



(19)中華民國智慧財產局

(12)發明說明書公告本 (11)證書號數：TW I669395 B

(45)公告日：中華民國 108 (2019) 年 08 月 21 日

(21)申請案號：103138081

(22)申請日：中華民國 103 (2014) 年 11 月 03 日

(51)Int. Cl. : C12N15/63 (2006.01)

C12N15/66 (2006.01)

C12N15/82 (2006.01)

C07H21/04 (2006.01)

(30)優先權：2013/11/04 美國

61/899,543

(71)申請人：美商陶氏農業科學公司 (美國) DOW AGROSCIENCES LLC (US)
美國(72)發明人：沙士翠 丹特 拉克希米 SASTRY-DENT, LAKSHMI (US)；伊凡斯 史帝芬 L
EVANS, STEVEN L. (US)；布魯 萊恩 C BLUE, RYAN C. (US)；曹澤惠 CAO,
ZEHUI (CN)

(74)代理人：惲軼群；陳文郎

(56)參考文獻：

US 2011/0191899A1

審查人員：黃教威

申請專利範圍項數：16 項 圖式數：17 共 213 頁

(54)名稱

一種用於基因標定的通用供體系統

A UNIVERSAL DONOR SYSTEM FOR GENE TARGETING

(57)摘要

一種通用的供體聚核苷酸係描述的，該者可以在植物基因組中於一標定位置插入，以促進一供體分子在一特定基因組位置中快速及高通量的併入。

A universal donor polynucleotide is described that can be inserted at targeted locations in plant genomes to facilitate rapid and high throughput integration of a donor molecule within a specific genomic location.

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

一種用於基因標定的通用供體系統

A UNIVERSAL DONOR SYSTEM FOR GENE TARGETING

【技術領域】

相關申請案之交互參照

[0001]此申請案依35 U.S.C. §119(e)以2013年11月4日提申之美國專利臨時申請案第61/899,543號主張其利益，該者之內容係以其整體合併至本申請案以做為參考。

參照電子方式提交之序列表

[0002]序列表的官方拷貝係經由EFS-網站，以2014年10月27日創建並具有100千位元組大小、命名為”74435_ST25.txt”檔案的ASCII格式之序列表電子提交地，且係與本說明書同時提申。含在此ASCII格式文件中之該序列表係為本說明書之一部分，且係以其整體合併於此以作為參考。

發明領域

[0003]本揭露內容有關植物精確轉形、基因標定、標定基因組併入及在植物中蛋白質表現的領域。在一實施例中，本揭露內容描述一種通用供體聚核苷酸分子，該者可以在植物基因組中於一標定位置插入，並在高通量且有效率的方式中檢測。

【先前技術】

發明背景

[0004]引入農藝學上重要之性狀至植物可以用來改良營養價值的質量、提高產量、提供害蟲或疾病抗性、提高乾旱及壓力耐受性、改良園藝品質(例如，呈色(pigmentation)及開花)、授予除草劑抗性、使得能夠從植物生產工業上有用之化合物及/或材料，及/或能夠生產醫藥品。轉基因植物典型地係藉由農桿菌介導的轉形技術生成。其它轉形技術諸如原生質體中PEG介導的DNA攝入、微粒轟擊法(microprojectile bombardment)及矽晶鬚(silicon whisker)介導的轉形作用，亦可以使用以產生轉基因植物。引入選殖基因至植物細胞中並再生穩定的可育性轉基因植物可以用於對植物做此種的修飾並允許植物用於作物改良之基因工程。使用這些方法，外來DNA係隨機併入到該真核植物細胞的細胞核或質粒體DNA內，繼之分離含有併入到該細胞DNA之外來DNA的細胞，以產生穩定的轉形植物細胞。

[0005]上文描述該等植物轉形方法引致轉基因在該植物基因組內之隨機位置的併入，並在不同的拷貝數。經常地，該轉基因係以重複體併入，不是整個轉基因就是其部分。複合併入模式可能影響該轉基因之表現量(例如藉由透過轉錄後基因靜默機制破壞該經轉錄RNA，或藉由誘導該引入DNA的甲基化)，從而下調該轉基因之轉錄。更進一步，該轉基因之表現量可以藉由在該植物基因組內之併入位置(例如，位置效應)而影響。這些因素之組合引致在不同的轉基因植物細胞及植物品系之間，該轉基因或感興趣的外來

DNA在表現量中一很大的變化。再者，該感興趣外來DNA的併入可能對該併入發生處之基因組區域上具有一破壞性效應，且可以影響或干擾該標靶區域的正常功能，從而常常導致非所欲的副作用。

[0006]前述迫使，每當一特定的外來DNA引入至植物的效果係調查時，大量的轉基因植物品系係生成並分析，以為了獲得顯著的結果。同樣地，在轉基因作物植物生成中，其中一感興趣的特定DNA係引入至植物中，以提供具所欲表型的轉基因植物，大族群的獨立創建轉基因植物品系係創建的，以允許選擇那些具轉基因最佳表現的植物品系，並在該轉基因植物之總表型上伴隨最少或沒有副作用的。特別在此領域中，更導向的轉基因方法係為所欲的，舉例而言，有鑑於繁重的法規要求及關聯於對消除不想要的轉基因事件所要求之重複提申試驗的高費用。更進一步，其將為清楚的是，標定DNA插入的可能性在轉基因堆疊(*transgene stacking*)的過程中亦將為有利的。

[0007]在努力控制植物中轉基因插入，數種方法已經發展的。參閱，例如，Kumar及Fladung於”(2001)*Trends Plant Sci.* 6: 155-9”。這些方法依賴基於同源重組的轉基因併入。此策略已成功應用於原核生物及低等真核生物。Paszkowski等人於”(1988)*EMBO J.* 7: 4021-6”。然 而，直到最近，對於植物該用於轉基因併入之主要機制係基於該重組DNA各股之間涉及很小同源性的不正常重組(*illegitimate recombination*)。所以，此領域之主要挑戰為，罕見同源重

組事件之偵測，其係由經由不正常重組之更有效率引入的外來DNA併入而遮蔽。

[0008]客製化鋅指核酸酶(ZFN)係為蛋白質其設計以在DNA中發動一標定的位點特異性雙股斷裂，伴隨該切開端隨後之重組。ZFN組合一限制性核酸內切酶之非特異性切割域，諸如舉例而言*FokI*，與鋅指DNA結合蛋白。參閱，例如，Huang等人於”(1996)J.Protein Chem.15：481-9”；Kim等人於”(1997)Proc.Natl.Acad.Sci.USA 94：3616-20”；Kim等人於”(1996) Proc.Natl.Acad.Sci.USA 93：1156-60”；Kim等人於”(1994)Proc Natl.Acad.Sci.USA 91：883-7”；Kim等人於”(1997b)Proc.Natl.Acad.Sci.USA 94：12875-9”；Kim等人於”(1997c)Gene 203 : 43-9”；Kim等人於”(1998)Biol.Chem.379 : 489-95”； Nahon 及 Raveh 於”(1998)Nucleic Acids Res.26 : 1233-9”； Smith 等人於”(1999)Nucleic Acids Res.27 : 674-81”。個別的鋅指模體可以設計以標定並結合至一大範圍的DNA位點。標準的Cys₂His₂以及非標準的Cys₃His鋅指蛋白質係藉由插入 α -螺旋至DNA雙螺旋之大溝槽而結合至DNA。藉由鋅指識別DNA係模組化的：每一指主要接觸該標靶中的三個連貫鹼基對，而在該蛋白質中之些關鍵殘基介導識別。其係已經顯示的是，*FokI*限制性核酸內切酶必需經由該核酸酶域二聚體化，以切割DNA，引入一雙股斷裂。類似地，ZFN亦要求該核酸酶域二聚體化，以切割DNA。Mani等人於”(2005)Biochem.Biophys.Res.Commun.334 : 1191-7”；

Smith等人於”(2000)Nucleic Acids Res. 28 : 3361-9”。ZFN 之二聚體化係藉由兩個鄰近、相反面向的結合位點促進的。
id.

[0009]前述相關技術之例子及關於其之限制係意欲例示且非排他性的。一旦閱讀本說明書，該相關技藝之其它限制對熟習該項技藝者將變得顯而易見的。

【發明內容】

發明概要

[0010]在一較佳實施例中，本揭露內容有關於一種聚核苷酸供體卡匣，其包含一位點特異性核酸酶結合域、一解析域(analytical domain)及一質體域。在一進一步實施例中，該聚核苷酸供體卡匣包含小於3 Kbp之一長度。進一步實施例包括一轉基因細胞其包含該聚核苷酸供體卡匣者。在一些觀點中，該轉基因細胞係為一轉基因植物細胞。在另一實施例中，本揭露內容有關於一轉基因植物其包含該轉基因植物細胞者。進一步觀點包括該轉基因植物，其中該植物係為一單子葉植物或一雙子葉植物。在其它實施例中該單子葉植物係選自由玉米植物、小麥植物及水稻植物所組成之該群組。在又其它實施例中該雙子葉植物係選自由大豆植物、棉花植物及油菜植物所組成之該群組。還進一步的，本揭露內容有關於從該轉基因植物產生的一轉基因種子。

[0011]在又另一實施例中，本主體揭露內容有關該聚核苷酸供體卡匣，其中該聚核苷酸供體卡匣係選自由下列所

組成之一群組：對序列辨識編號：132具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：133具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：134具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：135具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：136具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：137具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：138具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：139具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：140具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：141具至少90%序列一致性的一聚核苷酸序列。

[0012]在進一步實施例中，該聚核苷酸供體卡匣包含一位點特異性核酸酶結合域，該者係由一或多個位點特異性核酸酶結合序列構成。在一觀點中，該位點特異性核酸酶結合域包含一或多個由一鋅指結合蛋白、巨核酸酶結合蛋白、CRISPR結合蛋白或TALEN結合蛋白所結合的位點特異性核酸酶結合序列。

[0013]在其它實施例中，該聚核苷酸供體卡匣包含一解析域其選自由下列所組成之該群組：一聚核苷酸其對序列辨識編號：142具至少90%的序列一致性，及一聚核苷酸序列其對序列辨識編號：143具至少90%的序列一致性。在一些觀點中，該解析域包含一或多個限制酵素序列。在其它觀點中，該解析域包含40至60%百分比之鳥糞嘌呤與胞嘧

啶鹼基對。在進一步觀點中，該解析域包含一聚核苷酸其不含有數個重複序列者，其中該重複序列長度係至少9 Bp。在另外觀點中，該解析域包含一聚核苷酸其不含有長度大於9 Bp之一系列一致的鹼基對序列。在一觀點中，該解析域包含一聚核苷酸其具有大於-19 kcal/mol自由能(ΔG)之一二級結構。在各種其它觀點中，該解析域包含一或多個引子結合序列。

[0014]在一實施例中，該聚核苷酸供體卡匣包含一或多個同源臂序列。在其它觀點中，該一或多個同源臂序列長度係於50及100個鹼基對之間。在另一實施例中，該解析域不編碼一勝肽。在一又一實施例中，該解析域編碼一勝肽。在一觀點中，該解析域包含一基因表現卡匣其包含一轉基因者。在一進一步觀點中，該基因表現卡匣包含一啓動子。在又另一觀點中，該轉基因包含一報告基因。在一另外觀點中，該報告基因係選自由下列所組成之該群組：*yfp*基因、*gus*基因、*rfp*基因、*gfp*基因、卡那黴素抗性基因、*aad-1*基因、*aad-12*基因、*pat*基因及草甘膦耐受性基因。在一另外實施例中，該質體域包含一pUC19質體。在一觀點中，該質體域包含一高拷貝數複製起點。在一進一步觀點中，該高拷貝數複製起點包含一*cole1*複製起點。在一另外觀點中，該質體域包含一可選擇標記。在一進一步觀點中，該可選擇標記係選自由下列所組成之該群組：卡那黴素可選擇標記、安比西林可選擇標記、觀黴素可選擇標記、氯黴素可選擇標記、及利福平可選擇標記。

[0015]在一實施例中，本主體揭露內容有關於在一植物細胞之基因組內標定併入一聚核苷酸供體卡匣的方法。在一觀點中，該方法有關於表現一位點特異性DNA結合核酸酶，該者包含至少一DNA結合域及至少一核酸酶域，其中該至少一DNA結合域結合至該植物細胞基因組內的一標靶位點；將該植物細胞與一聚核苷酸供體卡匣接觸；以該位點特異性DNA結合核酸酶切開在該植物細胞基因組內的標靶位點；且將該聚核苷酸供體卡匣併入至該植物細胞基因組內的標靶位點中。在一進一步觀點中，該至少一DNA結合域係選自由下列所組成之該群組：一鋅指結合域、一亘核酸酶結合域、一TALEN結合域、一亘核酸酶及一CRISPR結合域。在另一觀點中，該核酸酶域係源自一IIS型限制性核酸內切酶。在另一觀點中，該IIS型限制性核酸內切酶係選自由FokI及StsI所組成之該群組。在一另外觀點中，該聚核苷酸供體卡匣表現一多勝肽。在又另一觀點中，該聚核苷酸供體卡匣包含一非編碼的核酸序列。在一觀點中，該聚核苷酸供體卡匣包含一或多個鋅指結合序列。在另一觀點中，該聚核苷酸供體卡匣包含一或多個引子結合序列。在一另外觀點中，該聚核苷酸供體卡匣之併入係經由一同源性導向修復機制發生的。在一另外觀點中，該聚核苷酸供體卡匣之併入係經由一非同源性末端接合導向修復機制發生的。在一實施例中，選擇用於併入該聚核苷酸供體卡匣之該植物細胞係為一單子葉或雙子葉植物細胞。在一實施例中，該植物細胞係為選自由玉米植物細胞、小麥植物

細胞及水稻植物細胞所組成之該群組的一單子葉植物細胞。在一實施例中，該植物細胞係為選自由大豆植物細胞、棉花植物細胞及油菜植物細胞所組成之該群組的一雙子葉植物細胞。

[0016]除了上文描述之該示範性觀點與實施例外，進一步的觀點與實施例藉由研讀下列說明將變成顯而易見的。

【圖式簡單說明】

[0017]圖1例示pDAB111845之質體圖譜。該等編號元件(意即，5、7、8、9、10、11、12、15、16、25、與26)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。

[0018]圖2例示pDAB111846之質體圖譜。該等編號元件(意即，1、2、5、6、11、12、15、16、21、22、29與30)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。

[0019]圖3例示pDAB117415之質體圖譜。該等編號元件(意即，ZFN51與ZFN52)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識

別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0020] 圖4例示pDAB117416之質體圖譜。該等編號元件(意即，ZFN54與ZFN53)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0021] 圖5例示pDAB117417之質體圖譜。該等編號元件(意即，ZFN55)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0022] 圖6例示pDAB117419之質體圖譜。該等編號元件(意即，ZFN59與ZFN60)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0023] 圖7例示pDAB117434之質體圖譜。該等編號元件(意即，ZFN66、ZFN67、ZFN68與ZFN69)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0024] 圖8例示pDAB117418之質體圖譜。該等編號元件(意即，ZFN56、ZFN57與ZFN58)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點

及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0025] 圖9例示pDAB117420之質體圖譜。該等編號元件(意即，ZFN61與ZFN62)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)包含該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

圖10例示pDAB117421之質體圖譜。該等編號元件(意即，PPL17配對3、PPL17配對1及PPL17配對2)對應於長度約20至35個鹼基對的鋅指核酸酶結合序列，其係由相應的鋅指核酸酶蛋白質所識別及切開。這些鋅指結合序列及該註釋的”UZI序列”(該者係為一100-150 bp之模板區域，其含有限制性位點及用於引子設計或編碼序列之DNA序列)構成該通用供體卡匣。進一步包括在此質體設計中的為該”104113重疊”，該者係為對該質體載體共享同源性的序列，用於該通用供體卡匣在一質體載體內之高通量組裝(意即，經由吉布森組裝)。

[0026] 圖11代表用於經由非同源性末端接合(NHEJ)併入該通用供體聚核苷酸序列。兩種推薦載體係提供的，其中一感興趣DNA(DNA X)在該感興趣DNA之任一端包含一或多個(意即，"1-N")鋅指結合位點(ZFN BS)。垂直箭頭顯示獨特的限制性位點而橫向箭頭代表潛在的PCR引子位點。

[0027] 圖12代表用於經由同源性導向修復(HDR)併入該通用供體聚核苷酸序列。一感興趣DNA(DNA X)包含兩個夾擊該感興趣DNA的同源序列區域(HA)，伴隨鋅指核酸酶結合位點(ZFN)包圍該DNA X與HA序列。垂直箭頭顯示獨特的限制性位點而橫向箭頭代表潛在的PCR引子位點。

[0028] 圖13A-13C例示使用於標定及驗證該通用供體聚核苷酸系統在玉蜀黍最佳基因組基因座內併入之標定與驗證的該構建體。圖13A)ZFN空間設計其具如先前在圖5之pDAB111845所示的ZFN對位置。該等ZFN對係標誌數值地，且對應於由ZFN蛋白特異性識別用於結合與切割的特異性ZFN結合序列。圖13B)ZFN蛋白質表現構建體之配置。該ZFN表現構建體含有一組成型植物啓動子(Zm Ubi1)，其係使用以驅動該ZFN蛋白的表現。該ZFN蛋白含有細胞核定位序列(NLS)、鋅指蛋白(ZFP-L與ZFP-R，其中L指出左手結合ZFN蛋白而R指出右手結合蛋白)、Fok-1核酸內切酶(Fok1)及該自水解2A(2A)。圖13C)用於NHEJ介導的玉蜀黍最佳基因組基因座標定之通用供體聚核苷酸。Z1-Z6代表對玉蜀黍最佳基因組基因座標靶特異性的ZFN結合位點。ZFN位點的數目可以從3-6變化。垂直箭頭顯示獨特的限制性位點而橫

向箭頭代表潛在的PCR引子位點。該通用供體聚核苷酸系統係為一個短(110 bp)序列，該序列對使用以併入玉蜀黍最佳基因組基因座內之供體為共同的。

[0029] 圖14例示pDAB8393之質體圖譜。

[0030] 圖15例示該ZFN於玉蜀黍選定的基因組基因座標靶的切割活性。切割活性係呈現為於該ZFN切割位點，每百萬高品質讀數，具插入/刪除(indels)(插入與刪除)的序列數目。

[0031] 圖16例示使用基於NHEJ之快速標定分析法(RTA)驗證玉蜀黍選定的基因組基因座標靶。

[0032] 圖17例示經由隨機併入轉形至玉蜀黍內的質體構建體，其包含用於夾擊序列分析及轉基因表現研究的事件。

【實施方式】

較佳實施例之詳細說明

I. 概述

[0033] 本揭露內容提供，在各種觀點，聚核苷酸供體卡匣其包含一鋅指核酸酶結合域、一解析域及一質體域者。在其它觀點中，本揭露內容提供方法，其提供用於在一植物細胞的基因組內標定併入聚核苷酸供體卡匣。根據該揭露方法，一位點特異性DNA結合核酸酶，其包含至少一DNA結合域及至少一核酸酶域，該者係表現的，使得至少一DNA結合域結合至該植物細胞基因組內的一標靶位點。此外，將該植物細胞與一聚核苷酸供體卡匣接觸。隨後，在該植

物細胞基因組內的標靶位點係以該位點特異性DNA結合核酸酶切開，且該聚核苷酸供體卡匣係併入至該植物細胞基因組內的該標靶位點。

II.術語

[0034]除非另有界定，於此使用的所有技術及科學術語具有如本揭露內容所相關之該技藝中之一般技藝人士普遍了解的相同意義。在衝突的情況下，本申請案包括的定義為準。除非上下文另有要求，單數術語應包括複數，且複數術語應包括單數。所有於此提及之公開案、專利及其它參考文獻為了一切目的係以其整體合併入以做為參考，宛若每一個別的公開案或專利申請案係具體地且個別地指出被合併入以做為參考，除非僅有專利或專利公開案的特定節被指出合併入以做為參考。

[0035]為了進一步闡明本揭露內容，下列術語、縮寫及定義係提供的。

[0036]如於此所使用的，術語”包含”、”包含著”、”包括”、”包括著”、”具有”、”具有著”、”含有”或”含有著”，或其等之任何其它變體，係意欲為非排他性或開放式的。舉例而言，包含一列出元件的一組成物、混合物、製程、方法、物品、或裝置係非必要地受限於僅那些元件，但可能包括其它未明確列出或固有於此種組成物、混合物、製程、方法、物品、或裝置的元件。進一步，除非明確陳述為相反的，”或”意指一包含性的或，而不是排他性的或。舉例而言，條件A或B係滿足下列任何一者：A為真實的(或

存在的)且B爲假的(或不存在的)、A爲假的(或不存在的)且B爲真實的(或存在的)、及A與B皆爲真實的(或存在)。

[0037]如於此所使用，該術語”發明”或”本發明”係爲一非限制性術語，且係非意欲意指本特定發明之任何單一實施例，而是含括在本申請案中所揭露之所有可能的實施例。

[0038]如於此所使用，術語”植物”包括整株植物及一植物之任何後代、細胞、組織，或其部分。該術語”植物部分”包括一植物的任何部分(等)，包括，舉例而言但不限於：種子(包括成熟種子、沒有種皮的不成熟胚及不成熟的種子)；植物插條；植物細胞；植物細胞培養；植物器官(例如，花粉、胚、花、果實、幼枝、葉、根、莖及相關的培植體)。一植物組織或植物器官可能爲一種子、癒傷組織、或者組織成一結構或功能單元的任何其它植物細胞族群。一植物細胞或組織培養可能能夠再生一植物其具有源自獲得該細胞或組織之植物的生理學與形態學特徵，且能夠再生具有基本上相同於該植物之基因型的一植物。相反的，一些植物細胞不能夠再生以產生植物。在一植物細胞或組織培養中可再生細胞可能爲胚、原生質體、分生細胞、癒傷組織、花粉、葉、花藥、根、根尖、花絲、花、核仁、穗、穗軸、果殼或稈。

[0039]植物部分包括可收穫部分及對後代植物繁殖爲有用的部分。對繁殖有用的植物部分包括，舉例而言且不限於：種子；果實；插條；幼苗；塊莖；及砧木。植物的

可收穫部分可能為植物的任何有用部分，包括，舉例而言且不限於：花；花粉；幼苗；塊莖；葉；莖；果實；種子；及根。

[0040]一植物細胞係為該植物之結構與生理單位。如於此所使用，植物細胞包括原生質體及具部分細胞壁的原生質體。一植物細胞可能為一分離單一細胞或一細胞聚集體的形式(例如，易碎的癒傷組織及培養的細胞)，並且可能為一更高的組織化單元的一部分(例如，一植物組織、植物器官及植物)。因此，一植物細胞可能為一原生質體、配子產生細胞、或可以再生為整株植物的細胞或細胞集合。因此，包含多種植物細胞且能夠再生為整株植物之一種子在於此之實施例中係視為一”植物部分”。

[0041]如於此所使用，該術語”原生質體”意指一植物細胞其完全或部分的細胞壁被移除，伴隨著其之脂雙層膜裸露的。典型地，原生質體係為具有再生為細胞培養或整株植物該潛力的一沒有細胞壁之分離植物細胞。

[0042]如於此所使用，”內源性序列”界定了一聚核苷酸、基因或多勝肽在該生物體中或在該生物體之基因組中於其天然位置中的自然形式。

[0043]如於此所使用，該術語”分離的”意謂已從其天然環境移除的。

[0044]如於此所使用，該術語”經純化”有關分離一分子或化合物，在基本上沒有於自然或天然環境中正常地與該分子或化合物相關聯之污染物之形式中，且意謂在純度中

已提高的，因為已從該原始組成物之其它組分分開來的結果。該術語”經純化核酸”於此係使用以描述已從其它化合物，包括但不限於多勝肽、脂質及碳水化合物，分離的一個核酸序列。

[0045]如於此所使用，該等術語”聚核苷酸”、”核酸”及”核酸分子”係互換使用地，且可能含括一單一核酸；數個核酸；一核酸片段、其變體或衍生物；及核酸構建體(例如，傳訊RNA(mRNA)與質體DNA(pDNA))。一聚核苷酸或核酸可能含有一全長cDNA序列的核苷酸序列，或其之片段，包括非轉譯5'及/或3'序列及編碼序列(等)。一聚核苷酸或核酸可能由任何聚核糖核苷酸或聚去氧核糖核苷酸構成，該二者可能包括未經修飾的核糖核苷酸或去氧核糖核苷酸，或經修飾的核糖核苷酸或去氧核糖核苷酸。舉例而言，一聚核苷酸或核酸可能由下者構成；單股及雙股DNA；DNA其係為單股與雙股區域之一混合物者；單股及雙股RNA；及RNA其係為單股與雙股區域之一混合物者。包含DNA與RNA的雜交分子可能為單股、雙股、或單股與雙股區域之一混合物。前述術語亦包括一聚核苷酸或核酸之經化學、酵素與代謝修飾形式。

[0046]其應理解的是，一特定的DNA亦意指其之補體，該者之序列係根據去氧核糖核苷酸鹼基配對規則決定的。

[0047]如於此所使用，該術語”基因”意指編碼一功能性產物(RNA或多勝肽/蛋白質)之一核酸。一基因可能包括於該編碼功能性產物序列之前(5'非編碼序列)及/或之後(3'非

編碼序列)的調控序列。

[0048]如於此所使用，該術語”編碼序列”意指編碼一特定胺基酸序列的一核酸序列。一”調控序列”意指位於一編碼序列上游(例如，5'非編碼序列)、之中、或下游(例如，3'非編碼序列)的一核苷酸序列，該者影響該關聯編碼序列之轉錄、RNA加工或穩定性、或轉譯。調控序列包括，舉例而言且不限於：啓動子；轉譯前導子序列；內含子；多聚腺苷酸化識別序列；RNA加工位點；效應子(effectors)結合位點；及莖環結構。

