物產生之所有DNA區域，無論該等調節序列是否鄰近於編碼及/或轉錄序列。因此，基因包括(但不必限於)啟動子序列、終止子、諸如核糖體結合位點及內部核糖體進入位點之轉譯調節序列、強化子、靜止子、隔離子、邊界元件、複製起點、基質連接位點及基因座控制區域。

"基因表現"係指包含在基因中之資訊轉變成基因產物。基因產物可為基因之直接轉錄產物(例如 mRNA、tRNA、

rRNA、反義RNA、核糖核酸酶、結構RNA或任何其它類型之RNA)或mRNA轉譯所產生之蛋白質。基因產物亦包括藉由諸如封端、多聚腺嘌呤化、甲基化及編輯之方法修飾之RNA，及藉由(例如)甲基化、乙醯化、磷酸化、泛素化、ADP-核糖基化、肉豆蔻酸化及糖基化修飾之蛋白質。

"調節"基因表現係指基因活性之改變。表現之調節可包括(但不限於)基因活化及基因抑制。

"植物"細胞包括(但不限於)單子葉(monocotyledonous, monocots)或雙子葉(dicotyledonous, dicots)植物之細胞。單子葉植物之非限制性實例包括穀類植物，諸如玉米、水稻、大麥、燕麥、小麥、高粱、黑麥、甘蔗、菠蘿、洋蔥、香蕉及椰子。雙子葉植物之非限制性實例包括煙草、番茄、向日葵、棉花、甜菜、馬鈴薯、萵苣、瓜、大豆、菜籽(油菜籽)及紫苜蓿。植物細胞可來自植物之任何部分及/或來自植物發育之任何階段。

"所關注之區域"為細胞染色質之任何區域，諸如基因或基因內或鄰近於基因之非編碼序列，其中希望結合外源性分子。結合可用於靶向DNA裂解及/或靶向重組之目的。所關注之區域可存在於(例如)染色體、離合染色小體、細胞器基因組(例如線粒體、葉綠體)或感染病毒基因組中。所關注之區域可在基因之編碼區域內部，在諸如引導序列、尾隨序列或內含子之轉錄非編碼區域內，或在編碼區域之上游或下游之非轉錄區域內部。所關注之區域長度可

低至單個核苷酸對或高達25,000個核苷酸對，或任何整數值之核苷酸對。

當提及兩個或兩個以上組分(諸如序列元件)之鄰接時，術語"操作性連接"及"經操作性連接"(或"可操作連接")可互換使用，其中組分經排列以使得組分正常行使功能且允許組分之至少一者可介導在其它組分之至少一者後發揮之功能的可能性。以說明之方式，若轉錄調節序列(諸如啟動子)回應一或多種轉錄調節因子之存在或不存在而控制編碼序列之轉錄量，則該轉錄調節序列係與編碼序列操作性連接。轉錄調節序列通常以順式與編碼序列操作性連接，但不需要直接與其相鄰。舉例而言，強化子為與編碼序列操作性連接之轉錄調節序列，即使其並不毗鄰。

關於融合多肽，術語"操作性連接"可指各組分在與其它組分之連接中執行與其在不如此連接之情況下執行之功能相同之功能。舉例而言，關於ZFP DNA結合域與裂解域融合之融合多肽，若在融合多肽中，ZFP DNA結合域部分能夠結合其標靶位點及/或其結合位點，而裂解域能夠裂解標靶位點附近之DNA，則ZFP DNA結合域與裂解域係呈操作性連接。

蛋白質、多肽或核酸之"功能性片段"為其序列不同於全長蛋白質、多肽或核酸，但保留與全長蛋白質、多肽或核酸相同之功能之蛋白質、多肽或核酸。功能性片段可具有更多、更少或與相應天然分子相同數量之殘基，及/或可含有一或多個胺基酸或核苷酸取代。測定核酸功能(例如

編碼功能，與另一核酸雜交之能力)之方法為此項技術中所熟知。類似地，測定蛋白質功能之方法為熟知的。舉例而言，多肽之DNA結合功能可(例如)藉由濾紙結合、電泳遷移率遷移或免疫沈澱檢定來測定。DNA裂解可藉由凝膠電泳來檢定。參見Ausubel等人，同上。蛋白質與另一蛋白質相互作用之能力可(例如)藉由共免疫沈澱、雙雜交檢定(two-hybrid assay)或遺傳及生物化學互補檢定來測定。參見(例如)，Fields等人(1989) *Nature* 340:245-246；美國專利第5,585,245號及PCT WO 98/44350。

鋅指結合域

本文中描述非標準鋅指結合域及編碼該等鋅指結合域之聚核苷酸。在某些實施例中，本文中描述之非標準鋅指結合域為C3H鋅指，其中2個保守鋅配位組胺酸殘基中之一者經轉變為半胱胺酸。在其他實施例中，最靠C末端之組胺酸殘基經轉變為半胱胺酸殘基，產生"CCHC蛋白質"。

鋅指結合域可包含一或多個鋅指(例如2、3、4、5、6、7、8、9個或9個以上鋅指)，且可經工程化以與任何標靶序列(例如，基因組序列)結合。鋅指結合域可與DNA、RNA及/或蛋白質結合。通常，單一鋅指域長度為約30個胺基酸。鋅指包括標準C₂H₂鋅指(亦即，鋅離子與2個半胱胺酸及2個組胺酸殘基配位者)及包括(例如)C₃H鋅指(鋅離子與3個半胱胺酸殘基及1個組胺酸殘基配位者)之非標準鋅指。亦參見美國專利申請第20030108880、20060246567及20060246588號；其揭示內容以引用之方式併入本文

中。

結構研究已證明，標準鋅指域(基元)含有2個 β 片(保持為含有2個不變半胱氨酸殘基之 β 轉折)及 α 螺旋(含有2個不變組氨酸殘基)，其藉由2個半胱氨酸及2個組氨酸經由鋅原子之配位而保持為特定構形。本文中揭示之非標準鋅指保留該 β - β - α 結構。

本文中描述之非標準鋅指可為天然存在之鋅指結合域。然而，更通常地，如本文中所述之非標準鋅指包括一或多種鋅指組分，其中鋅配位半胱氨酸或組氨酸殘基中之至少一者已經一或多個胺基酸置換。舉例而言，在某些實施例中，標準鋅指結合模組之C末端His殘基經Cys殘基置換。

本文中描述之CCHC鋅指亦可包含一或多個在不為鋅配位殘基之胺基酸殘基之序列中的變異(相對於天然存在之C₂H₂鋅指之序列而言)。該等變異可包含取代、缺失及/或插入。胺基酸可在鋅指中之任何位置變異。變異之非限制性實例包括：(1)變異鋅配位殘基周圍之單個殘基取代；(2)在變異鋅配位殘基之前或之後添加額外殘基(例如，在其中最靠C末端之His殘基經轉變為Cys之狀況下，添加額外胺基酸殘基可藉由補償較短半胱氨酸側鏈來促進鋅配位)；及/或(3)使位於天然存在CCHC鋅指之His與Cys殘基之間的殘基取代至非標準CCHC鋅指之相應區域中。

在某些實施例中，本文中描述之鋅指蛋白包括至少一個包含非標準(非C₂H₂)鋅指之鋅指，其中非標準鋅指具有牽涉於DNA結合中之螺旋部分且其中螺旋部分之鋅配位區域

包含胺基酸序列HX₁X₂RCX_L(SEQ ID NO:2)；且其中鋅指蛋白經工程化以與標靶序列結合。在某些實施例中，X₁為A或K或T；X₂為Q或E或R；且X_L為G。

在其他實施例中，本文中描述之非標準鋅指具有通用結構：
Cys-(X^A)₂₋₄-Cys-(X^B)₁₂-His-(X^C)₃₋₅-Cys-(X^D)₁₋₁₀(SEQ ID NO:3)，其中X^A、X^B、X^C及X^D表示任何胺基酸。在X^C包含3個殘基之實施例中，(i)該等殘基中之至少一者與標準CCHC鋅指相比經變異；及/或(ii)與標準CCHH鋅指相比，X^D包含至少一個缺失、取代或插入。在某些實施例中，X^D包含序列QLV或QKP。在其他實施例中，X^D包含一或多個(例如1、2、3、4、5、6、7、8、9或10個)Gly(G)殘基。

例示性非標準鋅指之部分胺基酸序列(包括第三鋅配位殘基之C末端)展示於表1、2、3及4中。在所有表中，2個最靠C末端(亦即第三及第四)鋅配位殘基(H及C)加下劃線表示。與"野生型"非標準指(表1及3之第2列)相比之變異(例如取代、插入、缺失)以雙下劃線展示。

表 1

C2H2	<u>H</u> TKI <u>H</u> LRGSQLV(野生型標準)(SEQ ID NO:12)
1=C3H	<u>H</u> TKI <u>C</u> LRGSQLV(野生型經變異至非標準)(SEQ ID NO:13)
2	<u>H</u> TK <u>G</u> <u>C</u> LRGSQLV (SEQ ID NO:14)
3	<u>H</u> TK <u>A</u> <u>C</u> LRGSQLV (SEQ ID NO:15)
4	<u>H</u> TK <u>V</u> <u>C</u> LRGSQLV (SEQ ID NO:16)
5	<u>H</u> TK <u>L</u> <u>C</u> LRGSQLV (SEQ ID NO:17)
6	<u>H</u> TK <u>S</u> <u>C</u> LRGSQLV (SEQ ID NO:18)

7	<u>HTKN</u> CLRGSQLV (SEQ ID NO:19)
8	<u>HTKK</u> CLRGSQLV (SEQ ID NO:20)
9	<u>HTKR</u> CLRGSQLV (SEQ ID NO:21)
10	<u>HTKIGG</u> CLRGSQLV (SEQ ID NO:22)
11	<u>HTKICG</u> LRGSQLV (SEQ ID NO:23)
12	<u>HTKICGG</u> LRGSQLV (SEQ ID NO:24)
13	<u>HTKIGCG</u> LRGSQLV (SEQ ID NO:25)
14	<u>HTKIGCGG</u> LRGSQLV (SEQ ID NO:26)
15	<u>HLKG</u> NCLRGSQLV (SEQ ID NO:27)
16	<u>HLKG</u> NCPAGSQLV (SEQ ID NO:28)
17	<u>HSEGG</u> CLRGSQLV (SEQ ID NO:29)
18	<u>HSEGGC</u> PGGSQLV (SEQ ID NO:30)
19	<u>HSSN</u> CLRGSQLV (SEQ ID NO:31)
20	<u>HSSNCT</u> IGSQLV (SEQ ID NO:32)

表 2

1	<u>HTKICGGG</u> LRGSQLV (SEQ ID NO:33)
2	<u>HTKIGCGG</u> LRGSQLV (SEQ ID NO:34)
3	<u>HTKIGG</u> CLRGSQLV (SEQ ID NO:35)
4	<u>HTKIGGCG</u> LRGSQLV (SEQ ID NO:36)
5	<u>HTKIGGCGG</u> LRGSQLV (SEQ ID NO:37)
6	<u>HTKRCG</u> LRGSQLV (SEQ ID NO:38)
7	<u>HTKRCGG</u> LRGSQLV (SEQ ID NO:39)
8	<u>HTKRCGGG</u> LRGSQLV (SEQ ID NO:40)
9	<u>HTKRG</u> CLRGSQLV (SEQ ID NO:41)
10	<u>HTKRGCG</u> LRGSQLV (SEQ ID NO:42)
11	<u>HTKRGCGG</u> LRGSQLV (SEQ ID NO:43)
12	<u>HTKRGCGGG</u> LRGSQLV (SEQ ID NO:44)
13	<u>HTKRGG</u> CLRGSQLV (SEQ ID NO:45)
14	<u>HTKRGGCG</u> LRGSQLV (SEQ ID NO:46)
15	<u>HTKRGGCGG</u> LRGSQLV (SEQ ID NO:47)

16	<u>HLKGN</u> <u>CGLR</u> GSQLV (SEQ ID NO:48)
17	<u>HLKGN</u> <u>CGGLR</u> GSQLV (SEQ ID NO:49)
18	<u>HLKGN</u> <u>CGGGLR</u> GSQLV (SEQ ID NO:50)
19	<u>HKERC</u> <u>G</u> LRGSQLV (SEQ ID NO:51)
20	<u>HTRRC</u> <u>G</u> LRGSQLV (SEQ ID NO:52)
21	<u>HAQRC</u> <u>G</u> LRGSQLV (SEQ ID NO:53)
22	<u>HKKFYC</u> <u>G</u> LRGSQLV (SEQ ID NO:54)
23	<u>HKKHYC</u> <u>G</u> LRGSQLV (SEQ ID NO:55)
24	<u>HKKYTC</u> <u>G</u> LRGSQLV (SEQ ID NO:56)
25	<u>HKKYYC</u> <u>G</u> LRGSQLV (SEQ ID NO:57)
26	<u>HKQYYC</u> <u>G</u> LRGSQLV (SEQ ID NO:58)
27	<u>HLLKK</u> <u>C</u> LRGSQLV (SEQ ID NO:59)
28	<u>HQKFPC</u> <u>G</u> LRGSQLV (SEQ ID NO:60)
29	<u>HQKKLC</u> <u>G</u> LRGSQLV (SEQ ID NO:61)
30	<u>HQIRGC</u> <u>G</u> LRGSQLV (SEQ ID NO:62)
31	<u>HIKROSC</u> <u>G</u> LRGSQLV (SEQ ID NO:63)
32	<u>HIRRYTC</u> <u>G</u> LRGSQLV (SEQ ID NO:64)
33	<u>HISSKKC</u> <u>G</u> LRGSQLV (SEQ ID NO:65)
34	<u>HKIQKAC</u> <u>G</u> LRGSQLV (SEQ ID NO:66)
35	<u>HKRIYTC</u> <u>G</u> LRGSQLV (SEQ ID NO:67)
36	<u>HLKGONC</u> <u>G</u> LRGSQLV (SEQ ID NO:68)
37	<u>HLKKDGC</u> <u>G</u> LRGSQLV (SEQ ID NO:69)
38	<u>HLKYTPC</u> <u>G</u> LRGSQLV (SEQ ID NO:70)
39	<u>HTKR</u> <u>CGR</u> GSQLV (SEQ ID NO:71)
40	<u>HTKIGC</u> <u>GGR</u> GSQLV (SEQ ID NO:72)
41	<u>HLKGN</u> <u>CGR</u> GSQLV (SEQ ID NO:73)
42	<u>HLKGN</u> <u>CGGGS</u> QLV (SEQ ID NO:74)
43	<u>HIRTCTGSQ</u> <u>KP</u> (SEQ ID NO:75)
44	<u>HIRTCGTGSQ</u> <u>KP</u> (SEQ ID NO:76)
45	<u>HIRTGCTGSQ</u> <u>KP</u> (SEQ ID NO:77)
46	<u>HIRTGCGTGSQ</u> <u>KP</u> (SEQ ID NO:78)

47	<u>H</u> <u>I</u> <u>R</u> <u>R</u> <u>C</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:79)
48	<u>H</u> <u>I</u> <u>R</u> <u>R</u> <u>G</u> <u>C</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:80)

表 3

wt	<u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>H</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:81)
1a	<u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:82)
2a	<u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:83)
3a	<u>H</u> <u>A</u> <u>Q</u> <u>R</u> <u>C</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:84)
4a	<u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:85)
5a	<u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:86)
6a	<u>H</u> <u>A</u> <u>Q</u> <u>R</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>S</u> <u>Q</u> <u>K</u> <u>P</u> (SEQ ID NO:87)

表 4

wt	<u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>H</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:88)
7a	<u>H</u> <u>A</u> <u>Q</u> <u>R</u> <u>C</u> <u>G</u> <u>G</u> (SEQ ID NO:89)
8a	<u>H</u> <u>A</u> <u>Q</u> <u>R</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> (SEQ ID NO:90)
9a	<u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> (SEQ ID NO:91)
10a	<u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> (SEQ ID NO:92)
11a	<u>H</u> <u>A</u> <u>Q</u> <u>R</u> <u>C</u> <u>G</u> (SEQ ID NO:93)

如上所述，ZFP可包括任何數量之鋅指結合域，例如至少3個鋅指。此外，1個，超過1個或所有鋅指可為如本文中所述之非標準鋅指。

在某些實施例中，多指鋅指蛋白之最靠C末端之指包含標準鋅指。在其他實施例中，多指鋅指蛋白之最靠C末端之指包含如本文中所述之CCHC指，例如在最靠C末端之鋅配位Cys殘基之C末端包含一或多個胺基酸插入之CCHC指。參見描述4指鋅指蛋白之實例1-5，其中指2(F2)及/或指4(F4)為如本文中所述之非標準鋅指。

鋅指結合域可經工程化以與所選序列結合。參見(例如)，Beerli等人(2002) *Nature Biotechnol.* **20**:135-141；Pabo等人(2001) *Ann. Rev. Biochem.* **70**:313-340；Isalan等人(2001) *Nature Biotechnol.* **19**:656-660；Segal等人(2001) *Curr. Opin. Biotechnol.* **12**:632-637；Choo等人(2000) *Curr. Opin. Struct. Biol.* **10**:411-416。工程化鋅指結合域與天然存在鋅指蛋白相比，可具有新穎結合特異性。工程化方法包括(但不限於)合理設計及各種類型之選擇(例如，針對單個標靶核苷酸序列篩選複數個不同鋅指序列之方法)。合理設計包括(例如)使用包含三聯組(或四聯組)核苷酸序列及個別鋅指胺基酸序列之資料庫，其中各三聯組或四聯組核苷酸序列係與結合特定三聯組或四聯組序列之鋅指之一或多個胺基酸序列相關。參見(例如)，共同擁有之美國專利6,453,242及6,534,261。其他設計方法揭示於(例如)美國專利6,746,838；6,785,613；6,866,997；及7,030,215中。已在(例如)共同擁有之美國專利第6,794,136號中描述鋅指結合域之結合特異性之增強。

包括噬菌體顯示及雙雜交系統之例示性選擇方法揭示於美國專利5,789,538；5,925,523；6,007,988；6,013,453；6,410,248；6,140,466；6,200,759；及6,242,568；以及WO 98/37186；WO 98/53057；WO 00/27878；WO 01/88197及GB 2,338,237中。

因為個別鋅指結合於3核苷酸(亦即三聯組)序列(或4核苷酸序列，其可與相鄰鋅指之4核苷酸結合位點重疊1個核苷

酸)，所以鋅指結合域經工程化所結合之序列(例如標靶序列)之長度將決定工程化鋅指結合域中之鋅指數。舉例而言，對鋅指基元不與重疊子位點結合之ZFP而言，6核苷酸標靶序列由2指結合域結合；9核苷酸標靶序列由3指結合域結合，等等。標靶位點中個別鋅指(亦即子位點)之結合位點無需毗鄰，而可由1或若干個核苷酸隔開，其視多指結合域中鋅指(亦即指間連接子)之間的胺基酸序列之長度及性質而定。參見(例如)，美國專利6,479,626；6,903,185及7,153,949及美國專利申請公開案第2003/0119023號；其揭示內容以引用之方式併入本文中。

在多指鋅指結合域中，相鄰鋅指可由具有約5個胺基酸之胺基酸連接子序列(所謂"標準"指間連接子)或者由一或多個非標準連接子隔開。參見(例如)，共同擁有之美國專利第6,453,242及6,534,261號。對包含超過3個指之工程化鋅指結合域而言，在某些鋅指之間插入更長("非標準")指間連接子序列可增加親和力及/或由結合域結合之特異性。參見(例如)，美國專利第6,479,626號及美國專利申請公開案第2003/0119023號；其揭示內容以引用之方式併入本文中。因此，多指鋅指結合域亦可關於非標準指間連接子之存在及位置來表徵。較長指間連接子之使用亦可促進鋅指蛋白與包含非鄰接核苷酸之標靶位點之結合。因此，鋅指結合域之標靶位點中之一或多個子位點可由1、2、3、4、5個或5個以上核苷酸彼此隔開。為提供(僅)一實例，4指結合域可與按序包含2個鄰接3核苷酸子位點、插

入核苷酸及2個鄰接三聯組子位點之13核苷酸標靶位點結合。

標靶子位點為由單個鋅指結合之核苷酸序列(通常3或4個核苷酸)。然而，標靶位點不必為3個核苷酸之倍數。舉例而言，在交叉鏈相互作用發生之狀況下(參見(例如)，美國專利6,453,242及6,794,136)，多指結合域之個別鋅指中之一或多者可與重疊四聯組子位點結合。亦參見美國專利6,746,838及6,866,997。為提供(僅)一實例，3指結合域可與包含3個重疊4核苷酸子位點之10核苷酸標靶位點結合。

細胞染色質中由鋅指域結合之序列(例如標靶位點)之選擇可(例如)根據揭示於共同擁有之美國專利第6,453,242號(2002年9月17日)中之方法來實現，該專利亦揭示設計與所選序列結合之ZFP之方法。對熟習此項技術者顯而易見，對於選擇標靶位點亦可使用核苷酸序列之簡單目視檢查。因此，用於標靶位點選擇之任何方式可用於本文中所述之方法中。

多指鋅指蛋白可藉由聯接(例如)藉由設計或選擇獲得之個別鋅指來建構。或者，由2個鋅指組成之結合模組可使用標準或較長、非標準指間連接子(參見上文)而彼此聯接，以產生4指及6指蛋白質。該兩指模組可(例如)藉由選擇2個結合多指蛋白質(通常3指)之上下文中之特定6核苷酸標靶序列的相鄰指而獲得。參見(例如)，WO 98/53057及美國專利申請公開案第2003/0119023號；其揭示內容以引用之方式併入本文中。或者，兩指模組可藉由個別鋅指之

組裝來建構。

因此，本文中描述之鋅指域可個別地使用或以各種組合使用，以建構與任何標靶位點結合之多指鋅指蛋白。

序列(例如標靶位點)之間的距離係指如自彼此最接近之序列之邊界量測，插入2個序列之間的核苷酸或核苷酸對之數目。

在使用 ZFN，例如裂解係依賴於2個鋅指域/裂解半域融合分子之結合以分隔標靶位點之實施例中，2個標靶位點可在相對DNA鏈上。在其他實施例中，2個標靶位點均在同一DNA鏈上。參見(例如)，WO 2005/084190；其揭示內容以引用之方式併入本文中。

編碼鋅指或鋅指蛋白之聚核苷酸亦係在本揭示案之範疇內。該等聚核苷酸可使用標準技術來建構且插入至載體中，且可將該載體引入細胞中(參見下文關於將聚核苷酸引入細胞中之載體及方法之額外揭示案)以使得所編碼之蛋白質在細胞中表現。

融合蛋白

亦提供包含一或多個本文中所述之非標準鋅指組分之融合蛋白。

融合分子係藉由熟習此項技術者所熟知之選殖及生物化學接合方法來建構。融合分子包含含CCHC之ZFP及(例如)裂解域、裂解半域、轉錄活化域、轉錄抑制域、染色質重塑複合物之組分、隔離子域、該等域中任一者之功能性片段；及/或兩個或兩個以上功能域或其片段之任何組合。

在某些實施例中，融合分子包含經修飾植物鋅指蛋白及至少2個功能域(例如，隔離子域或甲基結合蛋白質域及，另外，轉錄活化或抑制域)。

融合分子亦視需要包含核定位信號(諸如，來自SV40 T 抗原或玉米 Opaque-2 NLS者)及抗原決定基標記(諸如 FLAG或血球凝集素)。融合蛋白(及編碼其之核酸)經設計以使得轉譯閱讀框架保存於融合組分中。

設計及建構融合蛋白(及編碼其之聚核苷酸)之方法為熟習此項技術者所已知。舉例而言，設計及建構包含鋅指蛋白之融合蛋白(及編碼其之聚核苷酸)之方法描述於共同擁有之美國專利6,453,242及6,534,261中。

編碼該等融合蛋白之聚核苷酸亦在本揭示案之範疇內。該等聚核苷酸可使用標準技術來建構且插入至載體中，且可將該載體引入細胞中(參見下文關於將聚核苷酸引入細胞中之載體及方法之額外揭示案)。

用於與ZFP DNA結合域融合，欲用於抑制基因表現之例示性功能域為來自人類KOX-1蛋白質之KRAB抑制域(參見(例如)，Thiesen等人，New Biologist 2, 363-374 (1990)；Margolin等人，Proc. Natl. Acad. Sci. USA 91, 4509-4513 (1994)；Pengue等人，Nucl. Acids Res. 22:2908-2914 (1994)；Witzgall等人，Proc. Natl. Acad. Sci. USA 91, 4514-4518 (1994))。KOX域亦適於用作抑制域。另一適合抑制域為甲基結合域蛋白質2B(MBD-2B)(亦參見Hendrich等人(1999) *Mamm Genome* 10:906-912得到MBD蛋白質之

描述)。另一適用抑制域為與 v-ErbA 蛋白質相關者。參見(例如), Damm, 等人(1989) *Nature* 339:593-597; Evans (1989) *Int. J. Cancer Suppl.* 4:26-28; Pain 等人(1990) *New Biol.* 2:284-294; Sap 等人(1989) *Nature* 340:242-244; Zenke 等人(1988) *Cell* 52:107-119; 及 Zenke 等人(1990) *Cell* 61:1035-1049。其他例示性抑制域包括(但不限於)甲狀腺激素受體(TR)、SID、MBD1、MBD2、MBD3、MBD4、MBD樣蛋白質、DNMT家族之成員(例如DNMT1、DNMT3A、DNMT3B)、Rb、MeCP1及MeCP2。參見(例如), Zhang 等人(2000) *Ann Rev Physiol* 62:439-466; Bird 等人(1999) *Cell* 99:451-454; Tyler 等人(1999) *Cell* 99:443-446; Knoepfler 等人(1999) *Cell* 99:447-450; 及 Robertson 等人(2000) *Nature Genet.* 25:338-342。其他例示性抑制域包括(但不限於)ROM2及AtHD2A。參見(例如), Chern 等人(1996) *Plant Cell* 8:305-321; 及 Wu 等人(2000) *Plant J.* 22:19-27。

用於實現活化之適合域包括 HSV VP16 活化域(參見(例如), Hagmann 等人, *J. Virol.* 71, 5952-5962 (1997))核激素受體(參見(例如), Torchia 等人, *Curr. Opin. Cell. Biol.* 10:373-383 (1998)); 核因子 κ B 之 p65 子單元 (Bitko & Barik, *J. Virol.* 72:5610-5618 (1998) 及 Doyle & Hunt, *Neuroreport* 8:2937-2942 (1997); Liu 等人, *Cancer Gene Ther.* 5:3-28 (1998)), 或人工嵌合功能域, 諸如 VP64 (Seifpal 等人, *EMBO J.* 11, 4961-4968 (1992))。

其他例示性活化域包括(但不限於)p300、CBP、PCAF、SRC1、PvALF及ERF-2。參見(例如), Robyr等人(2000) *Mol. Endocrinol.* **14**:329-347; Collingwood等人(1999) *J. Mol. Endocrinol.* **23**:255-275; Leo等人(2000) *Gene* **245**:1-11; Manteuffel-Cymborowska (1999) *Acta Biochim. Pol.* **46**:77-89; McKenna等人(1999) *J. Steroid Biochem. Mol. Biol.* **69**:3-12; Malik等人(2000) *Trends Biochem. Sci.* **25**:277-283; 及Lemon等人(1999) *Curr. Opin. Genet. Dev.* **9**:499-504。其他例示性活化域包括(但不限於)OsGAI、HALF-1、C1、AP1、ARF-5、ARF-6、ARF-7及ARF-8、CPRF1、CPRF4、MYC-RP/GP及TRAB1。參見(例如), Ogawa等人(2000) *Gene* **245**:21-29; Okanami等人(1996) *Genes Cells* **1**:87-99; Goff等人(1991) *Genes Dev.* **5**:298-309; Cho等人(1999) *Plant Mol. Biol.* **40**:419-429; Ulmason等人(1999) *Proc. Natl. Acad. Sci. USA* **96**:5844-5849; Sprenger-Haussels等人(2000) *Plant J.* **22**:1-8; Gong等人(1999) *Plant Mol. Biol.* **41**:33-44; 及Hobo等人(1999) *Proc. Natl. Acad. Sci. USA* **96**:15,348-15,353。

其他功能域揭示於(例如)共同擁有之美國專利第6,933,113號中。另外,適用於融合分子中之隔離子域、諸如含ISWI之域之染色質重塑蛋白質及甲基結合域蛋白質描述於(例如)共同擁有之國際公開案WO 01/83793及WO 02/26960中。

在其他實施例中,融合蛋白為包含一或多個如本文中所

述之 CCHC 鋅指及裂解域(或裂解半域)之鋅指核酸酶(ZFN)。鋅指可經工程化以識別任何所選基因組區域中之標靶序列，且當引入細胞中時，將產生融合蛋白與其結合位點之結合及在該基因組區域中或附近之裂解。該裂解可造成在非同源末端聯接後，基因組區域之核苷酸序列產生變異(例如突變)。或者，若含有與基因組區域同源之序列之外源性聚核苷酸亦存在於該細胞中，則在由 ZFN 靶向裂解後，在基因組區域與外源性聚核苷酸之間以高速率發生同源重組。同源重組可造成外源性序列之靶向序列置換或靶向整合，其視外源性聚核苷酸之核苷酸序列而定。

本文中描述之非標準鋅指在併入至 ZFN 中時，提供改良之裂解功能。如實例中所述，含有至少一個如本文中所述之 CCHC 指之 4 指 ZFN 至少與只含有 CCHH 指之核酸酶一樣適當裂解。在某些實施例中，當 C 末端指包含非標準 CCHC 鋅指時，第三與第四鋅配位殘基之間(亦即 C 末端 His 與 Cys 殘基之間)的殘基不同於存在於標準 CCHH 鋅指中之彼等殘基，且一或多個甘胺酸殘基(例如 1、2、3、4、5 個或 5 個以上)在 C 末端 Cys 殘基後插入。

本文中揭示之 ZFN 之裂解域部分可自任何核酸內切酶或核酸外切酶獲得。裂解域所來源之例示性核酸內切酶包括(但不限於)限制性核酸內切酶及歸巢核酸內切酶(homing endonuclease)。參見(例如)，2002-2003 Catalogue, New England Biolabs, Beverly, MA；及 Belfort 等人(1997) *Nucleic Acids Res.* **25**:3379-3388。裂解 DNA 之其他酶為已

知的(例如 S1 核酸酶；綠豆核酸酶；胰腺 DNase I；小球菌核酸酶；酵母 HO 核酸內切酶；亦參見 Linn 等人(編) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993)。該等酶中之一或多者(或其功能性片段)可用作裂解域及裂解半域之來源。

類似地，裂解半域可來源於如上文所陳述之任何核酸酶或其部分，其限制條件為裂解半域需要二聚作用以實現裂解活性。一般而言，若融合蛋白包含裂解半域，則基因組 DNA 之靶向裂解需要兩種融合蛋白。或者，可使用包含 2 個裂解半域之單個蛋白質。2 個裂解半域可來源於相同核酸內切酶，或各裂解半域可來源於不同核酸內切酶。另外，兩種融合蛋白之標靶位點彼此相關安置，以使得兩種融合蛋白與其各別標靶位點之結合將裂解半域以彼此允許裂解半域(例如)藉由二聚化形成功能性裂解域之空間定向置放。因此，在某些實施例中，標靶位點之附近邊界由 5-8 個核苷酸對或 15-18 個核苷酸對隔開。在其他實施例中，標靶位點彼此在 10 個核苷酸對內。然而，可在 2 個標靶位點之間插入任何整數之核苷酸或核苷酸對(例如 2 至 50 個核苷酸或 50 個以上)。一般而言，裂解點位於標靶位點之間。

限制性核酸內切酶(限制酶)存在於許多物種中且能夠與 DNA(在識別位點)序列特異性結合，且在結合位點處或附近裂解 DNA。某些限制酶(例如 IIS 型)在遠離識別位點之位點處裂解 DNA 且具有可分離之結合及裂解域。舉例而言，

IIS型酶 *Fok* I 在一條鏈上離其識別位點9個核苷酸處及另一條鏈上離其識別位點13個核苷酸處催化DNA雙股裂解。參見(例如)，美國專利5,356,802；5,436,150及5,487,994；以及Li等人(1992) *Proc. Natl. Acad. Sci. USA* **89**:4275-4279；Li等人(1993) *Proc. Natl. Acad. Sci. USA* **90**:2764-2768；Kim等人(1994a) *Proc. Natl. Acad. Sci. USA* **91**:883-887；Kim等人(1994b) *J. Biol. Chem.* **269**:31,978-31,982。因此，在一實施例中，融合蛋白包含來自至少一種IIS型限制酶之裂解域(或裂解半域)及一或多個可或可不經工程化之鋅指結合域。

其裂解域可與結合域分離之例示性IIS型限制酶為 *Fok* I。該特定酶作為二聚體具有活性。Bitinaite等人(1998) *Proc. Natl. Acad. Sci. USA* **95**: 10,570-10,575。因此，對本揭示案之目的而言，所揭示融合蛋白中所用之 *Fok* I 酶部分視為裂解半域。因此，對使用包含鋅指-*Fok* I 融合體之ZFN來靶向雙股裂解及/或靶向置換細胞序列而言，可使用兩種各自包含 *Fok* I 裂解半域之融合蛋白來重構催化活性裂解域。或者，亦可使用含有鋅指結合域及2個 *Fok* I 裂解半域之單個多肽分子。使用鋅指-*Fok* I 融合體之靶向裂解及靶向序列變異之參數提供於本揭示案中之其它地方及(例如)美國專利申請公開案第2005/0064474號中；該專利之揭示內容以引用之方式併入本文中。

在其他實施例中，*Fok*I裂解半域可包括一或多個在影響二聚作用之任何胺基酸殘基處之突變。該等突變可適用於

防止 ZFP/*FokI* 融合體對中之一者經歷可導致在非所要序列處裂解之均二聚作用。舉例而言，在 *FokI* 之位置 446、447、479、483、484、486、487、490、491、496、498、499、500、531、534、537 及 538 處之胺基酸殘基均足夠接近影響二聚作用之二聚介面。因此，在上述位置中之一或多者處之胺基酸序列變異可用以改變裂解半域之二聚性質。該等改變可(例如)藉由建構含有(或編碼)該等位置處之不同胺基酸殘基之文庫且選擇具有所要性質之變體，或藉由合理設計個別突變體來引入。除防止均二聚作用外，與用 2 個野生型裂解半域獲得者比較，一些該等突變亦可能增加裂解效率。

因此，對使用 ZFP/*FokI* 融合體對之靶向裂解而言，融合蛋白中之一或兩者可包含一或多個胺基酸變異，其抑制自身二聚作用，但允許兩種融合蛋白之雜二聚作用發生，以使得裂解在所要標靶位點處發生。在某些實施例中，在兩種融合蛋白中均存在變異，且該等變異具有相加效應；亦即，與用野生型裂解半域獲得者相比，引起異常裂解之融合體之均二聚作用被最小化或消除，而兩種融合蛋白之雜二聚作用得以促進。

在某些實施例中，裂解域包含 2 個裂解半域，兩者均為包含結合域、第一裂解半域及第二裂解半域之單一多肽之部分。裂解半域可具有相同胺基酸序列或不同胺基酸序列，只要其起裂解 DNA 之作用即可。

裂解半域亦可提供於分離分子中。舉例而言，可在細胞

中表現兩種融合多肽，其中各多肽包含結合域及裂解半域。裂解半域可具有相同胺基酸序列或不同胺基酸序列，只要其起裂解DNA之作用即可。另外，結合域與通常以使得在融合多肽結合後，2個裂解半域以彼此允許裂解域重構(例如藉由半域之二聚作用)的空間定向提供，由此使半域相對於彼此定位以形成功能性裂解域，使得細胞染色質在所關注區域中裂解之方式安置的標靶序列結合。通常，重構裂解域導致之裂解發生在位於2個標靶序列之間的位點處。該等蛋白質中之一或兩者可經工程化以與其標靶位點結合。

兩種融合蛋白在細胞中之表現可經由將兩種蛋白質傳遞至細胞中；將一種蛋白質及一種編碼該等蛋白質中之一者之核酸傳遞至細胞中；將兩種各編碼該等蛋白質中之一者之核酸傳遞至細胞中；或藉由將編碼兩種蛋白質之單一核酸傳遞至細胞中來產生。在其他實施例中，融合蛋白包含單一多肽鏈，其包含2個裂解半域及鋅指結合域。在該狀況下，在細胞中表現單一融合蛋白，且不希望受理論約束，咸信該單一融合蛋白由於形成裂解半域之分子內二聚體而裂解DNA。

在某些實施例中，ZFN之組分經排列以使得鋅指域最接近融合蛋白之胺基末端，且裂解半域最接近羧基末端。此反映裂解域在諸如來源於*Fok I*酶之天然存在二聚裂解域中的相對定向，其中DNA結合域最接近胺基末端，且裂解半域最接近羧基末端。在該等實施例中，使裂解半域形成功

能性核酸酶之二聚作用係藉由融合蛋白與相對DNA鏈上之位點結合來產生，其中結合位點之5'末端彼此鄰近。

在該定向中，最靠C末端之鋅指鄰近於*FokI*裂解半域。先前已確定，當CCHC型鋅指作為最靠C末端之指存在時，非標準鋅指蛋白最有效地結合其DNA標靶。因此有可能先前所述之CCHC型鋅指接近*FokI*裂解半域之存在抑制其功能。若為該狀況，則當前揭示之最佳化CCHC鋅指顯然不顯示該假定之抑制活性。

在其他實施例中，融合蛋白(例如ZFP-*Fok I*融合體)之組分經排列以使得裂解半域最接近融合蛋白之胺基末端，且鋅指域最接近羧基末端。在該等實施例中，使裂解半域形成功能性核酸酶之二聚作用係藉由融合蛋白與相對DNA鏈上之位點結合來產生，其中結合位點之3'末端彼此鄰近。

在其他實施例中，第一融合蛋白含有最接近該融合蛋白之胺基末端之裂解半域，及最接近羧基末端之鋅指域，且第二融合蛋白經排列以使得鋅指域最接近融合蛋白之胺基末端，且裂解半域最接近羧基末端。在該等實施例中，兩種融合蛋白結合於同一DNA鏈，其中含有最接近羧基末端之鋅指域之第一融合蛋白的結合位點位於含有最接近胺基末端之鋅指域之第二融合蛋白之結合位點的5'側。亦參見WO 2005/084190；其揭示內容以引用之方式併入本文中。

鋅指域與裂解域(或裂解半域)之間的胺基酸序列表示"ZC連接子"。ZC連接子應區別於上文所討論之指間連接子。關於獲得最佳化裂解之ZC連接子之細節，請參見(例

如)，美國專利公開案 20050064474A1 及 20030232410 及國際專利公開案 WO 2005/084190；其揭示內容以引用之方式併入本文中。

表現載體

編碼一或多種 ZFP 或 ZFP 融合蛋白(例如 ZFN)之核酸可選殖至用於轉型至供複製及/或表現之原核或真核細胞中之載體中。載體可為原核或真核載體，包括(但不限於)質體、穿梭載體、昆蟲載體、二元載體(參見(例如)，美國專利第 4,940,838 號；Horsch 等人(1984) *Science* 233:496-498，及 Fraley 等人(1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803)及其類似物。編碼 ZFP 之核酸亦可選殖至供投與至植物細胞中之表現載體中。

為表現融合蛋白，通常將編碼 ZFP 或 ZFP 融合體之序列次選殖至含有指導轉錄之啟動子之表現載體中。適合之細菌及真核啟動子為此項技術中所熟知且描述於(例如)Sambrook 等人，*Molecular Cloning, A Laboratory Manual* (第 2 版，1989 年；第 3 版，2001 年)；Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990)；及 *Current Protocols in Molecular Biology*(Ausubel 等人，同上)中。用於表現 ZFP 之細菌表現系統可在(例如)大腸桿菌 (*E. coli*)、桿菌屬 (*Bacillus sp.*)及沙門氏菌 (*Salmonella*)中獲得(Palva 等人，*Gene* 22:229-235 (1983))。該等表現系統之套組可購得。用於哺乳動物細胞、酵母及昆蟲細胞之真核表現系統為熟習此項技術者所熟知且亦可購得。

用以指導編碼ZFP之核酸表現之啟動子視特定應用而定。舉例而言，對於ZFP之表現及純化通常使用適於宿主細胞之強組成性啟動子。

相反，當活體內投與ZFP以供植物基因調節時(參見，下文之"向植物細胞中傳遞核酸"部分)，根據ZFP之特定用途使用組成性或可誘導啟動子。植物啟動子之非限制性實例包括來源於擬南芥(*A. thaliana*)泛素-3(ubi-3)(Callis, 等人, 1990, J. Biol. Chem. 265-12486-12493); 根癌土壤桿菌(*A. tumefaciens*)甘露鹼合成酶(Δ mas)(Petolino等人, 美國專利第6,730,824號); 及/或木薯葉脈嵌紋病毒(Cassava Vein Mosaic Virus, CsVMV)(Verdaguer等人, 1996, Plant Molecular Biology 31:1129-1139)之啟動子序列。亦參見實例。

除啟動子外，表現載體通常含有包含在原核或真核宿主細胞中表現核酸所需之所有其他元件的轉錄單元或表現卡匣。因此，典型表現卡匣含有可操作連接於(例如)編碼ZFP之核酸序列之啟動子，及(例如)使轉錄物有效多聚腺嘌呤化、轉錄終止、核糖體結合位點或轉譯終止所需之信號。卡匣之其他元件可包括(例如)強化子及異源剪接信號。

用以將遺傳資訊傳輸至細胞中之特定表現載體係關於ZFP之所欲用途，例如在植物、動物、細菌、真菌、原生動物等中之表現來選擇(參見下文所述之表現載體)。標準細菌及動物表現載體為此項技術中所已知且詳細描述於

(例如)美國專利公開案20050064474A1及國際專利公開案WO 05/084190、WO 05/014791及WO 03/080809中；其揭示內容以引用之方式併入本文中。

標準轉染方法可用以產生表現大量可隨後使用標準技術純化之蛋白質之細菌、哺乳動物、酵母或昆蟲細胞系(參見(例如)，Colley等人，*J. Biol. Chem.* 264:17619-17622 (1989)；*Guide to Protein Purification, in Methods in Enzymology*，第182卷(Deutscher，編，1990))。真核及原核細胞之轉型係根據標準技術來執行(參見(例如)，Morrison，*J. Bact.* 132:349-351 (1977)；Clark-Curtiss & Curtiss，*Methods in Enzymology* 101:347-362(Wu等人，編，1983)。

可使用用於將外來核苷酸序列引入該等宿主細胞中之任何熟知程序。該等程序包括使用磷酸鈣轉染、凝聚胺(polybrene)、原生質體融合、電穿孔、超聲波法(例如聲破法)、脂質體、顯微注射、裸露DNA、質體載體、病毒載體、游離基因及整合基因，及用於將經選殖基因組DNA、cDNA、合成DNA或其它外來遺傳物質引入宿主細胞中之任何其它熟知方法(參見(例如)，Sambrook等人，同上)。唯一必要的為，所用特定遺傳工程化程序能夠成功將至少一種基因引入能夠表現所選蛋白質之宿主細胞中。

向植物細胞中傳遞核酸

如上所述，DNA構築體可藉由多種習知技術引入所要植物宿主中(例如，其基因組中)。關於該等技術之回顧，請

參見(例如), Weissbach & Weissbach *Methods for Plant Molecular Biology*(1988, Academic Press, N.Y.)第VIII節, 第421-463頁; 及 Grierson & Corey, *Plant Molecular Biology* (1988, 第2版), Blackie, London, 第7-9章。

舉例而言, DNA構築體可使用諸如電穿孔及植物細胞原生質體之顯微注射之技術引入植物細胞中, 或DNA構築體可使用基因槍法, 諸如DNA粒子轟擊而直接引入植物組織中(參見(例如), Klein等人(1987) *Nature* 327:70-73)。或者, DNA構築體可與適合T-DNA側接區域組合且引入習知根癌土壤桿菌(*Agrobacterium tumefaciens*)宿主載體中。包括卸甲(disarming)及使用二元載體之根癌土壤桿菌介導轉型技術充分描述於科學文獻中。參見(例如)Horsch等人(1984) *Science* 233:496-498, 及 Fraley等人(1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803。

另外, 基因轉移可使用諸如以下各物之非土壤桿菌屬細菌或病毒來實現: 根瘤菌(*Rhizobium sp.*)NGR234、苜蓿中華根瘤菌(*Sinorhizobium meliloti*)、百脈根中生根瘤菌(*Mesorhizobium loti*)、馬鈴薯病毒X、花椰菜嵌紋病毒及木薯葉脈嵌紋病毒及/或煙草嵌紋病毒, 參見(例如), Chung等人(2006) *Trends Plant Sci.* 11(1):1-4。

當細胞藉由使用二元T DNA載體之細菌(Bevan (1984) *Nuc. Acid Res.* 12:8711-8721)或共同培養程序(Horsch等人(1985) *Science* 227:1229-1231)感染時, 根癌土壤桿菌宿主之毒性功能將指導構築體及相鄰標記物插入至植物細胞

DNA中。通常，使用土壤桿菌轉型系統工程化雙子葉植物 (Bevan等人(1982) *Ann. Rev. Genet.* 16:357-384；Rogers等人(1986) *Methods Enzymol.* 118:627-641)。亦可使用土壤桿菌轉型系統將DNA轉型以及轉移至單子葉植物及植物細胞中。參見美國專利第 5,591,616 號；Hernalsteen等人(1984) *EMBO J* 3:3039-3041；Hooykass-Van Slogteren等人(1984) *Nature* 311:763-764；Grimsley等人(1987) *Nature* 325:1677-179；Boulton等人(1989) *Plant Mol. Biol.* 12:31-40.；及Gould等人(1991) *Plant Physiol.* 95:426-434。

替代性基因轉移及轉型方法包括(但不限於)經由裸露DNA之鈣、聚乙二醇(PEG)或電穿孔介導之吸收的原生質體轉型(參見Paszkowski等人(1984) *EMBO J* 3:2717-2722，Potrykus等人(1985) *Molec. Gen. Genet.* 199:169-177；Fromm等人(1985) *Proc. Nat. Acad. Sci. USA* 82:5824-5828；及Shimamoto (1989) *Nature* 338:274-276)及植物組織之電穿孔(D'Halluin等人(1992) *Plant Cell* 4:1495-1505)。用於植物細胞轉型之其他方法包括顯微注射、碳化矽介導之DNA吸收(Kaeppler等人(1990) *Plant Cell Reporter* 9:415-418)及微彈轟擊(參見Klein等人(1988) *Proc. Nat. Acad. Sci. USA* 85:4305-4309；及Gordon-Kamm等人(1990) *Plant Cell* 2:603-618)。

所揭示之方法及組合物可用以將外源性序列插入植物細胞基因組之預定位置中。其為適用的，此係由於引入至植物基因組中之轉殖基因之表現嚴重依賴於其整合位點。因

此，編碼(例如)營養、抗生素或治療分子之基因可藉由靶向重組而插入有利於其表現之植物基因組區域中。

藉由任何上述轉型技術產生之經轉型植物細胞可經培養以再生具有經轉型基因型及因此具有所要表型的完整植物。該等再生技術依賴於對組織培養生長培養基中某些植物激素之操作，通常依賴於已連同所要核苷酸序列一起引入之殺生物劑及/或除草劑標記物。由所培養原生質體再生植物描述於Evans，等人，*Handbook of Plant Cell Culture*中之"Protoplasts Isolation and Culture"，第124-176頁，Macmillian Publishing Company, New York, 1983；及Binding, *Regeneration of Plants, Plant Protoplasts*，第21-73頁，CRC Press, Boca Raton, 1985。再生亦可自植物愈傷組織、外植體、器官、花粉、胚胎或其部分獲得。該等再生技術通常描述於Klee等人(1987) *Ann. Rev. of Plant Phys.* 38:467-486中。

引入植物細胞中之核酸可用以對基本上任何植物賦予所要特性。多種植物及植物細胞系統可就本文中所述之所要生理及農藝特徵使用本揭示案之核酸構築體及上文提及之各種轉型方法來工程化。在某些實施例中，用於工程化之標靶植物及植物細胞包括(但不限於)彼等單子葉及雙子葉植物，諸如包括以下各物之作物：穀類作物(例如小麥、玉米、水稻、粟、大麥)，果實作物(例如番茄、蘋果、梨、草莓、橘)，飼料作物(例如紫苜蓿)，根菜作物(例如胡蘿蔔、馬鈴薯、甜菜、山藥)，葉菜作物(例如萵苣、菠

菜)；開花植物(例如矮牽牛、玫瑰、菊花)，針葉樹及松樹(例如松樹杉、雲杉)；用於植物修復之植物(例如重金屬積聚植物)；油料作物(例如向日葵、菜籽)及用於實驗目的之植物(例如芥菜屬)。因此，所揭示之方法及組合物在廣泛範圍之之植物中具有用途，該等植物包括(但不限於)來自下列屬之物種：蘆筍屬(Asparagus)、燕麥屬(Avena)、甘藍屬(Brassica)、橘屬(Citrus)、西瓜屬(Citrullus)、辣椒屬(Capsicum)、南瓜屬(Cucurbita)、胡蘿蔔屬(Daucus)、大豆屬(Glycine)、棉屬(Gossypium)、大麥屬(Hordeum)、萵苣屬(Lactuca)、番茄屬(Lycopersicon)、蘋果屬(Malus)、木薯屬(Manihot)、煙草屬(Nicotiana)、稻屬(Oryza)、酪梨屬(Persea)、豌豆屬(Pisum)、梨屬(Pyrus)、李屬(Prunus)、蘿蔔屬(Raphanus)、黑麥屬(Secale)、茄屬(Solanum)、高粱屬(Sorghum)、小麥屬(Triticum)、葡萄屬(Vitis)、豇豆屬(Vigna)及玉蜀黍屬(Zea)。

熟習此項技術者應認識到，在表現卡匣穩定併入轉殖基因植物中且確認為可操作之後，其可藉由有性雜交引入至其它植物中。可根據欲雜交之物種而使用許多標準育種技術中之任何技術。

經轉型植物細胞、愈傷組織、組織或植物可藉由針對存在於轉型DNA上之標記基因所編碼之特性選擇或篩選工程化植物材料來鑑別及分離。舉例而言，選擇可藉由使經工程化植物材料在含有抑制量之抗生素或除草劑之培養基上生長來執行，其中轉型基因構築體賦予抵抗該抗生素或除

草劑之抗性。另外，經轉型植物及植物細胞亦可藉由篩選可存在於重組核酸構築體上之任何可見標記基因(例如 β -葡糖醛酸酶、螢光素酶、B或C1基因)之活性來鑑別。該等選擇及篩選方法為熟習此項技術者所熟知。

亦可使用物理及生物化學方法鑑別含有插入基因構築體之植物或植物細胞轉型體。該等方法包括(但不限於)：1)用於偵測及測定重組DNA插入物之結構之南方分析或PCR擴增；2)用於偵測及檢查基因構築體之RNA轉錄物之北方墨點、S1 RNase保護、引子延伸或逆轉錄酶PCR擴增；3)用於偵測酶或核糖核酸酶活性之酶檢定，其中該等基因產物係由基因構築體編碼；4)蛋白質凝膠電泳、西方墨點技術、免疫沈澱或酶連接免疫檢定，其中基因構築體產物為蛋白質。諸如原位雜交、酶染色及免疫染色之其他技術亦可用以偵測重組構築體在特定植物器官及組織中之存在或表現。用於進行所有該等檢定之方法為熟習此項技術者所熟知。

使用本文中所揭示之方法之基因操作的效應可藉由(例如)自所關注組織分離之RNA(例如mRNA)之北方墨點來觀測。通常，若mRNA之量增加，則可假定相應內源基因正以比先前高之速率表現。可使用量測基因及/或CYP74B活性之其它方法。可根據所用底物及偵測反應產物或副產物增加或降低之方法使用不同類型之酶檢定。另外，所表現之基因及/或CYP74B蛋白質之量可以免疫化學方法量測，亦即ELISA、RIA、EIA及熟習此項技術者熟知之其他基於

抗體之檢定，諸如藉由電泳偵測檢定(連同染色或西方墨點)來量測。轉殖基因可選擇性表現於植物之一些組織中或在一些發育階段選擇性表現，或轉殖基因可在大體上所有植物組織中表現，大體上在其整個生命週期中表現。然而，任何組合表現模式亦可適用。

本揭示案亦涵蓋上文所述之轉殖基因植物之種子，其中該種子具有轉殖基因或基因構築體。本揭示案另外涵蓋上文所述之轉殖基因植物之子代、純系、細胞系或細胞，其中該子代、純系、細胞系或細胞具有該轉殖基因或構築體。

ZFP及編碼ZFP之表現載體可直接投與至植物以供靶向裂解及/或重組。

投與有效量係藉由通常用於引入ZFP以與欲處理植物細胞最後接觸之任何途徑來實現。ZFP以任何適合方式投與。投與該等組合物之適合方法可獲得且為熟習此項技術者所熟知，且儘管可使用一種以上途徑投與特定組合物，但特定途徑通常可提供比另一途徑更直接且更有效之反應。

亦可使用載劑且其部分由所投與之特定組合物以及用以投與組合物之特定方法確定。因此，存在多種可用醫藥組合物之適合調配物(參見(例如)，*Remington's Pharmaceutical Sciences*，第17版1985)。

應用

包含一或多個如本文中所述之非標準鋅指之鋅指蛋白適

用於當前使用標準C2H2 ZFP之所有基因組調節及編輯應用，其包括(但不限於)：基因活化；基因抑制；基因組編輯(裂解、靶向插入、置換或缺失)；及表觀基因組編輯(經由靶向組蛋白或DNA之共價修飾)。

包含如本文中所揭示之非標準鋅指之ZFN可用以在細胞染色質中所關注之區域處(例如，在基因組，例如在突變或野生型基因中之所要或預定位點處)裂解DNA。對該靶向DNA裂解而言，鋅指結合域經工程化以在預定裂解位點處或預定裂解位點附近結合標靶位點，且在細胞中表現包含經工程化鋅指結合域及裂解域之融合蛋白。在融合蛋白之鋅指部分結合於標靶位點後，DNA藉由裂解域在標靶位點附近裂解。裂解之確切位點可視ZC連接子之長度而定。

或者，兩種各自包含鋅指結合域及裂解半域之ZFN於細胞中表現，且結合於以使得功能性裂解域得以重構且DNA在標靶位點附近裂解之方式鄰接的標靶位點。在一實施例中，裂解發生在2個鋅指結合域之標靶位點之間。鋅指結合域中之一或兩者可經工程化。

對使用鋅指結合域-裂解域融合多肽之靶向裂解而言，結合位點可涵蓋裂解位點，或接近結合位點之邊界可與裂解位點相隔1、2、3、4、5、6、10、25、50個或50個以上核苷酸(或1與50之間的任何整數值之核苷酸)。結合位點相對於裂解位點之確切位置將視特定裂解域及ZC連接子之長度而定。對其中使用兩種各自包含鋅指結合域及裂解半域之融合多肽之方法而言，結合位點通常跨過裂解位點。因

此，接近第一結合位點之邊界可與裂解位點之一側相隔1、2、3、4、5、6、10、25個或25個以上核苷酸(或1與50之間的任何整數值之核苷酸)，且接近第二結合位點之邊界可與裂解位點之另一側相隔1、2、3、4、5、6、10、25個或25個以上核苷酸(或1與50之間的任何整數值之核苷酸)。用於活體外及活體內定位裂解位點之方法為熟習此項技術者所已知。

一旦引入標靶細胞中或在標靶細胞中表現後，融合蛋白即結合於標靶序列且在標靶序列處或標靶序列附近裂解。裂解之確切位點視裂解域之性質及/或結合域與裂解域之間的連接子序列之存在及/或性質而定。在其中使用兩種各自包含裂解半域之ZFN之狀況下，接近結合位點之邊界之間的距離可為1、2、3、4、5、6、7、8、9、10、25個或25個以上核苷酸(或1與50之間的任何整數值之核苷酸)。裂解之最佳程度亦可視兩種ZFN之結合位點之間的距離(參見(例如)，Smith等人(2000) *Nucleic Acids Res.* **28**:3361-3369；Bibikova等人(2001) *Mol. Cell. Biol.* **21**:289-297)及各ZFN中ZC連接子之長度而定。亦參見，美國專利公開案20050064474A1及國際專利公開案WO 05/084190、WO 05/014791及WO 03/080809，其揭示內容以引用之方式併入本文中。

兩種各自包含裂解半域之ZFN可以相同或相反極性結合於所關注之區域中，且其結合位點(亦即標靶位點)可隔開任何數量之核苷酸，例如0至50個核苷酸對或其間之任何

整數值。在某些實施例中，兩種各自包含鋅指結合域及裂解半域之融合蛋白之結合位點，可如自最接近另一結合位點之各結合位點之邊界所量測，其間相隔5個與18個之間的核苷酸對定位，例如相隔5-8個核苷酸對，或相隔15-18個核苷酸對，或相隔6個核苷酸對，或相隔16個核苷酸對，或彼此相隔10個核苷酸對以內，且裂解發生在結合位點之間。

裂解DNA之位點通常位於兩種融合蛋白之結合位點之間。DNA之雙股斷裂常常由2個單鏈斷裂或偏離1、2、3、4、5、6個或6個以上核苷酸之"缺口"產生，(例如，由天然 *Fok* I對雙股DNA之裂解由偏離4個核苷酸之單鏈斷裂產生)。因此，裂解不必發生在各DNA鏈之精確相對位點處。另外，融合蛋白之結構及標靶位點之間的距離可影響裂解是否在單核苷酸對附近發生，或裂解是否在若干位點處發生。然而，對包括靶向重組及靶向突變誘發之許多應用而言，在一定範圍核苷酸內之裂解通常已足夠，且不需要在特定鹼基對之間的裂解。

如上所述，融合蛋白可在將多肽及/或聚核苷酸引入細胞中後表現於細胞中。舉例而言，可將兩種各自包含編碼上述多肽中之一者之序列的聚核苷酸引入細胞中，且當表現多肽且各多肽結合於其標靶序列時，裂解在標靶序列處或標靶序列附近發生。或者，將包含編碼兩種融合多肽之序列之單一聚核苷酸引入細胞中。聚核苷酸可為DNA、RNA或DNA及/或RNA之任何經修飾形式或類似物。

在某些實施例中，基因組區域中由ZFN實現之靶向裂解在由非同源末端聯接(NHEJ)修復裂解事件後，產生該區域核苷酸序列之變異。

在其他實施例中，基因組區域中由ZFN實現之靶向裂解亦可為使基因組序列(例如，細胞染色質中所關注之區域)經同源非相同序列(亦即，藉由靶向重組)經由同源依賴性機制(例如，連同一或多個與預定基因組序列(亦即標靶位點)相同或同源但非相同之序列插入包含外源性序列之供體序列)置換之程序的部分。因為在裂解位點附近細胞DNA中雙股斷裂激活細胞修復機制達數千倍，故利用如本文中所述之ZFN實現之靶向裂解在基因組中幾乎任何位點允許序列變異或置換(經由同源引導修復)。

除本文中所述之ZFN外，所選基因組序列之靶向置換亦需要引入外源性(供體)聚核苷酸。供體聚核苷酸可在表現ZFN之前、與表現ZFN同時或在表現ZFN之後引入細胞中。供體聚核苷酸含有足夠的與基因組序列之同源性，以支持其與其具有同源性之基因組序列之間的同源重組(或同源引導修復)。約25、50、100、200、500、750、1,000、1,500、2,000個核苷酸或2,000個以上核苷酸之序列同源性(或10與2,000個之間的任何整數值或2,000個以上之核苷酸)將支持同源重組。供體聚核苷酸長度可在10至5,000個核苷酸(或其間之任何整數值之核苷酸)或更長範圍內。

顯而易見地，供體聚核苷酸之核苷酸序列通常與其所置

換之基因組序列之核苷酸序列不同。舉例而言，與基因組序列相比，供體聚核苷酸序列可含有一或多個取代、插入、缺失、倒置或重排，只要存在足夠與染色體序列之同源性即可。該等序列改變可為任何大小，且可少至單個核苷酸對。或者，供體聚核苷酸可含有側接有2個同源性區域之非同源序列(亦即，區別於外源性聚核苷酸之外源性序列)。另外，供體聚核苷酸可包含含有與細胞染色質中所關注區域之序列不同源之序列的載體分子。一般而言，供體聚核苷酸之同源區域與需要重組之基因組序列具有至少50%之序列一致性。在某些實施例中，存在60%、70%、80%、90%、95%、98%、99%或99.9%之序列一致性。視供體聚核苷酸之長度而定，可存在1%與100%之間的任何值之序列一致性。

供體分子可含有與細胞染色質具有同源性之若干不連續區域。舉例而言，對通常不存在於所關注區域中之序列之靶向插入而言，該等序列可存在於供體核酸分子中，且側接有與所關注區域中之序列具有同源性的區域。

為簡化用於確定來自供體聚核苷酸之序列成功插入的檢定(例如雜交、PCR、限制酶切割)，某些與基因組序列有差異之序列可存在於供體序列中。較佳地，若位於編碼區域中，則該等核苷酸序列差異將不改變胺基酸序列，或將產生絨默型胺基酸改變(亦即，不影響蛋白質之結構或功能之改變)。供體聚核苷酸可視需要含有對應於所關注區域中之鋅指域結合位點之序列的改變，以防止已藉由同源

重組引入細胞染色質中之供體序列裂解。

聚核苷酸可作為載體分子之部分而引入細胞中，該載體分子具有諸如複製起點、啟動子及編碼抗生素抗性之基因的額外序列。此外，供體聚核苷酸可以裸露核酸、與諸如脂質體或泊洛沙姆(poloxamer)之藥劑複合之核酸的形式來引入，或可由細菌或病毒(例如，土壤桿菌屬、根瘤菌NGR234、苜蓿中華根瘤菌、百脈根中生根瘤菌、煙草嵌紋病毒、馬鈴薯病毒X、花椰菜嵌紋病毒及木薯葉脈嵌紋病毒)傳遞。例如參見Chung等人(2006) *Trends Plant Sci.* 11(1):1-4。

對染色體序列之變異而言，並不須將供體之整個序列複製至染色體中，只要複製足夠之供體序列以實現所要序列變異即可。

藉由同源重組插入供體序列之效率與細胞DNA中雙股斷裂位置與需要重組之位點之間的距離反向相關。換言之，當雙股斷裂與需要重組之位點較近時，觀察到較高之同源重組效率。在未預先確定重組之精確位點之狀況下(例如，所要的重組可在基因組序列之整個間隔範圍內發生時)，連同裂解位點之供體核酸之長度及序列係經選擇以獲得所要的重組結果。在設計所要結果為改變基因組序列中單核苷酸對之序列的狀況下，細胞染色質在該核苷酸對之任一側10,000個核苷酸以內被裂解。在某些實施例中，裂解發生在其序列欲改變之核苷酸對之任一側1,000、500、200、100、90、80、70、60、50、40、30、20、

10、5或2個核苷酸以內，或2與1,000之間的任何整數值之核苷酸以內。

外源性序列靶向插入至基因組區域中係藉由在基因組區域中使用ZFN之靶向裂解來實現，遵循之限制條件為，外源性(供體)聚核苷酸含有外源性序列。供體聚核苷酸亦通常含有側接外源性序列之序列，其含有足夠與基因組區域之同源性以支持基因組序列中雙股斷裂之同源引導修復，進而將外源性序列插入至基因組區域中。因此，供體核酸可具有足以支持藉由同源依賴性修復機制(例如同源重組)整合外源性序列的任何大小。不希望受任何特定理論約束，認為側接外源性序列之具有同源性之區域為斷裂染色體末端提供模板，以供在雙股斷裂之位點處再合成遺傳資訊。

如上所述之外源性序列之靶向整合可用以在經選擇染色體位置處插入標記基因。標記基因包括(但不限於)編碼介導抗生素抗性(例如，安比西林(ampicillin)抗性、新黴素(neomycin)抗性、G418抗性、嘌呤黴素(puromycin)抗性)之蛋白質之序列，編碼有色或螢光或發光蛋白質(例如，綠色螢光蛋白質、增強綠色螢光蛋白質、紅色螢光蛋白質、螢光素酶)及介導增強之細胞生長及/或基因擴增之蛋白質(例如二氫葉酸還原酶)的序列。例示性標記基因因此包括(但不限於) β -葡糖醛酸酶(GUS)、草胺膦(phosphinothricin) N-乙醯基轉移酶(PAT, BAR)、新黴素磷酸轉移酶、 β -內醯胺酶、兒茶酚雙加氧酶、 α -澱粉酶、酪胺酸酶、 β -半乳糖

苷酶、螢光素酶、水母發光蛋白(aequorin)、EPSP合成酶、腓水解酶、乙醯乳酸合成酶(ALS)、二氫葉酸還原酶(DHFR)、茅草枯(dalapon)去鹵酶及鄰胺基苯甲酸合成酶。在某些實施例中，使用靶向整合來插入RNA表現構築體，例如負責微小RNA(micro RNA)或siRNA之調節表現之序列。啟動子、強化子及其他轉錄調節序列亦可併入RNA表現構築體中。

在包含鋅指/核酸酶融合分子及供體DNA分子之細胞中，靶向重組之效率之進一步增加係藉由將細胞阻滯於細胞週期之G₂期來實現，其時同源性驅動修復過程具有最大活性。該阻滯可以許多方式實現。舉例而言，可將細胞用(例如)影響細胞週期進程以便將細胞阻滯於G₂期之藥物、化合物及/或小分子來處理。該類型之例示性分子包括(但不限於)影響微管聚合之化合物(例如長春鹼、諾考達唑(nocodazole)、紫杉酚)，與DNA相互作用之化合物(例如順鉑(II)二胺二氯化物、順鉑、阿黴素(doxorubicin))及/或影響DNA合成之化合物(例如胸苷、羥基脲、L-含羞草鹼、依託泊苷(etoposide)、5-氟尿嘧啶)。重組效率之額外增加係藉由使用組蛋白脫乙醯基酶(HDAC)抑制劑(例如丁酸鈉、曲古菌素A(trichostatin A))來實現，該等抑制劑改變染色質結構以使基因組DNA更可由細胞重組機構所接近。

用於細胞週期阻滯之其他方法包括使抑制CDK細胞週期激酶之活性之蛋白質過度表現，例如，藉由將編碼該蛋白質之cDNA引入細胞中或藉由將活化編碼該蛋白質之基因

之表現的工程化ZFP引入細胞中。細胞週期阻滯亦係藉由(例如)使用RNAi方法(例如美國專利第6,506,559號)抑制細胞週期素及CDK之活性或藉由將抑制一或多種諸如細胞週期素及/或CDK基因之細胞週期進程中所涉及之基因的表現之工程化ZFP引入細胞中來實現。關於合成用於調節基因表現之工程化鋅指蛋白之方法，請參見(例如)共同擁有之美國專利第6,534,261號。

如上所述，所揭示用於靶向裂解之方法及組合物可用以誘導基因組序列之突變。靶向裂解亦可用以產生基因剔除(例如，用於功能性基因組或標靶驗證)及促進將序列靶向插入至基因組中(亦即基因敲入)。插入可藉助於染色體序列經由同源重組之置換，或藉由靶向整合，其中將側接有與染色體中所關注區域同源之序列的新序列(亦即不存在於所關注區域中之序列)插入預定標靶位點處。相同方法亦可用以將野生型序列用突變序列置換，或將一等位基因轉化成不同等位基因。

感染或整合植物病原體之靶向裂解可用以治療植物宿主中之病原體感染，例如，藉由裂解病原體之基因組以使得降低或消除其病原性。另外，靶向裂解編碼植物病毒受體之基因可用以阻斷該等受體之表現，由此防止植物中之病毒感染及/或病毒傳播。

例示性植物病原體包括(但不限於)植物病毒，諸如苜蓿花葉病毒(*Alfamoviruses*)、 α 潛隱病毒(*Alphacryptoviruses*)、桿狀DNA病毒(*Badnaviruses*)、 β 潛隱病毒