[0049]如於此所使用，該術語”解析域”界定含有協助核酸序列標定插入之功能性元件的一個核酸序列。舉例而言，一解析域可能含有特異性設計的限制酵素位點、鋅指結合位點、工程化的降落墊(landing pads)或工程化的轉基因併入平台，且可能包含或不包含基因調控元件或一開放閱讀框架。參閱，舉例而言，美國專利公開案第20110191899號，該者係以其整體合併於此以做為參考。

[0050]如於此所界定，”位點特異性核酸酶結合域”係為包含一或多個位點特異性結合序列的一核酸序列。如於此所使用，一位點特異性結合序列係為在一序列特異性交互作用中結合至一蛋白質的一核酸序列。

[0051]如於此所使用，該術語”多勝肽”包括一單一多勝肽、數個多勝肽、及其片段。此術語意指由醯胺鍵(也稱為勝肽鍵)線性鏈接之單體(胺基酸)所構成的一種分子。該術語”多勝肽”意指二或多種胺基酸之任意鏈或鏈等，且不意

指特定長度或大小的產物。相應地，胜肽、二胜肽、三胜肽、寡胜肽、蛋白質、胺基酸鏈、及任何其它使用以意指二種或多種胺基酸鏈或鏈等的術語，係包括在“多胜肽”定義內，且前述術語於此係與“多胜肽”互換使用的。多胜肽可能從一天然生物來源分離，或藉由重組技術生產的，但一特定的多胜肽不必要從一特定核酸轉譯的。一多胜肽可能在任何適當方式中生成，包括舉例而言且不限於，藉由化學合成。

[0052]相反地，該術語”異源的”意指一聚核苷酸、基因或多胜肽其正常地不在該參照(宿主)生物體中於其位置發現的。舉例而言，一異源核酸可能為正常地在該參照生物體中於一不同基因組位置發現的一個核酸。藉助於進一步例子，一異源核酸可能為正常地不在該參照生物體中發現的一個核酸。包含一異源聚核苷酸、基因或多胜肽之一宿主生物體可能藉由引入該異源聚核苷酸、基因或多胜肽至該宿主生物體內而產生。在特定例子中，一異源聚核苷酸包含一天然編碼序列或其部分，該者係於不同於該相應的天然聚核苷酸之一形式中再引入至一來源生物體。在特定例子中，一異源基因包含一天然編碼序列或其部分，該者係於不同於該相應的天然基因之一形式中再引入至一來源生物體。舉例而言，一異源基因可能包括一天然編碼序列，該者係為一嵌合基因的一部分，其包括再引入至該天然宿主內之一非天然調控區域。在特定例子中，一異源多胜肽係為一天然多胜肽，該者係於不同於該相應的天然多胜肽

之一形式中再引入至一來源生物體。

[0053]一異源基因或多勝肽可能為一基因或多勝肽其包含融合至另一基因或多勝肽的一功能性多勝肽或編碼一功能性多勝肽之核酸序列，以產生一嵌合或融合多勝肽，或編碼該者之基因。特定實施例之基因與蛋白質包括具體例舉的全長序列及這些序列之部分、區段、片段(包括連續片段或相較於該全長分子的內部及/或末端刪除)、變體、突變體、嵌合體及融合體。

[0054]如於此所使用，該術語”修飾”可以意指在於此揭露之一聚核苷酸中的一改變，該改變引致由該聚核苷酸編碼之一多勝肽降低、基本上消除或消除的活性，以及意指在於此揭露之一多勝肽中的一改變，該改變引致該多勝肽降低、基本上消除或消除的活性。或者，該術語”修飾”可以意指在於此揭露之一聚核苷酸中的一改變，該改變引致由該聚核苷酸編碼之一多勝肽提高或增強的活性，以及意指在於此揭露之一多勝肽中的一改變，該改變引致該多勝肽提高或增強的活性。此種改變可以藉由在該技藝中廣為知悉的方法製成，包括但不限於，刪除、突變(例如，自發突變作用、隨機突變作用、由致增變基因(mutator genes)造成的突變作用、或轉位子突變作用)、取代、插入、下調、改變細胞位置、改變聚核苷酸或多勝肽狀態(例如，甲基化、磷酸化或泛素化(ubiquitination))、移除一輔因子、引入一反義RNA/DNA、引入一干擾RNA/DNA、化學修飾、共價修飾、以UV或X射線照射、同源重組、有絲分裂重組、啓

動子置換方法、及/或其等之組合。決定哪一核苷酸或胺基酸殘基可以被修飾的指南可以藉由比較該特定聚核苷酸或多勝肽之序列與該同源聚核苷酸或多勝肽之序列者發現的，且最大化在高同源性(保守區)或共有序列區域中做成的修飾數目。

[0055]該術語”衍生物”，如於此所使用，意指在本揭露內容中陳述之一序列的修飾。此種修飾之例示性將為有關於在此揭露之一編碼序列的核酸序列之一或多個鹼基之取代、插入及/或刪除，其保存、略微改變、或提高於此揭露之一編碼序列在該作物物種中的功能。此種衍生物可以由熟習該項技藝者輕易地決定，舉例而言，使用電腦模擬技術(computer modeling techniques)用於預測與優化序列結構。該術語”衍生物”，從而亦包括與在此揭露之該編碼序列具有實質性序列一致性的核酸序列，使得它們能夠具有在製造本揭露內容之實施例中所使用之該經揭露功能性。

[0056]該術語”啓動子”意指能夠控制一核酸編碼序列或功能性RNA表現的一DNA序列。在一例子中，該受控制編碼序列位於一啓動子序列之3'。一啓動子可能以其整體衍自一天然基因，一啓動子可能由衍自於在自然界中發現之不同啓動子的不同元件構成的，或一啓動子甚至可能包含合理設計的DNA區段。對熟習該項技藝者，其係不言而明的，不同啓動子可以主導一基因在不同組織或細胞類型中，或在不同發育階段，或回應不同環境或生理條件的表現。前述啓動子的所有例子係已知的，且在該技藝中係使

用以控制異源核酸的表現。主導一基因在大多數細胞類型中於大部分時間之表現的啓動子通常意指為”組成型啓動子”。更進一步，雖然該項技藝者已嘗試(在許多情況下未成功)描繪調控序列的確切邊界，其係已理解的是，不同長度的DNA片段可能具有一致的啓動子活性。特定核酸的啓動子活性可能使用該項技藝者所熟悉之技術檢測。

[0057]該術語”操縱鏈接”意指在一單一核酸上核酸序列的關聯，其中，該核酸序列中之一者的功能係被另一者影響。舉例而言，一啓動子係與一編碼序列操縱鏈接地，當該啓動子能夠影響那個編碼序列的表現時(例如，該編碼序列係於該啓動子的轉錄控制之下)。一編碼序列可能在正義或反義面向中操縱鏈接至一調控序列。

[0058]如於此所使用，該術語”表現”可能意指轉錄並穩定聚積源自一DNA的正義(mRNA)或反義RNA。表現亦可能意指轉譯mRNA成多勝肽。如於此所使用，該術語”過度表現”意指一表現其係高於該基因或一相關基因之內源性表現者。因此，一異源基因係”過度表現的”，假若其表現較諸一可比較之內源基因者為高的話。

[0059]如於此所使用，該術語”轉形作用”或”轉形”意指一核酸或其片段轉移並併入至一宿主生物體，引致遺傳上穩定的傳承。含有一轉形核酸的宿主生物體係意指為”轉基因”、“重組”或”轉形”生物。已知的轉形方法包括，舉例而言：農桿腫瘤菌(*Agrobacterium tumefaciens*)及農桿根毛菌(*A. rhizogenes*)介導之轉形；磷酸鈣轉形；聚凝胺轉形；原

生質體融合；電穿孔；超音波震盪法(例如，聲致穿孔(sonoporation))；脂質體轉形；顯微注射；以裸露DNA轉形；以質體載體轉形；以病毒載體轉形；生物彈道學(微粒轟擊)；矽晶鬚介導轉形；氣霧劑集束(aerosol beaming)；及PEG介導之轉形。

[0060]如於此所使用，該術語”引入”(在上下文中引入核酸至細胞中)包括細胞之轉形，以及將包含該核酸之一植物與一第二植物雜交，使得第該二植物含有該核酸，可能如使用常規的植物育種技術執行。此種育種技術在該技藝中為已知的。對於植物育種技術的討論，參閱Poehlman於1995年之”*Breeding Field Crops, 4th Edition, AVI Publication Co., Westport CT oehlman*”。

[0061]回交：回交方法可能使用以將一核酸引至一植物中。此技術已經使用數十年，以將性狀引入至植物內。回交(及其它植物育種方法學)之說明之一例子可以發現於，舉例而言，Poelman(1995)，同上；及Jensen於”(1988) *Plant Breeding Methodology, Wiley, New York, NY*”。在一示範性回交實驗計畫中，一感興趣之基源植物(回交親代)係雜交至帶有該被引入之核酸的一第二植物(非回交親代)。來自此雜交所得之子代然後係再雜交至該回交親代，且該方法係重複，直到獲得該經轉換之植物，其中除了源自該非回交親代之核酸以外，基本上該回交親代之所有所欲的形態與生理特徵皆在該經轉換植物中重新找回。

[0062]“結合”意指大分子之間(例如，蛋白質與核酸之

間)之一序列特異性、非共價交互作用。並非所有結合交互作用之組分需要為序列特異性(例如，在DNA骨架中與磷酸根殘基之接觸)，只要該交互作用整體上為序列特異性即可。此種交互作用一般係藉由 10^{-6} M^{-1} 或更低之一解離常數(K_d)表徵的。”親和力”意指結合強度：提高的結合親和力關聯於一較低的 K_d 。

[0063]一”結合蛋白”係為能夠非共價地結合至另一分子的蛋白質。結合蛋白可結合至，舉例而言，DNA分子(DNA結合蛋白)、RNA分子(RNA結合蛋白)、及/或蛋白質分子(蛋白質結合蛋白)。在蛋白質結合蛋白之事例中，其可結合至自身(以形成同型二聚體、同型三聚體……等等)，及/或其可結合至一或多個不同的蛋白質或蛋白質等之分子。一結合蛋白可以具有多於一者之結合活性類型。舉例而言，鋅指蛋白質具有DNA結合、RNA結合及蛋白質結合活性。

[0064]“重組”意指兩個聚核苷酸之間遺傳訊息交換的過程，包括但不限於藉由非同源性末端接合(NHEJ)及同源重組所捕獲的供體。為了達成本揭露內容之目的，”同源重組(HR)”意指，舉例而言，在細胞中之雙股斷裂修復期間，經由同源導向修復機制發生的此種交換之特別形式。此過程要求核苷酸序列同源性，使用一”供體”分子以模板修復一”標靶”分子(亦即，經歷雙股斷裂者)，且係不同地稱為”非交換型基因轉換(non-crossover gene conversion)”或”短道基因轉換(short tract gene conversion)”，因為其導致遺傳訊息自該供體轉移至該標靶。不希望受任何特定理論束縛，

此種轉移可以涉及形成於該斷裂標靶與供體之間的異源雙鏈體DNA之錯配修正，及/或”合成依賴性股黏著”，其中該供體係使用以再合成將成為該標靶之部分的遺傳訊息，及/或相關過程。此種特殊化HR常引致該標靶分子之序列的變更，使得部分或全部的供體聚核苷酸係併入至該標靶聚核苷酸內。對於HR-導向併入，該供體分子對基因組含有至少一同源性區域(“同源臂”)，長度至少為50-100個鹼基對。參閱，例如，美國專利公開案第20110281361號。

[0065]在本揭露內容之方法中，一或多個標定核酸酶如於此所描述般在該標靶序列(例如，細胞染色質)中於一預決定位點創建一雙股斷裂，且對該斷裂區域中之該核苷酸序列具有同源性的一個”供體”聚核苷酸可以引入至該細胞內。該雙股斷裂的存在已經顯示了促進該供體序列的併入。該供體序列可能物理上併入或，或者，該供體聚核苷酸係使用作為一模板，用於經由同源重組修復該斷裂，引致在該供體中之全部或部分的核苷酸序列引入到該細胞染色質內。因此，在細胞染色質中的一第一序列可以被變更，且在某些實施例中，可以轉換為存在於一供體聚核苷酸中的一序列。因此，術語”以……置換”或”置換”可以使用以代表一核苷酸序列由另一者所置換，(意即，在訊息意義中一序列之置換)，且不一定要求一聚核苷酸由另一者物理或化學置換。

[0066]“切割”意指一DNA分子共價骨架之斷裂。切割可以藉由各種方法起始的，包括但不限於，磷酸二酯鍵之酵

素或化學水解。單股切割與雙股切割兩者皆為可能的，且雙股切割可以作為兩個不同單股切割事件之結果發生。DNA切割可以引致產生平端或交錯端。在某些實施例中，融合多勝肽係使用於標定雙股DNA切割。

[0067]如於此所使用，該術語”質體”及”載體”意指一染色體外元件，其可能帶有一或多個非該細胞中心代謝之部分的基因(等)。質體與載體典型地為環狀雙股DNA分子。然而，質體與載體可以為一單股或雙股DNA或RNA的線性或環狀核酸，且可能帶有基本上衍自任何來源的DNA，其中許多核苷酸序列已被接合或重組成一獨特的構建體，該者能夠將一啓動子片段與一編碼DNA序列，連同任何適當的3'非轉譯序列引入至一細胞中。在例子中，質體與載體可能包含用於在細菌宿主中繁殖的自主複製序列。

[0068]多勝肽及”蛋白質”於此係互換使用的，且包括透過勝肽鍵鏈接的二或多個胺基酸之一分子鏈。該術語不意指特定長度的產物。因此，”勝肽”與”寡勝肽”係包括在多勝肽的定義範圍內。該術語包括多勝肽之轉譯後修飾，舉例而言，糖基化、乙醯化、磷酸化及之類。此外，蛋白質片段、類似物、突變或變體蛋白、融合蛋白及之類，係包括在多勝肽的含義內。該術語亦包括其中一或多個胺基酸類似物或非正準或非天然胺基酸係包括的分子，該者可以使用已知蛋白質工程技術合成或重組表現地。此外，藉由廣為知悉的有機化學技術，本發明之融合蛋白可以如於此所描述般衍生的。

[0069]該術語”融合蛋白”指出該蛋白質包括衍自一個以上親本蛋白質或多勝肽的多勝肽組分。典型地，融合蛋白係從一融合基因表現的，在該基因中，編碼源自一蛋白之一多勝肽序列的核苷酸序列係與編碼源自不同蛋白質之一多勝肽序列的一個核苷酸序列附加在框架中，且任選地由一連接子分隔的。該融合基因然後可以由重組宿主細胞以單一蛋白質表現。

III.本發明之實施例等

[0070]在本揭露內容之一實施例中，包含一鋅指核酸酶結合域、一解析域及一質體域的一聚核苷酸供體卡匣係提供的。

[0071]在請求項1之該聚核苷酸供體卡匣一些實施例中，該聚核苷酸供體卡匣包含小於3 Kbp之一長度。示範性地，聚核苷酸供體卡匣之長度可能為約3Kbp、2.9Kbp、2.8Kbp、2.7Kbp、2.6Kbp、2.5Kbp、2.4Kbp、2.3Kbp、2.2Kbp、2.1Kbp、2.0Kbp、1.9Kbp、1.8Kbp、1.7Kbp、1.6Kbp、1.5Kbp、1.4Kbp、1.3Kbp、1.2Kbp、1.1Kbp、1.0Kbp、0.9Kbp、0.8Kbp、0.7Kbp、0.6Kbp、0.5Kbp、0.4Kbp、0.3Kbp、0.2Kbp、或0.1Kbp。該術語”鹼基對”(亦描述為”bp”)係意指一對核苷酸鹼基(核苷酸)，每一者在一分隔的單股核酸中，在該者中，該對之每一鹼基係非共價地鍵合至另一者(例如，經由氫鍵)。譬如，華生-克力克鹼基對通常含有一個嘌呤與一個嘧啶。鳥糞嘌呤核苷可以配對胞嘧啶(G-C)，腺嘌呤可以配對胸腺嘧啶(A-T)，且尿嘧啶可配對腺嘌呤(U-A)。在一鹼基對中之該

兩個鹼基係被稱爲彼此互補的。鹼基對可能以千鹼基對(Kbp)、百萬鹼基對(Mbp)及十億鹼基對(Gbp)被提及，其中，"kilo"表示一千的增量，且"mega"表示一百萬的增量，且"mega"表示十億的增量。

[0072]在其它實施例中，一細胞係以包含一位點特異性核酸酶結合域、一解析域及一質體域的該聚核苷酸供體卡匣轉形，且該聚核苷酸供體卡匣係併入到該宿主細胞的DNA中。如於此所意指，該術語"細胞"含括能夠自我複製的活的生物體，且可能爲真核或原核細胞。在一些實施例中該轉形細胞係爲一植物細胞。在一些實施例中，該植物細胞可以爲但不限於任何高等植物，包括雙子葉植物與單子葉植物兩者，及可消費的植物，包括作物植物及針對其油使用的植物。因此，如下文進一步說明，任何植物物種或植物細胞可以被選擇用於轉形。

[0073]在一些實施例中，按照本揭露內容轉形之一植物細胞包括但不限於，任何高等植物，包括雙子葉植物與單子葉植物兩者，且特別是可消費植物，包括作物植物。此等植物可以包括，但不限於，舉例而言：紫花苜蓿、大豆、棉花、油菜籽(亦說明爲油菜)、亞麻子、玉米、水稻、臂形草屬(brachiaria)、小麥、紅花、高粱、甜菜、向日葵、煙草及草皮草。因此，任何植物物種或植物細胞可以被選擇。在一些實施例中，於此使用的植物細胞，及由其生長或衍生的植物包括，但不限於從下列所獲得之細胞：油菜籽(甘藍型油菜(*Brassica napus*))；印度芥(芥菜型油菜(*B. juncea*))；

阿比西尼亞芥(甘藍芥(*Brassica carinata*))；大頭菜(蕪菁(*Brassica rapa*))；甘藍(甘藍(*Brassica oleracea*))；大豆(黃豆(*Glycine max*));亞麻子/亞麻(亞麻(*Linum usitatissimum*))；玉蜀黍(亦說明為玉米)(玉蜀黍(*Zea mays*)); 紅花(紅花(*Carthamus tinctorius*)); 向日葵(向日葵(*Helianthus annuus*))；菸草(煙草(*Nicotiana tabacum*))；擬南芥(*Arabidopsis thaliana*)、巴西堅果(巴西果(*Betholettia excelsa*))；蓖麻子(蓖麻(*Ricinus communis*))；椰子(椰子(*Cocos nucifera*))；香菜(芫荽(*Coriandrum sativum*))；棉花(棉屬物種(*Gossypium spp.*))；落花生(花生(*Arachis hypogaea*))；荷荷芭(希蒙得木(*Simmondsia chinensis*))；油棕(油棕(*Elaeis gdineois*))；橄欖(油橄(*Olea europaea*))；水稻(水稻(*Oryza sativa*))；南瓜(西洋南瓜(*Cucurbita maxima*))；大麥(大麥(*Hordeum vulgare*))；甘蔗(甘蔗(*Saccharum officinarum*))；水稻(水稻(*Oryza sativa*))；小麥(小麥物種包括 *Triticum durum* 及 *Triticum aestivum*))；及浮萍(浮萍科(*Lemnaceae*物種))。在一些實施例中，在一植物物種內的遺傳背景可能變化的。

[0074]對於生產經基因修飾的植物，用於植物基因工程之方法在該技藝中係廣為知悉的。譬如，各種用於植物轉形的方法已經發展了，包括用於雙子葉植物以及單子葉植物的生物及物理轉形實驗計畫(例如，Goto-Fumiyuki等人於”*Nature Biotech* 17: 282-286 (1999)”; Miki等人於Glick, B. R. 及 Thompson, J. E. 編輯之”*Methods in Plant Molecular Biology and Biotechnology*, CRC Press, Inc., Boca Raton, pp.

67-88 (1993)"。此外，用於植物細胞或組織轉形及植物再生的載體與體外培養方法係可獲得的，舉例而言，Gruber等人於Glick, B. R. 及 Thompson, J. E.編輯之"Methods in Plant Molecular Biology and Biotechnology, CRC Press, Inc., Boca Raton, pp. 89-119 (1993)"。

[0075]大量技術可用於將DNA插入至一植物宿主細胞內。這些技術包括使用農桿腫瘤菌或農桿根毛菌做為轉形劑以無害的T-DNA轉形、磷酸鈣轉染、聚凝胺轉形、原生質體融合、電穿孔、超音波震盪法(例如，聲致穿孔)、脂質體轉形、顯微注射、裸露DNA、質體載體、病毒載體、生物彈道學(微粒轟擊)、矽晶鬚介導轉形、氣霧劑集束、或聚乙二醇(PEG)、以及其他可能的方法。

[0076]舉例而言，該DNA構建體可能使用諸如植物細胞原生質體之電穿孔與顯微注射技術直接引入至該植物細胞之基因組DNA內，或該DNA構建體可以使用諸如DNA粒子轟擊之生物彈道方法直接引入至植物組織中(參閱，例如Klein等人於"(1987)Nature 327 : 70-73")。用於植物細胞轉形的額外方法包括經由矽晶鬚介導的DNA攝入的顯微注射(Kaeppler等人於"(1990) Plant Cell Reporter 9 : 415-418")。或者，該DNA構建體可以經由奈米顆粒轉形引入到植物細胞內(參閱，例如美國專利申請案第12/245,685號，該者係以其整體合併於此以做為參考)。

[0077]植物轉形之另一種已知方法係為微粒介導的轉形，其中DNA係攜帶於該微粒之表面上。在此方法中，該

載體係以加速該微粒至足以穿透植物細胞壁與細胞膜之速度的一個生物彈道裝置引入至植物組織內。Sanford等人於”Part. Sci. Technol. 5 : 27 (1987)”；Sanford J. C.於”Trends Biotech. 6 : 299 (1988)”；Sanford J. C.於”Physiol. Plant 79 : 206 (1990)”；Klein等人於”Biotechnology 10 : 268 (1992)”。

[0078]或者，基因轉殖及轉形之方法包括，但不限於，透過氯化鈣沈澱的原生質體轉形；聚乙二醇(PEG)或電穿孔介導的裸露DNA攝入(參閱 Paszkowski等人於”(1984) EMBO J 3 : 2717-2722”、Potrykus等人於”(1985) Molec. Gen. Genet. 199 : 169-177”；Fromm等人於”Proc. Nat. Acad. Sci. USA 82 : 5824-5828”；及Shimamoto於”(1989) Nature 338 : 274-276”)與植物組織的電穿孔(D'Halluin等人於”(1992) Plant Cell 4 : 1495-1505”)。

[0079]用於引入一表現載體至一植物內之廣泛利用的方法係為奠基於農桿菌的天然轉形系統。Horsch等人於”Science 227 : 1229 (1985)”。農桿腫瘤菌及農桿根毛菌係為已知對基因轉形植物細胞為有用的植物病原土壤細菌。農桿腫瘤菌及農桿根毛菌的Ti及Ri質體係分別攜帶負責植物基因轉形的基因。Kado, C. I於”Crit. Rev. Plant. Sci. 10 : 1 (1991)”。農桿菌載體系統的描述及針對農桿菌介導之基因轉殖的方法亦可獲得的，舉例而言，同上文Gruber等人、同上文Miki等人、Moloney等人於”Plant Cell Reports 8 : 238 (1989)”、及美國專利案第4,940,838與第5,464,763號案。

[0080]假若農桿菌係使用於該轉形，該被插入之DNA

應選殖至特定質體內，即至一中間載體或一二元載體內。中間載體在農桿菌中不能複製自身。該中間載體可以藉由使用一輔助質體轉殖至一農桿腫瘤菌(接合作用)。該Japan Tobacco Superbinary系統係為此一系統之例子(由Komari等人於2006年在K. Wang編輯之"Methods in Molecular Biology No. 343: Agrobacterium Protocols(第二版, 第一卷) Humana Press Inc., Totowa, NJ, pp.15-41"; 及Komori等人於2007年在"Plant Physiol. 145: 1155-1160"中所綜述)。二元載體在大腸桿菌與農桿菌兩者中皆可以複製自身。它們包含一選擇標記基因及由右與左T-DNA交界區所框架之一連接子或多連接子。它們可以直接轉形至農桿菌內(Holsters, 1978)。使用做為宿主細胞之該農桿菌係包含帶有一vir區之一質體。該Ti或Ri質體亦包含對T-DNA轉殖為必要之該vir區。該vir區對T-DNA轉殖到該植物細胞內為必要的。額外的T-DNA可能含有的。

[0081]該農桿腫瘤菌宿主之致病性功能將主導含有構建體的T股與鄰近標記插入到該植物細胞DNA內，當該細胞使用一二元T DNA載體(Bevan於"(1984) Nuc. Acid Res. 12: 8711-8721")或共培養程序由該細菌感染時(Horsch等人於"(1985) Science 227: 1229-1231")。通常，該農桿菌轉形系統係使用以工程化雙子葉植物(Bevan等人於"(1982) Ann. Rev. Genet. 16: 357-384"; Rogers等人於"(1986) Methods Enzymol. 118: 627-641")。該農桿菌轉形系統亦可以使用以轉形，以及轉殖，DNA至單子葉植物及植物細胞。參閱

美國專利案第 5,591,616 號；Hernalsteen 等人於”(1984) EMBO J 3 : 3039-3041”；Hooykass-Van Slogteren 等人於”(1984) Nature 311 : 763-764”；Grimsley等人於”(1987) Nature 325 : 1677-179”；Boulton等人於”(1989) Plant Mol. Biol. 12 : 31-40”；及 Gould 等人於”(1991) Plant Physiol. 95 : 426-434”。繼之引入該基因構建體至特定植物細胞後，植物細胞可以生長且一旦諸如幼枝及根之分化組織出現，成熟植物可以產生。在一些實施例中，數個植物可以產生的。用於再生植物的方法論對該技藝之一般技藝人士係為已知的，且可以在下列中找到，舉例而言：Vasil 及 Thorpe 編輯之”Plant Cell and Tissue Culture, 1994, Kluwer Academic Publishers”，及 Hall 編輯之”Plant Cell Culture Protocols (Methods in Molecular Biology 111, 1999 Humana Press)”。於此描述之該經基因修飾植物可以在一發酵培養基中培養，或在一適合介質中生長，諸如土壤。在一些實施例中，高等植物適合的生長介質可以包括針對植物的任何生長介質，包括但不限於，土壤、砂、任何其它支持根生長的顆粒介質(例如，蛭石、珍珠岩……等等)或水培，以及適合的光線、水份及優化高等植物生長的營養補充劑。