(*Betacryptoviruses*)、雙生病毒(*Bigeminiviruses*)、雀麥花葉病毒(*Bromoviruses*)、大麥花葉病毒(*Bymoviruses*)、毛細病毒(*Capilloviruses*)、香石竹潛病毒(*Carlaviruses*)、卡莫病毒(*Carmoviruses*)、花椰菜花葉病毒(*Caulimoviruses*)、線形病毒(*Closteroviruses*)、豇豆花葉病毒(*Comoviruses*)、黃瓜花葉病毒(*Cucumoviruses*)、草莓病毒(*Cytorhabdoviruses*)、香石竹環斑病毒(*Dianthoviruses*)、依那莫病毒(*Enamoviruses*)、蠶豆病毒(*Fabaviruses*)、斐濟病毒(*Fijiviruses*)、真菌傳棒狀病毒(*Furoviruses*)、大麥病毒(*Hordeiviruses*)、甜菜頂捲曲病毒(*Hybrigeminiviruses*)、埃達病毒(*Idaeoviruses*)、煙草線條病毒(*Ilarviruses*)、依普莫病毒(*Ipomoviruses*)、黃矮病毒(*Luteoviruses*)、玉米萎黃斑病毒(*Machlomoviruses*)、甘薯病毒(*Macluraviruses*)、瑪拉病毒(*Marafiviruses*)、單生病毒(*Monogeminiviruses*)、那納病毒(*Nanaviruses*)、壞死病毒(*Necroviruses*)、線蟲傳多面體病毒(*Nepoviruses*)、核桿狀病毒(*Nucleorhabdoviruses*)、稻病毒(*Oryzaviruses*)、歐爾密病毒(*Ourmiaviruses*)、植物呼腸弧病毒(*Phytoreoviruses*)、馬鈴薯X病毒(*Potexviruses*)、馬鈴薯Y病毒(*Potyvirus*)、瑞莫病毒(*Rymoviruses*)、衛星RNA、衛星病毒(*satelliviruses*)、塞奎病毒(*Sequiviruses*)、南方菜豆花葉病毒(*Sobemoviruses*)、細病毒(*Tenuiviruses*)、煙草花葉病毒(*Tobamoviruses*)、煙草脆裂病毒(*Tobraviruses*)、番茄叢矮病毒(*Tombusviruses*)、番茄斑萎

病毒 (*Tospoviruses*)、髮狀病毒 (*Trichoviruses*)、蕪菁黃花葉病毒 (*Tymoviruses*)、幽影病毒 (*Umbraviruses*)、腫脹病毒 (*Varicosaviruses*)及維卡病毒 (*Waikaviruses*)；真菌病原體，諸如黑穗病(例如黑穗菌目 (*Ustilaginales*))、鏽病(鏽菌目 (*Uredinales*))、麥角(黑麥麥角菌 (*Claviceps purpurea*))及徽病；徽菌(卵菌綱 (*Oomycetes*))，諸如馬鈴薯晚疫病菌 (*Phytophthora infestans*)(馬鈴薯疫病)；細菌性病原體，諸如伊文氏桿菌屬 (*Erwinia*)(例如草生伊文氏菌 (*E. herbicola*))，假單胞菌屬 (*Pseudomonas*)(例如綠膿桿菌 (*P. aeruginosa*)、丁香假單胞菌 (*P. syringae*)、螢光假單胞菌 (*P. fluorescense*)及惡臭假單胞菌 (*P. putida*)，青枯病菌屬 (*Ralstonia*)(例如番茄青枯病菌 (*R. solanacearum*))，土壤桿菌屬及黃單胞菌屬 (*Xanthomonas*)；蛔蟲(線蟲綱 (*Nematoda*))；及豆類根瘤病菌 (*Phytomyxea*)(多黏桿菌屬 (*Polymyxa*)及甘藍根瘤病菌屬 (*Plasmodiophora*))。

用於靶向重組之所揭示方法可用以用同源、非相同序列置換任何基因組序列。舉例而言，突變基因組序列可由其野生型對應物置換，由此提供用於治療植物疾病之方法；提供針對植物病原體之抗性；增加作物產量等等。以類似方式，基因之一等位基因可使用本文中揭示之靶向重組方法，由不同等位基因置換。

在許多該等狀況下，所關注區域包含突變，且供體聚核苷酸包含相應野生型序列。類似地，需要時可將野生型基因組序列由突變序列置換。實際上，以任何方式依賴於特

定基因組序列之任何病理學可使用本文中揭示之方法及組合物校正或舒緩。

靶向裂解及靶向重組亦可用以改造非編碼序列(例如調節序列, 諸如啟動子、強化子、起始子、終止子、剪接位點)以改變基因產物之表現量。該等方法可用於(例如)治療目的、細胞生理學及生物化學之改造、功能基因組學及/或標靶驗證研究。

本文中描述之方法及組合物亦可用於活化及抑制使用非標準鋅指結合域與功能域之間的融合體之基因表現。該等方法揭示於(例如)共同擁有之美國專利 6,534,261 ; 6,824,978及6,933,113中; 其揭示內容以引用之方式併入本文中。其他抑制方法包括使用靶向欲抑制之基因之序列的反義寡核苷酸及/或小干擾RNA(siRNA或RNAi)。

在其他實施例中, 除本文中揭示之鋅指-裂解域融合體外或替代本文中揭示之鋅指-裂解域融合體, 可使用鋅指結合域與重組酶(或其功能性片段)之間的一或多種融合體, 以促進靶向重組。參見(例如), 共同擁有之美國專利第 6,534,261 號及 Akopian 等人(2003) *Proc. Natl. Acad. Sci. USA* 100:8688-8691。

在其他實施例中, 使用所揭示之方法及組合物來提供 ZFP 結合域與由於其活性而需要二聚作用(均二聚或雜二聚)之轉錄活化或抑制域之融合體。在該等狀況下, 融合多肽包含鋅指結合域及功能域單體(例如, 來自二聚轉錄活化或抑制域之單體)。兩種該等融合多肽與適當定位之

標靶位點之結合允許二聚作用，以便重構功能性轉錄活化或抑制域。

實例

本發明在以下實例中進一步定義，其中除非另外規定，否則所有份及百分比係以重量計且度為攝氏度。應瞭解，雖然指示本發明之某些實施例，但是該等實例僅以說明之方式給出。自以上論述及該等實例，熟習此項技術者可確定本發明之基本特徵，且在不悖離其精神及範疇之情況下，可進行本發明之各種改變及修改以使其適於各種用途及條件。

實例1：ZFN表現載體

如下修飾包含編碼如美國專利公開案2005/0064474(其揭示內容以引用之方式併入本文中)之實例2及14中所述之4指ZFN(命名為"5-8"及"5-9")之序列(參見彼申請案之實例2)的表現載體。簡言之，將5-8及5-9 ZFN(包含經由4胺基酸ZC連接子融合於IIS型限制酶*FokI*之核酸酶域(Wah等人(1998) *Proc. Natl. Acad. Sci. USA* 95:10564-10569之胺基酸384-579)之4鋅指域)修飾成CCHC結構。亦對C末端His與Cys鋅配位結構之間的殘基及/或指2及/或指4之C末端Cys之C末端進行其他修飾(取代及插入)。

實例2：報導體細胞系中eGFP之基因校正

包含如本文中所述之CCHC鋅指之ZFN促進同源重組的能力係以描述於Urnov (2005) *Nature* 435(7042):646-51及美國專利公開案第20050064474號(例如實例6-11)中之GFP

系統測試。簡言之，按照Invitrogen Lipofectamine 2000方案，每個樣本使用2 μ l之Lipofectamine 2000，將50 ng各ZFN及500 ng缺啟動子GFP供體(Urnov (2005) *Nature*)轉染至500,000個報導體細胞中。

轉染後24小時，以0.2 μ M最終濃度添加長春鹼，且在轉染後72小時移除。

轉染後5天，藉由每次轉染在Guava桌上型FACS分析儀上量測40,000個細胞來檢定細胞之GFP表現。

如圖1中所示，包含如上文表1及2中所示之經改造CCHC鋅指之大多數ZFN在報導體(GFP)基因座促進同源重組，產生高於未經修飾CCHC鋅指之量之GFP表現，且所執行之若干者可與包含CCHH鋅指之ZFN相當。當定位於指4(F4)中時，最佳執行變體包含以下序列(包括His鋅配位殘基且位於His鋅配位殘基之C末端)：HAQRCGLRGSQVLV (SEQ ID NO:53)(表2中指定為#21且作為"2-21"展示於圖1中之鋅指)。當定位於指2(F2)中時，最佳執行變體包含以下序列(包括His鋅配位殘基且位於His鋅配位殘基之C末端)：HIRTCTGSQKP(SEQ ID NO:75)(表2中指定為#43且作為"2-43"展示於圖1中之鋅指)。

實例3：藉由靶向重組編輯染色體IL2R γ 基因

如本文所述之ZFN亦以Urnov (2005) *Nature* 435 (7042):646-51及美國專利公開案第20050064474號之實例2中所述之內源IL2R γ 檢定加以檢定。簡言之，使用Nucleofector(Amaxa)將2.5微克各ZFN表現構築體轉染至

500,000個K562細胞中。收集基因組DNA且使用Surveyor核酸內切酶套組檢定位於內源IL2R γ 基因座之基因中斷。

ZFN展示於圖2之左上方。詳言之，經改造之鋅指20係指包含序列HTRRCGLRGSQVLV之CCHC鋅指；鋅指21包含序列HAQRCGLRGSQVLV(SEQ ID NO:53)；鋅指43包含序列HIRTCTGSQKP(SEQ ID NO:75)；鋅指45包含序列HIRTGCTGSQKP；鋅指47包含序列HIRRCTGSQKP；且鋅指48包含序列HIRRGCTGSQKP。鋅指20及21用於4指ZFN之指4中且鋅指43、45、47及48用於4指ZFN之指2中。

所測試之ZFN對展示於圖2上方及圖之右側及表5中：

表 5

樣本#	5-8 ZFN	5-9 ZFN
1	無(GFP)	
2	野生型(CCHH)	野生型(CCHH)
3	43(指2)	43(指2)
4	43(指2)	20(指4)
5	43(指2)	21(指4)
6	43(指2)	45(指2)
7	43(指2)	47(指2)
8	20(指4)	43(指2)
9	21(指4)	43(指2)
10	45(指2)	43(指2)
11	47(指2)	43(指2)
12	48(指2)	43(指2)

為確定是否已在裂解位點誘發突變，使用Cel-1檢定分

析擴增產物，其中將擴增產物變性且複性，接著用錯配特異性 Cel-1 核酸酶處理。參見(例如)，Oleykowski 等人 (1998) *Nucleic Acids res.* 26:4597-4602；Qui 等人 (2004) *BioTechniques* 36:702-707；Yeung 等人 (2005) *BioTechniques* 38:749-758。

對各樣本之2個實驗之結果展示於圖2中。樣本8及9之實驗#2在泳道中具有顯著背景雜訊，其降低該等 ZFN 之表觀功效。

如圖2中所示，某些 CCHC 變體基本上等效於野生型 C2H2 ZFN。指4處之鋅指21(樣本5及9)產生比指4處之鋅指20(樣本4及8)更佳之結果。在指2中，鋅指43產生最佳結果。

實例4：報導體細胞系中 eGFP 之基因校正

基於圖1及2中所示之結果，產生上文表3及4中所示(指定為1a至10a)之 CCHC 鋅指。將該等鋅指併入5-8及5-9 ZFN 中且以上文實例2中所述之 GFP 基因校正檢定加以測試。各樣本中所測試之 ZFN 對展示於各條形圖下方，其中鋅指第20、21、43、45、47及48號為實例3中所述者，且 CCHC 鋅指1a至10a包含上文表3及4中所示之序列。鋅指20、21、7a、8a、9a及10a用於指4中；鋅指43、45、47、48、1a、2a、3a、4a、5a及6a用於指2中。

結果展示於圖3中。各條形圖下方之上列係指併入 ZFN 5-8中之鋅指，且各條形圖下方之下列係指併入 ZFN 5-9中之鋅指。舉例而言，圖3圖上左邊之第2個條形圖係指經5-

8及5-9 ZFN轉染之樣本，其中2個ZFN之F4包含鋅指20之序列。如所示，所執行之許多包含CCHC鋅指之ZFN可與野生型(CCHH)ZFN相當。

實例5：標靶載體之設計及產生

A.標靶序列之總體結構

用於煙草(雙子葉植物)之標靶構築體包括如圖4及5中所示之以下7種組分：i)濕黴素磷酸轉移酶(HPT)表現卡匣，其包含驅動由根癌土壤桿菌開放閱讀框架-24(orf-24)3'非轉譯區(UTR)(Gelvin等人，1987，EP222493)終止之大腸桿菌HPT基因(Waldron等人，1985，Plant Mol. Biol. 18:189-200)之擬南芥泛素-3(ubi-3)啟動子(Callis，等人，1990，J. Biol. Chem. 265:12486-12493)；ii)同源序列-1，其包含煙草(*N. tabacum*)RB7基質附著區(MAR)(Thompson等人，1997，WO 9727207)；iii)由經修飾根癌土壤桿菌甘露鹼合成酶(Δ mas)啟動子(Petolino等人，美國專利第6,730,824號)驅動之5'綠色螢光蛋白質(GFP)基因片段(Evrogen Joint Stock Company, Moscow, Russia)；iv) β -葡糖醛酸酶(GUS)表現卡匣，其包含驅動由根癌土壤桿菌胭脂鹼合成酶(nos)3'UTR(DePicker等人，1982，J. Mol. Appl. Genet. 1:561-573)終止之GUS基因(Jefferson，1987，Plant Mol. Biol. Rep. 5:387-405)之木薯葉脈嵌紋病毒(CsVMV)啟動子(Verdaguer等人，1996，Plant Molecular Biology 31:1129-1139)；v)由根癌土壤桿菌 orf-1 3'UTR(Huang等人，J. Bacteriol. 172:1814-1822)終止之3'GFP基因片段(Evrogen

Joint Stock Company, Moscow, Russia)；vi)同源序列-2，其包含擬南芥4-香豆醯基-CoA合成酶(4-CoAS)內含子-1(基因座 At3g21320，GenBank NC 003074)及；vii)由根癌土壤桿菌 ORF-25/26 3'UTR(Gelvin等人，1987，EP222493)終止之綠色產色鏈黴菌(*S. viridochromogenes*)草胺膦磷酸轉移酶(PAT)(Wohlleben等人，1988，Gene 70:25-37)3'基因片段。

將鋅指-FokI融合蛋白結合位點(IL-1-L0-FokI)(Urnov等人，2005，US 2005/0064474)插入CsVMV啟動子(Verdaguer等人，1996，Plant Molecular Biology 31:1129-1139)之下游，且與GUS編碼序列(Jefferson, 1987，Plant Mol. Biol. Rep. 5:387-405)在N末端融合。第二鋅指-FokI融合蛋白結合位點(Scd27-L0-FokI)(Urnov等人，2005，US 2005/0064474)之2個複本側接5'及3'GFP基因片段(Evrogen Joint Stock Company, Moscow, Russia)。各結合位點含有特定鋅指-FokI融合蛋白之識別序列之4串聯重複以便各結合位點大小為約200 bp(圖6A)。其經設計以確保識別序列在複雜染色質環境中可由鋅指-FokI融合蛋白接近。各識別序列包括單一鋅指-FokI融合蛋白以同源二聚體形式結合且裂解雙股DNA之反向重複序列(圖6B)。5'及3'GFP基因片段重疊540 bp從而在標靶序列內部提供同源性，且在5'GFP片段之3'末端插入終止密碼子，以確保不自標靶序列發生功能性GFP轉譯。

包含標靶序列之轉型載體係經由如下文所述之多步選殖

方法來產生。

B. 建構 HPT 二元載體 (pDAB1584)

將含有 GUS 表現卡匣之載體 pDAB1400 用作起始基本構築體(圖 7)，該 GUS 表現卡匣包含驅動由根癌土壤桿菌 orf-1 UTR(Huang 等人，J. Bacteriol. 172:1814-1822)終止之 GUS 基因(Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405)之擬南芥 ubi-3 啟動子(Callis, 等人，1990, J. Biol. Chem. 265-12486-12493)。

為在標靶構築體中避免任何不必要重複調節元件，用根癌土壤桿菌 orf-24 UTR(Gelvin 等人，1987, EP222493)置換 pDAB1400 中之根癌土壤桿菌 orf-1 UTR(Huang 等人，J. Bacteriol. 172:1814-1822)，該根癌土壤桿菌 orf-24 UTR 係作為 SacI/XbaI 片段自 pDAB782(圖 8)切離且選殖至 pDAB1400 中之相同位點中。所得構築體含有驅動由根癌土壤桿菌 orf-24 UTR(Gelvin 等人，1987, EP222493)終止之 GUS 基因(Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405)之擬南芥 ubi-3 啟動子(Callis, 等人，1990, J. Biol. Chem. 265-12486-12493)，且命名為 pDAB1582(圖 9)。

使用引子 P1 及 P2，將 HPT 編碼序列(Waldron 等人，1985, Plant Mol. Biol. 18:189-200)自 pDAB354 質體(圖 10)PCR 擴增。在引子 P1 之 5' 末端添加 BbsI 位點，且在引子 P2 之 3' 末端保留 SacI 位點。用 BbsI/SacI 消化 HPTII PCR 片段，且將其選殖至經 NcoI-SacI 消化之 pDAB1582 中，以用來自 PCR 片段之 HPT 基因置換 GUS 基因。所得質體命名為

pDAB1583(圖 11)。

隨後，藉由NotI消化自 pDAB1583切離擬南芥ubi-3/HPT/根癌土壤桿菌orf-24片段，且用T4 DNA聚合酶處理以產生鈍末端。將經鈍末端處理之HPT表現卡匣選殖至二元基本載體 pDAB2407(圖 12)之PmeI位點中，產生質體 pDAB1584(圖 13)。

C.建構包含同源序列及Scd27鋅指-FokI融合蛋白結合位點之載體(pDAB1580)

用根癌土壤桿菌orf25/26 UTR(Gelvin等人，1987, EP222493)置換pDAB2418(圖 14)中之根癌土壤桿菌orf-1 UTR(Huang等人，J. Bacteriol. 172:1814-1822)，以在標靶載體中避免重複調節序列。為進行UTR交換，使用引子P3及P4由pDAB4045質體(圖 15)PCR擴增根癌土壤桿菌orf25/26 UTR(Gelvin等人，1987, EP222493)。將SmaI及AgeI位點添加至PCR片段之3'末端，且在5'末端保留SacI位點。用SacI及AgeI消化含有PAT基因表現卡匣及煙草RB7 MAR序列(Thompson等人，1997, WO 9727207)之pDAB2418質體DNA，該PAT基因表現卡匣包含驅動由根癌土壤桿菌orf-1 UTR(Huang等人，J. Bacteriol. 172:1814-1822)終止之PAT基因(Wohlleben等人，1988, Gene 70:25-37)之擬南芥泛素-10(ubi-10)啟動子(Callis, 等人，1990, J. Biol. Chem. 265:12486-12493)，且回收2個最大片段。將該等片段與經SacI及AgeI消化之根癌土壤桿菌orf25/26 UTR(Gelvin等人，1987, EP222493)PCR產物連接。所得質

體命名為 pDAB1575(圖 16)。煙草 RB7 MAR(Thompson 等人, 1997, WO 9727207)用作標靶載體中之同源序列-1。

選擇擬南芥 4-CoAS 之內含子-1(基因座 At3g21320, GenBank NC 003074)以用作標靶載體中之同源序列-2。分析 PAT 基因(Wohlleben 等人, 1988, Gene 70:25-37)編碼序列, 且鑑別出起始密碼子下游之 299/300 bp 為插入內含子之位點, 以便形成適當 5'及 3'剪接位點。隨後, 藉由 DNA 合成(Picoscript Ltd., LLP, Houston, Texas), 將全長內含子與 3'部分 PAT 編碼序列之 253 bp 融合。分別將 NotI 及 SacI 位點添加至 DNA 片段之 5'及 3'末端。隨後, 用 NotI/SacI 消化所合成之 DNA 片段且插入至 pDAB1575 之相同位點中以置換全長 PAT 編碼序列。所得構築體命名為 pDAB1577(圖 17)。

合成(Picoscript Ltd., LLP, Houston, Texas)含有 Scd27-L0-FokI 識別位點之 4 串聯重複(圖 6)之 241 bp DNA 片段, 同時向片段之 5'及 3'末端添加 SmaI 位點。隨後, 用 SmaI 消化含所合成之鋅指-FokI 結合位點之片段, 且將其插入至 pDAB1577 之 MscI 位點中。所得載體命名為 pDAB1579(圖 18)。隨後, 將第二個含經 SmaI 消化之鋅指-FokI 結合位點之片段插入至 pDAB1579 之 SmaI 位點中。所得構築體命名為 pDAB1580(圖 19)。該載體含有同源序列 1 及 2(分別為煙草 RB7 MAR 及擬南芥 4-CoAS 內含子 1)及 2 個合成 Scd27 鋅指-FokI 結合位點, 其各含有 Scd27-L0-FokI 識別位點之 4 串聯重複。

D. 建構含有 2 個部分雙重複無功能 GFP 片段之載體 (pDAB1572)

GFP 基因 CopGFP 係購自 Evrogen Joint Stock Company (Moscow, Russia) 且使用引子 P5 及 P6 將全長編碼序列 PCR 擴增。分別將 BbsI 及 SacI 位點添加至 PCR 產物之 5' 及 3' 末端。隨後，用 BbsI/SacI 消化 CopGFP PCR 產物，且將其選殖至 pDAB3401 (圖 20) 之 NcoI/SacI 位點中以置換 GUS 基因，該 pDAB3401 包含驅動 GUS 基因 (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) 且由根癌土壤桿菌 orf-1 3'UTR (Huang 等人, J. Bacteriol. 172:1814-1822) 終止之經修飾根癌土壤桿菌 Δ mas 啟動子 (Petolino 等人, US6730824)。所得載體命名為 pDAB1570 (圖 21)。

為製造 2 個部分雙重複無功能 GFP 片段，使用引子 P9 及 P10 將含有大部分 CopGFP 編碼序列並在 5' 末端具有 47 bp 缺失之 DNA 片段 PCR 擴增。將 ApaI 位點添加至 5' 及 3' 末端，且將另一 StuI 位點添加至 ApaI 位點下游之 5' 末端。隨後，用 ApaI 消化 PCR 產物，且將其插入至 pDAB1570 之 ApaI 位點中，由此產生在同一載體中具有 540 bp 雙重複序列之 2 個無功能 GFP 片段。所得構築體命名為 pDAB1572 (圖 22)。

E. 建構含有 IL-1 鋅指-FokI 融合蛋白結合位點/GUS 基因融合體之載體 (pDAB1573)

由 Picoscript Ltd., LLP, (Houston, Texas) 合成含有 IL-1-L0-FokI 識別位點之 4 串聯重複 (圖 6) 之 233 bp DNA 片段，同時分別將 NcoI 及 AflIII 位點添加至 5' 及 3' 末端。隨後，用

NcoI/AflIII消化所合成之片段且將其插入至pDAB4003(圖23)之NcoI位點中，該pDAB4003含有由以根癌土壤桿菌orf-1 3'UTR(Huang等人，J. Bacteriol. 172:1814-1822)終止之CsVMV啟動子(Verdaguer等人，1996, Plant Molecular Biology 31:1129-1139)驅動之GUS基因(Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405)。隨後產生IL-1-Lo-FokI結合位點與GUS編碼序列之間的N末端融合體。所得載體命名為pDAB1571(圖24)。

為在標靶載體中避免重複3'UTR元件，將根癌土壤桿菌nos 3'UTR(DePicker等人，1982, J. Mol. Appl. Genet. 1:561-573)以SacI/PmeI片段形式自pDAB7204(圖25)切離，且選殖至經SacI/NaeI消化之pDAB1571中，以置換根癌土壤桿菌orf-1 3'UTR(Huang等人，J. Bacteriol. 172:1814-1822)。所得質體命名為pDAB1573(圖26)。

F. 建構最終標靶載體(pDAB1585)

為製造最終標靶載體，藉由NotI消化將具有IL-1-FokI融合蛋白標靶位點插入物之GUS表現卡匣切離，鈍末端處理且插入至pDAB1572之StuI位點中。所得中間載體命名為pDAB1574(圖27)。自pDAB1574切離整個卡匣，其含有經修飾 Δ mas啟動子(Petolino等人，US6730824)、5'部分雙重複GFP序列(Evrogen Joint Stock Company, Moscow, Russia)、CsVMV啟動子(Verdaguer等人，1996, Plant Molecular Biology 31:1129-1139)、IL-1-FokI融合蛋白標靶序列、GUS基因(Jefferson, 1987, Plant Mol. Biol. Rep.

5:387-405)編碼區域、根癌土壤桿菌 nos 3'UTR(DePicker等人, 1982, J. Mol. Appl. Genet. 1:561-573)、3'部分雙重複 GFP(Evrogen Joint Stock Company, Moscow, Russia)及根癌土壤桿菌 orf-1 3'UTR(Huang等人, J. Bacteriol. 172:1814-1822), 且將其插入至 pDAB1580之 NotI位點中。所得質體命名為 pDAB1581(圖 28)。隨後, 將 pDAB1581之 AgeI片段插入至 pDAB1584之 AgeI位點中, 由此產生最終標靶構築體 pDAB1585(圖 4及 5)。

實例 6：產生具有整合標靶序列之轉殖基因細胞系

使用實例 5 之標靶序列經由土壤桿菌轉型穩定整合至其中稱為 BY2 之煙草細胞懸浮培養物。基本細胞系 BY2 係自 Japan Tobacco(Iwata, Shizuoka, Japan)之 Jun Ueki 獲得。在 100-150 個細胞叢集中, 將該培養物增殖為直徑 5-10 μ 之細胞, 加倍時間約為 18 小時。將 BY2 細胞懸浮培養物維持在含有 LS 基本鹽 (PhytoTechnology Labs L689)、170 mg/L KH_2PO_4 、30 g/L 蔗糖、0.2 mg/L 2,4-D 及 pH 值為 6.0 之 0.6 mg/L 噻胺-HCL 之培養基中。每隔 7 天, 藉由將 50 mL 以 LS 為主之培養基添加至 0.25 mL PCV 中來次培養 BY2 細胞。在 25°C 及 125 RPM 下, 將 BY2 細胞懸浮培養物維持在旋轉搖動器上之 250 mL 燒瓶中。

為產生具有整合標靶序列之轉殖基因 BY2 細胞培養物, 將具有次培養後 4 天之煙草懸浮液之燒瓶分成 10-12 個 4 mL 等分試樣, 在 100×25 mm 皮氏培養皿 (Petri dish) 中, 將其與 100 μ L 生長隔夜 OD_{600} 值達到約 1.5 之含有土壤桿菌菌株

LBA4404的 pDAB1585 共同培養。用蠟膜包裝培養皿且在 25℃ 下在不搖動之情況下培養3天，其後移除過量液體且用 11 mL含有 500 mg/L卡本西林(carbenicillin)之以LS為主的基本培養基置換。

將煙草細胞再懸浮後，將 1 mL懸浮液分配至具有經 8 g/L TC瓊脂固化含有 500 mg/L卡本西林及 200 mg/L濕黴素之適當基本培養基之 100×25 mm培養板上，且在黑暗中，在 28℃ 下解開包裝培養。此產生 120-144個選擇培養板以供單次處理。塗鋪後 10-14天出現個別抗濕黴素分離株且將其轉移至個別 60×20 mm培養板(每個培養板一個分離株)，其中以 14天次培養時程將其維持為愈傷組織直至需要用於分析及後續再轉型實驗。

實例 7：篩選及表徵標靶轉殖基因事件

如下分析如實例 6中所述之由將標靶載體轉型至 BY2煙草細胞培養物中產生之抗濕黴素轉殖基因事件。

為篩選該等轉殖基因事件所進行之初始分析包括指示標靶序列之可達性之 GUS表現分析，確認標靶載體之存在及完整性之部分及全長標靶序列的 PCR分析，及測定整合標靶序列之複本數之南方墨點分析。展示 GUS表現之轉殖基因事件之子集含有全長標靶序列之單一複本；選擇該等轉殖基因事件用於重建懸浮液培養物，以產生用於後續再轉型之標靶細胞系。該等重建標靶細胞系亦經受進一步表徵，其包括更徹底之南方墨點分析，整個標靶插入物之定序確認及側接基因組序列分析。

藉由在 37°C 下，在 150 μ L 檢定緩衝液中，將 50 mg 樣本培養 24-48 小時來分析由所選事件起始之轉殖基因煙草愈傷組織組織或懸浮液培養物之 GUS 活性。檢定緩衝液由 0.2 M 磷酸鈉 (pH 8.0)、各 0.1 mM 之鐵氰化鉀及亞鐵氰化鉀、1.0 mM EDTA 鈉、0.5 mg/mL 5-溴-4-氯-3-吡啶基- β -葡萄糖苷酸及 0.6%(v/v) Triton X-100 (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) 組成。藍色區域之顯現係用作 GUS 基因表現之指示劑，其指示標靶序列插入物具有轉錄活性且因此在局部基因組環境中可接近。

藉由使用引子對 P15/P16 之 PCR 檢定 GUS 表現轉殖基因事件，該引子對 P15/P16 引起自標靶序列之 5' 末端之 HPT 表現卡匣的 3'UTR 延伸至標靶序列之 3' 末端之部分 PAT 基因卡匣的 3'UTR 之 10 kb DNA 片段擴增。因為所有事件係在濕黴素選擇下獲得，所以假定 HPT 表現卡匣在所有標靶事件中為完整的。因此，僅 HPT 表現卡匣之 3'UTR 涵蓋於全長 PCR 分析中。亦使用引子對 P15/P17 及 P18/P19 PCR 檢定事件之子集，以分別確定標靶序列之 5' 及 3' 末端之完整性。用 PCR 分析確認之所有標靶事件進一步藉由南方墨點分析檢定以測定整合標靶序列之複本數。

對通過 GUS 表現及全長 PCR 之篩選之所有標靶事件進行南方墨點分析。用 NsiI 消化 10 μ g 基因組 DNA，該 NsiI 為標靶序列內之唯一切割器。在 0.8% 瓊脂糖凝膠上分離所消化基因組 DNA 且將其轉移至耐綸膜上。交聯後，將膜上之經轉移 DNA 與 HPT 基因探針雜交以測定標靶序列 5' 末端之複

本數。隨後，洗提相同墨點且與PAT基因探針再雜交以測定標靶序列3'末端之複本數。

選擇展示GUS表現且含有全長標靶序列之單一複本之多次事件以供進一步表徵，其包括更徹底之南方墨點分析、完整標靶序列確認及側接基因組序列分析。基於分子表徵選擇稱為BY2-380之一事件。自該事件重建懸浮液培養物，以供用包含供體DNA及非C2H2鋅指-FokI融合蛋白基因之載體之後續再轉型。

為確保自標靶事件BY2-380建立之懸浮培養物含有所預期之完整標靶序列，使用引子對P15/P16將自標靶序列之5'末端之HPT表現卡匣的3'UTR至標靶序列之3'末端之部分PAT基因卡匣的3'UTR之主要標靶序列PCR擴增，且將其選殖至pCR2.1 TOPO載體(Invitrogen, Carlsbad, California)中。由Lark technology, Inc.(Houston, Texas)將插入TOPO載體中之PCR產物定序。序列結果指示，BY2-380具有所預期之完整標靶序列。

使用Universal GenomeWalker套組(Clontech, Mountain View, California)進一步分析BY2-380細胞系以獲得側接基因組序列。簡言之，以單獨反應用3種鈍末端限制酶EcoRV、DraI及StuI消化2.5 µg基因組DNA。經由苯酚/氯仿萃取純化經消化之DNA且與BD GenomeWalker轉接器連接。執行套式PCR擴增，以連接物作為模板且將引子P20(走向標靶序列插入物之5'末端之上游)及P21(走向標靶序列插入物之3'末端之下游)用於初級PCR反應，且將引子

P22(走向標靶序列插入物之5'末端之上游)及P23(走向標靶序列插入物之3'末端之下游)用於第二套式PCR反應。將自第二PCR反應擴增之片段選殖至pCR2.1 TOPO或pCR Blunt II TOPO載體(Invitrogen, Carlsbad, California)中且使用染料終止子循環定序套組(Dye Terminator Cycle Sequencing Kit)(Beckman Coulter, Fullerton, CA)定序。經由該方法自BY2-380標靶細胞系獲得側接基因組序列。隨後，基於側接基因組序列設計引子且將其用於擴增完整標靶序列。

自該標靶細胞系獲得之擴增片段具有預期大小。藉由定序確認擴增片段之2個末端。

實例8：供體DNA載體之設計及產生

供體DNA構築體包括同源序列-1(煙草RB7 MAR)(Thompson等人，1997，WO 9727207)、全長擬南芥ubi10啟動子(Callis，等人，1990，J. Biol. Chem. 265-12486-12493)、299 bp之5'部分PAT基因編碼序列(Wohlleben等人，1988，Gene 70:25-37)及同源序列-2(擬南芥4-CoAS內含子-1)(基因座At3g21320，GenBank NC 003074)。供體載體中之同源序列-1及序列-2與標靶載體(pDAB1585)中之相應同源序列-1及序列-2相同。

為建構供體載體，經由以Picoscript Ltd., LLP,(Houston, Texas)之DNA合成，將299 bp之5'部分PAT編碼序列與全長擬南芥4-CoAS內含子-1(基因座At3g21320，GenBank NC 003074)融合。分別將NcoI及XhoI位點添加至片段之5'及3'末端。隨後，用NcoI/XhoI消化該所合成DNA片段，且將

其插入至pDAB1575之相同位點中，以置換全長PAT基因編碼序列及其3'UTR。所得構築體命名為pDAB1576(圖29)。

隨後，用AgeI消化pDAB1576，且將含有側接有同源序列-1及同源序列-2之5'部分PAT表現卡匣之整個片段插入至二元基本載體pDAB2407之相同位點中。所得構築體命名為pDAB1600(圖30)且為用於植物細胞再轉型之供體載體之二元型式。

實例9：鋅指核酸酶表現載體之設計及產生

鋅指-FokI融合蛋白基因係由CsVMV啟動子及5'UTR(Verdaguer等人，1996, Plant Molecular Biology 31:1129-1139)驅動。根癌土壤桿菌開放閱讀框架-24(orf-24)3'非轉譯區(UTR)(Gelvin等人，1987, EP222493)亦包括於卡匣中。

為製造該等載體，用BbsI或NcoI將上文實例1至4中所述之IL-1-FokI及Scd27-FokI編碼序列之C2H2對照及其C3H變體自其原始設計PCR擴增，且分別將NcoI及SacI位點添加至PCR片段之5'及3'末端，且選殖至經NcoI-SacI消化之pDAB3731(圖31)中。所得質體命名為pDAB4322(圖32)、pDAB4331(圖33)、pDAB4332(圖34)、pDAB4333(圖35)、pDAB4334(圖36)、pDAB4336(圖37)及pDAB4339(圖38)。所有該等載體含有側接ZFN表現卡匣之attL1及attL2位點，且可與Gateway™選殖系統(Invitrogen, Carlsbad, California)相容。

對於IL-1-FokI融合蛋白建構兩組二元型式載體。一者含

有PAT選擇標記基因且另一者不含有PAT選擇標記基因。對SCd27-FokI融合蛋白而言，僅建構不具有PAT選擇標記基因之載體之二元型式。為製造具有PAT選擇標記基因之二元載體，經由使用LR Clonase™酶混合物(Invitrogen, Carlsbad, California)之LR重組反應，將pDAB4322、pDAB4331、pDAB4332、pDAB4333、pDAB4334及pDAB4336中之IL-1-FokI融合蛋白表現卡匣選殖至pDAB4321(圖39)中。所得質體命名為pDAB4323(圖40)、pDAB4341(圖41)、pDAB4342(圖42)、pDAB4343(圖43)、pDAB4344(圖44)、pDAB4346(圖45)。為製造不具有PAT選擇標記基因之二元載體，經由使用LR Clonase™酶混合物(Invitrogen, Carlsbad, California)之LR重組反應，分別將pDAB4331、pDAB4336及pDAB4339中之C2H2 IL-1-FokI、C3H IL-1-FokI及Scd27-FokI表現卡匣選殖至pDAB4330(圖46)中。所得質體分別命名為pDAB4351(圖47)、pDAB4356(圖48)及pDAB4359(圖49)。

為製造SCD27-FokI之C2H2對照，用包含CsVMV啟動子及5'UTR及煙草5'UTR驅動GUS之用HindIII/SacI自pDAB7025(圖51)切離的片段置換pDAB7002(圖50)中之包含CsVMV啟動子及5'UTR驅動PAT之HindIII/SacI片段。所得質體命名為pDAB1591(圖52)。使用引子對P13/P14，將Scd27-L0-FokI編碼序列自其原始載體pCDNA3.1-SCD27a-L0-FokI(圖53)PCR擴增。分別將BbsI及SacI位點添加至PCR片段之5'及3'末端。經由SacI/NcoI選殖，用鋅指融合

蛋白基因PCR片段置換pDAB1591中之PAT基因。所得質體命名為pDAB1594(圖54)。藉由自pDAB1594將鋅指融合蛋白基因表現卡匣以PmeI/XhoI片段形式切離，填充末端且將其選殖至pDAB2407之PmeI位點中來建構該載體之二元型式。所得質體命名為pDAB1598(圖55)。用於植物轉型之所有二元載體之細節概述於表6中。

表6：鋅指核酸酶表現載體

載體	ZFN	ZFP之類型	鋅指之指位置	胺基酸序列	SEQ ID NO:
pDAB4323	IL 1-FokI	C2H2	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKI <u>H</u>	94
pDAB4341	IL 1-FokI.	C2H2	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKI <u>H</u>	95
pDAB4342	IL 1-FokI*	C3H	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKI <u>C</u>	96
pDAB4343	IL 1-FokI*	C3H	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKRC <u>G</u> GG	97
pDAB4344	IL 1-FokI*	C3H	F4	<u>C</u> ... <u>C</u> ... <u>H</u> AQR <u>C</u> G	98
pDAB4346	IL 1-FokI*	C3H	F2	<u>C</u> ... <u>C</u> ... <u>H</u> IR <u>T</u> G <u>C</u>	99
pDAB4351	IL 1-FokI*	C2H2	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKI <u>H</u>	100
pDAB4356	IL 1-FokI*	C3H	F2	<u>C</u> ... <u>C</u> ... <u>H</u> IR <u>T</u> G <u>C</u>	101
pDAB1598	Scd27-FokI	C2H2	F4	<u>C</u> ... <u>C</u> ... <u>H</u> TKI <u>H</u>	102
pDAB4359	Scd27-FokI*	C3H	F4	<u>C</u> ... <u>C</u> ... <u>H</u> AQR <u>C</u> GG	103

* FokI域偏好植物密碼子。

實例10：陽性對照載體之設計及產生

為估算不正常重組頻率且用作陽性對照，使用含有PAT基因表現卡匣之載體。為可與最終重組體相當，在PAT編碼序列(Wohlleben等人，1988，Gene 70:25-37)之299/300 bp處插入擬南芥4-CoAS內含子-1(基因座At3g21320，

GenBank NC 003074)。為製造該構築體，將來自 pDAB1576之2559 bp SwaI/ClaI片段與經相同限制酶消化之 pDAB1577(圖 56)之骨架片段連接。所得載體含有 PAT 基因表現卡匣，該卡匣具有在 PAT 編碼序列(Wohlleben 等人，1988, Gene 70:25-37)中間插入之1743 bp之擬南芥4-CoAS 內含子-1(基因座 At3g21320，GenBank NC 003074)(基因座 At3g21320，GenBank NC 003074)。該載體命名為 pDAB1578(圖 57)。

為製造 pDAB1578 之二元型式，用 PmeI/XhoI 自 pDAB1578切離具有擬南芥內含子-1(基因座 At3g21320，GenBank NC 003074)之 PAT 基因表現卡匣。將片段之3'末端經鈍末端處理後，將其插入至二元基本載體 pDAB2407 之 PmeI位點中。所得載體命名為 pDAB1601(圖 58)，其包含含有由擬南芥 ubi10啟動子(Callis，等人，1990, J. Biol. Chem. 265-12486-12493)驅動且由根癌土壤桿菌 orf25/26 3'UTR(Gelvin 等人，1987, EP222493)終止之擬南芥4-CoAS 內含子-1(基因座 At3g21320，GenBank NC 003074)序列之 PAT 基因(Wohlleben 等人，1988, Gene 70:25-37)。

實例 11：藉由用 C3H 鋅指核酸酶基因再轉型標靶細胞培養物來示範染色體內同源重組

為驗證 C3H 鋅指核酸酶在刺激染色體內同源重組中之功能性，如圖 59 中所示，使 2 個具有 540 bp 重疊序列之非功能性 GFP 片段包括於標靶載體中。GUS 基因表現卡匣在該等 2 個片段之間。將 IL-1-FokI 融合蛋白結合序列與 GUS 編

碼序列在其N末端融合。不受理論約束，假定在IL-1-FokI融合蛋白存在下，將識別IL-1 ZFN結合序列且誘導雙股DNA斷裂，其將刺激內源DNA修復過程。在不存在供體DNA之情況下，2個部分同源GFP片段將經歷染色體內同源重組過程且將重構功能性GFP基因。

使用含有標靶序列之單一、全長整合複本之BY2-380轉殖基因細胞系以藉由將約250-500 mg愈傷組織組織置放於40-50 mL含有100 mg/L濕黴素以LS為主的基本培養基中且如上所述每隔7天次培養來再引發懸浮培養物。再轉型之前，將懸浮培養物轉移至不具有濕黴素之基本培養基中，歷時至少2個繼代。

如上所述執行標靶細胞培養物之土壤桿菌介導之轉型。對各實驗而言，如下產生8個共培養板：一個培養板包含與300 μ L基本土壤桿菌菌株LBA4404共培養之細胞；一個培養板包含與300 μ L含有pDAB1590(功能性GFP構築體)之土壤桿菌菌株共培養的細胞；6個培養板各自包含分別與300 μ L含有pDAB4323、pDAB4341、pDAB4342、pDAB4343、pDAB4344及pDAB4346之土壤桿菌菌株共培養的細胞。使用上文所述之方法共培養後，將細胞塗鋪於8個含有補充有500 mg/L卡本西林而不具有選擇試劑的以LS為主的基本培養基之培養板上。轉型後約5-8天，所建構之功能性GFP基因之表觀表現產生可見螢光。藉由檢視每個培養板5個'隨機'顯微鏡視場計數每個視場之綠色螢光基因座之數量，在各實驗中每個構築體8個培養板，且自6

個獨立實驗取平均值。

如表 7 中所概述，自 2 個 C3H 鋅指核酸酶 pDAB4346 及 pDAB4343，每個視場分別觀察到平均 9.50 及 7.57 個綠色螢光基因座。IL-1-FokI 之該等 2 個 C3H 設計比其 C2H2 對照 pDAB4341 (每個視場 6.37 個基因座) 及 pDAB4323 (每個視場 5.53 個基因座) 執行更佳。同時，與 C2H2 對照相比，IL-1-FokI 融合蛋白之其它 2 個 C3H 變體 pDAB4344 (每個視場 4.39 個基因座) 及 pDAB4342 (每個視場 0.25 個基因座) 之功能顯著削弱，尤其 pDAB4342，其中 C3H 轉化僅僅藉由在第四指中用組胺酸置換第二半胱胺酸來進行。在用基本土壤桿菌菌株 LBA4404 轉型之陰性對照中，未觀察到超過輕微背景之可量測螢光。

表 7

經由 IL-1-FokI 鋅指融合蛋白刺激之染色體內同源重組建構功能性 GFP

載體	ZFP 之類型	GFP 表現	塔基測試 (Tukey Test)**
pDAB4346	C3H	9.50	A
pDAB4343	C3H	7.57	B
pDAB4341	C2H2	6.37	C
pDAB4323*	C2H2	5.53	D
pDAB4344	C3H	4.39	E
pDAB4342	C3H	0.25	F

*含有偏好非植物密碼子之 FokI 域

**未由相同字母連接之方式在 0.05 水準下顯著不同

實例 12：藉由用 C3H 鋅指核酸酶基因及供體 DNA 序列再轉型標靶細胞培養物來示範染色體內同源重組

為驗證在例示性煙草系統中，C3H 鋅指-FokI 融合蛋白在模擬染色體間同源重組中之功能性，開發且測試兩種策略。

在策略 1 中，使鋅指-FokI 融合蛋白(IL-1-L0-FokI)之結合位點包括在標靶構築體(圖 61)中間。在該策略中，結合位點在兩側分別側接有約 3 kb 之非同源序列，隨後在上游及下游分別側接有同源序列-1(煙草 RB7 MAR)及同源序列-2(擬南芥 4-CoAS 內含子-1)。如先前(例如美國專利公開案第 20050064474 號)在 C2H2 IL-1 鋅指-FokI 融合蛋白存在下所證明，識別 IL-1-L0-FokI 結合序列且在該特異性位點處誘發雙股 DNA 斷裂，其刺激內源 DNA 修復過程。在含有與標靶序列中之序列相同之同源序列的供體 DNA 存在下，5' 部分 PAT 基因連同其啟動子一起經由同源重組置換標靶中之同源序列之間的約 6 kb 之完整 DNA 片段。經由該方法，2 個部分 PAT 基因序列與插入其間之擬南芥 4-CoAS 內含子-1 一起重構功能性 PAT 基因，產生 PAT 表現及除草劑抗性表現型。

在策略 2 中，使 2 個鋅指-FokI 結合位點(Scd27-L0-FokI)包括於標靶載體中：一個直接在煙草 RB7 MAR 下游且另一個直接在擬南芥 4-CoAS 內含子 1 上游(圖 62)。在 2 個鋅指-FokI 融合蛋白結合位點之間為約 6 kb 之序列，其包括 5' GFP 片段、GUS 表現卡匣及 3' GFP 片段。如先前(例如美國專利

公開案第20050064474號)在Scd27鋅指-FokI融合蛋白存在下所證明，識別2個結合序列且在2個位置處誘發雙股DNA斷裂，此移除該等2個結合序列之間的約6 kb DNA片段，且刺激內源DNA修復過程。類似於策略1，在含有與標靶序列中之序列相同之同源序列的供體DNA存在下，5'部分PAT基因連同其啟動子一起經由同源重組插入標靶序列之誘導雙股DNA斷裂之位點中。經由該方法，2個部分PAT基因序列與插入於其間之擬南芥4-CoAS內含子-1一起重構功能性PAT基因，產生PAT表現及除草劑抗性表型。

如上所述執行BY2-380標靶細胞培養物之土壤桿菌介導之轉型。對各實驗而言，如下產生12個共培養培養板：一個培養板包含與50 μ L含有pDAB1600(供體DNA)之土壤桿菌菌株及250 μ L土壤桿菌基本菌株LBA4404共培養之細胞；一個培養板包含與50 μ L含有pDAB1601(PAT可選擇標記)之土壤桿菌菌株及250 μ L土壤桿菌基本菌株LBA4404共培養之細胞；2個培養板包含與50 μ L含有pDAB1600(供體DNA)之土壤桿菌菌株及250 μ L含有pDAB4351(C2H2 IL-1 ZFP-FokI)之土壤桿菌菌株共培養的細胞；3個培養板包含與50 μ L含有pDAB1600(供體DNA)之土壤桿菌菌株及250 μ L含有pDAB4356(C3H IL-1 ZFP-FokI)之土壤桿菌菌株共培養的細胞；2個培養板包含與50 μ L含有pDAB1600(供體DNA)之土壤桿菌菌株及250 μ L含有pDAB1598(C2H2 Scd 27a ZFP-FokI)之土壤桿菌菌株共培養的細胞；3個培養板包含與50 μ L含有pDAB1600(供體DNA)之土壤桿菌菌株及

250 μ L含有pDAB4359(C3H Scd27a ZFP-Fok1)之土壤桿菌菌株共培養的細胞。使用上文所述之方法共培養後，將細胞塗鋪於含有500 mg/L卡本西林及15 mg/L Bialaphos[®]之以LS為主之基本培養基上。塗鋪後2-4週出現個別Bialaphos[®]抗性分離株且將其轉移至個別60×20 mm培養板(每個培養板一個分離株)，其中以14天次培養時程將其維持為愈傷組織直至為分析所需。

自C3H IL-1鋅指核酸酶(pDAB4356)及C3H Scd27鋅指核酸酶(pDAB4359)獲得多種Bialaphos[®]抗性分離株。藉由使用引子對P24/25之PCR分析該等分離株，該引子對P24/25擴增涵蓋重構PAT基因之DNA片段。引子P24與供體DNA中PAT編碼序列之5'末端同源，且引子P25與標靶DNA中PAT編碼序列之3'末端同源。若2個部分PAT編碼序列經由同源重組連接，則產生2.3 kb PCR片段。如圖63中所示，自多個所分析之分離株獲得2.3 kb PCR產物。該等分離株係自C3H IL-1鋅指-Fok1融合蛋白基因/供體DNA及C3H Scd27鋅指-Fok1融合蛋白基因/供體DNA之共轉型獲得。自瓊脂糖凝膠純化來自多種代表來源於C3H IL-1鋅指-Fok1及C3H Scd27鋅指-Fok1融合蛋白基因轉型之分離株之獨立分離株的2.3 kb PCR產物，且將其選殖至pCR2.1 TOPO載體(Invitrogen, Carlsbad, California)中。隨後，使用染料終止子循環序列套組(Beckman Coulter)將插入TOPO載體中之2.3 kb PCR產物定序。定序結果確認，選殖於TOPO載體中之所有PCR產物含有所預測之重組序列，

包括具有插入擬南芥4 CoAS內含子-1之5'及3'部分PAT基因序列。該等結果確認所測試之兩種策略之所預測染色體間重組，且經由C3H鋅指-FokI融合蛋白基因之表現例證基因靶向。

實例13：鑑別玉米細胞培養物中之標靶基因序列

A. 序列鑑別

在該實例中，選擇具有已知功能之內源玉米基因之DNA序列作為用於使用工程化鋅指核酸酶之基因組編輯的標靶。來源於專屬玉米自交系5XH751之稱為IPP2-K之該基因的基因組結構及序列已描述於WO 2006/029296中；其揭示內容以引用之方式併入本文中。

詳言之，使用IPP2-K基因組序列以使用BLAST算法查詢TIGR玉米基因組資料庫（可在網際網路上，在<http://www.tigr.org/tdb/tgi/maize/>站點得到）。若干其他基因組片段係用與IPP2-K具有重疊同源性的包括（但不限於）寄存物AZM515213及TC311535的區段來鑑別。基於該等寄存物序列以及IPP2-K序列，使用Primer3程式(Rozen, S. 及 Skaletsky, H.J.(2000)*Primer3 on the WWW for general users and for biologist programmers*。在：Krawetz S, Misener S（編）*Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ，第365-386頁中；亦可在網際網路上得到）設計多種短寡核苷酸以用作PCR引子。該等引子包括（但不限於）以下正向定向寡核苷酸：

1. 5'-ATGGAGATGGATGGGGTTCTGCAAGCCGC-3'(SEQ ID NO:104)

2. 5'-CTTGGCAAGGTACTGCGGCTCAAGAAGATTC-3' (SEQ ID NO:161)

3. 5'-ATGAAGAAAGACAGGGAATGAAGGAC-3'(SEQ ID NO:162)

4. 5'-ATGAAGAAAGACAGGGAATGAAGGACCGCCAC -3'(SEQ ID NO:163)

5. 5'-CATGGAGGGCGACGAGCCGGTGTAGCTG-3'(SEQ ID NO:164)

6. 5'-ATCGACATGATTGGCACCCAGGTGTTG-3'(SEQ ID NO:165)

另外，引子包括(但不限於)以下反向定向寡核苷酸：

7. 5'-TTTCGACAAGCTCCAGAAAATCCCTAGAAAC-3'(SEQ ID NO:166)

8. 5'-ACAAGCTCCAGAAAATCCCTAGAAACAC-3'(SEQ ID NO:167)

9. 5'-TTCGACAAGCTCCAGAAAATCCCTAGAAACAC-3'(SEQ ID NO:168)

10. 5'-TGCTAAGAACATTCTTTTCGACAAGCTCC-3'(SEQ ID NO:169)

11. 5'-GAACATTCTTTTCGACAAGCTCCAGAAAATCC-3'(SEQ ID NO:170)

所有寡核苷酸引子係由 Integrated DNA Technologies

(IDT, Coralville, IA)合成且自其購得。

B. Hi II玉米細胞培養物

為獲得用於愈傷組織培養物起始之不成熟胚胎，在溫室生長之Hi-II親本A與B(Armstrong, C., Green, C.及Phillips, R. (1991) Maize Genet. Coop. News Lett. 65: 92-93)之間執行F₁雜交。自健康穗收集大小約為1.0-1.2 mm之胚胎(授粉後約9-10天)，且藉由用Liqui-Nox®皂洗刷來進行表面殺菌，浸入70%乙醇中2-3分鐘，隨後浸入20%商業漂白劑(0.1%次氯酸鈉)中30分鐘。

在無菌蒸餾水中沖洗穗，且將不成熟合子胚無菌切離且在15Ag10培養基(N6培養基(Chu C.C., Wang C.C., Sun C.S., Hsu C., Yin K.C., Chu C.Y.及Bi F.Y. (1975) Sci. Sinica 18:659-668)、1.0 mg/L 2,4-D、20 g/L蔗糖、100 mg/L酪蛋白水解產物(酶促消化)、25 mM L-脯胺酸、10 mg/L AgNO₃、2.5 g/L水晶洋菜(Gelrite)，pH 5.8)上培養2-3週，其中子葉盤背離培養基。以雙週間隔將展示預期形態學(Welter, ME, Clayton, DS, Miller, MA, Petolino, JF. (1995) Plant Cell Rep: 14:725-729)之組織選擇性轉移至新鮮15Ag10培養基上歷時約6週，隨後以雙週間隔轉移至4培養基(N6培養基，1.0 mg/L 2,4-D、20 g/L蔗糖、100 mg/L酪蛋白水解產物(酶促消化)、6 mM L-脯胺酸、2.5 g/L水晶洋菜，pH 5.8)歷時約2個月。

為起始胚胎發生懸浮液培養，將來源於單個胚胎之約3 ml紅血球容積(PCV)之愈傷組織添加至約30 ml之H9CP+液

體培養基中(MS基本鹽混合物(Murashige T., & Skoog F. (1962) *Physiol. Plant.* 15:473-497)、含有少10倍之菸鹼酸及高5倍之噻胺-HCl的經改質MS維生素、2.0 mg/L 2,4-D、2.0 mg/L α -萘乙酸(NAA)、30 g/L蔗糖、200 mg/L酪蛋白水解產物(酸消化)、100 mg/L肌醇(myo-inositol)、6 mM L-脯胺酸、5% v/v椰子汁(恰在次培養之前添加), pH 6.0)。在28°C下, 在設定於125 rpm之溫度受控搖動器中之125 ml錐形瓶中, 在暗條件下維持懸浮液培養物。在細胞系建立期間(2-3個月), 每3.5天, 藉由使用寬孔吸管將3 ml PCV之細胞及7 ml條件培養基添加至20 ml新鮮H9CP+液體培養基中來次培養懸浮液。如由生長加倍所證明達到成熟後, 將懸浮液按比例增容且維持於500 ml燒瓶中, 由此將12 ml PCV細胞及28 ml條件培養基轉移至80 ml H9CP+培養基中。完成懸浮培養物之建立後, 將等分試樣低溫保存以備將來使用。參見, WO 2005/107437。

C. DNA分離及擴增

在250 ml燒瓶中, 使如上所述之玉米HiII細胞培養物在標準GN6培養基(N6培養基、2.0 mg/L 2,4-D、30 g/L蔗糖、2.5 g/L水晶洋菜, pH 5.8)中生長, 且按製造商之推薦, 使用Qiagen(Valencia, CA)Plant DNeasy提取套組提取基因組DNA。在以下條件下進行使用上文所述之呈所有可能組合之引子的PCR擴增反應: 含有20 ng gDNA模板、20 pmol之各引子、1% DMSO及於酶製造商之緩衝液中之10個單位的Accuprime™ Pf聚合酶(Invitrogen, Carlsbad, CA)

之25 μ l反應容積。以由以下步驟組成之擴增循環產生大小在500 bp至2 kb範圍內之擴增產物：95°C -1'，(95°C -30"、57-62°C -30"、72°C -1') $\times$ 30，72°C -5'，4°C -保持。按製造商之推薦，使用來自Invitrogen(Carlsbad, CA)之TA選殖套組將經擴增片段直接選殖至載體pCR2.1(Invitrogen, Carlsbad, CA)中。

D. 序列分析

玉米自交5XH751及HiII細胞培養物中之IPP2-K基因之先前分析已指示存在構成小基因家族之2-3種不同基因(Sun等人, *in press, Plant Physiology*; WO 2006029296)。因此，按製造商之推薦，用來自Beckman Coulter(Fullerton, CA)之CEQ染料終止子循環定序套組將所分離選殖之片段定序。多個純系之序列分析顯示，已自HiII細胞分離出來源於玉米基因組之2個不同且先前已表徵之基因座的2個不同基因片段。

自HiII培養細胞分離之2個序列之比較指示，在預測編碼區域中，在兩種旁系同源物之間存在諸如單核苷酸多態現象(SNP)之小差異，而內含子及非編碼區域在核苷酸水平上顯著變化。應注意兩種旁系同源物之間的該等差異，因為其突出可由諸如鋅指域之序列依賴性DNA結合蛋白質辨別之序列區域。熟習此項技術者可設計結合於一基因序列且不結合於另一高度相似之基因序列之鋅指DNA結合域。將具有相應於所關注之旁系同源物之1.2 kb的部分基因序列(圖66)選作鋅指核酸酶蛋白質設計之模板且隨後使

其經受上文所述之鋅指 DNA 結合域分析。

實例 14：IPP2-K 鋅指 DNA 結合域之設計

使用針對 IPP2-K 鑑別出之標靶位點，選擇用於 IPP2-K 鋅指之識別螺旋。鋅指設計展示於下表 8 中：

表 8：IPP2-K 鋅指設計

ZFN 名稱	F1	F2	F3	F4	F5	F6
IPP2-K-1072a1	DRSALSR (SEQ ID NO:105)	RNDDRKK (SEQ ID NO:106)	RSDNLST (SEQ ID NO:107)	HSHARIK (SEQ ID NO:108)	RSDVLSE (SEQ ID NO:109)	QSGNLAR (SEQ ID NO:110)
IPP2-K-1072b1	DRSALSR (SEQ ID NO:105)	RNDDRKK (SEQ ID NO:106)	RSDNLAR (SEQ ID NO:111)	TSGSLTR (SEQ ID NO:112)	RSDVLSE (SEQ ID NO:109)	QSGNLAR (SEQ ID NO:110)
IPP2-K-1072c1	DRSALSR (SEQ ID NO:105)	RNDDRKK (SEQ ID NO:106)	TSGNLTR (SEQ ID NO:113)	TSGSLTR (SEQ ID NO:112)	RSDVLSE (SEQ ID NO:109)	QSGNLAR (SEQ ID NO:110)
IPP2-K-r1065a1	RSDHLSE (SEQ ID NO:114)	QSATRKK (SEQ ID NO:115)	ERGTLAR (SEQ ID NO:116)	RSDALTQ (SEQ ID NO:117)	無	無
IPP2-K-r1149a2	RSDSLSA (SEQ ID NO:118)	RSAALAR (SEQ ID NO:119)	RSDNLSE (SEQ ID NO:120)	ASKTRTN (SEQ ID NO:121)	DRSHLAR (SEQ ID NO:122)	無
IPP2-K-1156a2	RSDHLST (SEQ ID NO:123)	QSGSLTR (SEQ ID NO:124)	RSDHLSE (SEQ ID NO:114)	QNHHRIN (SEQ ID NO:125)	TGSNLTR (SEQ ID NO:126)	DRSALAR (SEQ ID NO:127)

鋅指設計之標靶位點展示於下表 9 中：

表 9：IPP2-K 鋅指之標靶位點

ZFN 名稱	標靶位點(5'至3')
IPP2-K-1072a1	GAAGTGGTTGAGTCGGTC (SEQ ID NO:128)
IPP2-K-1072b1	GAAGTGGTTGAGTCGGTC (SEQ ID NO:129)
IPP2-K-1072c1	GAAGTGGTTGAGTCGGTC (SEQ ID NO:129)
IPP2-K-r1065a1	ATGGCCCCACAG (SEQ ID NO:130)
IPP2-K-r1149a2	GGCACCCAGGTGTTG (SEQ ID NO:131)
IPP2-K-1156a2	GTCGATGGTGGGGTATGG (SEQ ID NO:132)

將IPP2-K設計併入至編碼具有CCHC結構之蛋白質之鋅指表現載體中。參見，上文表1至4。隨後將非標準鋅指編碼序列融合於IIS型限制酶*FokI*之核酸酶域(Wah等人(1998) Proc. Natl. Acad. Sci. USA 95:10564-10569之序列之胺基酸384-579，經由4個胺基酸ZC連接子)以形成IPP2-K ZFN。

實例15：使用IPP2-K鋅指核酸酶之基因校正

以Urnov (2005) Nature 435 (7042) :646-51及美國專利公開案第20050064474號中所述之GFP系統(例如實例6-11)，測試如本文中所述之IPP2-K ZFN促進同源重組之能力。簡言之，按照Invitrogen Lipofectamine 2000方案，每個樣本使用2 μ l之Lipofectamine 2000，將50 ng各ZFN及500 ng少啟動子GFP供體(Urnov (2005) Nature)轉染至500,000個報導體細胞中。

轉染後24小時，以0.2 μ M最終濃度添加長春鹼，且在轉染後72小時移除。

轉染後5天，對於每次轉染藉由在Guava桌上型FACS分析儀上量測40,000個細胞來檢定細胞之GFP表現。結果展示於圖69中。

實例16：在玉米HiII細胞中表現C3H1 ZFN

A.載體設計

建構用於在玉米細胞中表現ZFN蛋白質之質體載體。為最佳化形成功能性鋅指核酸酶雜二聚體所需之兩種不同蛋白質之表現及相對化學計量，採用在單一載體上產生由單個啟動子驅動之兩種ZFN單體之開放閱讀框架的插入之表

現策略。該策略利用來源於Thesoa assigna病毒之2A序列 (Mattion, N.M., Harnish, E.C., Crowley, J.C. & Reilly, P.A. (1996) J. Virol. 70, 8124-8127)、來自 *opaque-2* 基因 (op-2) 之玉米核定位 (NLS) 信號 (Maddaloni, M., Di Fonzo, N., Hartings, H., Lazzaroni, N., Salamini, F., Thompson, R., & Motto M. (1989) Nucleic Acids Research Vol. 17(18):7532) 及來源於玉米泛素-1基因之啟動子 (Christensen A.H., Sharrock R.A., & Quail P.H. (1992) Plant Mol Biol. 18(4):675-89) 之功能性。設計逐步模組化選殖流程以開發用於選自保存或重新合成之庫之任何給定ZFN編碼基因對的該等表現載體。

首先，修飾pVAX載體(參見(例如)美國專利公開案2005-0267061；其揭示內容以引用之方式併入本文中)以涵蓋如圖65中圖A至E所示之N末端表現域。該經修飾質體(pVAX-N2A-NLSop2-EGFP-FokMono)(圖65A)之特徵包括編碼來源於玉米op-2(RKRKESNRESARRSRYRK, SEQ ID NO:133)之NLS之經再設計及合成的區段，及編碼利用偏好玉米密碼子之FokI核酸酶域之經再設計及合成的區段。另外，獨特XhoI位點下游之單核苷酸插入(C)產生一額外SacI位點以便於選殖。

其次，pVAX載體(參見(例如)美國專利公開案2005-0267061)亦經修飾以涵蓋C末端表現域。該經修飾質體(pVAX-C2A-NLSop2-EGFP-FokMono)(圖65B)之特徵包括編碼來源於玉米op-2(RKRKESNRESARRSRYRK, SEQ ID

NO:133)之NLS之經再設計及合成的區段，及編碼利用偏好玉米密碼子之FokI核酸酶域之經再設計及合成的區段。另外，為隨後連接兩種蛋白質編碼域之目的，在ZFN ORF之N末端引入來自 *Thomomys caligaris* 病毒之2A序列(EGRGSLTCTCGDVEENPGP，SEQ ID NO:134)。

經由連接反應，使用限制酶KpnI及BamHI，將編碼個別鋅指蛋白之ORF之基因卡匣選殖至N2A或C2A載體中，以產生相容末端。接著，經由相同限制位點，將來自C2A載體之BglII/XhoI片段插入N2A載體中，產生含有包括側接有NcoI及SacI限制位點之2 ZFN編碼域之卡匣的中間構築體。

最後，經由使用彼等酶之限制作用，切離來自該中間建構(圖65C)含有兩種ZFN基因之NcoI/SacI卡匣，且將其連接至質體骨架pDAB3872(圖65D)中。所得質體包括ZFN基因加相關啟動子及終止子序列，加用於質體維持之選擇標記。

在最終建構中，其實例展示於圖65E中，ZFN表現卡匣(包括啟動子及終止子元件)側接有attL位點，以便於使用來自Invitrogen(Carlsbad, CA)之Gateway系統加以操作。將使用該選殖流程產生之ZFN構築體之各者轉型至 *E. coli* DH5a細胞(Invitrogen, Carlsbad, CA)中且隨後在適當選擇下維持。

B. DNA傳遞及瞬間表現

如圖65E中所述建構之ZFN表現載體之質體製劑係按製

造商之推薦，使用來自 Qiagen (Valencia, CA) 之無核酸內切酶 Gigaprep 套組，自生長於 LB 培養基加抗生素中之大腸桿菌細胞之 2 L 培養物產生。質體 DNA 使用多種方法直接傳遞至玉米 HiII 培養細胞中。

在一實例中，使玉米細胞經受經由 Whiskers™ 之 DNA 傳遞。DNA 傳遞之前約 24 小時，將 3 ml PCV 之 HiII 玉米懸浮細胞加 7 ml 條件培養基於 125 ml 錐形瓶中次培養至 20 ml 之 GN6 液體培養基 (缺乏水晶洋菜之 GN6 培養基) 中，且在 28 °C 下，置放於 125 rpm 之搖動器上歷時 24 小時。移除 2 mL PCV 且添加至 125 mL 錐形瓶中之 12 ml GN6 S/M 滲透培養基 (N6 培養基、2.0 mg/L 2,4-D、30 g/L 蔗糖、45.5 g/L 山梨糖醇、45.5 g/L 甘露醇、100 mg/L 肌醇，pH 6.0) 中。在 28 °C 下，在適度攪動 (125 rpm) 下，在黑暗中將燒瓶培養 30-35 分鐘。在該時間期間，藉由將適當體積之 GN6 S/M 液體培養基添加至預稱重無菌鬚晶中來製備 50 mg/ml 碳化矽鬚晶 (Advanced Composite Materials, Inc., Eureka Springs, AK) 之懸浮液。在 GN6 S/M 中培養後，將各燒瓶之內含物傾倒至 15 mL 錐形離心管中。

細胞沈降後，轉移接近 1 mL GN6 S/M 且收集於 125 mL 燒瓶中以備將來使用。在最大速度下，將鬚晶之預濕潤懸浮液渦旋 60 秒，使用寬孔過濾吸管頭將 160 µL 添加至離心管中，且添加 20 µg DNA。將管用 '手指渦旋'，且立即置放於經改進以固持 17×100 mm 培養管之 Caulk 'Vari-Mix II' 牙科混汞器中，且隨後在中間速度下攪動 60 秒。攪動後，使