[0082] 藉由上文任何轉形技術產生的轉形植物細胞可以被培養以再生擁有該轉形基因型及從而該所欲表型之一完整植物。此種再生技術依靠在一組織培養生長培養基中某些植物荷爾蒙的操作，典型地依靠已與該所欲核苷酸序列一起引入之一殺生物劑及/或除草劑標記。從培養原生質

體的植物再生係描述於Evans等人在"Handbook of Plant Cell Culture, pp. 124-176, Macmillian Publishing Company, New York, 1983"於"中之"Protoplasts Isolation and Culture"；及"Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, 1985"。再生亦可以從植物癒傷組織、培植體、器官、花粉、胚或其部分獲得。此等再生技術係由Klee等人於"(1987) Ann. Rev. of Plant Phys. 38: 467-486"中一般描述地。

[0083]在其它實施例中，經轉形之該等植物細胞係不能夠再生以產生一植物。此種細胞係被稱為瞬時轉形的。瞬時轉形細胞可能產生以檢測一特定轉基因的表現及/或功能性。瞬時轉形技術在該技藝中為已知的，且包含上述該等轉形技術的少量修飾。在該技藝中具技術之人士可能選擇以利用瞬時轉形以快速檢測特定轉基因的表現及/或功能性，因為瞬時轉形係迅速地完成，且不要求如穩定轉形技術那麼多的來源(例如，植物培養用於整個植株的發育、自花授粉或植物雜交用於固定一轉基因在該基因組內……等等)。

[0084]在一實施例中，該通用供體聚核苷酸基本上可以引入到任何植物內。廣泛的植物與植物細胞系統可能工程化用於位點特異性併入本揭露內容之該通用供體聚核苷酸及上文提及之各種轉形方法。在一實施例中，用於工程化之標靶植物及植物細胞包括，但不限於：那些單子葉植物及雙子葉植物，諸如包括下列之作物：糧食作物(例如小麥、

玉蜀黍、稻米、穀子、大麥)、果實作物(例如番茄、蘋果、梨、草莓、橙)、飼料作物(例如紫花苜蓿)、根菜類作物(例如胡蘿蔔、馬鈴薯、甜菜、山藥)、葉菜類作物(例如萵苣、菠菜)；顯花植物(例如矮牽牛、玫瑰、菊花)、針葉植物及松樹(例如松櫟樹、雲杉)；在植物修復(phytoremediation)中所使用的植物(例如重金屬聚積植物)；油料作物(例如向日葵、油菜籽)及用於實驗目的之植物(例如阿拉伯芥)。因此，所揭露方法及組成物已使用一廣範圍之植物，包括但不限於來自下列各屬之物種：蘆筍屬、燕麥屬、芸苔屬、柑桔屬、西瓜屬、辣椒屬、南瓜屬、胡蘿蔔屬、飛蓬屬、大豆屬、棉屬、大麥屬、萵苣屬、黑麥草屬、西紅柿屬、蘋果屬、木薯屬、煙草屬、諸葛菜屬(*Orychophragmus*)、稻屬、鱸梨屬、菜豆屬、豌豆屬、犁屬、李屬、蘿蔔屬、黑麥屬、茄屬、高粱、小麥屬、葡萄屬、豇豆屬、及玉蜀黍。

[0085]在其它實施例中，該聚核苷酸供體卡匣包含一特異性的聚核苷酸序列。特異性聚核苷酸序列之例子包括，但不限於：序列辨識編號：132、序列辨識編號：133、序列辨識編號：134、序列辨識編號：135、序列辨識編號：136、序列辨識編號：137、序列辨識編號：138、序列辨識編號：139、序列辨識編號：140及序列辨識編號：141。在各種實施例中，對該等特異性聚核苷酸序列序列辨識編號：132、序列辨識編號：133、序列辨識編號：134、序列辨識編號：135、序列辨識編號：136、序列辨識編號：137、序列辨識編號：138、序列辨識編號：139、序列辨識編號：

140及序列辨識編號：141具至少90%序列一致性的序列係於此揭露的。在其它實施例中，對該等特異性聚核苷酸序列序列辨識編號：132、序列辨識編號：133、序列辨識編號：134、序列辨識編號：135、序列辨識編號：136、序列辨識編號：137、序列辨識編號：138、序列辨識編號：139、序列辨識編號：140及序列辨識編號：141具至少95%序列一致性的序列係於此揭露的。在其它實施例中，對該等特異性聚核苷酸序列序列辨識編號：132、序列辨識編號：133、序列辨識編號：134、序列辨識編號：135、序列辨識編號：136、序列辨識編號：137、序列辨識編號：138、序列辨識編號：139、序列辨識編號：140及序列辨識編號：141具至少97%序列一致性的序列係於此揭露的。在又其它實施例中，對該等特異性聚核苷酸序列序列辨識編號：132、序列辨識編號：133、序列辨識編號：134、序列辨識編號：135、序列辨識編號：136、序列辨識編號：137、序列辨識編號：138、序列辨識編號：139、序列辨識編號：140及序列辨識編號：141具至少99%序列一致性的序列係於此揭露的。在一實施例中，該聚核苷酸供體卡匣包含一序列其對選自由序列辨識編號：142與143所組成之該群組的一序列具有至少90、93、95、97或99%序列的一致性。

[0086]如該技藝中所知悉，術語”百分比一致性”(或”一致性%”)，係為二或多個多勝肽序列或二或多個聚核苷酸序列之間的關係，該者係藉由比較該等序列而決定。在該技藝中，”一致性”亦意謂多勝肽或聚核苷酸序列之間序列相

關性的程度，由於該情況可能藉由此等序列串之間的匹配而決定。”一致性”與”相似性”可以輕意地藉由已知方法計算，包括但不限於在下列中揭露的那些：1.)Lesk, A. M.編輯之”Computational Molecular Biology Oxford University : NY (1988)”；2.)Smith, D. W. 編輯之”Biocomputing : Informatics and Genome Projects Academic : NY (1993)”；3.)Griffin, A. M.及Griffin, H. G.編輯之”Computer Analysis of Sequence Data, Part I Humania : NJ (1994)”；4.)von Heinje, G. 編輯之”Sequence Analysis in Molecular Biology Academic (1987)”；及5.)Gribskov, M.及Devereux, J.編輯之”Sequence Analysis Primer Stockton : NY (1991)”。

[0087]用於測定核酸及胺基酸序列一致性之技術在該技藝中為已知的。典型地，此等技術包括：針對一基因測定該mRNA之核苷酸序列及/或測定由此編碼之胺基酸序列，且將這些序列與一第二核苷酸或胺基酸序列比較。基因組序列亦可以在此方式中測定及比較。一般而言，一致性意指兩個聚核苷酸或多勝肽序列分別之一確切核苷酸對核苷酸或胺基酸對胺基酸的對應性。二或多個序列(聚核苷酸或胺基酸)可以藉由測定其一致性百分比而比較。兩序列之一致性百分比，無論是核酸或胺基酸序列，係為兩比對序列之間之確切匹配數除以較短序列之長度再乘以100。參閱 Russell, R. 與 Barton, G. 於 ”J. Mol. Biol. 244, 332-350 (1994)”，在第337頁之 ”Structural Features can be Unconserved in Proteins with Similar Folds”，該者係以其整

體合併於此以做為參考。

[0088]此外，測定一致性與相似度的方法係編撰於公諸大眾的電腦程式。序列比對及百分比一致性計算可以執行，舉例而言使用 Vector NTI® 套裝軟體的 AlignX 程式 (Invitrogen, 加州卡斯伯)或 LASERGENE 生物資訊學計算套裝軟體的 MegAlign™ 程式 (DNASTAR 公司, 威斯康辛州麥迪遜)。序列的多重比對係使用“Clustal method of alignment”執行，其含括該演算法的數種變化，包括“Clustal V method of alignment”，其對應於標誌為 Clustal V 的比對方法 (由 Higgins 及 Sharp 於 “CABIOS. 5 : 151-153 (1989)”; Higgins, D.G. 等人於 “Comput. Appl. Biosci., 8 : 189-191 (1992)” 所揭露)，並在 LASERGENE 生物資訊學計算套裝軟體的 MegAlign™ 程式 (DNASTAR 公司) 中可以找到。針對多重比對，內設值對應於 GAP PENALTY=10 與 GAP LENGTH PENALTY=10。使用 Clustal 方法，對於蛋白質序列之成對比對與百分比一致性計算的內設參數為 KTUPLE=1、GAP PENALTY=3、WINDOW=5 而 DIAGONALS SAVED=5。對於核酸，這些參數為 KTUPLE=2、GAP PENALTY=5、WINDOW=4 及 DIAGONALS SAVED=4。用 Clustal V 程序進行序列比對後，有可能通過查看同一程序中的“序列距離”表來獲得“百分比一致性”。在使用 Clustal V 程式比對序列後，其係可能藉由觀看相同程式中的「序列距離」表獲得「百分比一致性」。此外，該“Clustal W method of alignment”係可獲得的，並對應於標誌為 Clustal W 的比對方法 (由 Higgins

及Sharp於”CABIOS. 5 : 151-153 (1989)”；Higgins, D.G.等人於”*Comput. Appl. Biosci.*, 8 : 189-191 (1992)”所揭露)，並在LASERGENE生物資訊學計算套裝軟體的MegAlign™ 6.1版程式(DNASTAR公司)中可以找到。對於多重比對，內設參數為(GAP PENALTY=10、GAP LENGTH PENALTY=0.2、Delay Divergent Seqs(%)=30、DNA Transition Weight=0.5、Protein Weight Matrix=Gonnet Series、DNA Weight Matrix=IUB)。在使用Clustal W程式比對序列後，其係可能藉由觀看相同程式中的「序列距離」表獲得「百分比一致性」。

[0089]該術語”序列分析軟體”意指對核苷酸或胺基酸序列之分析為有用的任何電腦演算法或軟體程式。”序列分析軟體”可能從市面上獲得的，或獨立開發的。典型的序列分析軟體將包括但不限於：1.)GCG套裝程式(威斯康辛套裝9.0版，基因電腦團隊(Genetics Computer Group)(GCG)，威斯康辛州麥迪遜)；2.)BLASTP、BLASTN及BLASTX(Altschul 等人於”(1990)J.Mol.Biol.215 : 403-10”)；3.)DNASTAR(DNASTAR公司，威斯康辛州麥迪遜)；4.)Sequencher(Gene Codes公司，密西根州安娜堡)；及5.)合併Smith-Waterman演算法的FASTA程式(W. R. Pearson於Suhai, Sandor編輯，會議日期1992年之”*Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), 111-20. Plenum : New York, NY”)。在此申請案之上下文中，其應理解的是，序列分析軟體使用於分析時，該分析結果將基於程式參照

的”內設值”，除非另有指明。如於此所使用，”內設值”意謂當第一次初始化時，原始加載的任意值組或參數組。

[0090] 當意指雜交技術時，一已知核苷酸序列的全部或部分可以使用做為一探針，其選擇性地雜交至存在於來自一經抉擇生物體之選殖基因組DNA片段或cDNA片段族群(意即，基因組或cDNA集合庫)中的其它相應核苷酸序列。該等雜交探針可能為基因組DNA片段、質體DNA片段、cDNA片段、RNA片段、PCR放大的DNA片段、寡核苷酸、或其它聚核苷酸，且可能以諸如³²P之一可偵測基團或其它任何可偵測標記標誌的。因此，舉例而言，用於雜交的探針可以藉由標誌基於本揭露內容實施例之DNA序列的合成寡核苷酸而製成。製備用於雜交及用於構建cDNA與基因組集合庫之探針的方法在該技藝中為普遍知悉的且係揭露的(Sambrook等人於1989年)。

[0091] 本揭露內容之實施例的該等核酸探針與引子在嚴格條件下雜交至一標靶DNA序列。任何常規的核酸雜交或放大方法可以使用，以從一樣品中的轉基因事件辨識DNA的存在。核酸分子或其片段在某些境況下能夠特異性雜交至其它核酸分子。如於此所使用，兩個核酸分子被稱為能夠互相特異性地雜交，假若該兩個分子能夠形成反平行、雙股核酸結構的話。一核酸分子被稱為是另一個核酸分子之「互補體」，假若該兩核酸分子展現完全互補性的話。如於此所使用，當分子中之一者的每一核苷酸係互補於另一者之核苷酸時，該等分子被稱為展現「完全互補性」。展

現完全互補性的分子一般伴隨充分的穩定性互相雜交，以允許它們在常規的”高嚴格度”條件下依舊互相黏著。常規的高嚴格度條件係由Sambrook等人於1989年描述。

[0092]兩個分子被稱為展現「最小互補性」，假若它們可以互相雜交，伴隨充分的穩定性以允許其在至少常規的”低嚴格度”條件下依舊互相黏著。常規的低嚴格度條件係如Sambrook等人於1989年所描述。為了使一核酸分子用做為一引子或探針，其僅需要展現最小的序列互補性，以能夠在採用的特定溶劑及鹽濃度下形成一穩定的雙股結構。

[0093]影響雜交嚴格度之因素對熟習該項技藝者係廣為知悉的，且包括但不限於溫度、pH值、離子強度及有機溶劑之濃度，諸如舉例而言甲醯胺及二甲亞碸。如熟習該項技藝者所知悉，雜交嚴格度係藉由較高溫度、較低離子強度及較低溶劑濃度而提高。

[0094]該術語”嚴格條件”或”嚴格度條件”係藉由參照至由Sambrook等人於1989年在9.52-9.55中所討論之該特異性雜交過程一核酸探針雜交至一標靶核酸(意即，一特定的感興趣核酸序列)而功能性界定的。亦參閱Sambrook等人於1989年在9.47-9.52及9.56-9.58。

[0095]典型地，嚴格條件為其中該鹽濃度低於約1.5M Na^+ 離子，典型地於pH 7.0至8.3為約0.01至1.0M Na^+ 離子濃度(或其它鹽類)，且該溫度對短探針(例如10至50核苷酸)為至少約30°C，而對長探針(例如大於50核苷酸)為至少約60°C。嚴格條件亦可以伴隨加入諸如甲醯胺之去穩定劑而達

成。示範性的低嚴格度條件包括以緩衝溶液30至35%甲醯胺、1.0 M NaCl、0.1% SDS(硫酸十二烷酯鈉)於37°C 雜交，且在1X至2X SSC中(20X SSC=3.0M NaCl/0.3M檸檬酸三鈉)於50°C 至55°C 下洗滌。示範性的中等嚴格度條件包括在40至45%甲醯胺、1.0 M NaCl、0.1% SDS中於37°C 雜交，且在0.5X至1X SSC中於55°C 至60°C 下洗滌。示範性的高嚴格度條件包括在50%甲醯胺、1.0 M NaCl、0.1% SDS中於37°C 下雜交，且在0.1X SSC中於60°C 至65°C 下洗滌。

[0096]特異性典型係爲雜交後洗滌之一函數，關鍵因素爲最終洗滌溶液之離子強度及溫度。對於DNA-DNA雜交體， T_m 可以由 $T_m = 81.5^{\circ}\text{C} + 16.6 (\log M) + 0.41(\%GC) - 0.61(\%form.) - 500/L$ 該方程式估算的，其中M爲一價陽離子之莫耳濃度，%GC爲該DNA中鳥糞嘌呤核苷及胞嘧啶核苷酸之百分比，%form係爲該雜交溶液中之甲醯胺百分比，且L爲雜交體長度，以鹼基對爲單位(Meinkoth及Wahl, 1984)。 T_m 係爲50%互補標靶序列雜交至一完美匹配探針之溫度(於所界定之離子強度及pH下)。對於每1%錯誤匹配， T_m 減少約1°C；因此 T_m 、雜交及/或洗滌條件可以針對雜交所欲之序列一致性而調整的。舉例而言，假若具90%一致性之序列係追求的，該 T_m 可以降低10°C。通常嚴格條件係選擇對特定序列及其互補體於界定的離子強度及pH，比熱熔融點(T_m)低約5°C。然而，嚴厲嚴格條件可以利用於比該熱熔融點(T_m)低1、2、3或4°C之一雜交及/或洗滌；中等嚴格條件可以利用於比該熱熔融點(T_m)低6、7、8、9或10°C之一雜交及/或洗滌；低

嚴格度條件可以利用於比該熱熔融點(T_m)低11至20°C之一雜交及/或洗滌。使用該方程式、雜交及洗滌組成物及所欲的 T_m ，一般技藝人士將了解的是，雜交嚴格度及/或洗滌溶液中的變化係本質上描述的。假若該所欲的錯誤匹配程度引致小於45°C(水溶液)或32°C(甲醯胺溶液)之一 T_m ，其係較佳的是提高該SSC濃度，使得一較高溫度可以使用。核酸雜交之廣泛指南係於1997年Ausubel等人第三版之"Short Protocols in Molecular Biology"，頁2-40(1997)，及Sambrook等人(1989)中發現的。

[0097]在一實施例中，該聚核苷酸供體卡匣包含一或多個鋅指核酸酶結合域。因此，該聚核苷酸供體卡匣之一實施例包含1鋅指核酸酶結合域、2鋅指核酸酶結合域、3鋅指核酸酶結合域、4鋅指核酸酶結合域、5鋅指核酸酶結合域、6鋅指核酸酶結合域、7鋅指核酸酶結合域、8鋅指核酸酶結合域、9鋅指核酸酶結合域、10鋅指核酸酶結合域、或更多鋅指結合域。

[0098]標靶位點之選擇；ZFP及用於設計及構建融合蛋白(及編碼該者的聚核苷酸)的方法對熟習該項技藝者為已知的，並於下者中詳細地說明：美國專利第6,140,081號；第5,789,538號；第6,453,242號；第6,534,261號；第5,925,523號；第6,007,988號；第6,013,453號；第6,200,759號；WO 95/19431；WO 96/06166；WO 98/53057；WO 98/54311；WO 00/27878；WO 01/60970；WO 01/88197；WO 02/099084；WO 98/53058；WO 98/53059；WO 98/53060；WO 02/016536

及WO 03/016496。

[0099]在隨後的實施例中，該鋅指核酸酶結合域包含由一鋅指結合蛋白、巨核酸酶結合蛋白、CRIPSR或TALEN結合蛋白所結合之一或多個鋅指結合序列。

[0100]在某些實施例中，於此描述之該組成物與方法採用一巨核酸酶(歸巢核酸內切酶(homing endonuclease))結合蛋白或巨核酸酶DNA結合域，用於結合至該供體分子及/或結合至在該細胞基因組中的感興趣區域。天然發生巨核酸酶識別15-40鹼基對的切割位點，且普遍地分成四個家族：LAGLIDADG家族、GIY-YIG家族、His-Cyst盒家族及HNH家族。示範性的歸巢核酸內切酶包括I-SceI、I-CeuI、PI-PspI、PI-Sce、I-Sce IV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII及I-TevIII。其識別序列為已知的。亦參閱：美國專利第5,420,032號；美國專利第6,833,252號；Belfort等人於”(1997)*Nucleic Acids Res.* 25 : 3379-3388”；Dujon等人於”(1989)*Gene* 82 : 115-118”；Perler等人於”(1994)*Nucleic Acids Res.* 22,1125-1127”；Jasin於”(1996)*Trends Genet.* 12 : 224-228”；Gimble等人於”(1996)*J. Mol. Biol.* 263 : 163-180”；Argast等人於”(1998)*J. Mol. Biol.* 280 : 345-353”及新英格蘭生物實驗室(New England Biolabs)產品目錄。

[0101]在某些實施例中，於此描述之該方法與組成物利用包含經工程的(非天然發生的)歸巢核酸內切酶(巨核酸酶)之一核酸酶。歸巢核酸內切酶及巨核酸酶諸如I-SceI、

I-CeuI、PI-PspI、PI-Sce、I-Sce IV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII及I-TevIII，之識別序列係知悉的。亦參閱：美國專利第5,420,032號；美國專利第6,833,252號；Belfort等人於”(1997)*Nucleic Acids Res.* 25 : 3379-3388”；Dujon等人於”(1989)*Gene* 82 : 115-118”；Perler等人於”(1994)*Nucleic Acids Res.* 22, 1125-1127”；Jasin於”(1996)*Trends Genet.* 12 : 224-228”；Gimble等人於”(1996)*J. Mol. Biol.* 263 : 163-180”；Argast等人於”(1998)*J. Mol. Biol.* 280 : 345-353”及新英格蘭生物實驗室產品目錄。此外，歸巢核酸內切酶與巨核酸酶之DNA結合特異性可以工程化以結合非天然的標靶位點。參閱，舉例而言，Chevalier等人於”(2002) *Molec. Cell* 10 : 895-905”；Epinat等人於”(2003) *Nucleic Acids Res.* 31:2952-2962”；Ashworth等人於”(2006) *Nature* 441 : 656-659”；Paques等人於”(2007) *Current Gene Therapy* 7 : 49-66”；美國專利公開案第20070117128號。該歸巢核酸內切酶與巨核酸酶之DNA結合域可能以一整體在該核酸酶之內容變更的(意即，使得該核酸酶包括該同族切割域)，或可能融合至一異源切割域。

[0102] 在其它實施例中，在於此描述之方法與組成物中所使用之一或多個核酸酶之DNA結合域包含一天然發生或經工程化(非天然發生)之TAL效應子(effectort)DNA結合域。參閱，例如，美國專利公開案第20110301073號，該者係以其整體合併於此以做為參考。黃單胞菌屬

(*Xanthomonas*)之植物致病性細菌已知在重要作物植物中造成許多疾病。黃單胞菌屬之致病性取決於一種保守性第III型分泌(T3S)系統，其注入超過25種不同的效應子蛋白質至該植物細胞內。在這些注入的蛋白質中，模擬植物轉錄活化子並操作該植物轉錄體的是類轉錄活化子(transcription activator-like)(TAL) 之 效 應 子 (參 閱 ， Kay 等 人 於”(2007)Science 318 : 648-651”)。這些蛋白質含有一DNA結合域與一轉錄活化域。最被良好表徵的TAL效應子中之一者為源自辣椒斑點黃單胞菌(*Xanthomonas campestris* pv.*Vesicatoria*)之AvrBs3(參閱，Bonas等人於”(1989)Mol Gen Genet 218 : 127-136”及WO2010079430)。TAL效應子含有一串聯重複之集中域，每一重複含有大約34個胺基酸，其對於這些蛋白質的DNA結合特異性為關鍵的。此外，他們含有一細胞核定位序列(nuclear localization sequence)及一酸性轉錄活化域(綜述參閱Schornack S等人於”(2006)J Plant Physiol 163(3) : 256-272”)。此外，在植物病原菌青枯菌(*Ralstonia solanacearum*)中之兩個基因，命名為brg11與hpx17，在青枯菌生物型1菌株GMI1000中與生物群4菌株RS1000中，已經發現同源於黃單胞菌屬的AvrBs3家族(參閱，Heuer 等 人 於 ”(2007)Appl and Envir Micro 73(13) : 4379-4384”)。這些基因在核苷酸序列中彼此係98.9%一致的，但藉由在hpx17重複域中一1,575bp的刪除而不同。然而，兩基因產物與黃單胞菌屬AvrBs3家族蛋白質具有小於40%的序列一致性。參閱，例如，美國專利公開案第20110239315

號、第20110145940號及第20110301073號，該等係以其整體合併於此以做為參考。

[0103] 這些TAL效應子之特異性取決於在該串聯重複中發現的序列。該重複序列包含大約102 bp且該等重複典型地彼此係91-100%同源的(Bonas等人，文獻同上)。該等重複之多型性通常係位於位置12與13，且在位置12與13該高變異二殘基之一致性與在該TAL-效應子之標靶序列中該連續核苷酸之一致性之間似乎有一對一的對應關係(參閱Moscou及Bogdanove於”(2009) *Science* 326：1501”及Boch等人於”(2009) *Science* 326：1509-1512”)。實驗地，對這些TAL-效應子之DNA識別的天然碼已測定了，使得在位置12與13的一HD序列導致與胞嘧啶(C)的結合，NG結合至T，NI至A、C、G或T，NN結合至A或G，而ING結合至T。這些DNA結合重複體已組裝成具新組合與數目之重複體的蛋白質，以製成能夠與新序列交互作用的人工轉錄因子，並活化一非內源性報告基因在植物細胞中的表現(Boch等人，同上)。工程化的TAL蛋白質已被鏈接至一Fok I 切割半域，以產生一TAL效應子域核酸酶融合體(TALEN)，該者在酵母菌報告檢測(基於質體之標靶)中展現活性。參閱，例如，美國專利公開案第20110301073號；Christian等人於((2010)<*Genetics* epub 10.1534/genetics.110.120717)。

[0104] 在其他實施例中，該核酸酶係為一系統其包含該CRISPR(群聚且有規律間隔之短回文重複體)/Cas(CRISPR相關聯)核酸酶系統。該CRISPR/Cas係為一

最近的工程化核酸酶系統，其係基於一細菌系統，可以用於基因組工程。該者係基於許多細菌與古細菌的部分適應性免疫回應。當病毒或質體侵入細菌時，入侵者之DNA區段係由”免疫”回應轉換成CRISPR RNA(crRNA)。此crRNA然後透過部分互補區域，與稱為tracrRNA的另一類型RNA關聯，以引導該Cas9核酸酶至一區域其同源於在該標靶DNA中、毗鄰一前間區序列鄰近基序(protospacer adjacent motif)(PAM)NGG的crRNA。Cas9在DSB於特異於該crRNA轉錄體中所含有的20個核苷酸引導序列之位點，切開該DNA以產生平端。對於位點特異性DNA識別及切割，Cas9要求crRNA與tracrRNA兩者。此系統現今已經工程化，使得crRNA與tracrRNA可以組合成一個分子(該“單一引導RNA”)，且該單一引導RNA之crRNA等效部分可以工程化以引導Cas9核酸酶標定鄰近PAM的任何所欲序列(參閱Jinek等人於”(2012) *Science* 337, p. 816-821”、Jinek等人於”(2013) *eLife* 2: e00471”及David Segal於”(2013) *eLife* 2: e00563”)。因此，CRISPR/Cas系統可以工程化，以在基因組中於所欲的標靶創建一DSB，且該DSB之修復可以藉由使用修復抑制劑以造成一提高的易出錯修復而影響的。

[0105] 在某些實施例中，用於體內切割及/或標定一細胞基因組中切割之該等核酸酶之一或多個DNA結合域包含一鋅指蛋白質。較佳地，該鋅指蛋白質為非天然發生的，因為其係經工程化以結合至抉擇之標靶位點。參閱，舉例而言，Beerli等人於”(2002) *Nature Biotechnol.* 20: 135-141”；

Pabo等人於”(2001)Ann.Rev.Biochem.70 : 313-340”；Isalan等人於”(2001)Nature Biotechnol.19 : 656-660”；Segal等人於”(2001)Curr.Opin.Biotechnol.12 : 632-637”；Choo等人於”(2000)Curr.Opin.Struct.Biol.10 : 411-416”；美國專利案第6,453,242號；第6,534,261號；第6,599,692號；第6,503,717號；第6,689,558號；第7,030,215號；第6,794,136號；第7,067,317號；第7,262,054號；第7,070,934號；第7,361,635號；第7,253,273號；及美國專利公開案第2005/0064474號；第2007/0218528號；第2005/0267061號，全部係以其整體合併於此以做為參考。

[0106] 一經工程化之鋅指結合域可以具有一新穎的結合特異性，相較於一天然發生的鋅指蛋白質。工程化方法包括但不限於，合理性設計與各種類型之選擇。合理性設計包括，舉例而言，使用包含三重(或四重)核苷酸序列與個別鋅指胺基酸序列之資料庫，其中每一三重或四重核苷酸序列係與鋅指之一或多個胺基酸序列關聯，其中該等鋅指係結合至特定之三重或四重序列。參閱，舉例而言，專利權共有之美國專利第6,453,242號與第6,534,261號，該者係以其整體合併於此以做為參考。

[0107] 示範性的選擇方法，包括噬菌體展示法(phage display)與雙雜交系統，係揭露於美國專利第5,789,538號；第5,925,523號；第6,007,988號；第6,013,453號；第6,410,248號；第6,140,466號；第6,200,759號；及第6,242,568號；以及WO 98/37186；WO 98/53057；WO 00/27878；WO 01/88197

與GB 2,338,237。此外，針對鋅指結合域結合特異性之強化已描述於，舉例而言，專利權共有之WO 02/077227。

[0108] 此外，如於這些及其它參考文獻中所揭露，鋅指域及/或多重指化之鋅指蛋白質可能使用任何適當之連接子序列鏈接在一起，包括舉例而言，5或更多個胺基酸長度之連接子。亦參閱，美國專利第6,479,626號；第6,903,185號；及第7,153,949號，針對6或更多個胺基酸長度的示範性連接子序列。於此描述之該等蛋白質在該蛋白質之個別鋅指之間可能包括任何適當的連接子組合。

[0109] 於此描述的係為組成物，特別是核酸酶，該者對在體內切割帶有一轉基因之供體分子，及核酸酶其用於切割一細胞之基因組係為有用的，使得該轉基因在一標定方式中併入至該基因組內。在某些實施例中，該等核酸酶之一或者係為天然發生的。在其它實施例中，該等核酸酶之一或者係為非天然發生的，意即，在該DNA-結合域及/或切割域中經工程化的。舉例而言，一天然發生核酸酶之DNA-結合域可能變更的，以結合至一選定的標靶位點(例如，一旦核酸酶已被工程化以結合至位點其不同於該同族結合位點者)。在其它實施例中，該核酸酶包含異源DNA結合及切割域(例如鋅指核酸酶；TAL-效應子域DNA結合蛋白；具異源切割域之巨核酸酶DNA結合域)。

[0110] 任何適合的切割域可以操縱地鏈接至一DNA結合域，以形成一核酸酶。舉例而言，ZFP DNA結合域已經融合至核酸酶域，以創造ZFN，ZFN係為能夠透過其工程化

(ZFP)DNA結合域識別其意欲之核酸標靶，並造成該DNA通過核酸酶活性在靠近該ZFP結合位點處被切斷之一功能實體。參閱，例如，Kim等人於”(1996) *Proc Natl Acad Sci USA* 93(3)：1156-1160”。最近，在多種生物體中，ZFN已用於基因組修飾。參閱，舉例而言，美國專利公開案第20030232410號；第20050208489號；第20050026157號；第20050064474號；第20060188987號；第20060063231號；及國際公開案WO 07/014275。同樣地，TALE DNA結合域已經融合至核酸酶域，以創造TALEN。參閱，例如，美國公開案第20110301073號。

[0111] 如上文所指出，該切割域可能異源於該DNA結合域，舉例而言一鋅指DNA結合域與來自於一核酸酶之一切割域，或一TALEN DNA結合域與一切割域，或巨核酸酶DNA結合域與來自一不同核酸酶的切割域。異源性切割域可從一特定的核酸內切酶或核酸外切酶獲得。切割域可由其衍生之示範性核酸內切酶包括但不限於，某些限制性核酸內切酶及歸巢性核酸內切酶。參閱，舉例而言，麻州比佛利新英格蘭生物實驗室2002-2003年產品目錄；及Belfort等人於”(1997) *Nucleic Acids Res.* 25：3379-3388”。切割DNA的額外酵素為已知的(例如，S1核酸酶；綠豆核酸酶；胰DNase I；微球菌核酸酶；酵母菌HO核酸內切酶；亦參閱Linn等人(編輯)於”*Nucleases, Cold Spring Harbor Laboratory Press, 1993*”)。這些酵素(或其功能性片段)之一或更多者可以使用做為切割域及切割半域之來源。

[0112] 類似地，一切割半域可以衍自於任一核酸酶或其一部份，如上文所陳述，該者對於切割活性要求二聚體化。一般而言，對於切割，二個融合蛋白質係要求的，假若該兩融合蛋白質包含切割半域。或者，包含兩切割半域之單一蛋白質可以使用的。該兩切割半域可以衍自於相同的核酸內切酶(或其功能性片段)，或每一切割半域可以衍自於一不同的核酸內切酶(或其功能性片段)。此外，對該兩融合蛋白質之標靶位點較佳地係參照於彼此布署的，使得該兩融合蛋白質至其分別的標定位點之結合放置該切割半域彼此在允許該兩切割半域形成一功能性切割域的一空間定位中，例如，二聚體化。因此，在某些實施例中，該等標定位點之近邊係由5-8個核苷酸或15-18個核苷酸分隔。然而，任一整數之核苷酸或核苷酸對可以介入二標定位點之間(例如，從2至50個核苷酸對或更多)。一般而言，切割位點位於該等標靶位點之間。

[0113] 限制性核酸內切酶(限制酵素)存在於許多物種中，且係能夠序列-特異性地結合至DNA(於一識別位點)，並於結合位點或附近切割DNA。某些限制酵素(例如，第IIS型)於遠離該識別位點處切割DNA，並具有可分離的結合與切割域。舉例而言，第IIS型酵素*Fok I*催化DNA之雙股切割，於一股上在距其識別位點9個核苷酸處，且於另一股上在距其識別位點13個核苷酸處。參閱，舉例而言，美國專利第5,356,802號；第5,436,150號及第5,487,994號；以及Li等人於”(1992) *Proc.Natl.Acad.Sci.USA* 89 : 4275-4279”；Li等人

於”(1993)*Proc.Natl.Acad.Sci.USA* 90 : 2764-2768”；Kim等人於”(1994a)*Proc.Natl.Acad.Sci.USA* 91 : 883-887”；Kim等人於”(1994b)*J.Biol.Chem.*269 : 31,978-31,982”。因此，在一實施例中，融合蛋白質包含源自至少一第IIS型限制酵素之切割域(或切割半域)，及一或多個鋅指結合域，其可經或未經工程化的。

[0114] 第IIS型限制酵素，其切割域係可從該結合域分隔的，之一範例為*Fok I*。此特定酵素做為一二聚體為活性的。Bitinaite等人於”(1998)*Proc.Natl.Acad.Sci.USA* 95 : 10,570-10,575”。相應地，為了達成本揭露內容之目的，在該揭露融合蛋白質中使用之*Fok I*酵素之部分可視為一切割半域。因此，使用鋅指-*Fok I*融合體用於標定雙股切割，及/或標定DNA序列之置換，二融合蛋白質，每一者包含一*Fok I*切割半域，可以使用以再組成一催化活性的切割域。或者，含有一鋅指結合域與兩*Fok I*切割半域之單一多勝肽分子亦可以使用。使用鋅指-*Fok I*融合體用於標定切割與標定序列變更之參數係於本揭露內容別處提供的。

[0115] 一切割域或切割半域可以為保留切割活性，或保留多體化(例如，二聚體化)以形成功能性切割域之蛋白質的任一部份。

[0116] 示範性第IIS型限制酵素係於國際公開案WO 07/014275號中描述的，該者係以其整體合併於此。額外的限制酵素亦含有可分離之結合域與切割域，且這些係由本揭露內容視為可能的。參閱，舉例而言，Roberts等人

於”(2003)*Nucleic Acids Res.*31 : 418-420”。

[0117] 在某些實施例中，該切割域包含一或多個經工程的切割半域(亦意指為二聚體化域突變(dimerization domain mutants))，其可最小化或預防同質二聚體化，舉例而言，如描述於美國專利公開案第20050064474號；第20060188987號；第20070305346號及第20080131962號，該等全部之揭露內容係以其整體合併於此以做為參考。*Fok I*於位置446、447、479、483、484、486、487、490、491、496、498、499、500、531、534、537與538之胺基酸殘基，皆為影響*Fok I*切割半域二聚體化之標靶。

[0118] 形成絕對異質二聚體之示範性的經工程*Fok I*切割半域包括一對切割半域，其中第一切割半域包括於*Fok I*胺基酸殘基位置490與538的突變，而第二切割半域包括於胺基酸殘基位置486與499之突變。

[0119] 因此，在一實施例中，於490之突變係以Lys(K)置換Glu(E)；於538之突變係以Lys(K)置換Iso(I)；於486之突變係以Glu(E)置換Gln(Q)；且於位置499之突變係以Lys(K)置換Iso(I)。具體地，於此描述之該經工程切割半域係製備的，藉由在一切割半域中突變位置490(E→K)與538(I→K)，以產生命名為”E490K : I538K”之一經工程切割半域，且藉由在另一切割半域中突變位置486(Q→E)與499(I→L)，以產生命名為”Q486E : I499L”之一經工程切割半域。於此描述之該經工程切割半域係為絕對異質二聚體突變物，在該者中異常之切割係最小化或消除的。參閱，例如，美國專利

公開案第2008/0131962號，為了一切目的，該者之揭露內容係以其整體合併於此以做為參考。在某些實施例中，該經工程之切割半域包含於位置486、499與496(相對於野生型FokI而編號)之突變，譬如以Glu(E)殘基置換於位置486之野生型Gln(Q)殘基，以Leu(L)殘基置換於位置499之野生型Iso(I)殘基，及以一Asp(D)或Glu(E)殘基置換於位置496之野生型Asn(N)殘基(亦分別意指為一”ELD”與”ELE”域)。在其它實施例中，該經工程之切割半域包含於位置490、538與537(相對於野生型FokI而編號)之突變，譬如以Lys(K)殘基置換於位置490之野生型Glu(E)殘基，以Lys(K)殘基置換於位置538之野生型Iso(I)殘基，及以Lys(K)殘基或Arg(R)殘基置換於位置537之野生型His(H)殘基(亦分別意指為”KKK”與”KKR”域)。在其它實施例中，該經工程之切割半域包含於位置490與537(相對於野生型FokI而編號)之突變，譬如以Lys(K)殘基置換於位置490之野生型Glu(E)殘基，且以Lys(K)殘基或Arg(R)殘基置換於位置537之野生型His(H)殘基(亦分別意指為”KIK”與”KIR”域)。(參閱美國專利公開案第20110201055號)。在其它實施例中，該經工程之切割半域包含”Sharkey”及/或”Sharkey”突變(參閱Guo等人於”(2010) *J. Mol. Biol.* 400(1) : 96-107”)。

[0120] 於此描述之該工程化切割半域可以使用任何適合的方法製備，舉例而言，藉由如在美國專利公開案第20050064474號；第20080131962號；及第20110201055號中所描述的野生型切割半域(Fok I)的定點突變作用。

[0121] 或者，核酸酶可能在體內於該核酸標定位點組裝的，使用所謂的”分裂酵素(split-enzyme)”技術(參閱例如美國專利公開案第20090068164號)。此種分裂酵素組分之表現可能在分開的表現構建體上，或是可以在其中該個別組分係分開的一開放閱讀框中鏈接，舉例而言，藉由一自我切割2A勝肽或IRES序列。組分可能為個別的鋅指結合域或一巨核酸酶核酸結合域之域。

[0122] 核酸酶可以在使用之前對活性篩選，舉例而言，如於WO 2009/042163與20090068164中所描述般在一基於酵母菌的染色體系統。表現核酸酶的構建體可以使用該技藝中所知悉的方法輕易地設計。參閱，例如，美國專利公開案第20030232410號；第20050208489號；第20050026157號；第20050064474號；第20060188987號；第20060063231號；及國際公開案WO 07/014275。核酸酶的表現可能在一組成型啓動子或一誘導型啓動子下控制，舉例而言，棉子糖及/或半乳糖存在下被活化(去壓制)而在葡萄糖存在下被壓制的半乳糖激酶啓動子。

[0123] 在一實施例中，該聚核苷酸供體卡匣包含一解析域。在其它實施例中，該解析域包含一特異性的聚核苷酸序列。特異性聚核苷酸序列之例子包括，但不限於序列辨識編號：142及序列辨識編號：143。在各種實施例中，對序列辨識編號序列：142及序列辨識編號143之該兩特異性聚核苷酸序列具至少90%序列一致性的序列係於此揭露的。在其它實施例中，對序列辨識編號序列：142及序列辨

識編號143之該兩特異性聚核苷酸序列具至少95%序列一致性的序列係於此揭露的。在其它實施例中，對序列辨識編號序列：142及序列辨識編號143之該兩特異性聚核苷酸序列具至少97%序列一致性的序列係於此揭露的。在又其它實施例中，對序列辨識編號序列：142及序列辨識編號143之該兩特異性聚核苷酸序列具至少99%序列一致性的序列係於此揭露的。

[0124] 在一實施例中，該解析域包含40至60%之一鳥糞嘌呤與胞嘧啶之比率。在其它實施例中，該解析域包含42.25至57.5%之一鳥糞嘌呤與胞嘧啶之比率。在其它實施例中，該解析域包含45至55%之一鳥糞嘌呤與胞嘧啶之比率。在另一實施例中，該解析域包含47.75至52.25%之一鳥糞嘌呤與胞嘧啶之比率。在包含一聚核苷酸序列之區域，鳥糞嘌呤與胞嘧啶含量的比率可以藉由計數鳥糞嘌呤與胞嘧啶殘基在該聚核苷酸序列中的數目並將這個數目除以在該聚核苷酸序列之延伸區中鹼基對總數而計算，且然後將該結果乘以100，以決定該聚核苷酸序列之鳥糞嘌呤與胞嘧啶百分比。包含40-60%鳥糞嘌呤與胞嘧啶比率之引子在PCR放大期間引致與模板DNA更穩定的結合。然而，具大於60%之鳥糞嘌呤與胞嘧啶比率的DNA分子係難以通過PCR放大，因此序列可能形成莖環、二聚物及其它二級結構。

[0125] 熔融溫度(T_m)係界定為該DNA股一半係於雙螺旋狀態而一半係於不規則螺旋狀態時的溫度。熔融溫度可

以取決於變量，包括該分子的長度、該分子的具體核苷酸序列組成物及該鹽與核酸在溶液中之莫耳濃度。雙鏈體之穩定性與熔融溫度在熱循環可能涉及的核酸放大中為重要的。在此種熱循環步驟中，溫度係上升到足夠高於該熔融溫度以上，以使該標靶核酸與其互補體的雙鏈體解離。在後續的再黏著步驟中，溫度係被帶到低於該熔融溫度，使得該標靶核酸與引子(等)及/或探針(等)的雙鏈體能夠形成，同時仍然維持足夠高以避免非特異性雜交。熔融溫度及如何測量它在該技藝中為充分描述的。參閱，舉例而言，"Nucleic Acids Research (1990) 18, 6409-6412" 及 "Bioinformatics, V. 12(12), pp. 1226-1227"。有關雙鏈體穩定性與核苷酸序列的理論或經驗模型可以使用以預測核酸的熔融溫度。舉例而言，"Proc. Natl. Acad. Sci : U.S.A. 1986, 83:3746-3750"; "Biochemistry 1996, 35:3555-3562"；及 "Proc. Natl. Acad Sci. U.S.A. 1998, 95 : 1460-1465" 描述了在該技藝中廣泛使用且被稱為"最近鄰模式(nearest neighbour model)"的一種熔融溫度預測模型。此模型考量了寡核苷酸的長度與在其內的鹼基序列兩者。

[0126] 該預測的熔融溫度意指尚未與任何部分(moieties)共軛的寡核苷酸之理論計算的熔融溫度。其係通常為一寡核苷酸之實際熔融溫度的一估計值，且係基於該寡核苷酸的序列，典型地不考慮任何標誌及任何關聯的連接子與之類的效果。給予一寡核苷酸的確實序列，該寡核苷酸的熔融溫度可以如上文討論般預測。根據本發明該預

測的熔融溫度係計算的，在一實施例中，使用最近鄰模型。在一進一步實施例中，根據本發明該預測的熔融溫度使用採用SantaLucia值(SantaLucia, J., Jr. (1998) *Proc. Natl. Acad. Sci. U.S.A.* (1998) 95, 1460-1465)之最近鄰熱力學理論計算。該SantaLucia值係含在各種可用於預測熔解溫度的電腦程式中。在一實施例中，該”Beacon Designer”套裝軟體係使用的。在一實施例中，該SantaLucia值係伴隨下列反應條件使用：0.25 nM之核酸濃度；50 mM之單價離子濃度；5 mM的游離鎂離子濃度；332.84mM的總Na⁺當量；及25°C的自由能計算溫度。不論所使用之特定軟體程式或方法論的抉擇，相同的軟體程式或方法論所欲地係適合伴隨其內設設定使用的，以確定在本發明中所使用之引子(等)與探針(等)的預測熔融溫度。

[0127] 該術語”最小自由能”(MFE)，”自由能”與”吉布斯自由能”於此係互換使用的。吉布斯自由能係為一熱力學數量，其係為焓與絕對溫度與一系統熵之乘積之間的差。吉布斯自由能係為一系統做非機械功的能力而ΔG測量在其上進行的非機械功。(Perrot, Pierre. *A to Z of Thermodynamics*. Oxford University Press (1998))。吉布斯自由能係界定為G=H-TS中，其中H是焓，T是溫度而S是熵(H=U+pV，其中p是壓力而V是體積)。

[0128] 其係普遍認為的是，所有系統力求實現一最小自由能。因此，當在吉布斯自由能(ΔG)中之變化為負時，那麼一反應係為有利且能量係釋放的。釋放能量之數量係

等於做為該特定化學反應的結果可以被執行的最大數量的功。當條件引致吉布斯自由能(ΔG)的變化為正時，那麼能量必須加入到該系統中以使反應展開。在等溫、等壓系統中，吉布斯自由能係為在一熱力學過程中涉及的焓與熵之競爭效果的代表性測量。因此，吉布斯自由能可以被視為一動態量。

[0129] 其應注意的是，自由能的測量可以藉由核苷酸序列的長度偏倚的。較長的核苷酸序列較諸短的核苷酸序列具有更大範圍的自由能。因此，可變長度的核苷酸序列可以藉由將自由能除以序列長度常態化該自由能計算(kcal/mol)而比較(意即，平均自由能= $\Delta G/\text{核苷酸序列長度}(\text{kcal/mol/鹼基對})$)。

[0130] 相應地，如於此所使用，最小自由能(意即自由能，吉布斯自由能)辨識由熱力學優化(意即祖克演算法的落實(M. Zuker 及 P. Stiegler於”*Nucleic Acids Research* 9 : 133-148(1981))所找到對於該結構具有最低自由能值的值(意即，吉布斯自由能； $\Delta G(\text{kcal/mol})$ ； $\Delta G/\text{核苷酸序列長度}(\text{kcal/mol/鹼基對})$)。一序列的吉布斯自由能可以計算，舉例而言，使用熟習該項技藝者所知悉的 RNAfold™ 或 PRIMER EXPRESS™(1.0版,(Applied Biosystems, 加州福斯特城)或mFOLD™軟體(現今之UNIFold™)(IDT, 加州聖何塞)程式(參見，例如，Id. Hofacker等人於”*Monatshefte f. Chemie* 125 : 167-188 (1994)”； McCaskill J S.於”*Biopolymers* 29 (6-7) : 1105-19. (1990)”；及 Hofacker等人於”*Bioinformatics*

22 (10) : 1172-6 (2006)”。

[0131] 在一實施例中，該解析域包含不含有數個重複序列的一聚核苷酸，其中該重複序列長度至少為9 Bp。在一些實施例中，數個重複序列包含2、3、4、5、6、7、8、9、10或更多個長度至少9 Bp的聚核苷酸片段。在一實施例中，一第一聚核苷酸序列構成一重複序列，當包含相同序列之一第二聚核苷酸序列係直接位於該第一聚核苷酸序列下游時(例如，一第一聚核苷酸序列包含5'-gaaaaataacaa-3'，此聚核苷酸序列構成一重複序列，當包含相同序列的一第二聚核苷酸序列(下劃線)係直接位於該第一聚核苷酸序列下游時，5'-gaaaaataacaagaaaaataacaa-3')。在一實施例中，該解析域包含不含有數個此種重複序列的一聚核苷酸。

[0132] 在一實施例中一解析域包含一聚核苷酸其不含有一系列長度大於9 Bp的一致鹼基對序列。一致的鹼基對為在空間或時間順序中一個接著一個的相同嘌呤(腺嘌呤或鳥糞嘌呤)或嘧啶(胞嘧啶或胸腺嘧啶)鹼基對之任意數量的鹼基對(大於9 Bp)。相應地，一核苷酸延伸區諸如5'-AAAAAAA-3'、5'-GGGGGGGG-3'、5'-CCCCCCCC-3'或5'-TTTTTTTT-3'將包含一系列的9個一致鹼基對。在一實施例中，一系列一致的鹼基對序列可能長度大於9 Bp、10 Bp、11 Bp、12 Bp、13 Bp、14 Bp、15 Bp或更多個Bp。

[0133] 在其它實施例中，該聚核苷酸供體卡匣之解析域包含一或多個限制酵素序列。一限制酵素係為能夠切割DNA成片段的酵素。由一限制酵素將大分子量的基因組

DNA切割成低分子量片段之片段允許該等DNA片段對後續分析更有效的分離及操作。

[0134] 限制酵素消化，也引用為限制性核酸內切酶消化，係為當一核酸酶酵素係使用於切開該聚核苷酸序列時所執行的。對熟習該項技藝者，其係有眾多可獲得的限制酵素。如於 www.neb.com/nebcomm/tech_reference/restriction_enzymes/overview.asp 中所描述，四種分類係使用以表徵限制酵素。這些分類係於次單元組成物、切割位置、序列特異性及輔因子要求之基礎上做成的。

[0135] I型酵素在距該識別/結合序列一段距離之位置隨機切斷DNA(>1000 bp遠處)。由一I型酵素酶所結合的識別位點為不對稱的。因此這些酵素不使用於基因選殖，因為這些酵素不產生離散的限制性片段或區隔的凝膠條帶圖樣。I型酵素為多功能的，且包含一I型限制酵素之該等不同次單元係負責不同活性(意即次單元HsdR編碼限制，次單元HsdM編碼DNA之甲基化，而次單元HsdS編碼識別序列的特異性)。

[0136] II型酵素在位於識別序列鄰近區內的位置消化DNA。這些酵素以做為二聚體起作用的，其中一次單元結合至該正義股而該次單元之一第二拷貝於一典型地4-8個核苷酸長度之間的回文序列結合至該反義股。結合至該DNA的II型二聚體可以為結合至對稱DNA序列的一同型二聚體，或是結合至非對稱DNA序列的一異源二聚體。這些酵素可以識別連續序列或不連續序列兩者任一的。II型酵素

係可從市場上購得的，且普遍地使用於DNA分析及基因選殖。這些酵素的廣泛使用是由於所產生且可以在瓊脂糖凝膠上分辨之區隔的限制片段的結果。

[0137] II型酵素係為不相關蛋白質之一集合，該等蛋白質在胺基酸序列相似度中為高度發散的。II型酵素已被分成子類別，該等係使用一字母後綴標誌。IIB型限制酵素為含有多於一次單元的多聚體。這些酵素切割該識別序列的兩側，從而引致該識別序列之移除。IIE型與IIF型限制酵素繼之與其之識別序列的兩個拷貝交互作用後切開DNA。IIG型限制酵素係由一單一次單元購成的。該酵素之N-末端部分擁有一DNA切割域及DNA修飾域。該酵素之C-末端部分擁有一DNA序列結合域。這些酵素切開其識別序列的外部。IIM型限制酵素識別並切斷經甲基化的DNA。IIS型限制酵素以二聚體作用，且在該非回文不對稱識別位點外部之一位置切開DNA。這些酵素係由兩個區隔的域構成，一域用於DNA結合，而另一域用於DNA切割。