細胞、培養基、鬚晶及DNA之混合物連同18 ml額外GN6液體培養基一起返回至錐形瓶。在28°C下，在黑暗中，使細胞在125 RPM之搖動器上恢復2小時。

使用連接至家用真空管線之玻璃槽收集單元將約5-6 mL之分散懸浮液過濾至Whatman #4濾紙(5.5 cm)上，以使得每個樣本獲得5-6張濾紙。將濾紙置放於具有GN6培養基之60×20 mm培養板上，且在28°C下，在黑暗條件下培養。24、48或72小時後，自2-5張濾紙刮下細胞，收集至管中，置放於乾冰上，且隨後在-80°C下冷凍。

在DNA傳遞之另一實例中，使用自儀器製造商之方案改適之微彈轟擊技術將經純化之無核酸內切酶質體製劑直接傳遞至玉米細胞。用Biolistic PDS-1000/He™系統(Bio-Rad Laboratories, Hercules, CA)進行所有轟擊。為進行粒子塗佈，用100%乙醇將3 mg之直徑1.0微米之金粒子洗滌一次，用無菌蒸餾水洗滌兩次，且再懸浮於矽化Eppendorf管中之50 µl水中。將5微克質體DNA、20 µl亞精胺(0.1 M)及50 µl氯化鈣(2.5 M)添加至金懸浮液中。在室溫下，將混合物培養10 min，在10K rpm下粒化10 s，再懸浮於60 µl冷100%乙醇中，且將8-9 µl分布於各巨載劑上。為製備用於轟擊之細胞，自次培養後3天之液體培養物移除細胞叢集且置放於皮氏培養皿中直徑為2.5 cm呈圓形之由生長培養基加各為0.256 M之甘露醇及山梨糖醇組成的滲透培養基上。在轟擊之前，將細胞在滲壓劑中培養4 h。轟擊在上文所述之儀器中進行，該儀器藉由在1100 psi及27 in

之Hg真空之條件下，將組織置放於中間支架上，且按照操作手冊來使用。在處理後24小時之時間點，收集經轟擊細胞叢集，於液態N₂中冷凍且在-80℃下儲存。

DNA傳遞及ZFN於玉米細胞中之瞬間表現之另一實例涉及利用原生質體製劑。使用自Mitchell及Petolino (1991) *J. Plant. Physiol.* 137: 530-536及Lyznik等人(1995) *Plant J.* 8(2): 177-186)改進之方法，自HiII玉米細胞培養物製備原生質體。次培養後48小時(中間對數期生長)，藉由在1000 rpm下離心5分鐘收集懸浮培養物。移除培養基且在10 ml W5培養基(154 mM NaCl₂ ; 125 mM CaCl₂·H₂O ; 5 mM KCl₂ ; mM葡萄糖 ; pH 5.8)中輕輕洗滌5 ml包裝PCV。

經由在100 rpm下離心5分鐘收集經洗滌細胞且隨後在於25 ml經過濾滅菌K3培養基(2.5 g KNO₃ ; 250 mg NH₄NO₃ ; 900 mg CaCl₂(二水合物) ; 250 mg Mg₂SO₄ ; 250 mg NH₄SO₄ ; 150 mg NaPO₄(一元) ; 250 mg木糖 ; 10 ml硫酸亞鐵/螯合物儲備液(F318) ; 1 ml B5微量營養素(1000×儲備液-750 mg碘化鉀 ; 250 mg鉬酸(鈉鹽)脫水物 ; 25 mg氯化鈷 ; 25 mg硫酸銅) ; 10 ml K3維生素(100×儲備液-1 g肌醇 ; 10 mg吡哆醇HCl ; 100 mg噻胺HCl ; 10 mg菸鹼酸) ; +0.6M 甘露醇 ; pH=5.8]中含有3%纖維素酶Y-C+0.3%果膠酶Y23(Karlan Research Products Corp., Cottonwood, AZ)之酶混合物中培養。在25℃下，在輕微攪動(50 rpm)下，將細胞培養5-6小時以消化二級植物細胞壁。

細胞壁降解後，經由100微米細胞濾器過濾酶-細胞混合物且用等體積之K3+0.6 M甘露醇培養基洗滌含有原生質體及細胞碎片之流過物(flow-through)。在800 rpm下，將原生質體離心5分鐘，棄去上清液且重複洗滌。將原生質體小球洗滌再懸浮於20 ml K3+0.6 M甘露醇+9% Ficoll 400溶液中。將10 ml該溶液分配至2個無菌塑膠管中且將2 ml之TM培養基(19.52 g MES；36.45 g甘露醇；40 ml 2 M $\text{CaCl}_2 \cdot \text{H}_2\text{O}$ 儲備液；pH=5.5)輕輕覆蓋於懸浮液上，形成不連續梯度。

經由在800 rpm下離心5分鐘將活原生質體自非活原生質體、細胞碎片及完整懸浮細胞分離。用吸管移除在梯度介面形成之不同原生質體帶且用10 ml新鮮TM溶液洗滌，接著在800 rpm下離心5分鐘。將所得原生質體小球再懸浮於1 ml之TM培養基中，且在血球計中用25 mg/mg螢光素二乙酸酯(FDA)染色來量化活原生質體之數量。將原生質體溶液調整至於每毫升TM培養基中 1×10^7 個原生質體之最終濃度。

將約 1×10^6 個原生質體(100 μl)轉移至含有10-80 μg 純化質體DNA之2 ml Eppendorf管中。逐滴添加100 μl 之40% PEG-3350(Sigma Chemical Co., St. Louis, MO)溶液，且輕輕混合懸浮液。在室溫下，將原生質體/DNA混合物培養30分鐘，接著用1 ml GN6生長培養基逐滴稀釋。在25℃下，在該培養基中將經稀釋原生質體培養24小時，且隨後收集，冷凍於液態 N_2 中且在-80℃下儲存。

實例 17：活體內 ZFN 功能性

在該實例中之 ZFN 之功能性應理解為包括(但不限於)ZFN 在作物物種細胞中表現之能力，及 ZFN 經由識別、結合且裂解其所要標靶來在彼作物之內源基因組中介導雙股斷裂的能力。在該實例中，亦應瞭解，ZFN 之標靶為作物基因組內之內源基因座及構形中之基因。

為評估工程化 ZFN 是否具有針對基因組環境中預測之標靶基因之功能性，使用基於 DNA 序列之檢定。預測 ZFN 誘導之雙股 DNA 斷裂以誘導諸如非同源末端聯接(NHEJ)之修復機制(由 Cahill 等人，(2006) *Mechanisms Front Biosci.* 1(11): 1958-76 回顧)。NHEJ 之一結果為一定比例之斷裂 DNA 鏈將以不完全方式修復，在裂解位點產生小的缺失、插入或取代。熟習此項技術者可經由多種方法偵測 DNA 序列中之該等改變。

A. 基於 PCR 之選殖及定序

在一實例中，在轉型後 24 小時分離表現 ZFN 蛋白質之玉米 HiII 培養細胞，冷凍且按製造商之推薦，使其經受使用 Qiagen (Valencia, CA) Plant DNeasy 提取套組之基因組 DNA 提取。使用對標靶基因具有特異性且側接 ZFN 之預測裂解位點之寡核苷酸引子來進行 PCR 擴增。組合使用對靶向 IPP2-K 基因旁系同源物具有特異性之正向定向 PCR 引子 (5'-GGAAGCATTATTCCAATTTGATGATAATGG-3') (SEQ ID NO:135) 及反向定向 PCR 引子 (5'-CCCAAGTGTCGAGGTTGTCAATATGTTAC-3') (SEQ ID NO:136)，以在以下條件下

擴增經純化之基因組DNA：含有20 ng gDNA模板、20 pmol之各引子、1% DMSO及於酶製造商之緩衝液中之10個單位的Accuprime Pf聚合酶(Invitrogen, Carlsbad, CA)之25 µl反應容積。由以下步驟組成之擴增循環產生具有預期大小之擴增產物：95°C -1'，(95°C -30"、61°C -30"、72°C -1')×30，72°C -5'，4°C -保持。

使用來自Invitrogen(Carlsbad, CA)之TA選殖套組將所擴增片段直接選殖至載體pCR2.1(Invitrogen, Carlsbad, CA)中。按製造商之推薦，以96孔格式，用來自Beckman Coulter(Fullerton, CA)之CEQ染料終止子循環定序套組將經分離選殖之片段定序。在該實驗中，如圖66中所示，預測ZFN蛋白質結合於2個短的IPP2-K基因特異性序列，以產生裂解ds-DNA之雜二聚核酸酶。

來自多個純系之定序結果之分析揭示，純系#127恰好在ZFN之預測裂解位點處含有小缺失，指示NHEJ機制已介導DNA序列在彼位點處之不完全修復(圖67)。

該等結果證明，該等工程化ZFN在作物物種內之內源基因基因座處以特異性方式誘導靶向、雙股斷裂之能力。

B.大規模平行定序分析

在另一實例中，應用PCR及大規模平行焦磷酸定序方法之組合來探查表現靶向針對該相同序列之不同ZFN蛋白質之多個細胞樣本的基因組。合成(IDT, Coralville, IA)其中XXX=GGG、CCC或GGC之正向定向PCR引子(5'-XXXCACCAAGTTGTATTGCCTTCTCA-3')(SEQ ID NO:137)

之3種變體及其中XXX=GCC、CCG或CGG之反向定向PCR引子(5'-XXXATAGGCTTGAGCCAAGCAATCTT-3')(SEQ ID NO:138)之3種變體。各引子之5'-末端處之3-bp標記用作標識符且指示擴增子(amplicon)來源於何細胞樣本。組合使用具有匹配標識符標記(鍵標)之引子對，以在以下條件下擴增來源於玉米細胞樣本之經純化基因組DNA：含有40 ng gDNA模板、20 pmol之各引子、1% DMSO及於酶製造商之緩衝液中之10個單位的Accuprime Pf聚合酶(Invitrogen, Carlsbad, CA)之50 µl反應容積。由以下步驟組成之擴增循環產生具有預期大小之擴增產物：95°C -1'，(95°C -30"、65°C -30"、72°C -1')×30，4°C -保持，且按製造商之推薦，使用Qiagen's(Valencia, CA)MinElute PCR純化套組來純化。

如(Margulies等人(2005) Nature 437: 376-380)中所述，藉由454 Life Sciences(Branford, CT)直接對PCR產物執行大規模平行焦磷酸定序反應(亦稱為454定序)。藉由鑑別在DNA分子內含有預期大小及位置之缺失的序列讀取結果來進行454定序結果之分析。

該等分析之結果指示，如圖68中所示，在該等ZFN之預期裂解位點存在多個小的缺失。該等缺失精確定位至ZFN標靶位點且指示由ZFN誘導之ds斷裂在基因組中產生且隨後由NHEJ修復。該等結果進一步證明，該等工程化ZFN在作物物種內之內源基因基因座處以特異性方式誘導靶向、雙股斷裂之能力。

實例 18：用於靶向整合之供體 DNA 設計

在該實例中，供體 DNA 理解為包括傳遞至植物細胞中且併入核基因組中之雙股 DNA 分子。該等併入所發生之機制可為經由不依賴於同源性之非同源末端聯接 (NHEJ；藉由 Cahill 等人，(2006) *Mechanisms Front Biosci.* 1: 1958-76 評論) 或在核 DNA 中之雙股斷裂位點處的另一類似機制。供體 DNA 於基因組中之該 NHEJ 驅動、連接樣併入稱為隨機整合，此係因為供體 DNA 之整合位置主要由雙股 DNA 斷裂之存在確定。在該機制中，供體 DNA 整合至基因組中不依賴於斷裂位點處基因組之核苷酸序列或供體自身之核苷酸序列。因此，在隨機整合期間，基因組中供體 DNA 併入之"地址"並非基於供體 DNA 之序列規定及預測。隨機整合為經由土壤桿菌或基因槍介導之 DNA 傳遞至活植物細胞中進行標準植物轉型期間供體 DNA 之基因轉殖發生之主要機制。

與隨機整合相對比，供體 DNA 亦可經由靶向整合併入基因組中。靶向整合理解為在雙股斷裂之位點(位置)處，經由諸如同源依賴性單鏈黏接或同源重組之同源依賴機制性發生(在 van den Bosch 等人(2002) *Biol Chem.* 383(6): 873-892 中回顧)。在同源依賴性 DNA 斷裂修復之狀況下，含有具有與斷裂位點處之 DNA 之一致性或相似性的核苷酸序列之供體 DNA 可在彼位點處併入。因此，供體 DNA 整合至基因組中之"地址"係依賴於基因組與供體 DNA 分子之間的核苷酸序列一致性或序列相似性。在植物系統中，已知 DNA

中雙股斷裂之修復係利用 NHEJ 及同源依賴性途徑(在 Puchta (2005) *J. Exp. Bot.* 56: 1-14 中回顧)。

在該實例中，吾人描述欲經由在由序列特異性 ZFN 蛋白質誘導之雙股斷裂位點處之靶向整合來整合至基因組中的供體 DNA 分子之設計及建構。不同 ZFN 蛋白質可在標靶基因序列中之不同核苷酸處誘導雙股斷裂；所誘導雙股斷裂之特異性位點稱為位置。

如實例 13 中所述，吾人已表徵標靶基因(來自玉米之 IPP2K)之核苷酸序列。隨後，吾人設計結合於彼標靶基因之特定鹼基之 ZFN 蛋白質(實例 14)且驗證其在兩種異源系統中標靶基因內之彼序列處及針對玉米細胞中內源基因之結合/裂解活性(實例 15-17)。在此，吾人描述經設計以在 IPP2K 基因中 ZFN 介導之雙股斷裂位置處，經由靶向整合併入至玉米基因組中之各種供體分子之建構。熟習此項技術者可建構經設計以在核苷酸序列已知且預測該序列含有雙股斷裂之任何基因組中之任何位置處，經由同源性驅動靶向整合併入至 ZFN 誘導雙股斷裂中之供體 DNA 分子。

在本文中所述之一實施例中，供體 DNA 分子包含由與所靶向位置處之標靶基因 IPP2K 之序列相同的核苷酸序列區段界定之自主除草劑耐受性基因表現卡匣。在該實施例中，自主除草劑耐受性卡匣理解為包括含有已知在植物細胞中具有功能性之啟動子、除草劑耐受性基因及終止子序列之完整啟動子轉錄單元(PTU)。熟習此項技術者可選擇任何啟動子、基因及終止子組合以建構自主 PTU。在所指

示位置處與玉米中標靶基因(IPP2K)之序列具有一致性的DNA片段亦包括於該質體構築體上。該等片段用作供體DNA之"同源性側接序列"且在規定位置處經由靶向整合將該供體直接併入至標靶基因中。同源性側接序列以相對於PTU之正確5'至3'定向，置放於PTU之上游及下游。熟習此項技術者可預想在供體DNA建構中，具有不同大小及定向之同源性側接序列。

在本文中所述之另一實施例中，供體DNA分子包含含有由與標靶位置處IPP2K之序列相同的核苷酸序列區段界定之非自主除草劑耐受性基因表現卡匣之質體建構。在該實施例中，非自主除草劑耐受性卡匣理解為包括缺乏功能性啟動子之不完整啟動子轉錄單元(PTU)。非自主PTU含有已知在植物細胞中具有功能性之除草劑耐受性基因及終止子序列。熟習此項技術者可選擇任何基因及終止子組合以建構非自主PTU。在非自主供體之該實例中，除草劑耐受性基因之表現係依賴於供體區段併入至鄰近可驅動彼基因表現之功能性啟動子的基因組位置中。可預想相對罕有之情況，其中供體將經由隨機整合併入至偶見啟動子駐留且可用以驅動除草劑耐受性基因表現之遺傳基因座中。或者，基於在供體DNA建構中存在具有適當長度、與玉米中規定位置處之標靶基因的序列具有一致性之DNA片段之同源性側接序列，在規定位置處將供體DNA精確靶向整合至標靶基因中可發生(如對自主供體所述)且因此利用該標靶基因之內源啟動子。在該實施例中，同源性側接序列以相

對於PTU之正確5'至3'定向，置放於PTU之上游及下游。熟習此項技術者可預想在供體DNA建構中，具有不同大小及定向之同源性側接序列。

在本文所述之2個實施例(自主及非自主供體設計)中，質體建構通常含有使得除草劑耐受性基因能夠選殖、表現且後續分析之其他元件。該等元件包括細菌複製起點、工程化限制位點等且如下文所述。熟習此項技術者可預想利用構成供體DNA分子之不同元件。

A. 細菌菌株及培養條件

在37°C下，使用Luria-Bertani肉湯(LB:10 g/L Bacto-胰化蛋白、10 g/L NaCl、5 g/L Bacto-酵母提取物)、LB瓊脂(LB肉湯加15 g/L Bacto-瓊脂)，或Terrific肉湯(TB:12 g/L Bacto-胰化蛋白、24 g/L Bacto-酵母提取物、0.4% v/v甘油、17 mM KH_2PO_4 、72 mM K_2HPO_4)，使大腸桿菌菌株(One Shot® Top 10化學感受態細胞；MAX Efficiency® DH5α™化學感受態細胞，Invitrogen Life Technologies, Carlsbad, CA)生長16小時。在200 rpm下搖動液體培養物。按需要將氯黴素(Chloramphenicol)(50 µg/ml)、康黴素(kanamycin)(50 µg/ml)或安比西林(100 µg/ml)添加至培養基中。用於該研究中之所有抗生素、培養基及緩衝劑均購自Sigma-Aldrich Corporation(St. Louis, MO)或Difco Laboratories(Detroit, MI)。

B. 質體骨架位置-1

將含有IPP2K之位置-1之同源性側接序列的質體骨架工

程化以允許將任何供體DNA序列整合至IPP2K基因之相應標靶位點中。熟習此項技術者可預想使用各種選殖位點、模組化設計元件及與所關注基因組內之任何標靶序列同源之序列的質體骨架。在此所例證之質體骨架係以基本質體載體pBC SK(-)噬質體(phagemid)(3.4 Kbp)(Stratagene, La Jolla, CA)開始來產生。使用如下文所述之4步合成建構位置-1質體骨架。

在步驟#1中，製備基本質體。在37°C下，使用10個單位之*Spe* I及10個單位之*Not* I(New England Biolabs, Beverly, MA)限制性核酸內切酶將3 µg pBC SK(-)線性化1 hr。在100 V下，在補充有0.5%溴化乙錠(Sigma-Aldrich Corporation, St. Louis, MO)之1.0% TAE(0.04 M Tris-乙酸鹽、0.002 M EDTA)瓊脂糖凝膠中，將限制性DNA電泳1 hr。用UV光使DNA片段可見且藉由與1 Kbp DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較來估算片段大小。凝膠切離3.4 Kbp經*Spe* I/*Not* I消化之次選殖載體pBC SK(-)，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)來加以純化。

在步驟#2中，分離來自IPP2K位置-1之5' & 3'-同源性側接序列。由Integrated DNA Technologies, Inc.(Coralville, IA)，在標準脫鹽條件下合成以下寡核苷酸引子，且用水稀釋至0.125 µg/µl之濃度：5'-GCGGCCGCGTCTCACC GCGGCTTGGGGATTGGATACGGAGCT-3'(SEQ ID NO:143)
5'-ACTAGTGATATGGCCCCACAGGAGTTGCTCATGACTTG-

3'(SEQ ID NO:144)

5'-ACTAGTCCAGAACTGGTTGAGTCGGTCAAACAAGATTGCT-

3'(SEQ ID NO:145)

5'-GTCGACCTTGATGCTACCCATTGGGCTGTTGT-3'(SEQ

ID NO:146)

PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japan提供且由以下各物組成之試劑來進行：5 μ l 10 $\times$ LA PCR™緩衝液II(Mg²⁺)、20 ng 雙股gDNA模板(玉米HiII)、10 pmol正向寡核苷酸引子、10 pmol反向寡核苷酸引子、8 μ l dNTP混合物(各2.5 mM)、33.5 μ l H₂O、0.5 μ l(2.5個單位)TaKaRa LA Taq™ DNA聚合酶、1滴礦物油。使用Perkin-Elmer Cetus, 48-樣本DNA熱循環儀(Norwalk, CT)，在以下循環條件下執行PCR反應：94℃，4 min/1次循環；98℃ 20 sec，65℃ 1 min，68℃ 1 min/30次循環；72℃，5 min/1次循環；4℃/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將15 μ l之各PCR反應電泳1 hr。用UV光使經擴增片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。藉由0.821 Kbp之DNA片段(5'-同源性側接序列)或0.821 Kbp之DNA片段(3'-同源性側接序列)之存在來診斷預期擴增產物。

凝膠切離該等片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。隨後，根據製造商之方案，使用TOPO TA Cloning®

套組(用於pCR®2.1載體)及One Shot® TOP10化學感受態大腸桿菌細胞(Invitrogen Life Technologies, Carlsbad, CA)，將經純化片段選殖至pCR2.1質體中。

將個別菌落接種至含有補充有50 µl/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm之搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000×g下粒化1 min。移除上清液，且如上所述使用NucleoSpin®質體套組(BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位*Spe* I及*Not* I消化3 µg之來自5'-同源性側接序列純系質體之經分離質體。用10個單位*Spe* I及20個單位*Sal* I(New England Biolabs, Beverly, MA)消化3種主要同源性側接序列純系質體。在37°C下，將所有質體消化產物培養1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。除3.9 Kbp pCR®2.1載體外，藉由0.821 Kbp之插入DNA片段(5'-同源性側接序列)或0.821 Kbp之插入DNA片段(3'-同源性側接序列)之存在來診斷預期質體純系。

如由製造商所述，使用CEQ™ DTCS-快速啟動套組(Beckman-Coulter, Palo Alto, CA)執行質體純系之雙股定序反應。如由製造商方案所述，使用Performa DTR凝膠過濾濾筒(Edge BioSystems, Gaithersburg, MD)純化反應。在

Beckman-Coulter CEQ™ 2000 XL DNA分析系統上分析序列反應，且使用 Sequencher™ 4.1.4版(Gene Codes Corporation, Ann Arbor, MI)執行核苷酸表徵。相應於來源於IPP2K之位置-1 5'-同源性側接序列之0.821 Kbp片段之序列展示於圖87(SEQ ID NO:171)中。相應於來源於IPP2K之位置-1 3'-同源性側接序列之0.821 Kbp片段之序列展示於圖88(SEQ ID NO:172)中。

在步驟#3中，將位置-1 5'-同源性側接序列連接至基本質體中。凝膠切離相應於含有正確位置-1 5'-同源性側接序列之純系之限制性片段，且根據製造商之指南，使用 QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)純化。隨後，以1:5載體:插入物比率，使用500個單位之T4 DNA連接酶(Invitrogen Life Technologies, Carlsbad, CA)，在20 µl之反應體積中，在於16°C水浴中之16 hr培養之條件下，將相應於位置-1 5'-同源性側接序列(0.821 Kbp)之片段連接至經*Spe* I/*Not* I消化(步驟#1)之經純化基本質體中。隨後以5 µl之連接反應轉型大腸桿菌One Shot® Top 10化學感受態細胞(Invitrogen Life Technologies, Carlsbad, CA)且在製造商所述之選擇條件下塗鋪。將個別菌落接種至含有補充有50 µl/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm搖動下培養16小時。

培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000×g下粒化

1 min。移除上清液，且如上所述使用 NucleoSpin® 質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 *Spe* I 及 *Not* I (New England Biolabs, Beverly, MA) 消化 3 µg 之經分離質體 DNA，且在 37°C 下培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。除 3.4 Kbp 基本質體外，由 0.821 Kbp 之插入 DNA 片段 (5'-同源性側接序列) 之存在來診斷預期質體純系。

在步驟 #4 中，將位置 -1 3'-同源性側接序列連接至步驟 #3 產物中。在 37°C 下，使用 10 個單位之 *Spe* I 及 20 個單位之 *Sal* I (New England Biolabs, Beverly, MA) 限制性核酸內切酶將 3 µg 步驟 #3 中所述之工程化產物線性化 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠 (Sigma-Aldrich Corporation, St. Louis, MO) 之 1.0% TAE (0.04 M Tris-乙酸鹽、0.002 M EDTA) 瓊脂糖凝膠中，將限制性 DNA 電泳 1 hr。用 UV 光使 DNA 片段可見且藉由與 1 Kbp DNA 梯級 (Invitrogen Life Technologies, Carlsbad, CA) 比較估算片段大小。凝膠切離來自步驟 #3 之約 4.2 Kbp 經 *Spe* I/*Sal* I 消化之產物，且根據製造商之指南，使用 QIAquick 凝膠提取套組 (QIAGEN Inc., Valencia, CA) 加以純化。

隨後，將步驟 #2 中產生之 3'-同源性側接序列供體之經分離片段 (0.821 Kbp) 與經 *Spe* I/*Sal* I 消化之步驟 #3 產物組合，且如上所述，在使用 1:5 載體:插入物比率及 500 個單位 T4

DNA連接酶(Invitrogen Life Technologies, Carlsbad, CA)之20 μ l連接反應中純化。在16°C水浴中，將連接反應培養16 hr。連接後，按製造商之推薦，將5 μ l連接反應轉型至MAX Efficiency® DH5 α [™]化學感受態細胞(Invitrogen Life Technologies, Carlsbad, CA)中。將個別菌落接種至含有補充有50 μ l/ml氯黴素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中。

在37°C下，在200 rpm搖動下，將培養物培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000 $\times$ g下粒化1 min。移除上清液，且如上所述使用NucleoSpin[®]質體套組(BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位*Sal* I及*Not* I.(New England Biolabs, Beverly, MA)消化3 μ g之經分離質體，且在37°C下培養1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。由具有1.64 Kbp(插入物)及3.33 Kbp(基本質體)之2個DNA片段之存在來診斷預期純系。給予所得質體名稱pDAB7471(圖70)。

C.質體骨架位置-2

將含有IPP2K之位置-2之同源性側接序列的質體骨架工程化以允許將任何供體DNA序列整合至IPP2K基因之相應標靶位點中。熟習此項技術者可預想使用各種選殖位點、模組化設計元件及與所關注基因組內任何標靶序列之序列

同源的質體骨架。在此所例證之質體骨架以基本質體載體 pBC SK(-)噬質體(3.4 Kbp)(Stratagene, La Jolla, CA)開始來產生。使用如下文所述之4步合成建構位置-2質體骨架。

在步驟#1中，製備基本質體。在37°C下，使用10個單位之 *Spe* I及10個單位之 *Not* I(New England Biolabs, Beverly, MA)限制性核酸內切酶將3 µg pBC SK(-)線性化1 hr。在100 V下，在補充有0.5%溴化乙錠(Sigma-Aldrich Corporation, St. Louis, MO)之1.0% TAE(0.04 M Tris-乙酸鹽、0.002 M EDTA)瓊脂糖凝膠中，將限制性DNA電泳1 hr。用UV光使DNA片段可見且藉由與1 Kbp DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較估算片段大小。凝膠切離3.4 Kbp經 *Spe* I/*Not* I消化之次選殖載體 pBC SK(-)，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)來加以純化。

在步驟#2中，分離來自IPP2K位置-2之5' & 3'-同源性側接序列。由Integrated DNA Technologies, Inc.(Coralville, IA)，在標準脫鹽條件下合成以下寡核苷酸引子，且用水稀釋至0.125 µg/µl之濃度：

5'-GCGGCCGCTAGATAGCAGATGCAGATTGCT-3'(SEQ ID NO: 147)

5'-ACTAGTATTGGCACCCAGGTGTTGGCTCA-3'(SEQ ID NO: 148)

5'-ACTAGTCATGTCGATGGTGGGGTATGGTTCAGATTCAG-

3'(SEQ ID NO:149)

5'-GTCGACGTACAATGATTTTCAGGTTACGGCCTCAGGAC-

3'(SEQ ID NO:150)

PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japan提供且由以下各物組成之試劑來進行：5 μ l 10 $\times$ LA PCR™緩衝液II(Mg²⁺)、20 ng 雙股gDNA模板(玉米HiII)、10 pmol正向寡核苷酸引子、10 pmol反向寡核苷酸引子、8 μ l dNTP混合物(各2.5 mM)、33.5 μ l H₂O、0.5 μ l(2.5個單位)*TaKaRa LA Taq*™ DNA聚合酶、1滴礦物油。使用Perkin-Elmer Cetus, 48-樣本DNA熱循環儀(Norwalk, CT)，在以下循環條件下執行PCR反應：94℃，4 min/1次循環；98℃ 20 sec，55℃ 1 min，68℃ 1 min/30次循環；72℃，5 min/1次循環；4℃/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將15 μ l之各PCR反應電泳1 hr。用UV光使經擴增片段可見且藉由與1 Kbp DNA梯級比較來估算片段大小。藉由0.855 Kbp之DNA片段(5'-同源性側接序列)或0.845 Kbp之DNA片段(3'-同源性側接序列)之存在來診斷預期擴增產物。凝膠切離該等片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。隨後，根據製造商之方案，使用TOPO TA Cloning®套組(用於pCR®2.1載體)及One Shot® TOP10 化學感受態大腸桿菌細胞(Invitrogen Life Technologies, Carlsbad, CA)，將經純化片段選殖至pCR2.1質體中。

將個別菌落接種至含有補充有50 μ l/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000 $\times$ g下粒化1 min。移除上清液，且如上所述使用NucleoSpin[®]質體套組(BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位*Spe* I及*Not* I消化3 μ g之來自5'-同源性側接序列純系質體之經分離質體。用10個單位*Spe* I及20個單位*Sal* I (New England Biolabs, Beverly, MA)消化3種主要同源性側接序列純系質體。在37°C，將所有質體消化產物培養1 hr。

在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。除3.9 Kbp pCR[®]2.1載體外，藉由0.855 Kbp之插入DNA片段(5'-同源性側接序列)或0.845 Kbp之插入DNA片段(3'-同源性側接序列)之存在來診斷預期質體純系。

如由製造商所述，使用CEQ[™] DTCS-快速啟動套組(Beckman-Coulter, Palo Alto, CA)執行質體純系之雙股定序反應。如由製造商方案所述，使用Performa DTR凝膠過濾濾筒(Edge BioSystems, Gaithersburg, MD)純化反應。在Beckman-Coulter CEQ[™] 2000 XL DNA分析系統上分析序列反應，且使用Sequencher[™] 4.1.4版(Gene Codes Corporation,

Ann Arbor, MI)執行核苷酸表徵。相應於來源於IPP2K之位置-2 5'-同源性側接序列之0.855 Kbp片段之序列展示於圖89(SEQ ID NO:139)中。相應於來源於IPP2K之位置-2 3'-同源性側接序列之0.845 Kbp片段之序列展示於圖90(SEQ ID NO:140)中。

在步驟#3中，將位置-1 5'-同源性側接序列連接至基本質體中。凝膠切離相應於含有正確位置-2 5'-同源性側接序列之純系之限制性片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。隨後，以1:5載體:插入物比率，使用500個單位之T4 DNA連接酶(Invitrogen Life Technologies, Carlsbad, CA)，在20 µl之反應體積中，在於16°C水浴中之16 hr培養之條件下，將相應於位置-1 5'-同源性側接序列(0.855 Kbp)之片段連接至經*Spe* I/*Not* I消化(步驟#1)之經純化基本質體中。

隨後以5 µl之連接反應轉型大腸桿菌One Shot® Top 10化學感受態細胞(Invitrogen Life Technologies, Carlsbad, CA)且在製造商所述之選擇條件下塗鋪。將個別菌落接種至含有補充有50 µl/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000×g下粒化1 min。移除上清液，且如上所述使用NucleoSpin®質體套組(BD Biosciences/Clontech/

Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位 *Spe* I及 *Not* I.(New England Biolabs, Beverly, MA)消化3 μ g之經分離質體DNA，且在37°C下培養1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。除3.4 Kbp基本質體外，藉由0.855 Kbp之插入DNA片段(5'-同源性側接序列)之存在來診斷預期質體純系。

在步驟#4中，將位置-2 3'-同源性側接序列連接至步驟#3產物中。在37°C下，使用10個單位之 *Spe* I及20個單位之 *Sal* I(New England Biolabs, Beverly, MA)限制性核酸內切酶將3 μ g步驟#3中所述之工程化產物線性化1 hr。在100 V下，在補充有0.5%溴化乙錠(Sigma-Aldrich Corporation, St. Louis, MO)之1.0% TAE(0.04 M Tris-乙酸鹽、0.002 M EDTA)瓊脂糖凝膠中，將限制性DNA電泳1 hr。用UV光使DNA片段可見且藉由與1 Kbp DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較估算片段大小。凝膠切離來自步驟#3之4.25 Kbp經 *Spe* I/*Sal* I消化之產物，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。

隨後，將步驟#2中產生之3'-同源性側接序列供體之經分離片段(0.845 Kbp)與經 *Spe* I/*Sal* I消化之步驟#3產物組合，且如上所述，在使用1:5載體:插入物比率及500個單位T4 DNA連接酶(Invitrogen Life Technologies, Carlsbad,

CA)之20 μ l連接反應中純化。在16°C水浴中，將連接反應培養16 hr。連接反應後，按製造商之推薦，將5 μ l之連接反應轉型至MAX Efficiency® DH5 α ™化學感受態細胞(Invitrogen Life Technologies, Carlsbad, CA)中。將個別菌落接種至含有補充有50 μ l/ml氯黴素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中。在37°C下，在200 rpm搖動下，將培養物培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000 $\times$ g下粒化1 min。移除上清液，且如上所述使用NucleoSpin®質體套組(BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位*Sal* I及*Not* I(New England Biolabs, Beverly, MA)消化3 μ g之經分離質體，且在37°C下培養1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使片段可見且藉由與1 Kbp DNA梯級比較估算片段大小。藉由約1.7 Kbp(插入物)及3.33 Kbp(基本質體)之2個DNA片段之存在來診斷預期純系。給予所得質體名稱pDAB7451(圖71)。

D. 自主除草劑耐受性基因表現卡匣建構

建構自主除草劑耐受性基因表現卡匣，其包含含有啟動子、除草劑耐受性基因及多聚腺嘌呤化(polyA)終止序列之完整啟動子轉錄單元(PTU)(圖72)。在該實施例中，啟動子序列係來源於水稻(*O. sativa*)肌動蛋白1[McElroy等人(Plant Cell 2, 163-171; 1990); GenBank寄存物S44221及