[0138] III型酵素酶為組合限制性與修飾酵素。這些酵素識別兩個分開的非回文序列且切開其識別序列外側。III型酵素在相同的DNA分子內要求兩個相反面向的識別序列以完成切割。

[0139] IV型酵素識別經甲基化的DNA。例子包括大腸桿菌之McrBC與Mrr系統。

[0140] 在一實施例中，該聚核苷酸供體卡匣之解析域包含一或多個引子結合序列。該術語”引子結合序列”意指

該解析域或任何其它聚核苷酸序列之一區域，典型地夾擊一標靶標區域及/或放大子之序列，其可以直接做為，或憑藉其互補體，做為模板，在其上一引子可以黏著用於該技藝中已知的任何適合的引子延伸反應或放大反應，舉例而言，但不限於，PCR。對熟習該項技藝者，將理解的是，當兩個引子結合位點係存在於一單一聚核苷酸上時，該兩個引子結合位點之面向係不同的。舉例而言，一引子對的一個引子係互補且可以雜交至該第一引子結合位點，該而引子對之對應引子係設計以雜交至該第二引子結合位點的互補體。換句話說，在一些實施例中，該第一引子結合位點可以為正義面向，而該第二引子結合位點可以在一反義面向。放大子之引子結合位點可能，但不需要包含該標靶夾擊序列或其互補體之相同序列或至少一些序列。

[0141] 一引子可以設計以結合至一解析域其包含經計算具有-1至-18 kcal/mole之自由能(ΔG)的二級結構。在一實施例中，一引子可以設計為具有一熔融溫度(T_m)其係至少大於包含具-1至-18 kcal/mole之自由能(ΔG)的二級結構之該解析域區域者10°C，該自由能(ΔG)係使用”Markham, N. R. & Zuker, M. (2008) UNAFold : software for nucleic acid folding and hybridization. In Keith, J. M., editor, *Bioinformatics, Volume II. Structure, Function and Applications*, number 453 in *Methods in Molecular Biology*, chapter 1, pages 3–31. Humana Press, Totowa, NJ. ISBN 978-1-60327-428-9”計算的。

[0142] 在一實施例中，該聚核苷酸供體卡匣之解析域包含一或多個同源臂序列。該一或多個同源臂序列之後續實施例係由50至100個鹼基對長度所構成。由核苷酸之隨機延伸區所構成的同源臂可以針對各種應用輕易地設計與合成。該通用聚核苷酸供體序列可能包含一或多個採用於同源併入一聚核苷酸至一宿主基因組內的同源臂。在一些實施例中，該同源臂係藉由在該同源臂之內或旁邊引入一雙股斷裂並在該同源臂區域之內併入可能包含一基因表現卡匣其包含一轉基因的第二聚核苷酸而標定。典型地，該第二聚核苷酸的併入係經由同源重組導向修復發生。

[0143] 在一實施例中，該同源臂區域與一最佳基因組基因座共享至少80、90、93、95、99或100%的序列一致性。在一實施例中，該實施例同源臂區域與一植物細胞的最佳基因組基因座共享至少80、90、93、95、99或100%的序列一致性。在一實施例中，該最佳基因組基因座係為一非基因單子葉或雙子葉序列，其係低甲基化、可表現的、重組的示範性證據，且在該植物之基因組中係位於一基因區域的近側位置。

[0144] 在一實施例中，該最佳基因組基因座係為具有下列性質或特徵的低甲基化非基因序列：

- a) 該非基因序列之甲基化程度為1%或更低；
- b) 該非基因序列與在該基因組中所含有的任意其它序列共享小於40%的序列一致性；
- c) 該非基因序列與在該基因組中所含有的任意其它序列

共享小於40%的序列一致性；

- d) 該非基因序列係位於一已知或預測的表現編碼序列之40 Kb區域之內；及
- e) 該非基因序列在該基因組之內展現大於0.00041 cM/Mb之一重組頻率。

[0145] 在一實施例中，該最佳基因組基因座係為具有下列1、2、3、4或5種性質或特徵的一低甲基化非基因序列：

- a) 在該非基因序列40 Kb之內具有一已知或預測的基因編碼序列；
- b) 在該非基因序列之一端的40 Kb之內具有一序列其包含一已知基因的上游2Kb及/或下游1 Kb；
- c) 該非基因序列內不含有大於1%的DNA甲基化；
- d) 不含有對該基因組中任意其它序列具有大於40%序列一致性的1 KB序列；及
- e) 大於0.00041 cM/Mb之重組頻率的重組示範性證據。

[0146] 按照一實施例，該經選擇的非基因序列包含下列2、3、4、5、6、7或8種特徵：

- a) 該非基因序列不含有一甲基化聚核苷酸；
- b) 該非基因序列在該基因組中展現0.00041至62.42 cM/Mb之一重組率；
- c) 該非基因序列展現該基因組0至0.962程度之一核小體佔用(nucleosome occupancy)；
- d) 該非基因序列與該基因組中所含有的任意其它1 Kb序列共享小於40%的序列一致性；

- e) 該非基因序列具有從0.00373至0.99908之一相對位置值，代表距一染色體著絲點的基因組距離之比率；
- f) 該非基因序列具有25.17至68.3%之一鳥糞嘌呤/胞嘧啶百分比含量範圍；
- g) 該非基因序列位於一基因序列近側；且，
- h) 基因組序列之1 Mb區域包含該非基因序列，包含一或多個非基因序列。

[0147] 在一實施例中，該聚核苷酸供體卡匣之解析域包含不編碼勝肽之一序列。

[0148] 在一實施例中，該聚核苷酸供體卡匣之解析域包含編碼勝肽之一序列。為了表現一勝肽，編碼該勝肽序列的核苷酸序列典型地係次選殖至含有啓動子以主導轉錄的一表現載體內。適合的細菌及真核啓動子在該項技藝中係廣為知悉的，且描述於，例如，Sambrook等人於”Molecular Cloning, A Laboratory Manual(第2版，1989；第3版，2001)”；Kriegler於”Gene Transfer and Expression : A Laboratory Manual(1990)”；及”Current Protocols in Molecular Biology”(Ausubel等人，上文)。用於表現一勝肽之細菌表現系統係可從下列中獲得：例如，大腸桿菌(*E. coli*)、芽孢桿菌(*Bacillus* sp.)及沙門氏菌(Palva等人於”Gene 22:229-235 (1983)”)。用於此等表現系統之套組係可從市場上購得。用於哺乳動物細胞、酵母菌及昆蟲細胞之真核表現系統為熟習該項技藝者所熟知且亦可從市場上購得。

[0149] 在一實施例中，該聚核苷酸供體卡匣包含一基

因表現卡匣其包含一轉基因者。該基因表現卡匣典型地含有一轉錄單元或表現卡匣，該者含有在宿主細胞中表現該核酸所要求的所有額外元件，無論是原核或真核。一典型的基因表現卡匣從而含有例如與編碼該蛋白質之一核酸序列操縱鏈接之一啓動子，及例如用於該轉錄體有效聚腺苷酸化、轉錄終止、核糖體結合位點或轉譯終止所要求之信號。該卡匣之額外元件可能包括，例如增強子與異源剪接信號。

[0150] 在一實施例中該基因表現卡匣亦將於該感興趣異源核苷酸序列之3'端包括在植物中起作用之一轉錄及轉譯終止區域。該終止區域可以為天然於本揭露內容之實施例啓動子的核苷酸序列，可以為天然於該感興趣DNA序列者，或可以衍自另一來源。便利的終止區域係可從農桿腫瘤菌之Ti-質體獲得，諸如章魚鹼合成酶與胭脂鹼合成酶(nos)終止區域(Depicker等人於”Mol. and Appl. Genet. 1 : 561-573 (1982)”及 Shaw 等人於 ”(1984) Nucleic Acids Research vol. 12, No. 20 pp7831-7846(nos)”)；亦參閱 Guerineau等人於”Mol. Gen. Genet. 262 : 141-144 (1991)“；Proudfoot於”Cell 64 : 671-674 (1991) ”；Sanfacon等人於”Genes Dev. 5 : 141-149 (1991) ”；Mogen等人於”Plant Cell 2 : 1261-1272 (1990) ”；Munroe等人於”Gene 91 : 151-158 (1990) ”；Ballas等人於”Nucleic Acid Res. 17 : 7891-7903 (1989) ”；”Joshi等人於”Nucleic Acid Res. 15 : 9627-9639 (1987)”。

[0151] 在其它實施例中，該基因表現卡匣可以額外地含有5'前導子序列。此種前導子序列可以作用以增強轉譯。轉譯前導子在該技藝中為已知的，且藉助於例子，包括小核糖核酸病毒前導子、EMCV前導子(腦心肌炎5'非編碼區域)、Elroy-Stein等人於"Proc. Nat. Acad. Sci. USA 86 : 6126-6130 (1989)"；馬鈴薯Y群病毒(potyvirus)前導子，舉例而言，TEV前導子(煙草蝕刻病毒)，Carrington及Freed於"Journal of Virology, 64 : 1590-1597 (1990)"；MDMV前導子(玉米矮化嵌紋病毒)，Allison等人於"Virology 154 : 9-20 (1986)"；人類免疫球蛋白重鏈結合蛋白(BiP)，Macejak等人於"Nature 353 : 90-94 (1991)"；源自苜蓿嵌紋病毒外殼蛋白mRNA的非轉譯前導子(AMV RNA4)，Jobling等人於"Nature 325 : 622-625 (1987)"；煙草嵌紋病毒前導子(TMV)，Gallie等人於"(1989) Molecular Biology of RNA"，第237-256頁；及玉米黃化斑紋病毒(maize chlorotic mottle virus)前導子(MCMV)，Lommel等人於"Virology 81 : 382-385 (1991)"。亦參閱Della-Cioppa等人於"Plant Physiology 84 : 965-968 (1987)"。該構建體亦可以含有增強轉譯及/或mRNA穩定性的序列，諸如內含子。此種內含子之一例子為阿拉伯芥之組蛋白H3.III變體的基因II之第一內含子。Chaubet等人於"Journal of Molecular Biology, 225 : 569-574 (1992)"。

[0152] 在一實施例中，該聚核苷酸供體序列之基因表現卡匣包含一啓動子。使用以主導一勝肽編碼核酸表現的

啓動子取決於特定應用。舉例而言，適合宿主細胞的一強大、組成型啓動子典型地係使用於蛋白質表現及純化。較佳植物啓動子之非限制性例子包括衍自阿拉伯芥泛素10(ubi-10)的啓動子序列(Callis等人於"1990, *J. Biol. Chem.*, 265 : 12486-12493")；農桿腫瘤菌甘露鹼合酶(Δmas)(Petolino等人於美國專利第6,730,824號)；及/或木薯葉脈嵌紋病毒(CsVMV)(Verdaguer等人於"1996, *Plant Molecular Biology* 31 : 1129-1139")。

[0153] 在於此揭露之方法中，主導一基因在植物中表現的許多啓動子可以被採用的。此種啓動子可以選自於組成型、化學調控型、誘導型、組織特異性及種子較佳啓動子等。

[0154] 組成型啓動子包括，舉例而言，核心花椰菜嵌紋病毒35S啓動子(Odell等人於"(1985) *Nature* 313 : 810-812")；水稻肌動蛋白啓動子(McElroy等人於"(1990) *Plant Cell* 2 : 163-171")；玉米泛素啓動子(美國專利第5,510,474號；Christensen等人於"(1989) *Plant Mol. Biol.* 12 : 619-632"及Christensen等人於"(1992) *Plant Mol. Biol.* 18 : 675-689")；pEMU啓動子(Last等人於"(1991) *Theor. Appl. Genet.* 81 : 581-588")；ALS 啓動子(美國專利第5,659,026號)；玉米組蛋白啓動子(Chabouté等人於"Plant Molecular Biology, 8 : 179-191 (1987)")；及之類。

[0155] 可獲得的植物相容啓動子範圍包括組織特異性與誘導型啓動子。誘導型調控元件係為一個能夠回應誘導

物，直接或間接地活化一或多個DNA序列或基因之轉錄者。在缺乏誘導物下，該DNA序列或基因將不會轉錄。典型地，特異性結合至一誘導型調控元件以活化轉錄的蛋白質因子係於不活化形式中存在，其然後直接或間接由該誘導物轉換為活化形式。該誘導物可以為化學試劑，諸如蛋白質、代謝物、生長調節劑、除草劑或酚類化合物，或直接由熱、冷、鹽、或有毒元素或間接透過諸如病毒之病原體或致病劑作用施加的生理壓力。典型地，特異性結合至一誘導型調控元件以活化轉錄的蛋白質因子係於不活化形式中存在，其然後直接或間接由該誘導物轉換為活化形式。該誘導物可以為化學試劑，諸如蛋白質、代謝物、生長調節劑、除草劑或酚類化合物，或直接由熱、冷、鹽、或有毒元素或間接透過諸如病毒之病原體或致病劑作用施加的生理壓力。含有誘導型調控元件的植物細胞可能藉由從外部施用該誘導物至細胞或植物而曝露至一誘導物，諸如藉由噴霧、澆水、加熱或類似的方法。

[0156] 任意誘導型啟動子可以在本揭露內容之實施例中使用。參閱 Ward 等人於 "Plant Mol. Biol. 22 : 361-366 (1993)"。示範性的誘導型啟動子包括蛻皮激素受器啟動子(美國專利第6,504,082號)；源自 ACEI 系統、對銅回應的啟動子(Mett 等人於 "PNAS 90 : 4567-4571 (1993)")；源自玉米、對苯礦醯胺除草劑安全劑回應的 In2-1 與 In2-2 基因(美國專利第5,364,780號；Hershey 等人於 "Mol. Gen. Genetics 227 : 229-237 (1991)" 及 Gatz 等人於 "Mol. Gen. Genetics 243 :

32-38 (1994)"; 源自 Tn10 之 Tet 抑制子 (Gatz 等人於 "Mol. Gen. Genet. 227 : 229-237 (1991)" ; 或源自類固醇荷爾蒙基因的啓動子，該者之轉錄活性係藉由一糖皮質類固醇荷爾蒙誘導，Schena 等人於 "(1991) Proc. Natl. Acad. Sci. USA 88 : 10421(1991)"，及 McNellis 等人於 "(1998) Plant J. 14(2) : 247-257"；玉米 GST 啓動子，其藉由被用作芽前除草劑的疏水性親電子化合物活化 (參閱美國專利第 5,965,387 號及國際專利申請案公開號 WO 93/001294)；及煙草 PR-1a 的啓動子，其藉由水楊酸活化 (參閱 Ono S、Kusama M、Ogura R、Hiratsuka K. 於 "Evaluation of the Use of the Tobacco PR-1a Promoter to Monitor Defense Gene Expression by the Luciferase Bioluminescence Reporter System, Biosci Biotechnol Biochem. 2011 Sep 23 ; 75(9) : 1796-800")。其它感興趣的化學調控啓動子包括四環素誘導型及四環素抑制型啓動子 (參閱，舉例而言，Gatz 等人於 "(1991) Mol. Gen. Genet. 227 : 229-237"，及美國專利第 5,814,618 號及第 5,789,156 號)。

[0157] 其它感興趣的可調控啓動子包括一冷回應調控元件或熱休克調控元件，該者之轉錄可以分別回應曝露於冷或熱而影響 (Takahashi 等人於 "Plant Physiol. 99 : 383-390, 1992")；醇脫氫酶基因之啓動子 (Gerlach 等人於 "PNAS USA 79 : 2981-2985 (1982)"；Walker 等人於 "PNAS 84(19) : 6624-6628 (1987)")，由厭氧條件誘導；及衍自豌豆 *rbcS* 基因或豌豆 *psaDb* 基因之光誘導型啓動子 (Yamamoto 等人

於”(1997) *Plant J.* 12(2) : 255-265”)；光誘導型調控元(Feinbaum等人於”*Mol. Gen. Genet.* 226 : 449, 1991”；Lam及Chua於”*Science* 248 : 471, 1990”；Matsuoka等人於”(1993) *Proc. Natl. Acad. Sci. USA* 90(20) : 9586-9590”；Orozco等人於”(1993) *Plant Mol. Bio.* 23(6) : 1129-1138”)；植物荷爾蒙誘導型調控元件(Yamaguchi-Shinozaki等人於”*Plant Mol. Biol.* 15 : 905, 1990”；Kares等人於”*Plant Mol. Biol.* 15 : 225, 1990”), 及之類。誘導型調控元件亦可以為玉米In2-1或In2-2基因之啓動子，該者回應苯磺醯胺除草劑安全劑(Hershey等人於”*Mol. Gen. Gene.* 227 : 229-237, 1991”；Gatz等人於”*Mol. Gen. Genet.* 243 : 32-38, 1994”), 及轉座子Tn10之Tet抑制子(Gatz等人於”*Mol. Gen. Genet.* 227 : 229-237, 1991”)。壓力誘導型啓動子包括鹽/水壓力誘導啓動子，諸如P5CS(Zang等人於”(1997) *Plant Sciences* 129 : 81-89”；冷誘導的啓動子，諸如cor15a(Hajela等人於”(1990) *Plant Physiol.* 93 : 1246-1252”)、cor15b(Wilhelm等人於”(1993) *Plant Mol. Biol.* 23 : 1073-1077”)、wsc120(Ouellet等人於”(1998) *FEBS Lett.* 423-324-328”)、ci7(Kirch等人於”(1997) *Plant Mol. Biol.* 33 : 897-909”)、ci21A(Schneider等人於”(1997) *Plant Physiol.* 113 : 335-45”); 乾旱誘導啓動子，諸如Trg-31(Chaudhary等人於”(1996) *Plant Mol. Biol.* 30 : 1247-57”)、rd29(Kasuga等人於”(1999) *Nature Biotechnology* 18 : 287-291”); 滲透壓誘導的啓動子，諸如Rab17(Vilardell等人於”(1991) *Plant Mol. Biol.* 17 : 985-93”)及滲透素

(osmotin)(Raghothama等人於”(1993) *Plant Mol Biol* 23 : 1117-28”)；及熱誘導的啓動子，諸如熱休克蛋白(Barros等人於”(1992) *Plant Mol.* 19 : 665-75”；Marrs等人於”(1993) *Dev. Genet.* 14 : 27-41”)、smHSP(Waters等人於”(1996) *J. Experimental Botany* 47 : 325-338”)，及來自歐芹泛素啓動之熱休克誘導元件(WO 03/102198)。其它壓力誘導的啓動子包括 rip2(美國專利第 5,332,808 號及美國公開案 2003/0217393)及 rd29a(Yamaguchi-Shinozaki等人於”(1993) *Mol. Gen. Genetics* 236 : 331-340”)。某些啓動子係可藉由傷口誘導的，包括膿桿菌pMAS啓動子(Guevara-Garcia等人於”(1993) *Plant J.* 4(3) : 495-505”)及膿桿菌ORF13之啓動子(Hansen等人於”(1997) *Mol. Gen. Genet.* 254(3):337-343”)。

[0158] 組織較佳啓動子可以利用以在一特定植物組織內標定增強的轉錄及/或表現。當意指一偏好表現時，意謂的是在該特定植物組織中較諸其它植物組織，表現係於一較高水平的。這些類型之啓動子的例子包括種子較佳表現，諸如由菜豆啓動子所提供之(Bustos等人於” 1989. *The Plant Cell* Vol. 1, 839-853”)，及玉米球蛋白-1基因(Belanger等人於”1991 *Genetics* 129 : 863-972”)。對於雙子葉植物，種子較佳啓動子包括，但不限於， β -菜豆蛋白、油菜籽蛋白(napin)、 β -伴大豆球蛋白(β -conglycinin)、大豆凝集素(lectin)、十字花科蛋白(cruciferin)、及之類。對於單子葉植物，種子較佳啓動子包括，但不限於玉米15 kDa玉米蛋白(zein)、22 kDa玉米蛋白、27 kDa玉米蛋白、 γ -玉米蛋白、

蠟樣蛋白(waxy)、皺縮蛋白(shrunken)1、皺縮蛋白2、球蛋白1……等等。種子較佳啓動子亦包括那些主導基因表現主要在該種子內之特定組織的啓動子，諸如，舉例而言， γ -玉米蛋白之胚乳較佳啓動子、來自煙草的潛隱性啓動子(Fobert 等人於 1994 年之 "T-DNA tagging of a seed coat-specific cryptic promoter in tobacco. Plant J. 4 : 567-577")、來自玉米的P-基因啓動子(Chopra等人於1996年之 "Alleles of the maize P gene with distinct tissue specificities encode Myb-homologous proteins with C-terminal replacements. Plant Cell 7:1149-1158"，Erratum 於"Plant Cell.1997, 1:109")、來自玉米的球蛋白-1啓動子(Belenger 及 Kriz 於 1991 年之 "Molecular basis for Allelic Polymorphism of the maize Globulin-1 gene. Genetics 129: 863-972")、及主導表現至玉米粒之種皮或殼的啓動子，舉例而言果皮特異性麩胺酸合成酶啓動子(Muhitch等人於 2002 年之 "Isolation of a Promoter Sequence From the Glutamine Synthetase₁₋₂ Gene Capable of Conferring Tissue-Specific Gene Expression in Transgenic Maize. Plant Science 163:865-872")。

[0159] 一基因表現卡匣可能含有一5'前導子序列。此種前導子序列可以作用以增強轉譯。轉譯前導子在該技藝中為已知的，且藉助於例子包括，包括小核糖核酸病毒前導子、EMCV前導子(腦心肌炎5'非編碼區域)、Elroy-Stein等人於"Proc. Nat. Acad. Sci. USA 86 : 6126-6130 (1989)"；馬

鈴薯Y群病毒(potyvirus)前導子，舉例而言，TEV前導子(煙草蝕刻病毒)，Carrington及Freed於"Journal of Virology, 64 : 1590-1597 (1990)"；MDMV前導子(玉米矮化嵌紋病毒)，Allison等人於"Virology 154 : 9-20 (1986)"；人類免疫球蛋白重鏈結合蛋白(BiP)，Macejak等人於"Nature 353 : 90-94 (1991)"；源自苜蓿嵌紋病毒外殼蛋白mRNA的非轉譯前導子(AMV RNA4)，Jobling等人於"Nature 325 : 622-625 (1987)"；煙草嵌紋病毒前導子(TMV)，Gallie等人於"(1989) Molecular Biology of RNA, pages 237-256"；及玉米黃化斑紋病毒(maize chlorotic mottle virus)前導子(MCMV)，Lommel等人於"Virology 81 : 382-385 (1991)"。亦參閱Della-Cioppa等人於"Plant Physiology 84 : 965-968 (1987)"。

[0160] 該構建體亦可以含有增強轉譯及/或mRNA穩定性的序列，諸如內含子。此種內含子之一例子為阿拉伯芥之組蛋白H3.III變體的基因II之第一內含子。Chaubet等人於"Journal of Molecular Biology, 225 : 569-574 (1992)"。

[0161] 在意欲讓該異源核苷酸序列之表現產物導向一特定的胞器，特別是質粒體、澱粉體，或至內質網，或分泌到細胞表面或細胞外的那些情況下，該表現卡匣可能進一步包含針對一轉運肽的編碼序列。此種轉運肽在該技藝中為係廣為知悉的，且包括但不限於，針對醯基載體蛋白的轉運肽、RUBISCO之小的次單元、植物EPSP合成酶及向日葵(*Helianthus annuus*)(參閱Lebrun等人於美國專

利第5,510,417號)、玉蜀黍Brittle-1葉綠體轉運胜肽(Nelson等人於"Plant Physiol 117(4) : 1235-1252 (1998)"；Sullivan等人於"Plant Cell 3(12) : 1337-48"；Sullivan等人於"Planta (1995) 196(3): 477-84"；Sullivan等人於"J. Biol. Chem. (1992) 267(26) : 18999-9004")，及之類。此外，嵌合葉綠體轉運胜肽在該技藝中為已知的，諸如該 Optimized Transit Peptide(參閱，美國專利第5,510,471號)。額外的葉綠體轉運胜肽先前已描述於美國專利第5,717,084號；第5,728,925號。熟習該項技藝者將輕易地理解在表現一產物至一特定胞器中可有許多選項。舉例而言，該大麥 α -澱粉酶序列經常使用以主導表現到內質網。Rogers於"J. Biol. Chem. 260 : 3731-3738 (1985)"。

[0162] 在一實施例中，該聚核昔酸供體卡匣包含一轉基因。於此之些實施例提供編碼一多胜肽的轉基因，該者包含一基因表現卡匣。此一轉基因在任何各式各樣產生轉基因植物之應用中可能為有用的。包含一基因表現卡匣之轉基因的特定例子係為了解示之目的於此提供的，且包括包含一性狀基因、RNAi基因、或報告/可選擇標記基因之基因表現。

[0163] 在工程化於植物中表現之一基因中，準宿主植物(等)的密碼子偏倚可能測定的，舉例而言，透過使用大眾可獲得的DNA序列資料庫，以找到關於各種植物基因的植物基因組或蛋白質編碼區域的密碼子分佈的資訊。一旦最優化(例如，一植物優化的)的DNA序列已設計於紙上，或

在電腦，實際的DNA分子可能在實驗室中合成，以在序列精確對應於該設計序列。此種合成的核酸分子可以選殖並另外操作，宛若它們係衍自於天然或自然來源的。

[0164] 在一實施例中，一被表現的轉基因係於本主體申請案中揭露的。該基因表現卡匣可能包含一報告/可選擇標記基因、性狀基因、或RNAi基因。可選擇標記基因、性狀基因及RNAi基因之例子係進一步於下文提供。在本申請案中揭露的方法為有利的，因為它們提供用於選擇生殖轉形體的方法，該方法不依賴該蛋白質產物的特定功能，或該轉基因之其它功能。