GenBank 寄存物 X63830]。除草劑耐受性基因包含賦予針對除草劑畢拉草(bialaphos)之抗性之PAT(草胺膦乙醯基轉移酶)基因(最初來源於綠色產色鏈黴菌(*Streptomyces viridochromogenes*)(GenBank 寄存物 M22827; Wohlleben 等人 Gene 70, 25-37; 1988)之PAT編碼區域之經修飾型式)。對M22827之最長開放閱讀框架之原始序列的修飾為實質性的，且包括改變密碼子利用模式以最佳化在植物中之表現。除用甲硫胺酸取代纈胺酸作為第一編碼胺基酸，且添加丙胺酸作為第二編碼胺基酸外，由pDAB3014之PAT開放閱讀框架編碼之蛋白質與由寄存物M22827之最長開放閱讀框架編碼之蛋白質相同。PAT之重建型式以GenBank寄存物I43995存在。終止子序列來源於玉米(*Z. mays* L.)脂肪酶[GenBank寄存編號L35913之玉米脂肪酶cDNA純系，除L35913之位置1093處之C經pDAB3014中位置2468處之G置換外。該玉米序列包含PAT基因之3'非轉譯區/轉錄終止子區域]。

由Integrated DNA Technologies, Inc.(Coralville, IA)，在標準脫鹽條件下合成以下寡核苷酸引子，且用水稀釋至0.125 µg/µl之濃度：

5'-ACTAGTTAACTGACCTCACTCGAGGTCATTCATATGCTTGA-3'(SEQ ID NO:151)

5'-ACTAGTGTGAATTCAGCACTTAAAGATCT-3'(SEQ ID NO:152)

PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta

3-4-1, Otsu, Shiga, 520-2193, Japan提供且由以下各物組成之試劑來進行：5 μ l 10 $\times$ LA PCR™緩衝液II(Mg²⁺)、20 ng 雙股模板(pDAB3014質體DNA)、10 pmol正向寡核苷酸引子、10 pmol反向寡核苷酸引子、8 μ l dNTP混合物(各2.5 mM)、33.5 μ l H₂O、0.5 μ l(2.5個單位)*TaKaRa LA Taq*™ DNA聚合酶、1滴礦物油。使用Perkin-Elmer Cetus, 48-樣本DNA熱循環儀(Norwalk, CT)，在以下循環條件下執行PCR反應：94℃，4 min/1次循環；98℃ 20 sec，55℃ 1 min，68℃ 3 min/30次循環；72℃，5 min/1次循環；4℃/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將15 μ l之各PCR反應電泳1 hr。

用UV光使經擴增片段可見且藉由與1 Kbp DNA梯級比較來估算片段大小。藉由2.3 Kbp之DNA片段之存在來診斷預期擴增產物。凝膠切離該等片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。隨後，使用TOPO TA Cloning®套組將經純化片段選殖至pCR2.1質體中且根據製造商之方案，轉型至One Shot® TOP10化學感受態大腸桿菌細胞(Invitrogen Life Technologies, Carlsbad, CA)中。

將個別菌落接種至含有補充有50 μ l/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37℃下，在200 rpm搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000 $\times$ g下粒化1 min。

移除上清液，且如上所述使用 NucleoSpin® 質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 *Spe* I 及 *Not* I 消化 3 µg 之經分離質體。在 37°C 下，將所有質體消化產物培養 1 hr。

在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。除 3.9 Kbp pCR®2.1 載體外，藉由 2.325 Kbp 之插入 DNA 片段之存在來診斷預期質體純系。如由製造商所述，使用 CEQ™ DTCS-快速啟動套組 (Beckman-Coulter, Palo Alto, CA) 執行質體純系之雙股定序反應。如由製造商方案所述，使用 Performa DTR 凝膠過濾濾筒 (Edge BioSystems, Gaithersburg, MD) 純化反應。在 Beckman-Coulter CEQ™ 2000 XL DNA 分析系統上分析序列反應，且使用 Sequencher™ 4.1.4 版 (Gene Codes Corporation, Ann Arbor, MI) 執行核苷酸表徵。

E. 自主除草劑耐受性基因卡匣插入至質體骨架中-自主供體

為產生供體質體，將實例 18D 中所述之自主除草劑耐受性基因卡匣插入實例 18B 及 18C 中所述之質體骨架建構中。凝膠切離來源於含有上文所述之預期 2.325 Kbp 序列之純系的限制性片段 (圖 72)，且根據製造商之指南，使用 QIAquick 凝膠提取套組 (QIAGEN Inc., Valencia, CA) 加以純化。

隨後，在連接反應中，將該片段與已經限制酶 *Spe* I 消化且隨後脫磷酸之經純化 pDAB7471 (位置 -1 質體骨架，圖 70)

或 pDAB 7451(位置-2質體骨架圖 71)組合。在以下條件下進行連接反應：在 16°C 水浴中 16 hr 培養之條件下，1:5 載體:插入物比率及於 20 μ l 之反應體積中之 500 個單位 T4 DNA 連接酶(Invitrogen Life Technologies, Carlsbad, CA)。隨後，將 5 μ l 之連接反應轉型至 50 μ l 大腸桿菌 MAX Efficiency® DH5 α [™] 化學感受態細胞 (Invitrogen Life Technologies, Carlsbad, CA) 中且在製造商所述之選擇條件下塗鋪。

將個別菌落接種至含有補充有 50 μ l/ml 氯黴素之 2 ml TB 之 14 ml Falcon 管 (Becton-Dickinson, Franklin Lakes, NJ) 中，且在 37°C 下，在 200 rpm 搖動下培養 16 小時。培養後，將 1.5 ml 細胞轉移至 1.7 ml Costar 微離心管 (Fisher Scientific, Pittsburgh, PA) 中，且在 16,000 $\times$ g 下粒化 1 min。移除上清液，且如上所述使用 NucleoSpin[®] 質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 *Spe* I (New England Biolabs, Beverly, MA) 消化 3 μ g 之經分離質體 DNA，且在 37°C 下培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。藉由 2.325 Kbp 及約 4.9 Kbp (pDAB7471 載體) 或 2.325 Kbp 及約 5.0 Kbp (pDAB7451 載體) 之 DNA 片段之存在來診斷預期質體純系。

所得質體分別命名為 pDAB7422(位置-1 自主供體)(圖 73)

及 pDAB7452(位置-2自主供體)(圖 74)。

F. 非自主除草劑耐受性基因表現卡匣建構

建構非自主除草劑耐受性基因表現卡匣，其包含不完整啟動子轉錄單位(PTU)(圖 75)。在該實施例中，使用利用來源於 *Thesoa assigna* 病毒之 2A 序列 (Mattion, N.M., Harnish, E.C., Crowley, J.C. & Reilly, P.A. (1996) J. Virol. 70, 8124-8127)、除草劑耐受性基因及多聚腺嘌呤化 (polyA) 終止序列之功能性，但無啟動子之策略。在該實施例中，2A 轉譯終止信號序列已經工程化以與除草劑耐受性基因在轉譯上同框。另外，2A/除草劑編碼序列已經工程化以與 IPP2K 基因標靶之轉譯閱讀框架一致。除草劑耐受性基因包含賦予針對除草劑畢拉草之抗性之 PAT(草胺磷乙醯基轉移酶) 基因(最初來源於綠色產色鏈黴菌 (GenBank 寄存物 M22827; Wohlleben 等人 Gene 70:25-37; 1988) 之 PAT 編碼區域之經修飾型式)。對 M22827 之最長開放閱讀框架之原始序列的修飾為實質性的，且包括改變密碼子利用模式以最佳化在植物中之表現。除用甲硫胺酸取代纈胺酸作為第一編碼胺基酸，且添加丙胺酸作為第二編碼胺基酸外，由 pDAB3014 之 PAT 開放閱讀框架編碼之蛋白質與由 M22827 之最長開放閱讀框架(以 M22827 之位置 244 處之 GTG 開始)編碼之蛋白質相同。PAT 之重建型式以 GenBank 寄存物 I43995 存在。終止子序列來源於玉米脂肪酶 [GenBank 寄存編號 L35913 之玉米脂肪酶 cDNA 純系，除 L35913 之位置 1093 處之 C 經 pDAB3014 中位置 2468 處之 G 置

換外]。該玉米序列包含PAT基因之3'非轉譯區/轉錄終止子區域。

由Integrated DNA Technologies, Inc.(Coralville, IA)，在標準脫鹽條件下合成以下寡核苷酸引子，且用水稀釋至0.125 µg/µl之濃度：

5'-ACTAGTGGCGGCGGAGAGGGGCAGAGGAAGTCTTCTAA
CATGCGGTGACGTGGAGGAGAATCCCGGCCCTAGGATGG
CTTCTCCGGAGAGGAGACCAGTTGA-3(SEQ ID NO:153)

5'-ACTAGTATGCATGTGAATTCAGCACTTAAAGATCT-3'(SEQ
ID NO:154)

PCR擴增反應係使用由TaKaRa Biotechnology Inc.(Seta 3-4-1, Otsu, Shiga, 520-2193, Japan)提供且由以下各物組成之試劑來進行：5 µl 10×LA PCR™緩衝液II(Mg²⁺)、20 ng雙股模板(pDAB3014質體DNA)、10 pmol正向寡核苷酸引子、10 pmol反向寡核苷酸引子、8 µl dNTP混合物(各2.5 mM)、33.5 µl H₂O、0.5 µl(2.5個單位)TaKaRa LA Taq™ DNA聚合酶、1滴礦物油。使用Perkin-Elmer Cetus，48-樣本DNA熱循環儀(Norwalk, CT)，在以下循環條件下執行PCR反應：94℃，4 min/1次循環；98℃ 20 sec，55℃ 1 min，68℃ 2 min/30次循環；72℃，5 min/1次循環；4℃/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將15 µl之各PCR反應電泳1 hr。用UV光使經擴增片段可見且藉由與1 Kbp DNA梯級比較來估算片段大小。藉由約1 Kbp之DNA片段之存在來診斷預期

擴增產物。凝膠切離該等片段，且根據製造商之指南，使用 QIAquick凝膠提取套組 (QIAGEN Inc., Valencia, CA) 加以純化。隨後，使用 TOPO TA Cloning®套組將經純化片段選殖至 pCR2.1質體中，根據製造商之方案，轉型至 One Shot® TOP10化學感受態大腸桿菌細胞 (Invitrogen Life Technologies, Carlsbad, CA) 中。

將個別菌落接種至含有補充有 50 µl/ml 康徵素之 2 ml TB 之 14 ml Falcon 管 (Becton-Dickinson, Franklin Lakes, NJ) 中，且在 37°C 下，在 200 rpm 搖動下培養 16 小時。培養後，將 1.5 ml 細胞轉移至 1.7 ml Costar 微離心管 (Fisher Scientific, Pittsburgh, PA) 中，且在 16,000×g 下粒化 1 min。移除上清液，且如上所述使用 NucleoSpin®質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 *Spe* I 消化 3 µg 之經分離質體。在 37°C 下，將所有質體消化產物培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。藉由約 1.0 Kbp 及 3.9 Kbp (pCR®2.1 載體) 之插入 DNA 片段之存在來診斷預期質體純系。如由製造商所述，使用 CEQ™ DTCS-快速啟動套組 (Beckman-Coulter, Palo Alto, CA) 執行質體純系之雙股定序反應。如由製造商方案所述，使用 Performa DTR 凝膠過濾濾筒 (Edge BioSystems, Gaithersburg, MD) 純化反應。在 Beckman-Coulter CEQ™ 2000 XL DNA 分析系統上分析序列反應，

且使用 Sequencher™ 4.1.4版 (Gene Codes Corporation, Ann Arbor, MI) 執行核苷酸表徵。

G. 非自主除草劑耐受性基因卡匣插入至質體骨架中-非自主供體

為產生供體質體，將實例 18F 中所述之非自主除草劑耐受性基因卡匣插入實例 18B 及 18C 中所述之質體骨架建構中。凝膠切離相應於含有正確 1 Kbp 序列之純系之限制性片段，且根據製造商之指南，使用 QIAquick 凝膠提取套組 (QIAGEN Inc., Valencia, CA) 加以純化。隨後，在連接反應中，將該片段與已經限制酶 *Spe* I 消化且隨後脫磷酸之經純化 pDAB7471 (位置 -1 質體骨架) (圖 70) 或 pDAB 7451 (位置 -2 質體骨架) (圖 71) 組合。在以下條件下進行連接反應：在 16°C 水浴中 16 hr 培養之條件下，1:5 載體:插入物比率及於 20 µl 之反應體積中之 500 個單位 T4 DNA 連接酶 (Invitrogen Life Technologies, Carlsbad, CA)。隨後，將 5 µl 之連接反應轉型至 50 µl 大腸桿菌 MAX Efficiency® DH5α™ 化學感受態細胞 (Invitrogen Life Technologies, Carlsbad, CA) 中，且在由製造商所述之選擇條件下塗鋪。

將個別菌落接種至含有補充有 50 µl/ml 氯黴素之 2 ml TB 之 14 ml Falcon 管 (Becton-Dickinson, Franklin Lakes, NJ) 中，且在 37°C 下，在 200 rpm 搖動下培養 16 小時。培養後，將 1.5 ml 細胞轉移至 1.7 ml Costar 微離心管 (Fisher Scientific, Pittsburgh, PA) 中，且在 16,000×g 下粒化 1 min。移除上清液，且如上所述使用 NucleoSpin® 質體套組 (BD

Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體 DNA。用 10 個單位 *Spe* I(New England Biolabs, Beverly, MA)消化 3 μ g 之經分離質體 DNA，且在 37°C 下培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。藉由 DNA 片段 1.0 Kbp 及約 4.96 Kbp(pDAB7471 載體)或 1.0 Kbp 及約 5.0 Kbp(pDAB7451 載體)之存在來診斷預期質體純系。所得質體分別命名為 pDAB7423(位置 -1 非自主供體)(圖 76)及 pDAB7454(位置 -2 非自主供體)(圖 77)。

H. 位置 1 ZFN+HR 供體序列：組合質體。

作為將 2 個分離質體傳遞至植物細胞中之替代策略(例如一個質體含有 ZFN 元件且第二質體含有除草劑耐受性供體序列)，將單質體工程化，從而含有本專利中所說明之所有必要元件。該實例中所述之組合質體含有 2 個經設計以在規定 IPP2K 基因座處靶向且產生雙股斷裂之 ZFN，以及自主 PAT PTU 及/或非自主 2A/PAT PTU，及經設計以整合至彼等斷裂位點中之供體側接序列。

使用基於 λ 噬菌體之位點特異性重組之 Gateway® 技術(Landy, A. (1989) Ann. Rev. Biochem. 58:913)用以將載體 pDAB7422 及 pDAB7423(描述於實例 6E 及 6G 中)轉變成 Gateway® 目的載體。轉變後，含有 ZFN 表現卡匣(容納於 Gateway® 進入載體中)之質體可容易地移動至目的載體，產生 ZFN/供體組合質體。在 37°C 下，用 10 個單位 *Not*

I(New England Biolabs, Beverly, MA)消化1 μ g之各個該質體歷時1 hr。在65°C下，將*Not* I限制性核酸內切酶熱失活15 min，且隨後在37°C下，使用3個單位之蝦鹼性磷酸酶(SAP)(Roche Diagnostics GmbH, Mannheim, Germany)將片段末端脫磷酸1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使載體片段(pDB7422=7.317 Kbp，pDAB7423=5.971 Kbp)可見，藉由與1 Kbp DNA梯級比較估算大小，凝膠切離且隨後根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。

隨後，在以下條件下進行之連接反應中，將該載體片段與含有Gateway® Technology元件attR1、*ccdB*、Cm^R及attR2之2.274 Kbp *Not* I片段組合：在16°C水浴中16 hr培養之條件下，1:5載體:插入物比率及於20 μ l之反應體積中之500個單位T4 DNA連接酶(Invitrogen Life Technologies, Carlsbad, CA)。隨後，將5 μ l之連接反應轉型至50 μ l大腸桿菌One Shot® *ccdB* Survival™化學感受態細胞(Invitrogen Life Technologies, Carlsbad, CA)中，且在由製造商所述之選擇條件下塗鋪。

將個別菌落接種至含有補充有50 μ l/ml氯黴素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000 $\times$ g下粒化1 min。

移除上清液，且如上所述使用 NucleoSpin® 質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 *EcoRI* I (New England Biolabs, Inc., Beverly, MA) 消化 3 µg 之經分離質體 DNA，且在 37°C 下培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使片段可見且藉由與 1 Kbp DNA 梯級比較估算片段大小。藉由自主 PAT PTU 位置 -1 HR 供體之 1.448 Kbp、1.946 Kbp 及 6.197 Kbp 之 DNA 片段及非自主 PAT 位置 -1 HR 供體之 5.807 Kbp 及 2.438 Kbp 之 DNA 片段之存在來診斷預期質體純系。所得質體分別命名為 pDAB7424 (Gateway® 改適位置 -1 自主供體) (圖 78) 及 pDAB7425 (Gateway® 改適位置 -1 非自主供體) (圖 79)。

作為該等選殖操作之結果，質體 pDAB7424 及 pDAB7425 指定為 Gateway® 目的載體。pDAB7412 具有作為含有以下元件之 Gateway® 進入載體之功能性：ZmUbi1v.2/ZFN12/ZmPer5 3' UTR。為將 ZFN 表現卡匣 (Gateway® 進入載體) 轉移至自主或非自主供體分子 (Gateway® 目的載體) 中，如由製造商所概述，以 50 ng (進入載體):150 ng/µl (目的載體) 之比率執行 LR Clonase™ II (Invitrogen Life Technologies, Carlsbad, CA) 反應。所得陽性組合質體命名為 pDAB7426 (位置 -1 自主 HR 供體 /ZFN12) (圖 80) 及 pDAB7427 (非自主 HR 供體 /ZFN12) (圖 81)。

實例 19：ZFN 及供體 DNA 傳遞至植物細胞中

為使供體DNA能經由靶向整合而ZFN介導整合至植物基因組中，應瞭解，需要傳遞編碼ZFN之DNA，接著在植物細胞中表現功能性ZFN蛋白質。亦需要同時將供體DNA傳遞至該植物細胞中，以使得功能性ZFN蛋白質可在標靶DNA處誘導雙股斷裂，隨後經由將供體DNA同源性驅動整合至標靶位置中修復標靶DNA。熟習此項技術者可預想，功能性ZFN蛋白質之表現可藉由包括(但不限於)編碼ZFN之建構的基因轉殖或編碼ZFN之建構的瞬間表現之若干方法來實現。在該兩種狀況下，功能性ZFN蛋白質之表現及供體DNA傳遞至植物細胞中係同時達成以驅動靶向整合。

在此所引用之實例中，吾人示範將編碼ZFN之DNA及供體DNA同時傳遞至植物細胞中之方法。熟習此項技術者可使用適於植物細胞之多種DNA傳遞方法中之任何方法，包括(但不限於)土壤桿菌介導轉型、基於基因槍之DNA傳遞或Whiskers™介導DNA傳遞。在此所述之一實施例中，使用供體DNA與編碼ZFN之DNA建構之各種組合進行Whiskers™介導DNA傳遞實驗。該等組合包括1)含有ZFN編碼序列及供體DNA兩者之單一質體，及2)2個不同質體，一個含有ZFN編碼序列且另一個含有供體DNA。在另一實施例中，使用供體DNA與編碼ZFN之DNA建構之各種組合進行基於基因槍之DNA傳遞。熟習此項技術者可推論，該等組合可包括1)含有ZFN編碼序列及供體DNA兩者之單一質體，及2)2個不同質體，一個含有ZFN編碼序列且另一個含有供體DNA。

A. Whiskers™介導DNA傳遞

如本文中先前所述，製備玉米之胚胎發生Hi-II細胞培養物，且將其用作證明靶向整合之活植物細胞之來源。熟習此項技術者可預想，利用來源於多種植物物種之細胞培養物，或來源於多種植物物種之分化植物組織作為證明靶向整合之活植物細胞的來源。

在該實例中，將來自先前低溫保存細胞系之12 ml PCV加28 ml條件培養基次培養至500 ml錐形瓶中之80 ml之GN6液體培養基(N6培養基(Chu等人，1975)、2.0 mg/L 2,4-D、30 g/L蔗糖，pH 5.8))中，且在28℃下，置放於125 rpm之搖動器上。使用相同細胞系將該步驟重複2次，以使得在所有3個燒瓶中分布總共36 ml PCV。24小時後，移除GN6液體培養基且用72 mL GN6 S/M滲透培養基(N6培養基、2.0 mg/L 2,4-D、30 g/L蔗糖、45.5 g/L山梨糖醇、45.5 g/L甘露醇、100 mg/L肌醇，pH 6.0)置換。在28℃下，在適度攪動(125 rpm)下，將燒瓶在黑暗中培養30-35分鐘。在培養期期間，藉由將8.1 ml之GN6 S/M液體培養基添加至405 mg之無菌、碳化矽鬚晶中來製備碳化矽鬚晶(Advanced Composite Materials, LLC, Greer, SC)之50 mg/ml懸浮液。

在GN6 S/M滲透培養基中培養後，將各燒瓶之內含物彙集至250 mL離心瓶中。在燒瓶中之所有細胞沈降至底部後，取出超過約14 ml之GN6 S/M液體之內含物體積，且收集於無菌1 L燒瓶中備用。在渦旋上，以最大速度將鬚晶

之預濕潤懸浮液混合60秒，且隨後添加至離心瓶中。

在將含有ZFN編碼序列加供體DNA兩者之單一質體傳遞至植物細胞中之一實例中，將170 μg 經純化環形質體DNA添加至瓶中。在將2個不同質體共同傳遞(一個含有ZFN編碼序列且另一個含有供體DNA)之一替代實例中，關於DNA量評估多種策略。一種策略利用85 μg 供體DNA及85 μg 編碼鋅指之DNA。其它修改形式利用以個別質體大小(以千鹼基對為單位)計相對於1倍鋅指DNA供體DNA呈10、5或1倍之莫耳比率，以使得每一瓶添加總共170 μg 之DNA。在所有共同傳遞之狀況下，在添加至離心瓶之前，將DNA預先彙集於管中。在添加DNA後，立即將瓶置放於改良Red Devil 5400商用油漆混合器(Red Devil Equipment Co., Plymouth, MN)中且攪動10秒。攪動後，將細胞、培養基、鬚晶及DNA之混合物連同125 ml新鮮GN6液體培養基一起添加至1 L燒瓶之內含物中，以減少滲透劑。使細胞在設定在125 rpm下之搖動器上回收2小時。使用連接至家用真空管線之玻璃槽收集單元將6 mL分散懸浮液過濾至Whatman #4濾紙(5.5 cm)上，以使得每一瓶獲得60張濾紙。將濾紙置放於具有GN6固體培養基(除具有2.5 g/L水晶洋菜膠凝劑外，與GN6液體培養基相同)之60×20 mm培養板上且在28°C下，在黑暗條件下培養1週。

B. 基因槍介導DNA傳遞

在此引用之實例中，如先前在本文中所述，在實驗之前，將玉米之胚胎發生懸浮液次培養至GN6液體培養基中

約24小時。移除過量液體培養基且將約0.4 PCV之細胞以直徑2.5 cm之圓圈稀疏地展布於含有經2.5 g/L水晶洋菜固化之GN6 S/M培養基之100×15 mm皮氏培養皿中心。在黑暗條件下將細胞培養4小時。為用DNA塗佈基因槍粒子，用100%乙醇將3 mg之直徑1.0微米之金粒子洗滌一次，用無菌蒸餾水洗滌兩次，且再懸浮於矽化Eppendorf管中之50 μ l水中。將總共5 μ g質體DNA、20 μ l亞精胺(0.1 M)及50 μ l氯化鈣(2.5 M)單獨添加至金懸浮液中且在渦旋上混合。在室溫下，將混合物培養10 min，在10,000 rpm下，在桌上型微離心機中粒化10秒，再懸浮於60 μ l冷的100%乙醇中，且將8-9 μ l分布於各巨載劑上。

使用Biolistic PDS-1000/He™系統(Bio-Rad Laboratories, Hercules, CA)進行轟擊。在1100 psi及27吋Hg真空之條件下，將含有細胞之培養板置放於中間支架上，且按照操作手冊轟擊。轟擊後16小時，將呈小凝塊之組織轉移至GN6 (1H)培養基且在28℃下，在黑暗條件下培養2-3週。每2-4週繼續轉移直至出現因供體DNA之整合產生之推定轉殖基因分離株。經由基因槍介導DNA傳遞產生之推定供體DNA整合事件之鑑別、分離及再生與用於經由Whiskers™介導DNA傳遞產生之推定供體DNA整合事件之方法相同且描述於下文。

C.推定靶向整合轉殖基因事件之鑑別及分離

DNA傳遞後一週，將濾紙轉移至具有GN6 (1H)選擇培養基(N6培養基、2.0 mg/L 2,4-D、30 g/L蔗糖、100 mg/L肌

醇、1.0 mg/L來自Herbiace(Meiji Seika, Japan)之畢拉草、2.5 g/L水晶洋菜，pH 5.8)之60×20 mm培養板。在28℃下，將該等選擇培養板在黑暗中培養1週。

在黑暗中選擇1週後，藉由自各培養板將一半細胞刮取至保持在37-38℃下含有3.0 mL之GN6瓊脂糖培養基(N6培養基、2.0 mg/L 2,4-D、30 g/L蔗糖、100 mg/L肌醇、7 g/L SeaPlaque®瓊脂糖，pH 5.8，在121℃下僅高壓釜蒸10分鐘)及1 mg/L來自Herbiace之畢拉草的管中，使組織嵌入新鮮培養基上。

用刮勺將瓊脂糖/組織混合物打碎，且隨後將3 mL瓊脂糖/組織混合物均勻傾倒至含有GN6 (1H)培養基之100×15 mm皮氏培養皿表面上。對各培養板之2個二等分物重複該方法。嵌入所有組織後，用Nescofilm®或Parafilm M®將培養板個別地密封，且在28℃下，在黑暗條件下培養至多10週。自經嵌入培養板移除在該等選擇條件下生長之推定轉型分離株，且轉移至於60×20 mm培養板中之新鮮選擇培養基。若約2週後顯然持續生長，則事件視為對所應用之除草劑(畢拉草)具有抗性，且隨後將細胞之等分試樣收集至2 mL Eppendorf管中以供基因型分析。

熟習此項技術者可利用編碼供體DNA中任何適當選擇標記之基因且將相當選擇條件應用於活細胞。舉例而言，可利用如WO 2005/107437 A2中所述諸如AAD-1之替代性選擇標記基因，作為用於如本文中所述之玉米細胞中整合事件的選擇及回收之供體。

實例 20：經由 PCR 基因定型 (genotyping) 篩選靶向整合事件

在該實例中，PCR 基因定型理解為包括 (但不限於) 來源於經預測含有嵌入基因組中之供體 DNA 之分離玉米愈傷組織的基因組 DNA 之聚合酶鏈反應 (PCR) 擴增，接著為 PCR 擴增產物之標準選殖及序列分析。PCR 基因定型之方法已充分描述 (例如，Rios, G. 等人 (2002) *Plant J.* 32:243-253) 且可應用於來源於任何植物物種或組織類型 (包括細胞培養物) 之基因組 DNA。

熟習此項技術者可設計用於 PCR 基因定型之策略，其包括 (但不限於) 植物基因組中特異性序列之擴增，植物基因組中多特異性序列之擴增，植物基因組中非特異性序列之擴增，或其組合。如在該實例中所述，擴增可接著選殖及定序，或接著為擴增產物之直接序列分析。熟習此項技術者可預想用於分析本文中所產生之擴增產物之替代方法。

在本文所述之一實施例中，在 PCR 擴增中使用對基因標靶具有特異性之寡核苷酸引子。在本文所述之另一實施例中，在 PCR 擴增中使用對供體 DNA 序列具有特異性之寡核苷酸引子。另一實施例包括結合於基因標靶序列及供體 DNA 序列兩者之寡核苷酸引子之組合。熟習此項技術者可設計引子及擴增反應之其他組合以探查基因組。

A. 基因組 DNA 提取

自實例 19 中所述之經分離、除草劑耐受性玉米細胞提取基因組 DNA (gDNA)，且用作 PCR 基因定型實驗之模板。自約 100-300 μ l 紅血球容積 (PCV) 之如上所述根據詳述於

DNeasy® 96植物套組(QIAGEN Inc., Valencia, CA)中之製造商方案來分離的除草劑耐受性 HiII 愈傷組織提取 gDNA。將基因組 DNA 溶離於 100 μ l 套組供應之溶離緩衝液中，產生 20-200 ng/ μ l 之最終濃度，且隨後經由下文概述基於 PCR 之基因定型方法來分析。

B. 用於 PCR 基因定型之引子設計

熟習此項技術者可使用設計及實施基於 PCR 之基因定型之多種策略。經設計以與基因標靶、供體 DNA 序列及/或兩者之組合黏接之寡核苷酸引子為可行的。為設計可與未由建構至供體 DNA 分子中之同源性側接序列涵蓋之區域中的 IPP2K 基因標靶黏接之寡核苷酸引子，經由 DNA 定序表徵含有其他基因標靶序列資料之質體純系。如製造商所述，使用 CEQ™ DTCS-快速啟動套組(Beckman-Coulter, Palo Alto, CA)執行質體純系之雙股定序反應。如由製造商方案所述，使用 Performa DTR 凝膠過濾濾筒(Edge BioSystems, Gaithersburg, MD)純化反應。在 Beckman-Coulter CEQ™ 2000 XL DNA 分析系統上分析序列反應，且使用 Sequencher™ 4.1.4 版(Gene Codes Corporation, Ann Arbor, MI)執行核苷酸表徵。該等序列相應於 ZFN 靶向區域之 IPP2K 基因上游(5'-)及下游(3'-)之區域，且描述於圖 91(SEQ ID NO:141)及圖 92(SEQ ID NO:142)中。

在此所呈現之實例中，由 Integrated DNA Technologies, Inc.(Coralville, IA)，在標準脫鹽條件下合成以下寡核苷酸引子，且用水稀釋至 100 μ M 之濃度。以下組之正向及反向

寡核苷酸引子經設計以與位於供體DNA序列邊界外部對IPP2K基因標靶具有特異性之gDNA序列黏接。該等寡核苷酸如下：

5'-TGGACGGAGCGAGAGCCAGAATTCGACGCT G-3'(SEQ ID NO:153)

5'-GTGCAAGAATGTATTGGGAATCAACCTGAT G-3'(SEQ ID NO:154)。

第二組正向及反向寡核苷酸引子亦經設計以與在供體DNA序列邊界外部，但套於第一對內部之對IPP2K基因標靶具有特異性之gDNA序列黏接：

5'-CTGTGGTACCAGTACTAGTACCAGCATC-3'(SEQ ID NO:155)

5'-TCTTGGATCAAGGCATCAAGC ATTCCAATCT-3'(SEQ ID NO:156)。

正向及反向寡核苷酸引子另外經設計以與相應於除草劑耐受性基因之編碼區域之供體DNA特異性黏接：

5'-TGGGTAAGTGGCCTAAGTGG-3'(SEQ ID NO:157)

5'-TGGAAGGCTAGGAACGCTTA-3'(SEQ ID NO:158)

5'-CCAGTTAGGCCAGTTACCCA-3'(SEQ ID NO:159)

5' TAAGCGTTCCTAGCCTTCCA-3'(SEQ ID NO:160)。

C. 供體DNA-特異性PCR擴增

初級PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japan提供且由以下各物組成之試劑來進行：2.5 µl 10×Ex Taq PCR™緩衝液、40-

200 ng雙股基因組DNA模板、10 μ M正向寡核苷酸引子、10 μ M反向寡核苷酸引子、2 μ l dNTP混合物(各2.5 mM)、16 μ l H₂O、0.5 μ l(2.5個單位)*Ex Taq*TM DNA聚合酶。使用Bio-Rad, 96樣本DNA Engine Tetrad2, Peltier熱循環儀(Hercules, CA), 在以下循環條件下執行PCR反應：94°C, 3 min/1次循環；94°C 30 sec, 64°C 30 sec, 72°C 5 min/35次循環；72°C, 10 min/1次循環；4°C/保持。

隨後，在包含以下各物之二級PCR反應中將初級PCR反應之擴增產物再擴增：2.5 μ l 10×*Ex Taq* PCRTM緩衝液、2 μ l模板(1° PCR反應於H₂O中之1:100稀釋液)、10 μ M正向寡核苷酸引子、10 μ M反向寡核苷酸引子、2 μ l dNTP混合物(各2.5 mM)、16 μ l H₂O、0.5 μ l(2.5個單位)*Ex Taq*TM DNA聚合酶。使用Bio-Rad, 96樣本DNA Engine Tetrad2, Peltier熱循環儀(Hercules, CA), 在以下循環條件下執行PCR反應：95°C, 1 min/1次循環；94°C 15 sec, 61°C 30 sec, 72°C 30 sec/30次循環；72°C, 1 min/1次循環；4°C/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將10 μ l之各經擴增產物電泳1 hr。用UV光使經擴增片段可見且藉由與1 Kbp Plus DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較估算片段大小。如圖82中所示藉由DNA片段0.317 Kbp之存在來診斷含有預期片段之PCR產物。

實例21：偵測靶向整合事件

在含有編碼除草劑耐受性基因卡匣之整合供體DNA分子

之除草劑耐受性事件中，預期一些比例之該等事件為供體DNA靶向整合至ZFN誘導雙股斷裂之位點中之產物。為將該等靶向整合事件與來源於除草劑耐受性基因卡匣之隨機整合之彼等事件區分，利用使用基因組特異性及後續基因組特異性加供體特異性PCR引子之組合的基於PCR之基因定型策略。

A. 基因組特異性及後續基因組/供體特異性擴增

在該實施例中，初級PCR反應利用對供體整合區域之IPP2K基因標靶上游及下游之區域(例如圖92及93)具有特異性之寡核苷酸引子。初級PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japan提供且由以下各物組成之試劑來進行：2.5 μ l 10 $\times$ Ex Taq PCR™緩衝液、40-200 ng雙股玉米gDNA模板、10 μ M正向寡核苷酸引子、10 μ M反向寡核苷酸引子、2 μ l dNTP混合物(各2.5 mM)、16 μ l H₂O、0.5 μ l(2.5個單位)Ex Taq™ DNA聚合酶。使用Bio-Rad，96樣本DNA Engine Tetrad2，Peltier熱循環儀(Hercules, CA)，在以下循環條件下執行PCR反應：94℃，3 min/1次循環；94℃ 30 sec，64℃ 30 sec，72℃ 5 min/35次循環；72℃，10 min/1次循環；4℃/保持。

隨後，將初級PCR反應產物以1:100稀釋於H₂O中且用作兩個不同二級PCR反應之模板DNA。在該實施例中，二次反應利用結合於IPP2K基因組區域及供體分子中之引子，產生跨越基因組與供體之間的整合邊界之擴增子。第一反

應著重於基因組與供體之間的5'邊界。第二反應著重於供體與基因組之間的3'邊界。兩個反應由以下各物組成：2.5 μ l 10 $\times$ Ex Taq PCR™緩衝液、2 μ l模板[1° PCR反應之1:100稀釋液]、10 μ M正向寡核苷酸引子、10 μ M反向寡核苷酸引子、2 μ l dNTP混合物(各2.5 mM)、16 μ l H₂O、0.5 μ l(2.5個單位)Ex Taq™ DNA聚合酶。使用Bio-Rad，96樣本DNA Engine Tetrad2，Peltier熱循環儀(Hercules, CA)，在以下循環條件下執行PCR反應：94°C，3 min/1次循環；94°C 30 sec，60°C 30 sec，72°C 2 min/35次循環；72°C，10 min/1次循環；4°C/保持。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將20 μ l各2° PCR反應電泳1 hr。

用UV光使經擴增片段可見且藉由與1 Kbp Plus DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較估算片段大小。藉由DNA片段1.65 Kbp(5'邊界)(圖83)或1.99 Kbp(3'邊界)(圖84)之存在來診斷來源於供體靶向整合至IPP2K基因中之PCR產物。凝膠切離該等片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)加以純化。隨後，根據製造商之方案，使用TOPO TA Cloning®套組(用於pCR®2.1載體)及One Shot® TOP10化學感受態大腸桿菌細胞(Invitrogen Life Technologies, Carlsbad, CA)，將經純化片段選殖至pCR2.1質體中。

將個別菌落接種至含有補充有50 μ l/ml康黴素之2 ml TB

之 14 ml Falcon 管 (Becton-Dickinson, Franklin Lakes, NJ) 中，且在 37°C 下，在 200 rpm 搖動下培養 16 小時。培養後，將 1.5 ml 細胞轉移至 1.7 ml Costar 微離心管 (Fisher Scientific, Pittsburgh, PA) 中，且在 16,000×g 下粒化 1 min。移除上清液，且如上所述使用 NucleoSpin® 質體套組 (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA) 分離質體 DNA。用 10 個單位 Eco RI (New England Biolabs, Beverly, MA) 將 3 µg 之經分離質體消化。在 37°C，將所有質體消化產物培養 1 hr。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將限制性 DNA 電泳 1 hr。用 UV 光使經擴增片段可見且藉由與 1 Kbp Plus DNA 梯級 (Invitrogen Life Technologies, Carlsbad, CA) 比較估算片段大小。

除 3.9 Kbp pCR®2.1 載體外，藉由具有適當大小之插入 DNA 片段之存在來診斷預期質體純系。如由製造商所述，使用 CEQ™ DTCS-快速啟動套組 (Beckman-Coulter, Palo Alto, CA) 執行質體純系之雙股定序反應。如由製造商方案所述，使用 Performa DTR 凝膠過濾濾筒 (Edge BioSystems, Gaithersburg, MD) 純化反應。在 Beckman-Coulter CEQ™ 2000 XL DNA 分析系統上分析序列反應，且使用 Sequencher™ 4.1.4 版 (Gene Codes Corporation, Ann Arbor, MI) 執行核苷酸表徵。使用 Vector NTi 10.1 版 (Invitrogen Life Technologies, Carlsbad, CA) 執行核苷酸比對。

如下進行來自靶向整合事件 (事件 #073) 之序列資料之分

析。使跨越基因組之整個整合位點之初級PCR產物經受著重於基因組與供體之間的5'邊界或3'邊界之二級擴增。相應於該等二級擴增產物之經選殖片段與野生型IPP2K基因組序列以及靶向整合事件之預期序列之比對清楚地指示，發生供體DNA在標靶位點處之精確整合。

IPP2K基因組基因座之核苷酸序列、基因組/供體邊界、相應於IPP2K同源性側接序列之供體區域之核苷酸序列及除草劑耐受性卡匣之核苷酸序列全部保存於來源於該事件之多個選殖PCR產物中。因此，該事件代表ZFN介導雙股斷裂之同源性驅動修復及供體DNA之靶向整合在特定基因標靶處發生之基因組。已獲得代表獨特靶向整合發生之其他轉型事件，證明本文中教示之方法可在玉米愈傷組織中再現。熟習此項技術者可將該等方法應用於認為需要靶向整合之任何植物物種中之任何基因標靶。

B. 套式基因組/供體特異性擴增

在該實施例中，初級及後續二級PCR反應利用對供體整合區域之IPP2K基因標靶上游或下游之區域(附錄V及VI)具有特異性的寡核苷酸引子與對供體序列具有特異性的寡核苷酸引子之組合。在該實例中，初級PCR擴增反應係使用由TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japan提供由以下各物組成之試劑來進行：2.5 μ l 10 $\times$ Ex Taq PCR™緩衝液、40-200 ng雙股玉米gDNA模板、10 μ M正向寡核苷酸引子、10 μ M反向寡核苷酸引子、2 μ l dNTP混合物(各2.5 mM)、16 μ l H₂O、0.5 μ l(2.5個單位)Ex

*Taq*TM DNA 聚合酶。使用 Bio-Rad，96 樣本 DNA Engine Tetrad2，Peltier 熱循環儀 (Hercules, CA)，在以下循環條件下培養 PCR 反應：94°C，3 min/1 次循環；94°C 30 sec，52°C 或 64°C 30 sec，72°C 2 min/35 次循環；72°C，10 min/1 次循環；4°C / 保持。

隨後，將初級 PCR 反應產物 1:100 稀釋於 H₂O 中且用作二級 PCR 反應之模板 DNA。在該實施例中，二次反應亦利用結合於 IPP2K 基因組區域及供體分子中之引子，產生跨越基因組與供體之間的整合邊界之擴增子。所用特異性引子決定擴增著重於基因組與供體之間的 5' 邊界還是 3' 邊界。用於該等反應之試劑由以下各物組成：2.5 μl 10×Ex *Taq* PCRTM 緩衝液、2 μl 模板 [1° PCR 反應之 1:100 稀釋液]、10 μM 正向寡核苷酸引子、10 μM 反向寡核苷酸引子、2 μl dNTP 混合物各 2.5 mM、16 μl H₂O、0.5 μl (2.5 個單位) *Ex Taq*TM DNA 聚合酶。使用 Bio-Rad，96 樣本 DNA Engine Tetrad2，Peltier 熱循環儀 (Hercules, CA)，在以下循環條件下執行 PCR 反應：94°C，3 min/1 次循環；94°C 30 sec，54°C 或 60°C 30 sec，72°C 2 min/35 次循環；72°C，10 min/1 次循環；4°C / 保持。在 100 V 下，在補充有 0.5% 溴化乙錠之 1.0% TAE 瓊脂糖凝膠中將 20 μl 之各 2° PCR 反應電泳 1 hr。

用 UV 光使經擴增片段可見且藉由與 1 Kbp Plus DNA 梯級 (Invitrogen Life Technologies, Carlsbad, CA) 比較估算片段大小。藉由 DNA 片段 1.35 Kbp (5' 邊界) (圖 85) 或 1.66 Kbp (3' 邊界) (圖 86) 之存在來診斷來源於供體靶向整合至

IPP2K基因中之PCR產物。凝膠切離該等片段，且根據製造商之指南，使用QIAquick凝膠提取套組(QIAGEN Inc., Valencia, CA)具有純化。隨後，根據製造商之方案，使用TOPO TA Cloning®套組(用pCR®2.1載體)及One Shot®TOP10化學感受態大腸桿菌細胞(Invitrogen Life Technologies, Carlsbad, CA)，將經純化片段選殖至pCR2.1質體中。

C. 基因定型PCR產物之核苷酸序列分析

將實例21B中所述之個別菌落接種至含有補充有50 µl/ml康徽素之2 ml TB之14 ml Falcon管(Becton-Dickinson, Franklin Lakes, NJ)中，且在37°C下，在200 rpm搖動下培養16小時。培養後，將1.5 ml細胞轉移至1.7 ml Costar微離心管(Fisher Scientific, Pittsburgh, PA)中，且在16,000×g下粒化1 min。移除上清液，且如上所述使用NucleoSpin®質體套組(BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA)分離質體DNA。用10個單位*Eco* RI(New England Biolabs, Beverly, MA)將3 µg之經分離質體消化。在37°C，將所有質體消化產物培養1 hr。在100 V下，在補充有0.5%溴化乙錠之1.0% TAE瓊脂糖凝膠中將限制性DNA電泳1 hr。用UV光使經擴增片段可見且藉由與1 Kbp Plus DNA梯級(Invitrogen Life Technologies, Carlsbad, CA)比較估算片段大小。

除3.9 Kbp pCR®2.1載體外，藉由插入DNA片段之存在來診斷質體純系。如由製造商所述，使用CEQ™ DTCS-快

速啟動套組(Beckman-Coulter, Palo Alto, CA)執行質體純系之雙股定序反應。如由製造商方案所述，使用Performa DTR凝膠過濾濾筒(Edge BioSystems, Gaithersburg, MD)純化反應。在Beckman-Coulter CEQ™ 2000 XL DNA分析系統上分析序列反應，且使用Sequencher™ 4.1.4版(Gene Codes Corporation, Ann Arbor, MI)執行核苷酸表徵。使用Vector NTi 10.1版(Invitrogen Life Technologies, Carlsbad, CA)執行核苷酸比對。

亦獲得涵蓋來源於多個靶向整合事件之上游(5'-)IPP2K基因組序列與供體DNA之間的邊界之序列資料，包括涵蓋來源於多個靶向整合事件之供體DNA與下游(3'-)IPP2K基因組序列之間的邊界之序列資料以及包括來源於單個轉型愈傷組織事件(#114)之上游(5'-)邊界序列之序列資料。轉型靶向整合事件(#114)為自主供體整合至IPP2K基因標靶中之結果。

在該等分析中，初級及二級PCR擴增反應著重於基因組與供體之間的5'邊界或3'邊界。相應於該等二級擴增產物之經選殖片段與野生型IPP2K基因組序列以及靶向整合事件之預期序列之比對揭示，發生供體DNA在標靶位點處之整合。IPP2K基因組基因座之核苷酸序列、基因組/供體邊界、相應於IPP2K同源性側接序列之供體區域之核苷酸序列及除草劑耐受性卡匣之核苷酸序列全部保存於來源於該事件之多個選殖PCR產物中。

因此，該事件代表ZFN介導雙股斷裂之同源性驅動修復

在特定基因標靶處已發生之基因組。已獲得代表獨特靶向整合發生之其他轉型事件，證明本文中教示之方法可在玉米愈傷組織中再現。熟習此項技術者可將該等方法應用於認為需要靶向整合之任何植物物種中之任何基因標靶。

實例22：自玉米愈傷組織再生可孕之完整植物

來源於HiII細胞培養物之除草劑耐受性玉米細胞之經分離愈傷組織可再生成完整、可孕性玉米植株。熟習此項技術者可自多種胚胎發生玉米細胞培養物再生完整的可孕性玉米植株。

在該實例中，經分離、抗畢拉草HiII愈傷組織之再生係藉由將經分離愈傷組織轉移至基於細胞分裂素之誘導培養基28 (1H)中來引發，該培養基含有MS鹽及維生素、30.0 g/L蔗糖、5 mg/L苄胺基嘌呤、0.25 mg/L 2,4-D、1 mg/L畢拉草及2.5 g/L水晶洋菜；pH 5.7。使細胞在低光照(13 $\mu\text{Em-2s-1}$)下生長1週，隨後轉移至較高光照(40 $\mu\text{Em-2s-1}$)之條件下歷時1週。隨後將細胞轉移至再生培養基36 (1H)，其除缺乏植物生長調節劑外與誘導培養基相同。用手動工具切離小的(3-5 cm)幼苗且置放於含有SHGA培養基(Schenk及Hildebrandt基本鹽及維生素，1972, Can. J. Bot 50:199-204；1 g/L肌醇、10 g/L蔗糖、2.0 g/L水晶洋菜，pH 5.8)之無菌150×25-mm玻璃培養管中。

幼苗發育足夠大且分化出根及莖系統後，將其移植至含有Metro-Mix 360生長培養基(Sun Gro Horticulture Canada Ltd.)之4吋罐中且置放於溫室中。用透明塑膠杯將幼苗完

全或部分覆蓋2-7天，隨後移植至含有由95% Metro-Mix 360生長培養基及5%黏土/壤土組成之混合物之5加侖罐中，且生長至成熟。植物可自身授粉或與自交系交叉授粉以分別產生T1或F1種子。熟習此項技術者可使再生植物自身授粉或將再生植物與多種胚質交叉授粉以使玉米能育種。

與靶向裂解、靶向重組及靶向整合有關之其他資訊可見於美國專利申請公開案 US-2003-0232410；US-2005-0026157；US-2005-0064474；US-2005-0208489；及 US-2006-0188987中；及2006年7月26日申請之美國專利申請案第11/493,423號中，該等專利之揭示內容係以引用之方式全部併入本文中用於所有目的。

本文中提及之所有專利、專利申請案及公開案由此以引用之方式全部併入本文中用於所有目的。

儘管已以說明及實例之方式略微詳細地提供揭示內容以便清楚理解，但對熟習此項技術者顯而易見，可在不悖離本揭示案之精神或範疇之情況下實施各種改變和修改。因此，上述描述及實例不應解釋為具有限制性。

【圖式簡單說明】

圖1為描述在如美國專利第2005/0064474號及下文所述之GFP細胞報導體檢定系統中，如藉由表現GFP之細胞之百分比所量測的基因校正率之圖。指定ZFN變體為"X-Y"，其中"X"係指表編號且"Y"係指在經特定選擇之表中給銻指之編號。舉例而言，"2-21"係指具有包含表2中編號

21 列 中 所 示 之 序 列 ， 亦 即 HAQRCGLRGSQVLV(SEQ ID NO:53)之指的 ZFN。

圖 2 為 描 述 因 使 用 各 種 ZFN 變 體 對 裂 解 產 生 之 Cel-1 信 號 的 百 分 比 之 圖 。 藉 由 參 考 樣 本 編 號 ， 展 示 各 ZFN 對 之 2 個 實 驗 的 結 果 。 對 各 樣 本 所 用 之 變 體 對 展 示 於 右 上 方 角 之 方 框 中 ， 其 中 "wt 5-8" 及 "wt 5-9" 係 指 美 國 專 利 申 請 案 第 2005/0064474 號 之 實 例 14(表 17)中 所 揭 示 之 標 準 ZFN 對 。 在 樣 本 3-12 中 ， 標 準 ZFN 5-8 或 5-9 之 指 2 或 指 4 之 識 別 螺 旋 的 C 末 端 區 域 係 經 非 標 準 序 列 置 換 。 樣 本 3-12 中 指 定 為 20、21、43、45、47 及 48 之 非 標 準 ZFN 變 體 之 部 分 序 列 及 該 等 變 體 在 4 指 ZFN 內 之 指 位 置 展 示 於 該 圖 上 方 之 左 上 角 。 描 述 來 自 樣 本 8 及 9 之 實 驗 2 之 結 果 的 條 形 圖 上 方 之 星 號 指 示 泳 道 中 導 致 對 ZFN 功 效 低 估 之 背 景 。

圖 3 為 描 述 美 國 專 利 第 2005/0064474 號 及 本 文 中 所 述 之 GFP 細 胞 報 導 體 檢 定 系 統 中 之 基 因 校 正 率 的 圖 。 各 樣 本 中 所 測 試 之 ZFN 對 展 示 於 各 條 形 圖 下 方 ， 其 中 鋅 指 編 號 20、21、43、45、47 及 48 為 描 述 於 實 例 3 中 者 ， 且 CCHC 鋅 指 1a 至 10a 包 含 表 3 及 4 中 所 示 之 序 列 。 鋅 指 20、21、7a、8a、9a 及 10a 用 於 指 4 中 ； 鋅 指 43、45、47、48、1a、2a、3a、4a、5a 及 6a 用 於 指 2 中 。

圖 4 為 質 體 pDAB1585， 煙 草 之 標 靶 載 體 之 線 性 示 意 圖 。

圖 5 為 質 體 pDAB1585， 煙 草 之 標 靶 載 體 之 示 意 圖 。

圖 6A 及 6B 描 述 鋅 指 核 酸 酶 (ZFN)。 圖 6A 為 描 述 ZFN 結 合 之 示 意 圖 。 圖 6B 展 示 標 靶 序 列 之 序 列 。

圖 7 為質體 pDAB1400 之示意圖。

圖 8 為質體 pDAB782 之示意圖。

圖 9 為質體 pDAB1582 之示意圖。

圖 10 為質體 pDAB354 之示意圖。

圖 11 為質體 pDAB1583 之示意圖。

圖 12 為質體 pDAB2407 之示意圖。

圖 13 為質體 pDAB1584 之示意圖。

圖 14 為質體 pDAB2418 之示意圖。

圖 15 為質體 pDAB4045 之示意圖。

圖 16 為質體 pDAB1575 之示意圖。

圖 17 為質體 pDAB1577 之示意圖。

圖 18 為質體 pDAB1579 之示意圖。

圖 19 為質體 pDAB1580 之示意圖。

圖 20 為質體 pDAB3401 之示意圖。

圖 21 為質體 pDAB1570 之示意圖。

圖 22 為質體 pDAB1572 之示意圖。

圖 23 為質體 pDAB4003 之示意圖。

圖 24 為質體 pDAB1571 之示意圖。

圖 25 為質體 pDAB7204 之示意圖。

圖 26 為質體 pDAB1573 之示意圖。

圖 27 為質體 pDAB1574 之示意圖。

圖 28 為質體 pDAB1581 之示意圖。

圖 29 為質體 pDAB1576 之示意圖。

圖 30 為質體 pDAB1600 之示意圖。

圖 31 為 質 體 pDAB3731 之 示 意 圖 。

圖 32 為 質 體 pDAB4322 之 示 意 圖 。

圖 33 為 質 體 pDAB4331 之 示 意 圖 。

圖 34 為 質 體 pDAB4332 之 示 意 圖 。

圖 35 為 質 體 pDAB4333 之 示 意 圖 。

圖 36 為 質 體 pDAB4334 之 示 意 圖 。

圖 37 為 質 體 pDAB4336 之 示 意 圖 。

圖 38 為 質 體 pDAB4339 之 示 意 圖 。

圖 39 為 質 體 pDAB4321 之 示 意 圖 。

圖 40 為 質 體 pDAB4323 之 示 意 圖 。

圖 41 為 質 體 pDAB4341 之 示 意 圖 。

圖 42 為 質 體 pDAB4342 之 示 意 圖 。

圖 43 為 質 體 pDAB4343 之 示 意 圖 。

圖 44 為 質 體 pDAB4344 之 示 意 圖 。

圖 45 為 質 體 pDAB4346 之 示 意 圖 。

圖 46 為 質 體 pDAB4330 之 示 意 圖 。

圖 47 為 質 體 pDAB4351 之 示 意 圖 。

圖 48 為 質 體 pDAB4356 之 示 意 圖 。

圖 49 為 質 體 pDAB4359 之 示 意 圖 。

圖 50 為 質 體 pDAB7002 之 示 意 圖 。

圖 51 為 質 體 pDAB7025 之 示 意 圖 。

圖 52 為 質 體 pDAB1591 之 示 意 圖 。

圖 53 為 質 體 pcDNA3.1-SCD27a-L0-FokI ， 用 於 PCR 擴 增
Scd27 ZFN 之 DNA 模 板 的 示 意 圖 。

圖 54 為質體 pDAB1594 之示意圖。

圖 55 為質體 pDAB1598 之示意圖。

圖 56 為質體 pDAB1577 之示意圖。

圖 57 為質體 pDAB1578 之示意圖。

圖 58 為質體 pDAB1601，PAT 基因對照載體之示意圖。

圖 59 為描述由 IL-1-FokI 融合蛋白刺激之預測染色體內同源重組之示意圖。

圖 60 為質體 pDAB1590，陽性 GFP 表現對照之圖示。

圖 61 為描述由 IL-1 鋅指-FokI 融合蛋白刺激之預測染色體內同源重組之示意圖。

圖 62 為描述由 Scd27 鋅指-FokI 融合蛋白刺激之預測染色體內同源重組之示意圖。

圖 63 為描述重組體之 PCR 分析之凝膠。左邊前 4 個泳道標記於凝膠上方。標記為 1-5 之泳道展示因用 C3H IL-1-FokI 融合蛋白基因轉型 BY2-380 導致之 HR 事件，且標記為 6-7 之泳道展示因用 C3H SCD27-FokI 融合蛋白基因轉型 BY2-380 導致之 HR 事件。

圖 64 展示玉米 IPP2K 基因序列 (SEQ ID NO:6)，其來源於 HiII 細胞培養物且用作用於工程化靶向玉米 IPP2K 之 ZFN 之設計模板。

圖 65，圖 A 至 E，描述 ZFN 表現載體選殖流程。使用逐步選殖策略產生 ZFN 表現構築體。將個別編碼 ZFN 之基因選殖至載體 pVAX-N2A-NLSop2-EGFP-FokMono(A) 及 pVAX-C2A-NLSop2-EGFP-FokMono(B) 中以產生雙蛋白質卡匣

(C)。將該卡匣連接至pDAB3872(D)中以產生用於表現ZFN雜二聚體之最終質體(E)。

圖 66 描述玉米 IPP2K 基因中之 ZFN 結合。需要 2 個 ZFN 蛋白質執行 DNA 之雙股裂解。展示圍繞裂解位點(用向下箭頭指示)之序列(SEQ ID NO:7)。一蛋白質(8705)係結合於序列 CTGTGGGGCCAT(上鏈)(SEQ ID NO:8)，其中另一蛋白質(8684、8685 或 8686)結合於下游序列(CTTGACCAACTCAGCCAG，下鏈)(SEQ ID NO:9)。

圖 67 描述野生型之序列(上部序列，SEQ ID NO:10)及 ZFN 純系 127 之序列(下部序列，SEQ ID NO:11)。該 ZFN 之裂解標靶以灰色框標出。

圖 68 展示如由 454 定序所偵測，由玉米 IPP2K 基因中 ZFN 介導 dsDNA 斷裂之非同源末端聯接(NHEJ)產生的多個缺失之比對。該 ZFN 之裂解標靶以灰色框標出。

圖 69 為描述美國專利第 2005/0064474 號及本文中所述之 GFP 細胞報導體檢定系統中之基因校正率的圖。各樣本中測試之 ZFN 對展示於各條形圖下方。

圖 70 描述如實例 18B 中所述建構之質體 pDAB7471。

圖 71 描述如實例 18C 中所述建構之質體 pDAB7451。

圖 72 為描述例示性自主除草劑耐受性基因表現卡匣之示意圖。該建構包含含有啟動子、除草劑耐受性基因及如實例 18D 中所述之多聚腺嘌呤化(polyA)終止序列之完整啟動子轉錄單元(PTU)。

圖 73 描述如實例 18E 中所述建構之質體 pDAB7422。質體

包括含有啟動子、除草劑耐受性基因及插入位置1質體骨架中之多聚腺嘌呤化(polyA)終止序列之完整啟動子轉錄單元(PTU)。

圖74描述如實例18E中所述建構之質體pDAB7452。質體包括含有啟動子、除草劑耐受性基因及插入位置2質體骨架中之多聚腺嘌呤化(polyA)終止序列之完整啟動子轉錄單元(PTU)。

圖75為描述例示性非自主除草劑耐受性基因表現卡匣之示意圖。該建構包含含有除草劑耐受性基因及如實例18F中所述之多聚腺嘌呤化(polyA)終止序列之不完整啟動子轉錄單元(PTU)。

圖76描述如實例18G中所述建構之質體pDAB7423。該質體包括含有除草劑耐受性基因及插入位置1質體骨架中之多聚腺嘌呤化(polyA)終止序列之不完整啟動子轉錄單元(PTU)。

圖77描述如實例18G中所述建構之質體pDAB7454。質體包括含有除草劑耐受性基因及如實例18G中所述插入位置2質體骨架中之多聚腺嘌呤化(polyA)終止序列之不完整啟動子轉錄單元(PTU)。

圖78描述如實例18H中所述建構之質體pDAB 7424(例示性經Gateway®改適之位置1自主供體)。

圖79描述如實例18H中所述建構之質體pDAB 7425(例示性經Gateway®改適之位置1自主供體)。

圖80描述如實例18H中所述建構之質體pDAB 7426。

pDAB 7426為含有位置1自主供體與ZFN表現卡匣之組合質體。

圖 81描述如實例 18H中所述建構之質體 pDAB 7427。pDAB 7427為含有位置1自主供體與ZFN表現卡匣之組合質體。

圖 82描述自基因組DNA擴增供體-DNA特異性序列。317 bp產物之存在為含有如實例 20C中所述插入至玉米 calli系# 61-72之基因組中之PAT基因的供體DNA存在的診斷特徵。HiII指示野生型陰性對照。

圖 83描述供體-DNA與對IPP2K具有特異性之玉米基因組序列之間5'邊界的擴增。因供體靶向整合至IPP2K基因中獲得之第二PCR產物係如實例 21A中所述，藉由具有1.65 Kbp之DNA片段之存在來診斷。HiII指示野生型陰性對照。

圖 84描述供體-DNA與對IPP2K具有特異性之玉米基因組序列之間3'-邊界的擴增。因供體靶向整合至IPP2K基因中獲得之第二PCR產物係如實例 21A中所述，藉由具有1.99 Kbp之DNA片段之存在來診斷。HiII指示野生型陰性對照。

圖 85描述基因組與供體之間上游(5'-)邊界的擴增。因供體靶向整合至IPP2K基因(5'-邊界)中獲得之PCR產物係如實例 21B中所述，藉由大小為1.35 Kbp之DNA片段之存在來診斷。HiII指示野生型陰性對照。

圖 86描述供體與基因組之間下游(3'-)邊界的擴增。因供

體靶向整合至IPP2K基因(3'-邊界)中獲得之PCR產物係如實例21B中所述，藉由大小為1.66 Kbp之DNA片段之存在來診斷。HiII指示野生型陰性對照。

圖 87 描述位置 1 5'同源側接序列之序列(SEQ ID NO:171)。

圖 88 描述位置 1 3'同源側接序列之序列(SEQ ID NO:172)。

圖 89 描述位置 2 5'同源側接序列之序列(SEQ ID NO:139)。

圖 90 描述位置 2 3'同源側接序列之序列(SEQ ID NO:140)。

圖 91 描述 ZFN 靶向區域之上游(5'-)IPP2K 基因組序列之序列(SEQ ID NO:141)。

圖 92 描述 ZFN 靶向區域之下游(3'-)IPP2K 基因組序列之序列(SEQ ID NO:142)。

五、中文發明摘要：

本文中揭示包含CCHC鋅配位殘基之鋅指。本發明亦描述包含該等CCHC鋅指之鋅指蛋白及融合蛋白以及編碼該等蛋白質之聚核苷酸。本發明亦描述使用該等蛋白質用於基因編輯及基因調節之方法。

六、英文發明摘要：

Disclosed herein are zinc fingers comprising CCHC zinc coordinating residues. Also described are zinc finger proteins and fusion proteins comprising these CCHC zinc fingers as well as polynucleotides encoding these proteins. Methods of using these proteins for gene editing and gene regulation are also described.

七、指定代表圖：

(一)本案指定代表圖為：第 (2) 圖。

(二)本代表圖之元件符號簡單說明：

(無元件符號說明)

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

103年9月10日修(更)正本

第096147782號專利申請案
中文序列表(103年9月)

序列表

<110> 美商道禮責任有限公司
美商聖加莫生物科技公司

<120> 最佳化之非標準鋅指蛋白

<130> 8325-4002.40

<140> TW 096147782

<141> 2007-12-13

<150> 60/874,911

<151> 2006-12-14

<150> 60/932,497

<151> 2007-05-30

<160> 199

<170> PatentIn version 3.5

<210> 1

<211> 25

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<220>

<221> 其他特徵

<222> (2) .. (3)

<223> Xaa = 任何胺基酸

<220>

<221> MISC_FEATURE

<222> (4) .. (5)

<223> Xaa = 任何胺基酸

Xaa 可存在或不存在

<220>

<221> MISC_FEATURE

<222> (7) .. (18)

<223> Xaa = 任何胺基酸

<220>

<221> MISC_FEATURE

<222> (20) .. (22)

<223> Xaa = 任何胺基酸

i ;

<220>
<221> MISC_FEATURE
<222> (23) .. (24)
<223> Xaa = 任何胺基酸
Xaa 可存在或不存在

<400> 1

Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
20 25

<210> 2
<211> 6
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<220>
<221> MISC_FEATURE
<222> (2) .. (2)
<223> Xaa = 任何胺基酸(A, K, 或T為較佳)

<220>
<221> MISC_FEATURE
<222> (3) .. (3)
<223> Xaa = 任何胺基酸(Q, E, 或R為較佳)

<220>
<221> MISC_FEATURE
<222> (6) .. (6)
<223> Xaa = 任何胺基酸(G為較佳)

<400> 2

His Xaa Xaa Arg Cys Xaa
1 5

<210> 3
<211> 35
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<220>
<221> MISC_FEATURE
<222> (2) .. (3)
<223> Xaa = 任何胺基酸

<220>

<221> MISC_FEATURE
<222> (4) .. (5)
<223> Xaa = 任何胺基酸
Xaa 可存在或不存在

<220>
<221> MISC_FEATURE
<222> (7) .. (18)
<223> Xaa = 任何胺基酸

<220>
<221> MISC_FEATURE
<222> (20) .. (22)
<223> Xaa = 任何胺基酸

<220>
<221> MISC_FEATURE
<222> (23) .. (24)
<223> Xaa = 任何胺基酸
Xaa 可存在或不存在

<220>
<221> MISC_FEATURE
<222> (26) .. (26)
<223> Xaa = 任何胺基酸

<220>
<221> MISC_FEATURE
<222> (27) .. (35)
<223> Xaa = 任何胺基酸
Xaa 可存在或不存在

<400> 3

Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa
35

<210> 4
<211> 5
<212> PRT
<213> 人工

<220>
<223> ZC 連接子

<400> 4

Gly Leu Arg Gly Ser
1 5

<210> 5

<211> 6
<212> PRT
<213> 人工

<220>
<223> ZC 连接子

<400> 5

Gly Gly Leu Arg Gly Ser
1 5

<210> 6
<211> 1199
<212> DNA
<213> 玉米

<400> 6
atggagatgg atggggttct gcaagccgcg gatgccaagg actgggttta caagggggaa 60
ggcgccgcga atctcatcct cagctacacc ggctcgtcgc cctccatggg aagcgctgag 120
taggttctta ctgagcgtgc acgcatcgat cacttgactt taggggctca atgtgtgatt 180
cacgggtgcc gcggcgccat tcgagctcca gatccagtac cgctcgagca agtgataaaa 240
catggagcag ggacgatcac gtggtcactt gaaaattacg tgaggtccgg ggcgacgatg 300
tacggcgcgg cgaactctca aacactcaca caacaaaac cgcttcgtgt tcgtctttgt 360
tccaagcgac tgtgtgagtg tttgagagtt cgccagcgcg acatcgcccg atctgacaaa 420
ttaagctttc gttgcttttc catgattgtg cattttgtga gcatgcactg aatactatga 480
tggatatgtt tggaggaagc attattccaa tttgatgata aggggtgttat ttacacttgt 540
tttcagcttg gcaagggtact gcggctcaag aagattctaa aaaacaagtc gcagcgggca 600
ccgagttgta ttgtattctc aagtcatgag caactcctgt ggggccatat cccagaactg 660
gttgagtcgg tcaaacaaga ttgcttggct caagcctatg cagtgcattg tatgagccaa 720
cacctgggtg ccaatcatgt cgatggtggg gtatggttca gattcagttc atttatgtcc 780
tgttattgtg attttgattg gtaacatatt gacaacctcg acacttggga tcagattcag 840
ttcacttatg gaagaaattg gagaattgtt ataatttatc tataatcacc cctactgaaa 900
tagaaataac atggcatcaa tgtgcatgct attggatttt gacacgaata tgctttattc 960
tatcatatgt tggtaattcc agcaggcagc aggcactact ctttggatcc acgtgacttg 1020
acaaagaaat catgccatct ttccacaatg caggtccgtg tacgtgtttc tagggatttt 1080
ctggagcttg tcgaaaagaa tgttcttagc agccgtcctg ctgggagagt aaatgcaagt 1140
tcaattgata aactgctga tgccgctctt ctaatagcag accactcttt attttctgg 1199

<210> 7

<211> 50
<212> DNA
<213> 玉米

<400> 7
caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaaga 50

<210> 8
<211> 12
<212> DNA
<213> 玉米

<400> 8
ctgtggggcc at 12

<210> 9
<211> 18
<212> DNA
<213> 玉米

<400> 9
cttgaccaac tcagccag 18

<210> 10
<211> 59
<212> DNA
<213> 玉米

<400> 10
aagtcatgag caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaag 59

<210> 11
<211> 53
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 11
aagtcatgag caactcctgt ggggccaaga actgggtgag tcggtcaaac aag 53

<210> 12
<211> 12
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 12

His Thr Lys Ile His Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 13
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 13

His	Thr	Lys	Ile	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 14
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 14

His	Thr	Lys	Gly	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 15
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 15

His	Thr	Lys	Ala	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 16
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 16

His	Thr	Lys	Val	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 17
 <211> 12
 <212> PRT
 <213> 人工