[0165] 賦予對害蟲或疾病抗性之轉基因或編碼序列

(A)植物疾病抗性基因。植物防禦通常係藉由在植物中一疾病抗性基因(R)的產物與在該病原體中一相應無毒(Avr)基因之產物間的特異性交互作用而活化。一植物品種可以藉由選殖抗性基因而轉形，以工程化對特定病原體菌株具抗性的植物。此種基因之例子包括，用於抵抗葉黴病菌(*Cladosporium fulvum*)的番茄Cf-9基因(Jones等人於"1994 *Science* 266 : 789")、番茄*Pto*基因，其編碼對細菌性葉斑病(*Pseudomonas syringae* pv. Tomato)抗性之一蛋白質激酶(Martin等人於"1993 *Science* 262 : 1432")，及對楊桃細菌性斑點病(*Pseudomonas syringae*)抗性的阿拉伯芥RSSP2基因(Mindrinos等人於"1994 *Cell* 78 : 1089")。

(B)一種蘇雲金芽孢桿菌(*Bacillus thuringiensis*)蛋白質，其之衍生物，或以其為模組之合成性多勝肽，諸如Bt δ-內

毒素基因之核苷酸序列(Geiser等人於”(1986)Gene 48: 109”，及一植物殺蟲(VIP)基因(參閱，例如Estruch等人於”(1996)Proc. Natl. Acad. Sci. 93: 5389-94”)。再者，編碼δ-內毒素基因之DNA分子可購自American Type Culture Collection(Rockville, Md.)，在ATCC登錄號40098; 67136; 31995；與31998下。

(C)凝集素，諸如數個君子蘭(*Clivia miniata*)甘露糖結合凝集素基因之核苷酸序列(Van Damme等人於”(1994)Plant Molec. Biol. 24: 825”)。

(D)維生素結合蛋白，諸如抗生物素(avidin)及抗生物素同源物，該等做為殺幼蟲劑對抗昆蟲害蟲為有用的。參閱美國專利第5,659,026號。

(E)一酵素抑制劑，例如，蛋白酶抑制劑或澱粉酶抑制劑。此種基因之例子包括水稻半胱氨酸蛋白酶抑制劑(Abe等人於”(1987)J. Biol. Chem. 262: 16793”、菸草蛋白酶抑制劑I(Huub等人於”(1993)Plant Molec. Biol. 21: 985”、及 α -澱粉酶抑制劑(Sumitani等人於”(1993)Biosci. Biotech. Biochem. 57: 1243”)。

(F)一昆蟲-特異性賀爾蒙或費落蒙，諸如，蛻皮類固醇與保幼激素，或其變異物，或基於其之模擬物，或其拮抗劑或協同劑，諸如桿狀病毒表現的經選殖保幼激素酯酶，該者為一保幼激素之去活性劑(Hammock等人於1990年之”Nature 344: 458”。

(G)一昆蟲-特異性胜肽或神經胜肽，該者在表現時，破壞受影響害蟲之生理(J. Biol. Chem. 269: 9)。此種基因之例

子包括一昆蟲利尿激素受器(Regan, 1994)、在太平洋折翅蠬(*Diploptera punctata*)中辨識之異史他汀(allostatin) (Pratt, 1989)、及昆蟲-特異性、麻痺性神經毒素(美國專利第5,266,361號)。

(H)由蛇、黃蜂……天然產生之昆蟲-特異性毒液，諸如蠍子昆蟲毒素胜肽(Pang等人於”(1992)Gene 116 : 165”)。

(I)負責單萜、倍半萜類、類固醇、異羥肟酸、苯丙烷類衍生物或其它具有殺蟲劑活性之非蛋白質分子之過度累積的一酵素。

(J)涉及生物活性分子修飾，包括轉譯後修飾之酵素；舉例而言，醣類分解酵素、蛋白質分解酵素、脂肪分解酵素、核酸酶、環化酶、氨基轉移酶、酯酶、水解酶、去磷酸酶、激酶、磷酸化酶、聚合酶、彈性蛋白酶、幾丁質酶及葡聚醣酶，不論是天然或合成的。此種基因之例子包括胼胝質酶(callase)基因(PCT公開申請案WO 93/02197)、幾丁質酶-編碼序列(舉例而言，該者可在登錄號39637與67152下自 ATCC 獲得)、煙草鉤蟲幾丁質酶(Kramer 等人於”(1993)Insect Biochem.Molec.Biol.23 : 691”、及荷蘭芹ubi4-2 聚泛素基因(Kawalleck 等人於”(1993)Plant Molec.Biol.21 : 673”)。

(K)一刺激訊息傳遞的分子。此種分子之例子包括綠豆鈣調蛋白cDNA選殖體的核苷酸序列(Botella等人於1994年之”Plant Molec.Biol.24 : 757”，及玉米鈣調蛋白cDNA選殖體核苷酸序列(Griess等人於”(1994)Plant Physiol.104 :

1467")。

(L)一疏水力矩(hydrophobic moment)胜肽，參閱美國專利第5,659,026號及第5,607,914號；後者教示賦予疾病抗性之合成抗微生物胜肽。

(M)一細胞膜穿透酶，一種通道形成物或通道阻斷物，諸如給予轉基因煙草植物抗青枯假單胞菌(*Pseudomonas solanacearum*)的天蠶蛾素-β(cecropin-β)裂解胜肽類似物(Jaynes等人於"(1993)Plant Sci 89 : 43")。

(N)病毒-侵入蛋白質或衍自其之複合毒素。舉例而言，病毒外殼蛋白在轉形植物細胞中的積累傳授抵抗病毒感染及/或由該外殼蛋白基因所衍自之病毒，以及由相關病毒，所推導的疾病發展。外殼蛋白介導的抗性已經賦予轉形植物抗性對抗苜蓿嵌紋病毒、黃瓜嵌紋病毒、煙草條紋病毒、馬鈴薯X群病毒、馬鈴薯Y群病毒、煙草蝕刻病毒、煙草脆裂病毒及煙草嵌紋病毒。參閱，舉例而言，Beachy等人於1990年之"Ann. Rev. Phytopathol. 28 : 451"。

(O)一昆蟲特異性抗體或衍自其的免疫毒素。從而，標定在昆蟲腸道中一重要代謝功能之抗體將去活化一受影響的酵素，殺死該昆蟲。舉例而言，Taylor等人於1994年摘要#497，第七屆國際分子植物微生物交互作用研討會顯示了在轉基因煙草中經由單鏈抗體片段的生產之酵素去活化。

(P)一病毒特異性抗體。參閱，舉例而言，Tavladoraki等人於"(1993)Nature 366 : 469"，該者顯示的是，表現重組抗體基因之轉基因植物係免於病毒攻擊的。

(Q)由病原體或寄生蟲天然製造之發育阻滯蛋白質。從而，藉由溶解植物細胞壁之同- α -1,4-D-半乳糖醛酸酶，真菌內 α -1,4-D-聚半乳糖醛酸酶促進真菌聚落形成與植物養分釋放(Lamb等人“(1992)Bio/Technology 10：1436”)；選殖與表徵編碼豆類內聚半乳糖醛酸酶抑制蛋白之基因係由Toubart等人描述於“(1992 Plant J. 2：367)”。

(R)由植物天然製造之發育阻滯蛋白質，諸如對真菌疾病提供一提高抗性的荷蘭芹核糖體-去活化基因(Logemann等人於1992年之”Bio/Technology 10：305”)。

(S)RNA干擾，其中RNA分子係用以抑制一標靶基因的表現。在一例子中，RNA分子係部分或全部雙股的，該者觸發一靜默回應，引致dsRNA切割成小型干擾RNA，該者然後係併入至摧毀同源mRNA的靶定複合體。參閱，例如Fire等人於美國專利第6,506,559號；Graham等人於第6,573,099號。

[0166] 對一除草劑賦予抗性之基因

(A)基因其編碼對抑制生長點或分生組織之除草劑一抗性或耐受性者，諸如咪唑啉酮(imidazalinone)、磺醯苯胺或磺醯脲除草劑。在此一類中的示範性基因編碼一突變ALS酵素(Lee等人於1988年之”EMBOJ. 7：1241”)，該者亦被稱為AHAS酵素(Miki等人於1990年之”Theor. Appl. Genet. 80：449”)。

(B)一或多個編碼對草甘膦抗性或耐性的額外基因，該抗性或耐受性係給予的，藉由突變的EPSP合成酶及aroA基

因，或透過諸如GAT(草甘膦乙醯轉移酶)或GOX(草甘膦氧化酶)基因，及其它諸如固殺草(glufosinate)(pat及bar基因；DSM-2)、及芳氧基苯氧基丙酸與環己二酮(ACCase抑制劑編碼基因)膦化合物的代謝去活化。參閱，舉例而言，美國專利第4,940,835號，該者揭露可以賦予草甘膦抗性之一種形式EPSP之核苷酸序列。編碼突變aroA基因之DNA分子可以在ATCC登錄號39256下獲得，且該突變基因之核苷酸序列係揭露於美國專利第4,769,061號。歐洲專利申請案號0 333 033與美國專利第4,975,374號揭露了麴胺酸合成酶基因之核苷酸序列，其賦予對諸如L-草丁膦(L-phosphinothricin)除草劑之抗性。一草丁膦乙醯基轉移酶基因之核苷酸序列係於歐洲申請案號0 242 246中提供的。DeGreef等人於”(1989)Bio/Technology 7 : 61”中描述了表現嵌合bar基因其編碼草丁膦乙醯基轉移酶活性之轉基因植物的生產。賦予對芳氧基苯氧基丙酸與環己二酮，諸如西殺草(sethoxydim)與合氯氟(haloxyfop)，抗性之示範性基因係由Marshall等人於”(1992)Theor.Appl.Genet.83 : 435”中所描述之Acc1-S1、Acc1-S2及Acc1-S3基因。

(C)編碼對抑制光合成之除草劑抗性或耐受性的基因，諸如三嗪(psbA與gs+基因)及莖膚(膚酶基因)。Przibilla等人於”(1991) Plant Cell 3 : 169”中描述使用編碼突變psbA基因的質體轉形衣藻(*Chlamydomonas*)。膚酶基因之核苷酸序列係於美國專利第4,810,648號中揭露，且含有這些基因的DNA分子可在ATCC登錄號53435、67441及67442下獲得。

編碼麴胱甘胺酸S-轉移酶之DNA的選殖及表現係由Hayes等人描述於”(1992) *Biochem. J.* 285 : 173”。

(D)編碼對結合至羥基苯丙酮酸雙加氧酶(HPPD)之除草劑的抗性或耐受性之基因，HPPD酵素催化對-羥基苯丙酮酸(HPP)轉化為尿黑酸(homogentisate)的反應。此包括除草劑諸如異噁唑類(EP418175、EP470856、EP487352、EP527036、EP560482、EP682659、美國專利第5,424,276號)，特別是異噁唑草酮(isoxaflutole)，其係為對玉米的一種選擇性除草劑，二酮基腈類(EP496630、EP496631)，特別是2-氰基-3-環丙基-1-(2-SO₂CH₃-4-CF₃苯基)丙烷-1,3-二酮及2-氰基-3-環丙基-1-(2-SO₂CH₃-4-2,3Cl₂苯基)丙烷-1,3-二酮、三酮類(EP625505、EP625508、美國專利第5,506,195號)，特別是礦草酮(sulcotrione)，及吡唑啉酸類(pyrazolinates)。於植物體內產生過度豐富HPPD的基因可以提供對此種除草劑的耐受性或抗性，包括舉例而言，於美國專利第6,268,549號及6,245,968號及美國專利申請案公開號20030066102中所描述的基因。

(E)編碼對苯氧基生長素除草劑諸如2,4-二氯苯氧基乙酸(2,4-D)抗性或耐受性，且亦可能賦予對芳氧基苯氧基丙酸酯(AOPP)除草劑抗性或耐受性之基因。此等基因之例子包括 α -酮戊二酸依賴性雙加氧酶酵素(aad-1)基因，描述於美國專利第7,838,733號。

(F)編碼對苯氧基生長素除草劑諸如2,4-二氯苯氧基乙酸(2,4-D)抗性或耐受性，且亦可能賦予對吡啶氧基生長素

除草劑，諸如氯氟吡氧乙酸(fluroxypyrr)或三氯比(triclopyr)抗性或耐受性之基因。此等基因之例子包括 α -酮戊二酸-依賴性雙加氧酶酵素基因(aad-12)，描述於WO 2007/053482 A2。

(G)編碼對麥草畏(dicamba)抗性或耐受性之基因(參閱美國專利公開案第20030135879號)。

(H)提供對抑制原卟啉原氧化酶(protoporphyrinogen oxidase)(PPO)之除草劑抗性或耐受性之基因(參閱美國專利第5,767,373號)。

(I)提供對三嗪除草劑(諸如草脫淨(atrazine)及尿素衍生物(諸如敵草隆(diuron))除草劑抗性或耐受性之基因，該等除草劑係結合至光系統II反應中心(PS II)之核心蛋白(參閱Brussian等人於”(1989) EMBO J.1989,8_4 : 1237-1245”)。

[0167] 賦予或貢獻附加價值之性狀之基因

(A)經修飾之脂肪酸代謝，舉例而言，藉由以硬脂醯基-ACP去飽和酶之反義基因轉形玉米或蕓苔屬(Brassica)，以提高植物之硬脂酸含量(Knultzon等人於1992年)之”Proc.Natl.Acad.Sci.U.S.A.89 : 2624”。

(B)降低植酸含量

(1)引入一植酸酶編碼基因，諸如該黑黴菌(*Aspergillus nige*)植酸酶基因(Van Hartingsveldt等人於”1993 Gene 127 : 87”)，增強植酸之降解，加入更多游離磷酸鹽至該轉形植物內。

(2)降低植酸含量之一基因可以引入的。在玉米中，

舉例而言，此可藉由選殖且然後再引入DNA其關聯於負責表徵低含量植酸之玉米突變的單一對偶基因而達成(Raboy等人於”1990 Maydica 35：383”)。

(C)經修飾碳水化合物組成物，舉例而言藉由以編碼變更澱粉分枝圖樣之酵素的一基因轉形植物而作用。此種酵素之例子包括，黏液鏈球菌(*Streptococcus mucus*)果糖基轉移酶基因(Shiroza等人於1988年)之”J.Bacteol.170：810”、枯草桿菌(*Bacillus subtilis*)果聚糖蔗糖酶(*levansucrase*)基因(Steinmetz等人於1985年之”Mol.Gen.Genet.20：220”)、地衣芽胞桿菌(*Bacillus licheniformis*) α -澱粉酶(Pen等人於”1992 Bio/Technology 10：292”)、蕃茄轉化酶基因(Elliott等人於1993)、大麥澱粉酶基因(Sogaard等人於1993年之”J.Biol.Chem.268：22480”、及玉米胚乳澱粉分支酵素II(Fisher等人於”1993 Plant Physiol.102：1045”)。

[0168] 在隨後之實施例中，該轉基因包含一報告基因。在各種實施例中，該報告基因係選自由下列所組成之該群組：*yfp*基因、*gus*基因、*rfp*基因、*gfp*基因、卡那黴素抗性基因、*aad-1*基因、*aad-12*基因、*pat*基因及草甘膦耐受性基因。用於選擇轉形細胞或組織或植物部分或植物的報告或標記基因可能包括在該轉形載體中。可選擇標記之例子包括那些對抗代謝物，諸如除草劑或抗生素賦予抗性者，舉例而言，二氫葉酸還原酶，其賦予對胺甲喋呤之抗性(Reiss於”Plant Physiol. (Life Sci. Adv.) 13：143-149, 1994”；亦參閱Herrera Estrella等人於” Nature 303：209-213, 1983”；

Meijer等人於"Plant Mol. Biol. 16: 807-820, 1991")；新黴素磷酸轉移酶，其賦予對胺基糖苷類新黴素、卡那黴素及帕洛黴素(paromycin)之抗性(Herrera-Estrella於"EMBO J. 2: 987-995, 1983"及Fraley等人於"Proc. Natl. Acad. Sci USA 80: 4803 (1983)")；潮黴素磷酸轉移酶，其賦予對潮黴素之抗性(Marsh於"Gene 32: 481-485, 1984"；亦參閱Waldron等人於"Plant Mol. Biol. 5: 103-108, 1985"；Zhijian等人於"Plant Science 108: 219-227, 1995")；trpB，其允許細胞利用吲哚代替色氨酸；hisD，其允許細胞利用組胺醇代替組胺酸(Hartman於"Proc. Natl. Acad. Sci., USA 85: 8047, 1988")；甘露糖-6-磷酸異構酶，其允許細胞利用甘露糖(WO 94/20627)；鳥胺酸脫羧酶(ornithine decarboxylase)，其賦予對鳥胺酸脫羧酶抑制劑，2-(二氟甲基)-DL-鳥胺酸，之抗性(DFMO；McConlogue於1987年在冷泉港實驗室編輯之"Current Communications in Molecular Biology")；及源自從土麴菌(*Aspergillus terreus*)之脫胺酶，其賦予對殺稻瘟菌素S(Blasticidin S)之抗性(Tamura於"Biosci. Biotechnol. Biochem. 59: 2336-2338, 1995")。

[0169]額外的可選擇標記包括，舉例而言，突變的乙醯乳酸合成酶，其賦予對咪唑啉酮或磺醯脲之抗性(Lee等人於"EMBO J. 7: 1241-1248, 1988")；突變psbA，其賦予對草脫淨之抗性(Smeda等人於"Plant Physiol. 103: 911-917, 1993")；或一突變原卟啉原氧化酶(參閱美國專利第5,767,373號)，或其它賦予對除草劑，諸如固殺草，抗性之

標記。適合的可選擇標記基因之例子包括，但不限於，編碼對下列抗性的基因：氯黴素(Herrera Estrella等人於"EMBO J. 2 : 987-992, 1983")；鏈黴素(Jones等人於" Mol. Gen. Genet. 210 : 86-91, 1987")；觀黴素(Bretagne-Sagnard等人於"Transgenic Res. 5 : 131-137, 1996")；博來黴素(Hille等人於"Plant Mol. Biol. 7 : 171-176, 1990")；礦醯胺(Guerineau等人於"Plant Mol. Biol. 15 : 127-136, 1990")；溴苯腈(bromoxynil)(Stalker等人於"Science 242 : 419-423, 1988")；草甘膦(Shaw等人於"Science 233 : 478-481, 1986")；草丁膦(DeBlock等人於"EMBO J. 6 : 2513-2518, 1987")，及之類。

[0170] 對於一選擇性基因使用之選項係為固殺草抗性編碼DNA，且在一實施例中可以為該草丁膦乙醯基轉移酶(*pat*)，在木薯葉脈嵌紋病毒啓動子控制下的玉米優化*pat*基因或*bar*基因。這些基因賦予對畢拉草(Bialaphos)之抗性。參閱，(見 Wohlleben等人於"(1988) Gene 70 : 25-37")；Gordon-Kamm等人於"Plant Cell 2 : 603 ; 1990"；Uchimiya等人於" BioTechnology 11 : 835, 1993"；White等人於"Nucl. Acids Res. 18 : 1062, 1990"；Spencer等人於"Theor. Appl. Genet. 79 : 625-631, 1990"；及Anzai等人於" Mol. Gen. Gen. 219 : 492, 1989")。*pat*基因之一版本係為於美國專利第6,096,947號中描述之該玉米優化*pat*基因。

[0171] 此外，促進辨識含有編碼該標記之聚核苷酸之植物細胞的標記可能被採用的。可刻痕或可篩選標記為有

用的，其中該序列的存在產生一可測量的產物，且可以生產該產物而不破壞該植物細胞。例子包括 β -葡萄糖醛酸酶或uidA基因(GUS)，其編碼各種顯色受質係知悉的一酵素(舉例而言，美國專利第5,268,463號及第5,599,670號)；氯黴素乙醯轉移酶(Jefferson等人於”The EMBO Journal vol. 6 No. 13 pp. 3901-3907”)；及鹼性磷酸酶。在一較佳實施例中，所使用的標記係為 β -胡蘿蔔素或原維生素A(Ye等人於”Science287 : 303-305-(2000)”)。該基因已被用於增強水稻的營養，但在這種情況下其係採用代替做為一可篩選標記，而連接至感興趣基因之該基因的存在係藉由所提供的金色而偵測。不像該基因被用於其對植物營養貢獻的情況，一較小量的蛋白質足夠用於標記目的。其它可篩選標記包括一般的花青素/類黃酮基因(參閱Taylor及Briggs於”The Plant Cell (1990)2: 115-127”中之討論)，其包括，舉例而言，R-基因座基因，該者編碼調控花青素色素(紅色)在植物組織中之產生的一種產物(Dellaporta等人於Appels與Gustafson編輯之”Chromosome Structure and Function, Kluwer Academic Publishers, pp. 263-282 (1988)”)；控制類黃酮色素生合成之基因，諸如玉米的C1基因(Kao等人於”Plant Cell (1996) 8 : 1171-1179”；Scheffler等人於”Mol. Gen. Genet. (1994) 242 : 40-48”)及玉米C2(Wienand等人於”Mol. Gen. Genet. (1986) 203 : 202-207”)；B基因(Chandler等人於”Plant Cell (1989) 1 : 1175-1183”)、p1基因(Grotewold等人於”Proc. Natl. Acad. Sci USA (1991) 88 : 4587-4591”；Grotewold等人

於”Cell (1994) 76 : 543-553”； Sidorenko等人於”Plant Mol. Biol. (1999) 39 : 11-19”； 青銅基因座基因(Ralston等人於”Genetics (1988) 119 : 185-197”； Nash等人於”Plant Cell (1990) 2(11) : 1039-1049)”，其它等等。

[0172] 適合標記之進一步例子包括靛色螢光蛋白(CYP)基因(Bolte等人於”(2004) J. Cell Science 117 : 943-54”及Kato等人於”(2002) Plant Physiol 129 : 913-42”)、黃色螢光蛋白基因(源自 Evrogen 之 PHIYFP™； 參閱 Bolte 等人於”(2004) J. Cell Science 117 : 943-54”)； lux基因，其編碼螢光素酶，該者之存在可能使用舉例而言，X光片、閃爍計數(scintillation counting)、螢光光度法、低光視訊攝影機、光子計攝相機、多孔冷光儀偵測(Teeri等人於”(1989) EMBO J. 8 : 343”)； 綠色螢光蛋白(GFP)基因(Sheen等人於”Plant J. (1995) 8(5) : 777-84”)； 及DsRed2，其中轉形以該標記基因的植物細胞為紅色的，且因此可視覺選擇的(Dietrich等人於”(2002) Biotechniques 2(2) : 286-293”)。額外例子包括 β -內醯胺酶基因(Sutcliffe於”Proc. Nat'l. Acad. Sci. U.S.A. (1978) 75:3737”)，其編碼各種顯色受質係知悉的一酵素(例如，PADAC，顯色頭孢菌素)； xylE基因(Zukowsky等人於”Proc. Nat'l. Acad. Sci. U.S.A. (1983) 80 : 1101”)，其編碼可以轉換顯色兒茶酚的兒茶酚雙加氧酶； α -澱粉酶基因(Ikuta等人於”Biotech. (1990) 8 : 241”)； 及酪胺酸酶基因(Katz等人於”J. Gen. Microbiol. (1983) 129 : 2703”)，其編碼能夠氧化酪胺酸為DOPA與多巴醌之一酵素，其轉而縮合以

形成容易偵測的化合物黑色素。顯然地，此種標記許多為可獲得的，且對熟習該項技藝者為已知的。

[0173] 在隨後實施例中，該聚核苷酸供體之解析域包含序列辨識編號：142或序列辨識編號：143。序列辨識編號：142或序列辨識編號143之修飾及衍生物被視為在主體揭露內容之視界中。序列辨識編號：142或序列辨識編號143之此種修飾與衍生物可以引致共享特定程度序列一致性的序列。在各種實施例中，對序列辨識編號142具1、2、3、4、5、6、7、8、9或10個Bp修飾的序列係於此揭露的。在其它實施例中，對序列辨識編號143具1、2、3、4、5、6、7、8、9或10個Bp修飾的序列係於此揭露的。

[0174] 在其它實施例中，該通用聚核苷酸供體卡匣之質體域包含質體序列。一質體或載體可以描述或意指為原核載體、穿梭載體、昆蟲載體或真核載體。典型地，質體係為染色體外元件，常常環狀DNA，其係由複製起點及一可選擇標記基因構成。有時，其係可能較佳地具有在大腸桿菌中起作用的一質體(例如，DNA序列分析、插入物構建、獲得大量核酸)。使用於該通用聚核苷酸供體的特定質體可以相對於該意欲用途而選擇(例如，在植物、動物、細菌、真菌及原生動物中表現)。標準的細菌與動物表現質體在該技藝中為已知的，並詳細描述於，舉例而言，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。

[0175] 在其它實施例中，該聚核苷酸供體卡匣之質體

域包含pUC19質體序列。pUC19係為一小的雙股DNA圓環。該質體含有夠細菌性複製的高拷貝數複製起點，且含有一多重選殖位點。參閱Yanisch-Perron, C.、Vieira, J.及Messing, J.於”(1985). *Gene*. 33, 103-119”。其它質體係為已知且在該技藝中為普遍使用的。舉例而言，pUC18、pBR322、pBR325、pBR328、pACYC184、pAT153、pUC118及pUC119係為該技藝中普遍知悉的質體。

[0176] 在其它實施例中，該聚核苷酸供體卡匣之質體域包含一高拷貝數複製起點。高拷貝數複製來點之隨後實施例包括ColE1複製起點。如於此所使用，該術語”複製起點”或”ORI”係意欲含括對複製一質體所必需的核苷酸區域。複製起點之些例子：pBR322之1766-3148核苷酸；ColE1之1667-2892核苷酸；及pACC184之580-1407核苷酸。在一實施例中，該聚核苷酸供體卡匣之質體域可能包含二或多個不同的複製起點。