<220>

<223> 鋅指基序

<400> 17

His	Thr	Lys	Leu	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 18

<211> 12

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 18

His	Thr	Lys	Ser	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 19

<211> 12

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 19

His	Thr	Lys	Asn	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 20

<211> 12

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 20

His	Thr	Lys	Lys	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10		

<210> 21

<211> 12

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 21

His Thr Lys Arg Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 22
<211> 14
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 22

His Thr Lys Ile Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 23
<211> 13
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 23

His Thr Lys Ile Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 24
<211> 14
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 24

His Thr Lys Ile Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 25
<211> 14
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 25

His Thr Lys Ile Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 26

<211> 15
 <212> PRT
 <213> 人工

<220>
 <223> 锌指基序

<400> 26

His	Thr	Lys	Ile	Gly	Cys	Gly	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5				10					15	

<210> 27
 <211> 13
 <212> PRT
 <213> 人工

<220>
 <223> 锌指基序

<400> 27

His	Leu	Lys	Gly	Asn	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5				10				

<210> 28
 <211> 13
 <212> PRT
 <213> 人工

<220>
 <223> 锌指基序

<400> 28

His	Leu	Lys	Gly	Asn	Cys	Pro	Ala	Gly	Ser	Gln	Leu	Val
1				5				10				

<210> 29
 <211> 13
 <212> PRT
 <213> 人工

<220>
 <223> 锌指基序

<400> 29

His	Ser	Glu	Gly	Gly	Cys	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5				10				

<210> 30
 <211> 13
 <212> PRT
 <213> 人工

<220>

<223> 锌指基序

<400> 30

His Ser Glu Gly Gly Cys Pro Gly Gly Ser Gln Leu Val
1 5 10

<210> 31

<211> 13

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 31

His Ser Ser Ser Asn Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 32

<211> 13

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 32

His Ser Ser Ser Asn Cys Thr Ile Gly Ser Gln Leu Val
1 5 10

<210> 33

<211> 15

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 33

His Thr Lys Ile Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 34

<211> 16

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 34

His Thr Lys Ile Gly Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val

1 5 10 15

<210> 35
<211> 14
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 35

His Thr Lys Ile Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 36
<211> 15
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 36

His Thr Lys Ile Gly Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 37
<211> 16
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 37

His Thr Lys Ile Gly Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 38
<211> 13
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 38

His Thr Lys Arg Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 39
<211> 14

<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 39

His Thr Lys Arg Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 40
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 40

His Thr Lys Arg Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 41
<211> 13
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 41

His Thr Lys Arg Gly Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 42
<211> 14
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 42

His Thr Lys Arg Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 43
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 43

His Thr Lys Arg Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 44

<211> 16

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 44

His Thr Lys Arg Gly Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 45

<211> 14

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 45

His Thr Lys Arg Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 46

<211> 15

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 46

His Thr Lys Arg Gly Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 47

<211> 16

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 47

His Thr Lys Arg Gly Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 48
 <211> 14
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 48

His	Leu	Lys	Gly	Asn	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				

<210> 49
 <211> 15
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 49

His	Leu	Lys	Gly	Asn	Cys	Gly	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				15	

<210> 50
 <211> 16
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 50

His	Leu	Lys	Gly	Asn	Cys	Gly	Gly	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10					15	

<210> 51
 <211> 13
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 51

His	Lys	Glu	Arg	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10			

<210> 52
 <211> 13
 <212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 52

His	Thr	Arg	Arg	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10			

<210> 53

<211> 13

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 53

His	Ala	Gln	Arg	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10			

<210> 54

<211> 14

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 54

His	Lys	Lys	Phe	Tyr	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				

<210> 55

<211> 14

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 55

His	Lys	Lys	His	Tyr	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				

<210> 56

<211> 14

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 56

His Lys Lys Tyr Thr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 57

<211> 14

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 57

His Lys Lys Tyr Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 58

<211> 14

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 58

His Lys Gln Tyr Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 59

<211> 14

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 59

His Leu Leu Lys Lys Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 60

<211> 14

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 60

His Gln Lys Phe Pro Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 61
 <211> 14
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 61

His Gln Lys Lys Leu Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 62
 <211> 14
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 62

His Gln Ile Arg Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 63
 <211> 15
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 63

His Ile Lys Arg Gln Ser Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 64
 <211> 15
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 64

His Ile Arg Arg Tyr Thr Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 65
 <211> 15
 <212> PRT
 <213> 人工

<220>
<223> 鋅指基序

<400> 65

His	Ile	Ser	Ser	Lys	Lys	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				15	

<210> 66
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 66

His	Lys	Ile	Gln	Lys	Ala	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				15	

<210> 67
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 67

His	Lys	Arg	Ile	Tyr	Thr	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				15	

<210> 68
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 68

His	Leu	Lys	Gly	Gln	Asn	Cys	Gly	Leu	Arg	Gly	Ser	Gln	Leu	Val
1				5					10				15	

<210> 69
<211> 15
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 69

His Leu Lys Lys Asp Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 70
<211> 15
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 70

His Leu Lys Tyr Thr Pro Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 71
<211> 12
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 71

His Thr Lys Arg Cys Gly Arg Gly Ser Gln Leu Val
1 5 10

<210> 72
<211> 14
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 72

His Thr Lys Ile Gly Cys Gly Gly Arg Gly Ser Gln Leu Val
1 5 10

<210> 73
<211> 13
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 73

His Leu Lys Gly Asn Cys Gly Arg Gly Ser Gln Leu Val
1 5 10

<210> 74
 <211> 13
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 74

His	Leu	Lys	Gly	Asn	Cys	Gly	Gly	Gly	Ser	Gln	Leu	Val
1				5					10			

<210> 75
 <211> 11
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 75

His	Ile	Arg	Thr	Cys	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10	

<210> 76
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 76

His	Ile	Arg	Thr	Cys	Gly	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10		

<210> 77
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 77

His	Ile	Arg	Thr	Gly	Cys	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10		

<210> 78
 <211> 13
 <212> PRT
 <213> 人工

<220>

<223> 锌指基序

<400> 78

His Ile Arg Thr Gly Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

<210> 79

<211> 11

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 79

His Ile Arg Arg Cys Thr Gly Ser Gln Lys Pro
1 5 10

<210> 80

<211> 12

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 80

His Ile Arg Arg Gly Cys Thr Gly Ser Gln Lys Pro
1 5 10

<210> 81

<211> 11

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 81

His Thr Lys Ile His Thr Gly Ser Gln Lys Pro
1 5 10

<210> 82

<211> 11

<212> PRT

<213> 人工

<220>

<223> 锌指基序

<400> 82

His	Thr	Lys	Ile	Cys	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10	

<210> 83
 <211> 11
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 83

His	Thr	Lys	Arg	Cys	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10	

<210> 84
 <211> 11
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 84

His	Ala	Gln	Arg	Cys	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10	

<210> 85
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 85

His	Thr	Lys	Ile	Cys	Gly	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10		

<210> 86
 <211> 12
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指基序

<400> 86

His	Thr	Lys	Arg	Cys	Gly	Thr	Gly	Ser	Gln	Lys	Pro
1				5					10		

<210> 87

<211> 12
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 87

His Ala Gln Arg Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

<210> 88
<211> 12
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 88

His Thr Lys Ile His Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 89
<211> 7
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 89

His Ala Gln Arg Cys Gly Gly
1 5

<210> 90
<211> 8
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 90

His Ala Gln Arg Cys Gly Gly Gly
1 5

<210> 91
<211> 8
<212> PRT
<213> 人工

<220>

<223> 鋅指基序

<400> 91

His Thr Lys Ile Cys Gly Gly Gly
1 5

<210> 92

<211> 8

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 92

His Thr Lys Arg Cys Gly Gly Gly
1 5

<210> 93

<211> 6

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 93

His Ala Gln Arg Cys Gly
1 5

<210> 94

<211> 7

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 94

Cys Cys His Thr Lys Ile His
1 5

<210> 95

<211> 7

<212> PRT

<213> 人工

<220>

<223> 鋅指基序

<400> 95

Cys Cys His Thr Lys Ile His

1 5

<210> 96
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 96

Cys Cys His Thr Lys Ile Cys
1 5

<210> 97
<211> 10
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 97

Cys Cys His Thr Lys Arg Cys Gly Gly Gly
1 5 10

<210> 98
<211> 8
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 98

Cys Cys His Ala Gln Arg Cys Gly
1 5

<210> 99
<211> 8
<212> PRT
<213> 人工

<220>
<223> 鋅指基序

<400> 99

Cys Cys His Ile Arg Thr Gly Cys
1 5

<210> 100
<211> 7

<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 100

Cys Cys His Thr Lys Ile His
1 5

<210> 101
<211> 8
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 101

Cys Cys His Ile Arg Thr Gly Cys
1 5

<210> 102
<211> 7
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 102

Cys Cys His Thr Lys Ile His
1 5

<210> 103
<211> 9
<212> PRT
<213> 人工

<220>
<223> 锌指基序

<400> 103

Cys Cys His Ala Gln Arg Cys Gly Gly
1 5

<210> 104
<211> 29
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 104
atggagatgg atggggttct gcaagccgc

29

<210> 105
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 105

Asp Arg Ser Ala Leu Ser Arg
1 5

<210> 106
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 106

Arg Asn Asp Asp Arg Lys Lys
1 5

<210> 107
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 107

Arg Ser Asp Asn Leu Ser Thr
1 5

<210> 108
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 108

His Ser His Ala Arg Ile Lys
1 5

<210> 109
 <211> 7
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指結合域

<400> 109

Arg Ser Asp Val Leu Ser Glu
 1 5

<210> 110
 <211> 7
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指結合域

<400> 110

Gln Ser Gly Asn Leu Ala Arg
 1 5

<210> 111
 <211> 7
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指結合域

<400> 111

Arg Ser Asp Asn Leu Ala Arg
 1 5

<210> 112
 <211> 7
 <212> PRT
 <213> 人工

<220>
 <223> 鋅指結合域

<400> 112

Thr Ser Gly Ser Leu Thr Arg
 1 5

<210> 113
 <211> 7
 <212> PRT
 <213> 人工

<220>
<223> 鋅指結合域

<400> 113

Thr Ser Gly Asn Leu Thr Arg
1 5

<210> 114
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 114

Arg Ser Asp His Leu Ser Glu
1 5

<210> 115
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 115

Gln Ser Ala Thr Arg Lys Lys
1 5

<210> 116
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 116

Glu Arg Gly Thr Leu Ala Arg
1 5

<210> 117
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 117

Arg Ser Asp Ala Leu Thr Gln
1 5

<210> 118
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 118

Arg Ser Asp Ser Leu Ser Ala
1 5

<210> 119
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 119

Arg Ser Ala Ala Leu Ala Arg
1 5

<210> 120
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 120

Arg Ser Asp Asn Leu Ser Glu
1 5

<210> 121
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 121

Ala Ser Lys Thr Arg Thr Asn
1 5

<210> 122

<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 122

Asp Arg Ser His Leu Ala Arg
1 5

<210> 123
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 123

Arg Ser Asp His Leu Ser Thr
1 5

<210> 124
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 124

Gln Ser Gly Ser Leu Thr Arg
1 5

<210> 125
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 125

Gln Asn His His Arg Ile Asn
1 5

<210> 126
<211> 7
<212> PRT
<213> 人工

<220>

<223> 鋅指結合域

<400> 126

Thr Gly Ser Asn Leu Thr Arg
1 5

<210> 127
<211> 7
<212> PRT
<213> 人工

<220>
<223> 鋅指結合域

<400> 127

Asp Arg Ser Ala Leu Ala Arg
1 5

<210> 128
<211> 18
<212> DNA
<213> 人工

<220>
<223> IPP2K鋅指標靶序列

<400> 128
gaactggttg agtcggtc 18

<210> 129
<211> 18
<212> DNA
<213> 人工

<220>
<223> IPP2K鋅指標靶序列

<400> 129
gaactggttg agtcggtc 18

<210> 130
<211> 12
<212> DNA
<213> 人工

<220>
<223> IPP2K鋅指標靶序列

<400> 130
atggccccac ag 12

<210> 131
<211> 15
<212> DNA

<213> 人工
<220>
<223> IPP2K鋅指標靶序列

<400> 131
ggcaccagg tggtg 15

<210> 132
<211> 18
<212> DNA
<213> 人工

<220>
<223> IPP2K鋅指標靶序列

<400> 132
gtcgatggtg gggtatgg 18

<210> 133
<211> 18
<212> PRT
<213> 人工

<220>
<223> 衍生自玉米op-2之NLS

<400> 133

Arg Lys Arg Lys Glu Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Tyr
1 5 10 15

Arg Lys

<210> 134
<211> 18
<212> PRT
<213> 人工

<220>
<223> 來自明脈扁刺蛾β四體病毒之2A序列

<400> 134

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15

Gly Pro

<210> 135
<211> 30
<212> DNA
<213> 人工

240

tgagtgtgca cgcacgatc acttgacttt aggggctcaa tgtgtgattc acgggtgccg	300
ccattcgagc tccagatcca gtatcgctcg agcaagtgat aaaacatgga gcagggacga	360
tcacgtggtc acttgaaaat tatgtgaggt ccggggcgac gatgtacggc gcggcgaact	420
ctcaaact cacaagcca aaaccgcttc gtgttcgtct ttgttccaag cgaccgtgtg	480
gtgtgtttgt agtagttcgc cggcgccgca catcgctgcc ccgatctga caaattaagc	540
tttcgttgct tttccacgat tgtgcatttt ctgagcatgc actgaatact atgatggata	600
tgtttggagg aagcattatt ccaatttgat gataatggtg ttatttacac ttgttttcag	660
cttggcaagg tactgcggct caagaagatt ctaaaaaaca agttgcagcg ggcaccaagt	720
tgtattgcct tctcaagtca tgagcaactc ctgtggggcc atatcccaga actggttgag	780
tcggtcaaac aagattgctt ggctcaagcc tatgcagtgc atgttatgag ccaacacctg	840
ggtgccaata ctagt	855

<210> 140
 <211> 845
 <212> DNA
 <213> 人工

<220>
 <223> 位置-2之3'-同源性側接序列

<400> 140	
actagtcatg tcgatggtag ggtatggctc agattcagtt catttatgtc ctgttattgt	60
gattttgatt ggtaacatat tgacaacctc gacacttggg atcagattca gttcacttat	120
ggaagaaatt ggagaattgt gataatttat ctataatcac ccctactgaa atagaaataa	180
catgacatca atgtgcatgc tattggattt tgacacgaat atgctttatt ctatcatatg	240
ttggtaattc cagcaggcag caggcactac tctttggatc cacgtgactt gacaaagaaa	300
tcatgccatc tttccacaat gcagggtcgt gtacgtgttt ctagggtatt tctggagctt	360
gtcgaaaaga atgttcttag cagccgtcct gctgggagag taaatgcaag ttcaattgat	420
aacactgctg atgccgtctt tctaatagca gaccactctt tattttctgg tacgtactct	480
atccctcttc ttaccataat ctgaatcttg ttaagggtta aaatatacga ttgattaagt	540
aaaatccaga gctctattca tatctcacgc actgatgttt tgatgaaacg cttgcagcaa	600
gacggttgcc tgttatttct atttgcatta gacaaacagt cacctttggt tataaagggtc	660
tttgaatttg cagttcttat aggtttaagt ttgcaactgt tacttacaac agcccaatgg	720
gtagcatcaa gattgttttt ttcagtgatt cataacttaa ctcttggtta aaccgctaga	780
acatggttgg tgtcttaaaa tgcaactggc cctgaggccg taacctgaaa tcattgtacg	840
tcgac	845

<210> 141
<211> 484
<212> DNA
<213> 人工

<220>
<223> ZFN標靶區域之上游5'-IPP2K基因體序列

<400> 141
tggacggagc gagagccaga attcgacgct ggcggcggcg cgtcgccaat acgcagcgcg 60
gatgtggagc cacatgcaaa cgtgtgtccg cccgcgtggc gtccactctc cctccacgtt 120
tcggcgtcct cgtcgccctc ctgggaaatc tccagctact gcccaactgcc ccttcctctc 180
agtccttttc cccgggctgt ggtaccagta ctagtaccag catctcttca ggctccacca 240
agcgcagaca ccgcagcagc ggcagcagca cgatccggtg accccccgcc gcgtccagcc 300
tgctcctccg gtgatcgccg gactggcggg gtaggaacca gcggagcgca gcccgccctc 360
ttccgctggg aagagtgacg cccgcccgct cctcccttcg ctcgcttctt tgctcttccg 420
attctggcgt accagtctca ccgcggcttg gggatttgat gcggagctag ttaaccagca 480
gagc 484

<210> 142
<211> 729
<212> DNA
<213> 人工

<220>
<223> ZFN標靶區域之下游3'-IPP2K基因體序列

<400> 142
attgtttttt tcagtgattc ataacttaac tcttggttaa accgctagaa catggttggg 60
gtcttaaaat gcaactggtc ctgaggccgt aacctgaaat cattgtactt ttctctcatt 120
tcttttagata tttccaaaac tctacattag atgatttatg tttgcttact tagtctttct 180
taatctcagg caatcctaag ggtagcagct gcatagctgt agagataaag gtactttgca 240
agcttcctct tttattctta tttttcattt cttatgtata tttctcctca accatttgac 300
ttcttttcgg catgctctac cttgcaggcc aaatgtgggt ttctgccatc atcagaatat 360
atatcagaag ataatactat caagaaacta gtaacgagat ataagatgca tcagcacctc 420
aaattttatc agggtgaggg gtgtagattg gaatgcttga tgccttgatc caagataaaa 480
ttccactctc ttttgcgcac ttaaaaaaca tccatcgatg atacaaactt gatcaaaata 540
ccttaaggct tggtatttac ggcactgttg taatattata ccgtctcttg ctttttgaca 600
tcaggttgat tccaatata ttcttgaca catttcagat atcgaagact agtgagtaca 660

atcctcttga tctattttct gggtcaaaag agagaatatg catggccatc aagtccttt 720
tctcaactc 729

<210> 143
<211> 42
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 143
gcggccgcgt ctcaccgcgg cttggggatt ggatacggag ct 42

<210> 144
<211> 38
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 144
actagtata tggccccaca ggagttgctc atgacttg 38

<210> 145
<211> 40
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 145
actagtcag aactggttga gtcggtcaaa caagattgct 40

<210> 146
<211> 32
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 146
gtcgaccttg atgctacca ttgggctgtt gt 32

<210> 147
<211> 30
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 147
gcggccgcta gatagcagat gcagattgct 30

<210> 148
<211> 29
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 148
actagtattg gcacccaggt gttggctca 29

<210> 149
<211> 38
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 149
actagtcattg tcgatgggtgg ggtatgggtc agattcag 38

<210> 150
<211> 37
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 150
gtcgacgtac aatgatttca gggtacggcc tcaggac 37

<210> 151
<211> 41
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 151
actagttaac tgacctcact cgaggtcatt catatgcttg a 41

<210> 152
<211> 29
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 152

actagtgtga attcagcact taaagatct

29

<210> 153

<211> 101

<212> DNA

<213> 人工

<220>

<223> 引子

<400> 153

actagtggcg gcggagaggg cagaggaagt cttctaacat gcggtgacgt ggaggagaat 60

cccggcccta ggatggcttc tccggagagg agaccagttg a 101

<210> 154

<211> 35

<212> DNA

<213> 人工

<220>

<223> 引子

<400> 154

actagtatgc atgtgaattc agcacttaaa gatct 35

<210> 155

<211> 28

<212> DNA

<213> 人工

<220>

<223> 引子

<400> 155

ctgtggtacc agtactagta ccagcatc 28

<210> 156

<211> 31

<212> DNA

<213> 人工

<220>

<223> 引子

<400> 156

tcttgatca aggcataag cattccaatc t 31

<210> 157

<211> 20

<212> DNA

<213> 人工

<220>

<223> 引子

<400> 157	
tggttaactg gcctaactgg	20
<210> 158	
<211> 20	
<212> DNA	
<213> 人工	
<220>	
<223> 引子	
<400> 158	
tggaaggcta ggaacgctta	20
<210> 159	
<211> 20	
<212> DNA	
<213> 人工	
<220>	
<223> 引子	
<400> 159	
ccagttaggc cagttaccca	20
<210> 160	
<211> 20	
<212> DNA	
<213> 人工	
<220>	
<223> 引子	
<400> 160	
taagcggtcc tagccttcca	20
<210> 161	
<211> 31	
<212> DNA	
<213> 人工	
<220>	
<223> 引子	
<400> 161	
cttggaagg tactgcggct caagaagatt c	31
<210> 162	
<211> 26	
<212> DNA	
<213> 人工	
<220>	
<223> 引子	
<400> 162	

atgaagaaag acagggaatg aaggac 26

<210> 163
<211> 32
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 163
atgaagaaag acagggaatg aaggaccgcc ac 32

<210> 164
<211> 28
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 164
catggagggc gacgagccgg ttagctg 28

<210> 165
<211> 27
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 165
atcgacatga ttggcaccca ggtgttg 27

<210> 166
<211> 31
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 166
tttcgacaag ctccagaaaa tccctagaaa c 31

<210> 167
<211> 28
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 167
acaagctcca gaaaatccct agaaacac 28

<210> 168
<211> 32
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 168
ttcgacaagc tccagaaaat ccctagaaac ac 32

<210> 169
<211> 29
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 169
tgctaagaac attcttttcg acaagctcc 29

<210> 170
<211> 32
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 170
gaacattctt ttcgacaagc tccagaaaat cc 32

<210> 171
<211> 30
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 171
tggacggagc gagagccaga attcgacgct 30

<210> 172
<211> 31
<212> DNA
<213> 人工

<220>
<223> 引子

<400> 172
gtgcaagaat gtattgggaa tcaacctgat g 31

<210> 173
 <211> 59
 <212> DNA
 <213> Zea 玉米

<400> 173
 aagtcatgag caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaag 59

<210> 174
 <211> 53
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 174
 aagtcatgag caactcctgt ggggccaaaga actgggtgag tcggtcaaac aag 53

<210> 175
 <211> 55
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 175
 aagtcatgag caactcctgt ggggccataa gaactgggtg agtcggtcaa acaag 55

<210> 176
 <211> 54
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 176
 aagtcatgag caactcctgt ggggccatag aactgggtga gtcggtcaaa caag 54

<210> 177
 <211> 56
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 177
 aagtcatgag caactcctgt ggggccatac agaactgggtt gagtcggtca aacaag 56

<210> 178
 <211> 57
 <212> DNA

<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 178
aagtcatgag caactcctgt ggggccatac cagaactggt tgagtcgggc aaacaag 57

<210> 179
<211> 52
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 179
aagtcatgag caactcctgt ggggccagaa ctgggtgagt cgggtcaaaca ag 52

<210> 180
<211> 57
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 180
aagtcatgag caactcctgt ggggccatat cagaactggt tgagtcgggc aaacaag 57

<210> 181
<211> 54
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 181
aagtcatgag caactcctgt ggggccatag aactggttga gtcggtcaaa caag 54

<210> 182
<211> 54
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 182
aagtcatgag caactcctgt ggggccatag aactggttga gtcggtcaaa caag 54

<210> 183
<211> 57
<212> DNA
<213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 183
 aagtcacgag caactcctgg tggggccata cagaactggt tgagtcgggc aaacaag 57

<210> 184
 <211> 54
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 184
 aagtcacgag caactcctgt gggggccatag aactgggtga gtcgggtcaaa caag 54

<210> 185
 <211> 53
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 185
 aagcatgagc aactcctgtg gggccataga actgggttgag tcgggtcaaac aag 53

<210> 186
 <211> 56
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 186
 aagtcacgag caactcctgt gggggccatac agaactgggt gagtcgggtca aacaag 56

<210> 187
 <211> 57
 <212> DNA
 <213> 人工

<220>
 <223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 187
 aagtcacgag caactcctgt gggggccatat cagaactggt tgagtcgggc aaacaag 57

<210> 188
 <211> 54
 <212> DNA
 <213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 188
aagtcatgag caactcctgt ggggccatag aactggttga gtcggtcaaa caag 54

<210> 189
<211> 54
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 189
aagtcatgag caactcctgt ggggccacag aactggttga gtcggtcaaa caag 54

<210> 190
<211> 54
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 190
aagtcatgag caactcctgt ggggccatag aactggttga gtcggtcaaa caag 54

<210> 191
<211> 57
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 191
aagtcatgag caactcctgt ggggccatac cagaactggt tgagtcggtc aaacaag 57

<210> 192
<211> 54
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 192
aagtcatgag caactcctgt ggggccatag aactggttga gtcggtcaaa caag 54

<210> 193
<211> 55
<212> DNA
<213> 人工

<220>



<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 193
aagtcattgag caactcctgt ggggccataa gaactgggtg agtcgggtcaa acaag 55

<210> 194
<211> 56
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 194
aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag 56

<210> 195
<211> 56
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 195
aagtcattgag caactcctgt ggggccatat agaactgggt gagtcgggtca aacaag 56

<210> 196
<211> 56
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 196
aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag 56

<210> 197
<211> 57
<212> DNA
<213> 人工

<220>
<223> 具ZFN媒介突變的玉米IPP2K基因序列

<400> 197
aagtcattgag caactcctgt ggggccatac cagaactgggt tgagtcgggtc aaacaag 57

<210> 198
<211> 821
<212> DNA
<213> 人工

<220>
<223> 位置-1之5'-同源性側接序列

```

<400> 198
gcggccgcgt ctcaccgcgg cttgggggatt ggatacggag ctagttaacc agcagagcta      60
gatagcagac gcagatcgct tgcttctctg gtttgatttt tggagtcacc atttctgttt      120
ggttcgtgtg cctcagtgtc tgacagcagc agatcctcga tggagatgga tgggggttctg      180
caagccgcgg atgccaaagga ctgggtttac aagggggaag gcgccgcgaa tctcatcctc      240
agctacaccg gcacgtcgcc ctccatggta agcgtgagt aggttcttac tgagtgtgca      300
cgcatcgatc acttgacttt aggggctcaa tgtgtgattc acgggtgccg ccattcgagc      360
tccagatcca gtatcgctcg agcaagtgat aaaacatgga gcagggacga tcacgtggtc      420
acttgaaaat tatgtgaggt ccggggcgac gatgtacggc gcggcgaaact ctcaaact      480
cacacagcca aaaccgcttc gtgttcgtct ttgttccaag cgaccgtgtg gtgtgtttgt      540
agtagttcgc cggcgccgca catcgctgcc ccggatctga caaattaagc tttcgttgct      600
tttccacgat tgtgcatttt ctgagcatgc actgaatact atgatggata tgtttggagg      660
aagcattatt ccaatttgat gataatgggtg ttatttacac ttgttttcag cttggcaagg      720
tactgcggt caagaagatt ctaaaaaaca agttgcagcg ggcaccaagt tgtattgcct      780
tctcaagtca tgagcaactc ctgtggggcc atatcactag t      821

```

```

<210> 199
<211> 821
<212> DNA
<213> 人工

```

```

<220>
<223> 位置-1之3'-同源性侧接序列

```

```

<400> 199
actagtccag aactggttga gtcggtcaaa caagattgct tggctcaagc ctatgcagtg      60
catgttatga gccaacacct gggtgccaat catgtcgatg gtggggtatg gttcagattc      120
agttcattta tgtcctgtta ttgtgatttt gattggtaac atattgacaa cctcgacact      180
tgggatcaga ttcagttcac ttatggaaga aattggagaa ttgtgataat ttatctataa      240
tcaccctac tgaaatagaa ataacatgac atcaatgtgc atgctattgg attttgacac      300
gaatatgctt tattctatca tatgttggtg attccagcag gcagcaggca ctactctttg      360
gatccacgtg acttgacaaa gaaatcatgc catctttcca caatgcaggt ccgtgtacgt      420
gtttctaggg attttctgga gcttgtcgaa aagaatgttc ttagcagccg tcctgctggg      480
agagtaaatg caagttcaat tgataaact gctgatgccg ctcttctaag agcagaccac      540
tctttatttt ctggtacgta ctctatccct cttcttacca taatctgaat cttgttaagg      600
tttaaaatat atgattgatt aagtaaaatc cagagctcta ttcatatctc acgcactgat      660

```

gttttgatga aacgcttgca gcaagacggt tgcttggtat ttctatttgc attagacaaa	720
cagtcacctt tgtttataaa ggtctttgaa tttgcagttc ttatagggtt aagtttgcaa	780
ctgttactta caacagccca atgggtagca tcaaggtcga c	821

十、申請專利範圍：

1. 一種鋅指蛋白，其包含非標準(非 C_2H_2)鋅指，其中該非標準鋅指具有與DNA結合有關之螺旋部分，且其中至少一個鋅指包含序列 $Cys-(X^A)_{2-4}-Cys-(X^B)_{12}-His-(X^C)_{3-5}-Cys-(X^D)_{1-10}$ (SEQ ID NO:3)，其中 X^A 、 X^B 、 X^C 及 X^D 可為任何胺基酸，且其中
 - (i) 該鋅指蛋白係經工程化以與標靶序列結合；
 - (ii) $(X^D)_{1-10}$ 包含選自由下列所組成之群之胺基酸序列：QLV、QKP及直接接在C末端Cys殘基之C末端的G殘基；
且
 - (iii) 當 $(X^D)_{1-10}$ 包含胺基酸序列QKP且 $(X^C)_{3-5}$ 由3個胺基酸所組成時，該殘基係選自由下列所組成之群：IRT、IRR、TKI及AQR。
2. 如請求項1之鋅指蛋白，其包含SEQ ID NO:13-87及89-93所示之序列。
3. 如請求項1之鋅指蛋白，其中 $(X^D)_{1-10}$ 包含序列QLV、QKP、G、GG或GGG。
4. 一種鋅指蛋白，其包含複數個鋅指，其中至少一個鋅指包含如請求項1至3中任一項之CCHC鋅指。
5. 如請求項1至4中任一項之鋅指蛋白，其中該鋅指蛋白包含4、5或6個鋅指蛋白域，該等鋅指蛋白域包含下列F1至F4、F1至F5或F1至F6順序的識別螺旋區域：
 - (i) F1: DRSALASR (SEQ ID NO: 105)；
F2: RNDDRKK (SEQ ID NO: 106)；

F3: RSDNLST (SEQ ID NO: 107)或 RSDNLAR (SEQ ID NO: 111)或 TSGNLTR (SEQ ID NO: 113) ;

F4: HSHARIK (SEQ ID NO: 108)或 TSGSLTR (SEQ ID NO: 112) ;

F5: RSDVLSE (SEQ ID NO: 109) ; 及

F6: QSGNLAR (SEQ ID NO: 110) ;

(ii) F1: RSDHLSE (SEQ ID NO: 114) ;

F2: QSATRKK (SEQ ID NO: 115) ;

F3: ERGTLAR (SEQ ID NO: 116) ; 及

F4: RSDALTQ (SEQ ID NO: 117) ;

(iii) F1: RSDSLSA (SEQ ID NO: 118) ;

F2: RSAALAR (SEQ ID NO: 119) ;

F3: RSDNLSE (SEQ ID NO: 120) ;

F4: ASKTRTN (SEQ ID NO: 121) ; 及

F5: DRSHLAR (SEQ ID NO: 122) ; 或

(iv) F1: RSDHLST (SEQ ID NO: 123) ;

F2: QSGSLTR (SEQ ID NO: 124) ;

F3: RSDHLSE (SEQ ID NO: 114) ;

F4: QNHHRIN (SEQ ID NO: 125) ;

F5: TGSNLTR (SEQ ID NO: 126) ; 及

F6: DRSALAR (SEQ ID NO: 127) ;

- 且其中該鋅指蛋白進一步係經工程化以與IPP2-K基因中之標靶序列結合。
6. 一種融合蛋白，其包含如請求項1至5中任一項之鋅指蛋白及一或多個功能域。
 7. 如請求項6之融合蛋白，其中該功能域包含裂解半域，且其中該融合蛋白包含插入於該裂解半域與該鋅指蛋白之間的ZC連接子。
 8. 如請求項7之融合蛋白，其中該ZC連接子為5至6個胺基酸長。
 9. 一種聚核苷酸，其編碼如請求項1至5中任一項之鋅指蛋白或如請求項6至8中任一項之融合蛋白。
 10. 一種植物細胞，其包含如請求項1至5中任一項之鋅指蛋白、如請求項6至8中任何一項之融合蛋白或如請求項9之聚核苷酸。
 11. 如請求項10之植物細胞，其中IPP2-K係部分或完全失活且該細胞中之植酸含量降低。
 12. 一種製備包含至少一種如請求項1至5中任一項之鋅指蛋白、至少一種如請求項6至8中任一項之融合蛋白或至少一種如請求項9之聚核苷酸之種子的方法，該方法包含製備含有至少一種如請求項9之聚核苷酸及/或穩定表現至少一種如請求項1至5中任一項之鋅指蛋白或至少一種如請求項6至8中任一項之融合蛋白的轉殖基因植物，及繁殖該轉殖基因植物。
 13. 如請求項12之方法，其中該種子中之IPP2-K係部分或完

全失活且該種子中之植酸含量係降低。

14. 一種靶向裂解植物細胞中細胞染色質之方法，該方法包含在該細胞中表現一對如請求項7之融合蛋白或至少一種編碼該一對融合蛋白之聚核苷酸；其中：

(a) 該等融合蛋白之標靶序列彼此在10個核苷酸以內；且

(b) 該等融合蛋白二聚化並裂解位於該等標靶序列之間的DNA。

15. 一種在宿主植物細胞中靶向基因重組之方法，該方法包含：

(a) 在該宿主細胞中表現一對如請求項7之融合蛋白或至少一種編碼該一對融合蛋白之聚核苷酸，其中該等融合蛋白之標靶序列存在於所選的宿主標靶基因座中；

(b) 鑑別在該宿主標靶基因座中具有序列變異之重組宿主細胞；及

(c) 視情況將外源性聚核苷酸引入該宿主細胞中，其中，當外源性聚核苷酸存在時，其係嵌入該宿主植物細胞之基因體中。

16. 如請求項15之方法，其中該序列變異係選自由遺傳物質缺失、遺傳物質插入、遺傳物質取代及其任何組合組成之群的突變。

17. 一種降低植物種子中植酸含量或使磷在植物種子中更具有代謝可利用性之方法，該方法包含使如請求項15之IPP2-K基因失活或變異。