[0177] 在實施例中，該質體域包含一可選擇標記。質體域可選擇標記之後續實施例係選自由下列所組成之該群組：卡那黴素可選擇標記、安比西林可選擇標記、觀黴素可選擇標記、氯黴素可選擇標記及利福平可選擇標記。適合的可選擇標記係為抗生素抗性基因。其它可選擇標記可以包括可刻痕標記基因，舉例而言 β -葡萄糖醛酸酶(GUS)(Kononov等人於”*Plant J*. 11, 945-957, 1997”)，其可以提供偵測骨架DNA存在之一手段，但不提供針對含有它們之細胞選擇的一手段，且該檢測為組織破壞性的。條件致死的

負向可選擇標記基因亦可以在骨架DNA中使用。其它條件致死基因產物之代表性例子包括：大腸桿菌鳥糞嘌呤磷酸核糖轉移酶，其將硫代黃嘌呤成轉換成有毒的硫代黃嘌呤磷酸(Besnard等人於”Mol. Cell. Biol. 7: 4139-4141, 1987”)；鹼性磷酸酶，其將無活性的磷酸化化合物，諸如絲裂黴素磷酸鹽 (mitomycin phosphate) 與 阿黴素磷酸 (doxorubicin-phosphate) 轉換成有毒的去磷酸化化合物；真菌(例如尖鐮孢菌(*Fusarium oxysporum*))或細菌胞嘧啶脫胺酶(codA)，將5-氟胞嘧啶轉換成有毒的化合物5-氟尿嘧啶 (Mullen於”PNAS 89: 33, 1992”)；羧肽酶G2將從對-N-雙(2-氯乙基)胺基苯甲醯基麩胺酸切開麩胺酸，從而生成一有毒的苯甲酸芥子；及青黴素V醯胺酶，其將轉換阿黴素與黴法蘭(melphalan)之苯氧基解氟靈(phenoxyacetabide)衍生物成為有毒化合物(一般參閱，Vrudhula等人於”J. of Med. Chem. 36(7) : 919-923, 1993”；Kern等人於”Canc. Immun. Immunother. 31(4) : 202-206, 1990”)；及膦單酯水解酶，pehA(美國專利第5,254,801號)。然而，外源性受質必須加入以提供對含有該骨架DNA之細胞為致死的有毒產物。本發明不要求添加額外的受質至該培養基或以對條件致死基因產物為需要的受質外源處理該植物培養細胞。

[0178] 在本揭露內容的另一實施例中，用於標定併入該通用聚核苷酸供體卡匣至一植物細胞基因組內的方法係揭露的。在某些實施例中，包含至少一DNA結合域與至少一核酸酶域之位點特異性DNA結合核酸酶係表現的，其

中該至少一DNA結合域結合至在該植物細胞基因組內的一標靶位點。在其它實施例中，該植物細胞係與一通用聚核苷酸供體卡匣接觸。在進一步實施例中，在該植物細胞基因組內之該標靶位點係以該位點特異性DNA結合核酸酶切開的。在又一實施例中，該通用聚核苷酸供體卡匣係併入至該植物細胞基因組內的標靶位點。

[0179] 在一實施例中，該通用聚核苷酸供體卡匣在該植物細胞基因組內經由一同源性導向修復機制的標定併入係揭露的。在另一實施例中，該通用聚核苷酸供體卡匣在該植物細胞基因組內經由一非同源性末端接合導向修復機制的標定併入係揭露的。

[0180] 於此揭露之該供體分子係經由標定、同源性無關的方法併入至一細胞的基因組內。對於此種標定併入，該基因組係使用一核酸酶在一所欲位置(或位置等)切開的，舉例而言，一DNA結合域(例如，鋅指結合域或TAL效應子域係工程化以在該預決定切割位點處或附近結合一標靶位點)與核酸酶域(例如，切割域或切割半域)之間的融合體。在某些實施例中，該兩個融合蛋白，每一者包含一DNA結合域與一切割半域，係於細胞中表現，並結合至以這樣方式並置的標靶位點，使得一功能性切割域係重建的且DNA係於該標靶位點近處切開。在一實施例中，切割發生在該兩DNA結合域之標靶位點之間。該等DNA結合域之一或二者可以工程化。亦參閱美國專利第7,888,121號；美國專利公開案第20050064474號及國際專利公開案WO 05/084190、

WO 05/014791及WO 03/080809。

[0181] 於此描述之該核酸酶可以以多勝肽及/或聚核苷酸引入。舉例而言，兩聚核苷酸，每一者包含編碼前述該等多勝肽中之一者之序列，可以引入至一細胞內，且當該等多勝肽係表現且每一者結合至其標靶序列時，切割於該標靶序列處或附近發生。或者，包含編碼兩融合多勝肽序列的一單一聚核苷酸係引入至細胞內。聚核苷酸可以為DNA、RNA或DNA及/或RNA之任何經修飾的形式或類似物。

[0182] 繼之在該感興趣區域引入一雙股斷裂，該轉基因係經由非同源依賴性方法(例如，非同源性末端接合(NHEJ))在一標定方式中併入至該感興趣區域內，繼之如於此所描述的該雙股供體分子之線性化之後。該雙股供體較佳地係於體內以一核酸酶線性化，舉例而言一或多個相同或不同的核酸酶係使用以在該基因組中引入雙股斷裂。染色體與供體在該細胞中的同步切割可能會限制供體DNA的降解(相較於該供體分子在引入到細胞內之前線性化而言)。使用於線性化該供體的核酸酶標靶位點較佳地不破壞該(等)轉基因之序列(等)。

[0183] 該轉基因可能在由核酸酶突出端之簡單連接所預期之該方向中併入至該基因組內(命名”正向”或”AB”面向)，或在交替方向(命名”反向”或”BA”面向)。在某些實施例中，該轉基因繼之供體與染色體突出端的準確連接後係併入的。在其它實施例中，無論是在BA或AB取向中併入該

轉基因引致數個核苷酸的刪除。

IV. 用於偵測該通用供體聚核苷酸之檢測

[0184] 各種檢測可以採用以偵測在本揭露內容之某些實施例中所描述的該通用供體聚核苷酸。下列技術在各種情況中為有用的，且在一實施例中，在偵測植物細胞內一核酸分子及/或編碼一轉基因之該多勝肽的存在為有用的。舉例而言，分子的存在可以在各種方式中測定，包括使用該序列之引子或探針、ELISA檢測以偵測該編碼蛋白質、西方墨點法以偵測該蛋白質、或北方墨點法或南方墨點法以偵測RNA或DNA。酵素檢測用於偵測該通用供體聚核苷酸可以採用的。進一步，可以偵測該通用供體聚核苷酸的蛋白質存在的一抗體可以使用該技藝公認的程序生成。額外技術，諸如原位雜交、酵素染色法及免疫染色法，亦可以使用以偵測該重組構建體於特定植物器官與組織中的存在或表現。在該通用供體聚核苷酸中所含有的轉基因可能於植物的某些組織中或在某個發育階段選擇性表現，或者，在該通用供體聚核苷酸中所含有的轉基因可能在實質上全部的植物組織中，實質上連同其整個生命週期中表現的。然而，任何組合的表現模式亦可以適用。

[0185] 南方分析法係為常用的偵測方法，其中DNA係使用限制性核酸內切酶切割，並在瓊脂糖凝膠上分選而藉由分子量分開該DNA，且然後轉移至尼龍膜。該者然後係與以³²P(或其它探針標誌)放射性標誌的探針片段雜交並於SDS溶液中洗滌。

[0186] 同樣地，北方墨點法分析展開相似的實驗計畫，其中RNA係以限制性核酸內切酶切割並在瓊脂糖凝膠上分選而藉由分子量分開該RNA，且然後轉移至尼龍膜。該者然後係與以³²P(或其它探針標誌)放射性標誌的探針片段雜交並於SDS溶液中洗滌。分析從感興趣組織分離的RNA(例如mRNA)可以指出相對表現程度。典型地，假若該mRNA係存在或該mRNA之數量已提高，其可以假定該相對應的轉基因係表現的。北方分析法或其它mRNA分析實驗計畫，可以使用以測定一引入的轉基因或天然基因的表現程度。

[0187] 在西方墨點法分析中，替代經分離DNA/RNA，該感興趣蛋白質係萃取並置於一丙烯醯胺凝膠上。該蛋白質然後係吸漬於膜上並與一標誌物質接觸。參閱，例如Hood等人於“Commercial Production of Avidin from Transgenic Maize; Characterization of Transformants, Production, Processing, Extraction and Purification, *Molecular Breeding* 3: 291-306 (1997)”；Towbin等人於1979年之“Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, Proc Natl Acad Sci USA 76(9): 4350-4354”；Renart等人於“Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure, Proc Natl Acad Sci USA 76(7): 3116-3120”。

[0188] 該通用供體聚核苷酸或其區段，可以使用做為引子用於PCR放大。在執行PCR放大中，引子與模板間一定

程度的錯配可以被可容忍。所以，該示範性引子之突變、刪除及插入(尤其加入核苷酸至5'端)落在本主體揭露內容之發明範圍內。藉由一般技藝人士所知的方法，突變、插入與刪除可以在一給定引子中產生。

[0189] 偵測方法之另一例子係為由Winge所描述之焦磷酸定序 (Pyrosequencing) 技術 (Innov.Pharma.Tech.00 : 18-24, 2000)。在此方法中，一寡核苷酸係設計成重疊該相鄰基因組DNA及插入的DNA接合點。該寡核苷酸係雜交至源自該感興趣區域的單股PCR產物(一引子在該插入的序列中，而一引子在該夾擊的基因組序列中)，且在DNA聚合酶、ATP、硫酸化酶、螢光素酶、三磷酸腺苷雙磷酸酶(apyrase)、腺苷5'磷醯硫酸及螢光素存在下培育。dNTP係個別地加入，且該合併引致一可測量之光訊號。光訊號指出該轉基因插入/夾擊序列之存在，歸因於成功的放大、雜交及單一或多鹼基延伸。(此技術係使用於初始定序，非針對一已知的特定基因偵測)。

[0190] 對於在序列偵測中使用之分子信標(Molecular Beacons)已經描述了。簡言之，一FRET寡核苷酸探針係設計重疊該夾擊基因組及插入DNA接合點。FRET探針之獨特結構引致其含有使螢光及淬滅部分保持極接近之二級結構。該FRET探針及PCR引子(一引子在該插入DNA序列中，而一引子在該夾擊基因組序列中)係於熱安定性聚合酶與dNTP存在下循環。繼之成功的PCR放大，雜交該(等)FRET探針至該標靶序列，引致該探針二級結構之移除，及該螢光與

淬滅部分空間上之分隔。螢光訊號指出該夾擊基因組/轉基因插入序列之存在，歸因於成功的放大與雜交。

[0191] 水解探針檢測，另稱為TAQMAN®(Life Technologies, Foster City, Calif.)，係為一種偵測及量化一DNA序列之存在的方法。簡言之，FRET寡核苷酸探針係設計成一寡核苷酸在該轉基因之內，而一寡核苷酸在該夾擊基因組序列中，用於事件特異性偵測。該FRET探針及PCR引子(一引子在該插入DNA序列中，而一引子在該夾擊基因組序列中)係於熱安定性聚合酶與dNTP存在下循環。雜交該FRET探針引致從該FRET探針上之該淬滅部分切割並釋放該螢光部分。螢光訊號指出該夾擊/轉基因插入序列之存在，歸因於成功的放大與雜交。

[0192] ELISA或酵素聯結免疫檢測自1971年起為已知。一般而言，溶解於緩衝液中的抗原係塗覆在一塑膠表面上。當加入血清時，抗體可以附接至固相上的抗原。當共軛至一酵素時，這些抗體的存在或不存在可以被驗證。加入適當的受質將偵測出可以定量的結合共軛物之數量。常見的ELISA檢測為使用生物素化抗(蛋白質)多株抗體與鹼性磷酸酶共軛物者。舉例而言，用於量化決定蟲漆酶(laccase)水平的ELISA可以為一抗體三明治檢測，其利用市售的多株兔抗體。該抗體係共軛至用於偵測的鹼性磷酸酶。於另一例子中，偵測胰蛋白酶或胰蛋白酶原的ELISA檢測使用生物素化抗胰蛋白酶或抗胰蛋白酶原多株抗體及卵白素-鹼性磷酸酶共軛物。

[0193] 本主體揭露內容之實施例係進一步於下列例子中示範。其應理解的是，這些例子係僅供例示而給予。由上文之實施例及下列的例子，熟習該項技藝者可以確定此揭露內容之主要特徵，且不悖離其精神及發明範圍，可以做出本揭露內容該等實施例之各種改變與修飾，以順應其至各種用途及狀況。因此，除了於此所顯示與描述者外，本揭露內容實施例之各種修飾對熟習該項技藝者從前文說明將為顯然易明的。此等修飾亦意欲落入於隨附之申請專利範圍之發明範圍內。下文係用於例示而非意欲限制本發明之發明範圍而提供。

例子

例子1：設計鋅指以結合玉蜀黍中之基因組基因座

[0194] 對該可標定玉蜀黍基因組基因座之DNA序列導向的鋅指蛋白質係如先前描述般設計的。參閱，例如，Urnov等人於”(2005)Nature 435：646-651”。示範性之標靶序列與識別螺旋係於表1(識別螺旋區域設計)與表2(標定位點)中顯示。在表2中，由該ZFP識別螺旋所接觸的標定位點中的核苷酸係以大寫字母指出，而非接觸的核苷酸係於小寫中指出。鋅指核酸酶(ZFN)標靶位點係針對玉蜀黍中該72個選定的基因組基因座全部設計的。眾多ZFP設計係開發並測試，以辨識那些在酵母代理系統中與在玉蜀黍中辨識並選擇的72種不同的代表性基因組基因座標靶位點伴隨最高效率水準結合的手指。具體的ZFP識別螺旋(表1)，其對該鋅指識別序列伴隨最高效率水準結合，係使用以標定並併入一供

體序列至該玉蜀黍基因組內。

[0195] 表1. 鋅指其針對玉蜀黍選定基因組基因座設計者(N/A表示“不適用”）。

PDAB號	ZFP號	F1	F2	F3	F4	F5	F6
111879	111879 ZFN5	序列辨識 編號：1 QSGDLTR	序列辨識 編號：2 RKDQLV	序列辨識 編號：3 RSDDLTR	序列辨識 編號：4 TSSNRKT	序列辨識 編號：5 RSDTLSE	序列辨識 編號：6 ARSTRTN
	111879 ZFN7	序列辨識 編號：7 RSDSLTV	序列辨識 編號：8 DRSNRK	序列辨識 編號：9 QSSHLTR	序列辨識 編號：10 RSDALA	序列辨識 編號：11 RSDDLTR	序列辨識 編號：12 DPSALRK
111885	111885 ZFN1	序列辨識 編號：13 RSDNLSQ	序列辨識 編號：14 ASNDRK	序列辨識 編號：15 ERGTLAR	序列辨識 編號：16 RSDHLSR	序列辨識 編號：17 ERGTLAR	序列辨識 編號：18 QSGHLSR
	111885 ZFN2	序列辨識 編號：19 RSANLA	序列辨識 編號：20 DRSDLSR	序列辨識 編號：21 RSDTLSQ	序列辨識 編號：22 RSADLSR	序列辨識 編號：23 DRSNLSR	序列辨識 編號：24 NSRNLRN
117404	SIG115737_31v1	序列辨識 編號：25 RSDSLTV	序列辨識 編號：26 DRSHLA	序列辨識 編號：27 DRSNLSR	序列辨識 編號：28 RRSDLKR	序列辨識 編號：29 RSDTLSE	序列辨識 編號：30 QNATRIN
	SIG115737_32v1	序列辨識 編號：31 QSGSLTR	序列辨識 編號：32 QSGDLTR	序列辨識 編號：33 RSDVLSE	序列辨識 編號：34 TRNGLK	N/A	N/A
117408	SIG120523_11v1	序列辨識 編號：35 RSDNLSR	序列辨識 編號：36 DNSNRK	序列辨識 編號：37 QNAHRK	序列辨識 編號：38 QKATRIT	序列辨識 編號：39 DRSHLTR	序列辨識 編號：40 RSDDRKK
	SIG120523_12v1	序列辨識 編號：41 ASKTRTN	序列辨識 編號：42 QSGSLTR	序列辨識 編號：43 LRHHLTR	序列辨識 編號：44 QSAHLK	N/A	N/A
117400	SIG115246_5	序列辨識 編號：45 QSGDLTR	序列辨識 編號：46 ASHNLRT	序列辨識 編號：47 DRSNLTR	序列辨識 編號：48 QSSDLTR	序列辨識 編號：49 DAGNRN	N/A
	SIG115246_6	序列辨識 編號：50 DRSDLSR	序列辨識 編號：51 RSDNLTR	序列辨識 編號：52 DRSHLSR	序列辨識 編號：53 TSGNLTR	序列辨識 編號：54 QSSDLTR	N/A
117402	SIG115636_1v1	序列辨識 編號：55 QSSDLTR	序列辨識 編號：56 HRSTRNR	序列辨識 編號：57 RSDDLTR	序列辨識 編號：58 DRSNLK	序列辨識 編號：59 DRSHLTR	序列辨識 編號：60 QRSTLKS
	SIG115636_2v1	序列辨識 編號：61 RSDALSR	序列辨識 編號：62 RSDDLTR	序列辨識 編號：63 DRSHLTR	序列辨識 編號：64 TSSNRKT	序列辨識 編號：65 RSDTLSE	序列辨識 編號：66 DRSHLAR

117406	SIG120417_11v1	序列辨識 編號：67 DRSARTR	序列辨識 編號：68 QSGHLSR	序列辨識 編號：69 QSGNLA	序列辨識 編號：70 RSDVLST	序列辨識 編號：71 RYAYLTS	序列辨識 編號：72 RRWTLVG
	SIG120417_12v1	序列辨識 編號：73 RSDNLSQ	序列辨識 編號：74 ASNDRK	序列辨識 編號：75 QSGDLTR	序列辨識 編號：76 LKDTLRR	序列辨識 編號：77 QSGNLA	N/A
117411	SIG120621_15v1	序列辨識 編號：78 QSGDLTR	序列辨識 編號：79 MQNYLS	序列辨識 編號：80 RSDHLSE	序列辨識 編號：81 QNANRK	序列辨識 編號：82 RSADLTR	N/A
	SIG120621_16v1	序列辨識 編號：83 RSDNLSE	序列辨識 編號：84 QSANRT	序列辨識 編號：85 RSDALSR	序列辨識 編號：86 DRSALA	序列辨識 編號：87 RSDHLSE	序列辨識 編號：88 DSQNRIK
117413	SIG12078_11v1	序列辨識 編號：89 QSGDLTR	序列辨識 編號：90 DKGNLT	序列辨識 編號：91 RSADLTR	序列辨識 編號：92 DRSHLA	序列辨識 編號：93 RSDTLSE	序列辨識 編號：94 DRSNRKT
	SIG12078_12v1	序列辨識 編號：95 DRSNLSR	序列辨識 編號：96 LRQDLK	序列辨識 編號：97 RSDHLSE	序列辨識 編號：98 DRSALA	序列辨識 編號：99 DRSALSR	序列辨識 編號：100 NRRGRWS
117429	SIG157315_1v1	序列辨識 編號：101 RPTYTLRL	序列辨識 編號：102 HRSSLRR	序列辨識 編號：103 RSDSLLR	序列辨識 編號：104 WLSSLSA	序列辨識 編號：105 QSGDLTR	序列辨識 編號：106 DRSHLAR
	SIG157315_2v1	序列辨識 編號：107 DRSNLSR	序列辨識 編號：108 LKQHLN	序列辨識 編號：109 LRHHLTR	序列辨識 編號：110 QSGNLH	序列辨識 編號：111 TSGHLSR	N/A

[0196] 表2.玉蜀黍選定之基因組基因座的標定位點。

基因座ID	名稱	pDAB編號	ZFP編號及結合位點(5'→3')	標定位點 序列辨識 編號：	最佳基因 組基因座 序列辨識 編號：
最佳基因座 _204637	OGL1	pDAB111879	111879ZFN5： ctACTCCGTATGCGAAGGCAcg	112	210
			111879ZFN7： taTTCGCGGTGGGACACTTGet	113	
最佳_基因座 _204726	OGL2	pDAB111885	111885ZFN1： ccGGAGCCGGGCCTCCAGgc	114	211
			111885ZFN2： atCGCGACGCGACGcGACGAGac	115	
最佳_基因座 _156393	OGL 12	pDAB117404	SIG115737_31v1： TGCATGCGCAGTA	116	212
			SIG115737_32v1： ACACCGGCGCACGGCACG	117	

最佳_基因座 _198387	OGL 15	pDAB117408	SIG120523_11v1： AGAGGTGTAACC	118	213
			SIG120523_12v1： TCGGGCACAAGAAACGAG	119	
最佳_基因座 _31710	OGL 08	pDAB117400	SIG115246_5： TACGCTGACAATGCA	120	214
			SIG115246_6： CCAGCTGATGGAGAGGAC	121	
最佳_基因座 _64542	OGL 11	pDAB117402	SIG115636_1v1： AGAGCAGGCGAG	122	215
			SIG115636_2v1： AGCAAAGTGAGTAGTT	123	
最佳_基因座 _197372	OGL14	pDAB117406	SIG120417_11v1： TGGATGGAAGGAATC	124	216
			SIG120417_12v1： GAAGCTACATCCAG	125	
最佳_基因座 _232228	OGL 16	pDAB117411	SIG120621_15v1： TACGCGCAACGGAACGCA	126	217
			SIG120621_16v1： CACCGGTGTCGTGTAACAG	127	
最佳_基因座 _285621	OGL17	pDAB117413	SIG12078_11v1： CCCGGACGACGCCGAG	128	218
			SIG12078_12v1： GACATGGCACGCGCATCGAG	129	
最佳_基因座 _157315	OGL 13	pDAB117429	SIG157315_1v1： GCATGTGTGGTTTG	130	219
			SIG157315_2v1： GGTCAAGGTAGTGAC	131	

[0197] 玉蜀黍代表性基因組基因座鋅指設計係合併至鋅指表現載體其編碼具有至少一個手指有CCHC結構之蛋白質。參閱，美國專利公開案2008/0182332。特別的，在每一蛋白質中，最後的手指具有一CCHC骨架用於識別螺旋。該非正準鋅指編碼序列係融合至IIS型限制酵素Fok I 之核酸酶域(Wah等人於"(1998) Proc. Natl. Acad. Sci. USA 95 : 10564-10569"該序列的384-579胺基酸)，經由一4胺基酸的

連接子及衍自於玉蜀黍之一不透明-2(opaque-2)細胞核定位訊號，以形成鋅指核酸酶(ZFN)。參閱，美國專利第7,888,121號。針對各種功能域的鋅指係選擇的，用於體內使用。在眾多設計、生產並測試以結合至該推定基因組標靶位點的ZFN中，在上文表2中描述之該等ZFN係辨識為具有體內活性，並表徵為能夠有效地結合並切割植物界中獨特的玉蜀黍基因組聚核苷酸標靶位點。

[0198] ZFN構建體組裝

[0199] 含有ZFN基因表現構建體的質體載體，該等係如先前描述般辨識的，係使用該技藝中普遍知悉的技能與技術設計並完成的(參閱，舉例而言Ausubel或Maniatis)。每一ZFN編碼序列係融合至編碼不透明-2細胞核定位訊號之一序列(Maddaloni等人於”(1989) *Nuc. Acids Res.* 17(18): 7532”)，其位於該鋅指核酸酶上游。該非正準鋅指編碼序列係融合至IIS型限制酵素Fok I 之核酸酶域(Wah等人於”(1998) *Proc. Natl. Acad. Sci. USA* 95 : 10564-10569”該序列的384-579胺基酸)。該融合蛋白質之表現係藉由源自玉蜀黍泛素基因的強大組成型啓動子驅動的(其包括該5'非轉譯區域(UTR)(Toki等人於”(1992) *Plant Physiology* 100 : 1503-07”)。該表現卡匣亦包括源自玉蜀黍過氧化酶5基因(Per5)之該3'UTR(包含轉錄終止及多聚腺苷酸化位點)(美國專利公開案第2004/0158887號)。編碼源自刺蛾(*Thosea asinna*)病毒該核苷酸序列的自水解2A(Szymczak等人於”(2004) *Nat Biotechnol.* 22 : 760-760”)係加入至選殖到該構建體內的兩

個鋅指核酸酶融合蛋白之間。

[0200] 該等質體載體係使用IN-FUSION™ Advantage Technology(Clontech,Mountain View,CA)組裝。限制性核酸內切酶係自新英格蘭生物實驗室(麻州易普斯維奇)獲得，而T4 DNA連接酶(Invitrogen)係使用於DNA連接。質體製備係使用NUCLEOSPIN®質體套組(Macherey-Nagel公司，賓州伯利恆)或Plasmid Midi Kit(Qiagen)，並遵照供應商之說明執行。在瓊脂三(羥甲)胺甲烷-醋酸鹽凝膠電泳後，DNA片段係使用QIAquick Gel Extraction Kit™(Qiagen)分離。藉由迷你製備DNA之限制消化作用，所有連接反應的選殖株係初步篩選。經選定選殖株的質體DNA係由一商業定序廠商(Eurofins MWG Operon，阿拉巴馬州亨茨維爾)定序。定序數據係使用SEQUENCHER™軟體(Gene Codes公司，密西根州安娜堡)組裝並分析。

[0201] 經由限制酵素消化並經由DNA定序，質體係構建並確認。

[0202] 通用供體構建體組裝

[0203] 為了支持快速測試大量的標靶基因座，一新穎、彈性的通用供體系統序列係設計並構建。該通用供體聚核苷酸序列與高通量載體構建方法及分析為相容的。該通用供體系統係由至少三個模組化域：一非可變的ZFN結合域、一解析及使用者界定的性質域，及一簡單的質體骨架用於載體量產。該非可變的通用供體聚核苷酸序列對所有供體為共同的，且允許一有限檢測組的設計，其可以橫跨所有

的玉蜀黍標定位點使用，從而在標定評估中提供統一性並降低分析週期時間。這些域之模組化天性允許用於高通量供體組裝。此外，該通用供體聚核苷酸序列具有其它旨在簡化下游分析並增強結果解釋的獨特性質。它含有一不對稱的限制性位點序列，允許PCR產物消化成診斷的預測大小。序列其包含預期在PCR放大中將有問題的二級結構者係移除。該通用供體聚核苷酸序列大小係小的(小於3.0 Kb)。最後，該通用供體聚核苷酸序列係建立於該高拷貝pUC19骨架上，其允許大量的測試DNA在適時的方式中被擴大。

[0204] 做為一實施例，包含一通用聚核苷酸供體卡匣序列的一例子質體係以序列辨識編號：132及圖1提供的。在一額外實施例中，一聚核苷酸供體卡匣序列係提供為：pDAB111846，序列辨識編號：133，圖2；pDAB117415，序列辨識編號：134，圖3；pDAB117416，序列辨識編號：135，圖4；pDAB117417，序列辨識編號：136，圖5；pDAB117419，序列辨識編號：137，圖6；pDAB117434序列辨識編號：138，圖7；pDAB117418，序列辨識編號：139，圖8；pDAB117420，序列辨識編號：140，圖9；及pDAB117421，序列辨識編號：141，圖10。在另一實施例中，包含該通用供體聚核苷酸序列伴隨功能性表現編碼序列或非功能性(無啓動子的)表現編碼序列的額外序列可以被構建的。構成該通用供體系統的各種域(一非可變的ZFN結合域、一解析及使用者界定性質域、及一簡單的質體骨架)係如上文所描述於表3中註釋用於該構建體。

[0205] 表3通用供體系統載體註釋以辨識該非可變ZFN結合域、分析域及使用者界定性質域及質體骨架。

載體名稱	ZFN結合域	解析域	同源臂區域	質體骨架
pDAB111845	2244-144 Bp	145-254 Bp	---	255-2243 Bp
pDAB111846	2243-143 Bp	144-253 Bp	---	254-2242 Bp
pDAB117415	1961-2069 Bp	2081-2190 Bp	1920-1954 Bp, 2191-2225 Bp	2226-1919 Bp
pDAB117416	51-155 Bp	171-280 Bp	1-35 Bp, 281-315 Bp	316-2234 Bp
pDAB117417	51-86 Bp	102-211 Bp	1-35 Bp, 212-246 Bp	247-2165 Bp
pDAB117419	51-119 Bp	201-310 Bp	1-35 Bp, 311-345 Bp	345-2264 Bp
pDAB117434	1970-2213 Bp	2229-2338 Bp	1920-1954 Bp, 2339-2373 Bp	1-1919 Bp
pDAB117418	51-162 Bp	178-287 Bp	1-35 Bp, 288-322 Bp	323-2241 Bp
pDAB117420	37-116 Bp	132-241 Bp	1-35 Bp, 242-276 Bp	277-2195 Bp
pDAB117421	51-143 Bp	159-268 Bp	1-35 Bp, 269-303 Bp	304-2222 Bp

[0206] 在另一實施例中，該通用供體聚核苷酸序列係爲以一質體遞送之一小的2-3Kb模組化供體系統。此係爲一最小的供體，其包含任何數目的ZFN結合位點、一稱爲”DNA X”或”UZI序列”或”解析域”之短的100-150 bp模板區域(序列辨識編號：142及序列辨識編號：143)，該者帶有限制性位點及用於引子設計(引子係設計爲Tm大於任何計算之二級結構10°C)或編碼序列的DNA序列、及一簡單的質體骨架(圖11)。在一實施例中，該解析域係設計成：含有40至60%百分比的鳥糞嘌呤與胞嘧啶鹼基對；不含有超過9 Bp的重複序列(例如，5'-gtatttcatgtatttcat-3')；不含有一系列大於9 BP的一致鹼基對；及，無二級結構，其中該二級結構

如由Markham, N. R. & Zuker, M.於”(2008) UNAFold : software for nucleic acid folding and hybridization. In Keith, J. M., editor, *Bioinformatics, Volume II. Structure, Function and Applications*, number 453 in *Methods in Molecular Biology*, chapter 1, pages 3–31. Humana Press, Totowa, NJ. ISBN 978-1-60327-428-9”所計算，自由能係小於-18 kcal/mole。參閱表4。該整個質體係透過繼之在適當的ZFN 結合位點DNA雙股斷裂後NHEJ而插入；該ZFN結合位點可以串聯併入的。通用供體聚核苷酸序列之此實施例係最適合用於標定位點與ZFN的快速篩選，且難以放大的序列在該供體中係最小化的。

[0207] 表4.針對ΔG自由能、一致鹼基對9 Bp運行數、超過9 Bp的重複序列數目、及鳥糞嘌呤/胞嘧啶百分比分析解析域組成物。

序列辨識編號：	ΔG自由能	一致鹼基對9 Bp運行數	超過9 Bp的重複序列數目	GC%
序列辨識編號：142	-12.42 kcal/mol	無	無	50.9%
序列辨識編號：143	-12.78 kcal/mol	無	無	47.5%

[0208] 在一進一步實施例中，該通用供體聚核苷酸序列係由至少四個模組構成，且帶有局部的ZFN結合位點、同源臂、具大約100bp解析件或編碼序列任一的DNA X。該通用供體聚核苷酸序列之此實施例係適合用於在各種聚核苷酸標定位點、藉由數個ZFN，訊問NHEJ介導的基因插入。(圖12)。

[0209] 該通用供體聚核苷酸序列可以與所有具界定

DNA結合域的標定分子使用，藉由兩種模式的標定供體插入(NHEJ/ HDR)。因此，當該通用供體聚核苷酸序列係與該適當的ZFN表現構建體共同遞送時，該供體載體與該玉米基因組兩者皆在由特定ZFN結合所指定的一特異性位置切開的。一旦線性化，該供體可以藉由NHEJ或HDR摻入至該基因組內。在載體設計中不同的解析考量然後可以利用以測定最大化標定併入之有效遞送的鋅指。(圖13)。

例子 2：玉蜀黍轉形程序

[0210] 在遞送至玉蜀黍變種HI-II原生質體之前，針對每一ZFN構建體的質體DNA係使用PURE YIELD PLASMID MAXIPREP SYSTEM®(Promega公司，威斯康辛州麥迪遜)或Plasmid Maxi Kit® (Qiagen,，加州瓦倫西亞)，遵照供應商的說明從大腸桿菌製備的。

[0211] 原生質體分離

[0212] 玉蜀黍變種Hi-II懸浮液細胞係於3.5天繼代排程中繼代，4 mL緊縮細胞體積(PCV)的細胞係收集並轉移至含有 20 mL 酵素溶液 (0.6% PECTOLYASE™ 、 6% CELLULASE™ (“Onozuka” R10 ; Yakult Pharmaceuticals, Japan)、4mM MES、0.6M甘露醇、15mM MgCl₂)的50 mL 無菌圓錐形管(Fisher Scientific)。該培養物係加蓋並包裹在PARAFILM™ 中，並放置在一平台式振動器(Thermo Scientific, Vari Mix platform Rocker)上，速度設定為10在室溫下培育達16-18小時，直至原生質體釋放。繼之培育之後，該細胞係用顯微鏡評估消化的品質。該經消化的細胞係通

過100 μm 的細胞濾網過濾，以10 mL的W5培養基[2mM MES (pH5.7), 205 mM NaCl, 167 mM CaCl₂, 6.7mM KCl]清洗，繼之通過70 μm 與40 μm 的細胞濾網過濾。該100 μm 與40 μm 濾網係以10mL的W5培養基清洗。該經過濾的原生質體連同清洗培養基係收集在50ml離心管中，且終體積大約為40 mL。然後，8 mL的”重梯度溶液”[500 mM蔗糖、1mM CaCl₂、5mM MES(pH6.0)]然後係緩慢地加入到該原生質體/酵素的底部，在具旋翼式轉子之一離心機中於300-350 $\times\text{g}$ 離心15分鐘。繼之離心後，約7-8 mL的原生質體帶係移除，以25 mL的W5清洗，並在180-200 $\times\text{g}$ 下離心15分鐘。該原生質體然後係再懸浮於10mL的MMG溶液[4 mM MES(pH 5.7)、0.6 M甘露醇、15 mM MgCl₂]。原生質體係使用血球計數器或流式細胞儀計數，並用MMG稀釋至每毫升167萬。

[0213] 使用PEG轉形玉蜀黍變種Hi-II懸浮培養衍生的原生質體

[0214] 大約50萬的原生質體(在MMG溶液中300 μL)係轉移至2 mL管子，與40 μL 的DNA混合，並在室溫下培育達5-10分鐘。接著，300 μL 新鮮製備的PEG溶液(36%PEG 4000、0.3M甘露糖醇、0.4M CaCl₂)係加入，且該混合物係於室溫下培育達15-20分鐘，伴隨藉由倒置週期性的混合。培育後，1mL的W5洗滌液係緩慢加入，該細胞係溫和地混合，而原生質體係藉由在180-200 $\times\text{g}$ 下離心15分鐘而成丸粒。該丸粒係以1 ml的WI培養基[4mM MES(pH 5.7)、0.6 M甘露醇、20 mM KCl]再懸浮，以鋁箔包裹管子並在室溫下培育過夜達

約16小時。

[0215] ZFN與供體之轉形

[0216] 對於每一選定的基因組基因座，玉蜀黍原生質體係轉染以*yfp*基因表現對照、ZFN單獨、供體單獨、及ZFN與供體在1:10比率(以重量計)之混合物。用於轉染50萬原生質體的DNA總量為80 μ g。所有處理係於三或六重複下進行。該使用的*yfp*基因表現對照為pDAB8393(圖14)，其含有玉蜀黍泛素1啓動子-黃色螢光蛋白編碼序列-玉蜀黍Per5 3'UTR及水稻肌動蛋白1啓動子-*pat*編碼序列-玉蜀黍脂肪酶3'UTR基因表現卡匣。在一典型的標定實驗，單獨4 μ g的ZFN，或與36 μ g的供體係共轉染的，40 μ g的*yfp*報告基因構建體係加入每一處理。納入一貫數量的*yfp*基因表現質體做為填充物允許橫跨多重基因座與重複處理的轉染品質評估。此外，使用一貫數量的*yfp*基因表現質體允許在該供體插入的快速標定分析中任何技術問題的快速故障排除。

例子 3：經由鋅指核酸酶切割玉蜀黍中的基因組基因座

[0217] ZFN轉染的玉蜀黍變種Hi-II原生質體係於轉染後24小時收穫，藉由在2mL Eppendorf™管子中於1600 rpm離心，且該上清液係移除。基因組DNA係使用QIAGEN PLANT DNA EXTRACTION KIT™ (Qiagen, 加州瓦倫西亞)從原生質體丸粒萃取。該分離的DNA係再懸浮在50 μ L水中，而濃度係藉由NANODROP® (Invitrogen, 紐約州格蘭島(Grand Island))測定。DNA之完整性係藉由在0.8%瓊脂糖凝膠電泳上運行樣品而估計。所有樣品均常態化(20-25 ng/ μ L)

用於PCR放大，以產生放大子用於定序(Illumina公司，加州聖地牙哥)。用於放大含括源自處理及對照樣品之每一測試 ZFN識別序列區域的條形碼PCR引子係設計並購自IDT(愛荷華州克拉爾維爾，HPLC純化)。最佳放大條件係藉由梯度PCR辨識的，其使用0.2 μ M適當的條形碼引子、ACCUPRIME PFX SUPERMIXTM(Invitrogen，加州卡斯伯)及100ng的模板基因組DNA在一23.5 μ L反應液裡。循環參數為初始變性於95°C(5分鐘)，繼之35個循環：變性(95°C，15秒)、黏著(55-72°C，30秒)、延伸(68°C，1分鐘)，及一最後延伸(68°C，7分鐘)。放大產物係於3.5%TAE瓊脂糖凝膠上分析，且對每一引子組合，適當的黏著溫度係測定並用於放大如上文所描述源自對照及ZFN處理樣品的放大子。所有的放大子係於3.5%瓊脂糖凝膠上純化，在水中洗脫，且濃度係藉由NanodropTM測定。對於次世代定序(Next Generation Sequencing)，大約100 ng來自ZFN處理及相應的玉米原生質體對照的PCR放大子係匯集在一起，並使用 Illumina次世代定序(NGS)定序的。

[0218] 於每一玉蜀黍選定的基因組基因座，適當的ZFN之切割活性係檢測。含括該等ZFN切割位點的短放大子係從基因組DNA放大，並從ZFN處理及對照的原生質體經送至Illumina NGS。ZFN誘導的切割或DNA雙股斷裂係藉由細胞NHEJ修復途徑，通過於切割位點核苷酸的插入或刪除(插入/刪除)而解決的，且因此在該切割位點插入/刪除的存在係為ZFN活性測量，且係藉由NGS測定。該標靶特異性ZFN

的切割活性係使用NGS分析軟體(專利公開案2012-0173,153,DNA序列之數據分析)，以每百萬高品質序列，具插入刪除序列之數目而估計(圖15)。對玉蜀黍選定的基因組基因座標靶，對照組5-100倍該範圍中的活性係觀察到的，且係進一步藉由序列比對證實，其在每一ZFN切割位點顯示插入刪除的多樣化足跡。此數據表明，玉蜀黍選定的基因組基因座係順適於由ZFN切割。在每一標靶差異的活性係為其染色質狀態與切割的順應性以及每一ZFN表現效率的反射。

例子4：聚核苷酸供體併入的快速標定分析

[0219] 經由鋅指核酸酶定序在玉蜀黍中的基因組基因座

[0220] 該通用供體聚核苷酸序列在玉蜀黍選定的基因組基因座標靶內經由非同源性末端接合(NHEJ)介導供體插入之標定的驗證，係使用一基於原生質體的半高通量快速標定分析方法執行。對每一玉蜀黍選定的基因組基因座標靶，三至六個ZFN設計係測試，且標定係通過藉由次世代定序方法測量ZFN介導的切割(圖15)，而供體插入係透過接合性內-外PCR(圖16)評估。在兩個檢測中均呈陽性的玉蜀黍選定基因組基因座係被辨識為一可標定基因座。

[0221] ZFN供體插入之快速標定分析

[0222] 為了測定一玉蜀黍選定的基因組基因座標靶是否可以被標定用於供體插入，一ZFN構建體與通用供體聚核苷酸構建體係共遞送至玉米原生質體，其在該基因組DNA萃取用於分析之前係培養24小時。假若該表現的ZFN

能夠在該玉蜀黍選定的基因組基因座標靶與該供體兩者中切開該標靶結合位點，該線性化供體然後將經由非同源性末端接合(NHEJ)途徑插入到玉米基因組中該經切開的標靶位點。於玉蜀黍選定基因組基因座標靶之標定併入的確認係基於一“內-外”PCR策略完成，其中一”外”引子識別於該天然基因組基因座的序列，而一“內”引子結合至在該供體DNA的序列。該等引子係在一方式中設計，只有當供體DNA插入到玉蜀黍選定的基因組基因座標靶時，該PCR檢測將產生一預期大小的放大產物。該內-外PCR檢測係於該插入接合之5'-與3'-兩端執行。用於分析併入聚核苷酸供體序列的引子係於表5中提供。

[0223] 使用巢式“內-外”PCR於標靶基因座的ZFN供體插入

[0224] 所有的PCR放大係使用TAKARA EX TAQ HSTM套組(Clonetech，加州山景城)進行。該第一內-外PCR係於20μL的終反應體積中實施，其含有1X TaKaRa Ex Taq HSTM緩衝液、0.2 mM dNTP、0.2μM ”外”引子(表5)、0.05μM ”內”引子(從上述通用供體卡匣設計的)、0.75單位的TAKARA EX TAQ HSTM聚合酶、及10 ng萃取的玉米原生質體DNA。該反應然後係使用PCR程式實施，該程式係由94°C達2分鐘，20個循環的98°C達12秒與68°C達2分鐘，繼之72°C達10分鐘，並保持在4°C。最終PCR產物係於瓊脂糖凝膠運行，連同1KB PLUS DNA LADDERTM(Life Technologies, 紐約州格蘭島)用於觀測。

[0225] 該巢式內-外PCR係於20μL的終反應體積中實施，其含有1X TaKaRa Ex Taq HSTM緩衝液、0.2 mM dNTP、

0.2 μ M "外"引子(表5)、0.1 μ M "內"引子(從上述通用供體卡匣設計的，表6)、0.75單位的TAKARA EX TAQ HS™聚合酶、及1 μ L的該第一PCR產物。該反應然後係使用PCR程式實施，該程式係由94°C達2分鐘，31個循環的98°C達12秒、66°C達30秒與68°C達45秒，繼之72°C達10分鐘，並保持在4°C。最終PCR產物係於瓊脂糖凝膠運行，連同1KB PLUS DNA LADDER™(Life Technologies，紐約州格蘭島)用於觀測。

[0226] 表5.用於巢式內-外PCR分析最佳基因組基因座的所有"外"引子之列表。

OGL1	第一 PCR	5'-端	APL02-5PriF1	序列辨識編號：144 CGCCACAAATCTGAACCAGCA
			Spec-PriR1	序列辨識編號：145 CCACGATCGACATTGATCTGGCTA
	3'-端	APL02-3PriR1	序列辨識編號：146 GCGACATATCAGGCCAACAGG	
			Uzi-PriF1	序列辨識編號：147 GGGATATGTGTCCTACCGTATCAGG
	巢式 PCR	5'-端	APL02-5nstPriF1	序列辨識編號：148 CCAGCATAACAGTTAGGGCCCA
			Spec-nstPriR1	序列辨識編號：149 GTTGCCTTGGTAGGTCCAGC
		3'-端	APL02-3nstPriR1	序列辨識編號：150 CGAAAACTCAGCATGCGGGAA
			Uzi-nstPriF1	序列辨識編號：151 GAGCCATCAGTCCAACACTGC
OGL2	第一 PCR	5'-端	APL01-5PriF1	序列辨識編號：152 ACAGGCGTACAGCAACACCA
		3'-端	APL01-3PriR1	序列辨識編號：153 GACCCTATGGTGTGGATCCCA
	巢式 PCR	5'-端	APL01-5nstPriF1	序列辨識編號：154 CGGGAGCTAGGCAACAAATCG
		3'-端	APL01-3nstPriR1	序列辨識編號：155 TCTGACTAACGGGTGGATGCTG

OGL8	第一 PCR	5'-端	OGL08-5nstPriF2	序列辨識編號：156 CGGATCAGTTGATTGCTCACTTCA
		3'-端	OGL08-3PriR	序列辨識編號：157 GCCGAAAAGCAGCAACTGGAA
	巢式 PCR	5'-端	OGL08-5nstPriF	序列辨識編號：158 GATTGCTACGCAGACCGCCTA
		3'-端	OGL08-3nstPriR	序列辨識編號：159 CACTATTCCCTCCGGCATGCAG
OGL11	第一 PCR	5'-端	OGL11-5PriF	序列辨識編號：160 TGACCTATTGATCGGTCGGCTC
		3'-端	OGL11-3PriR2	序列辨識編號：161 TGCCCTGAATCTCAGGGATGCA
	巢式 PCR	5'-端	OGL11-5nstPriF	序列辨識編號：162 GCCGAAGCTAACTAGCGGACA
		3'-端	OGL11-3nstPriR2	序列辨識編號：163 CATGGAGTAGCAGCTGTGCTG
OGL12	第一 PCR	5'-端	OGL12-5PriF	序列辨識編號：164 GAAAAGCAGTCACCGGCTCTG
		3'-端	OGL12-3PriR	序列辨識編號：165 CCATGGACATGAATTGGCACG
	巢式 PCR	5'-端	OGL12-5nstPriF	序列辨識編號：166 CTTTGCACCAACGGAGCAGAC
		3'-端	OGL12-3nstPriR	序列辨識編號：167 GCTAGCAAAACTTGAAGCTCGCTC
OGL13	第一 PCR	5'-端	OGL13-5PriF	序列辨識編號：168 GAGGTCCCTTACGGGTACCG
		3'-端	OGL13-3PriR	序列辨識編號：169 ACCAGGTCTATCTTGCAGAC
	巢式 PCR	5'-端	OGL13-5nstPriF	序列辨識編號：170 AATAGCGTGGTCGGGTCTAG
		3'-端	OGL13-3nstPriR	序列辨識編號：171 ACGAACGATCCAAGGTGCAGT
OGL14	第一 PCR	5'-端	OGL14-5PriF	序列辨識編號：172 TAGAGACGAGGACTCTGGGCT
		3'-端	OGL14-3PriR	序列辨識編號：173 AAGTCCAACATGGGCACAAACC
	巢式 PCR	5'-端	OGL14-5nstPriF	序列辨識編號：174 CCTCGTTAAGGGTGCAGGTTG
		3'-端	OGL14-3nstPriR	序列辨識編號：175 CCAAGTCAGCTTAAGCCATCAA

OGL15	第一 PCR	5'-端	OGL15-5PriF	序列辨識編號：176 AACCTAGACTTCTGCCCTGGTG
		3'-端	OGL15-3PriR	序列辨識編號：177 GCTCACTTACGAGCAGATCCCA
	巢式 PCR	5'-端	OGL15-5nstPriF	序列辨識編號：178 GGTGCACGCATGTTCTCATGT
		3'-端	OGL15-3nstPriR	序列辨識編號：179 TGTTTACCGCAGCCATGCTTG
OGL16	第一 PCR	5'-端	OGL16-5PriF	序列辨識編號：180 GTTGTATACGGCATCCATCCGCT
		3'-端	OGL16-3PriR	序列辨識編號：181 GAATGAAACTGGTGGTCTGCTCC
	巢式 PCR	5'-端	OGL16-5nstPriF	序列辨識編號：182 CCGACGAGGTACAAGTAGCAGG
		3'-端	OGL16-3nstPriR	序列辨識編號：183 CCCGTAGTCCAGATTCTTGTGGT
OGL17	第一 PCR	5'-端	OGL17-5PriF	序列辨識編號：184 GTCGTTGTTCCGGAAAGGGGAG
		3'-端	OGL17-3PriR	序列辨識編號：185 CGTAGTTGTCGGCATGTCCT
	巢式 PCR	5'-端	OGL17-5nstPriF	序列辨識編號：186 TGTATCCCTTCGGTGAGCACG
		3'-端	OGL17-3nstPriR	序列辨識編號：187 TGAATCGACTCGCTGACAGGTG

[0227] 表6用於巢式內-外PCR分析最佳基因組基因座的所有“內”引子列表。

所有 反應	第一 PCR	5'-端	Spec-PriR1	序列辨識編號：188 CCACGATCGACATTGATCTGGCTA
		3'-端	Uzi-PriF1	序列辨識編號：189 GGGATATGTGTCCTACCGTATCAGG
	巢式 PCR	5'-端	Spec-nstPriR1	序列辨識編號：190 GTTGCCTTGGTAGGTCCAGC
		3'-端	Uzi-nstPriF1	序列辨識編號：191 GAGCCATCAGTCCAACACTGC

[0228] 表7.用於ZFN切割活性的引子

OGL 1	對照/ZFN 111879	序列辨識編號：192 TGGCACTAATCTCACCGGCT
		序列辨識編號：193 AGTCTTAGAAGTACGCTACCGT
OGL 2	對照/ZFN 111885	序列辨識編號：194 TACTTGGCTTCGGCGGCGA
		序列辨識編號：195 GGGTGACTTTACGCGTCTCG
OGL 11	對照/ZFN 117402	序列辨識編號：196 GGTCACGACGCATGGCCTAA
		序列辨識編號：197 AGGATGCATGGATCACCGTC
OGL 12	對照/ZFN 117404	序列辨識編號：198 GCTCTGTTGTCAGCCGTAC
		序列辨識編號：199 CGTTGCAGATAACCACAGTGTAC
OGL 13	對照/ZFN 117429	序列辨識編號：200 GCTAGTAGCTGTTACACGGCGTCT
		序列辨識編號：201 AGGTCGAGACAACCAAGTAGAG
OGL 14	對照/ZFN 117406	序列辨識編號：202 ACAGGACATCGAGCTTGCAT
		序列辨識編號：203 CAGAAGAAAGGCATCAACTCATG
OGL 15	對照/ZFN 117408	序列辨識編號：204 CTCTTCACCTCTACTTTACTTCAG
		序列辨識編號：205 ATTGAACCGTTGTCAAAGCCA
OGL 16	對照/ZFN 117411	序列辨識編號：206 CACAGCGTCAGGGCGGTAAC
		序列辨識編號：207 GGCACGCACCTGTCACTGAC
OGL 17	對照/ZFN 117413	序列辨識編號：208 GTACGCGCCCGGAACTCCT
		序列辨識編號：209 CCTGCGGCCACGTGCATCT

[0229] 在原生質體中標定系統中，該內-外PCR檢測之開展係特別具挑戰性的，由於大量的質體的DNA係使用於轉染，而大量的質體DNA依舊在該原生質體標定系統中，並隨後跟著細胞基因組DNA萃取。殘餘的質體DNA可能稀釋為基因組DNA之相對濃度，並降低偵測的總靈敏度，且亦可以為非特異性、異常PCR反應之一顯著原因。該ZFN誘導、基於NHEJ的供體插入典型地在正向或反向方向發生。對正向插入，DNA之內-外PCR分析常常展現假的陽性條帶，

這可能是由於在該標靶與供體中 ZFN 結合位點周圍的同源性共享區域，其可能引致未併入供體 DNA 在放大過程中的引動與延伸。假陽性在針對反向插入產物探查的分析中係未看到的，因此，所有的標定供體併入分析係實施以在該快速標定分析中訊問反向供體插入。為了進一步提高特異性並降低背景，巢式 PCR 策略亦採用。該巢式 PCR 策略使用一第二 PCR 放大反應，其放大在該第一 PCR 反應之第一內放大產物內的較短區域。使用不對稱數量的”內”與”外”引子優化該接合 PCR，進一步用於在一選定的基因位點快速標定分析。

[0230] 該內-外 PCR 分析結果係於一瓊脂糖凝膠上觀測。對於所有的玉蜀黍選定基因組基因座，”ZFN+供體處理”在 5' 與 3' 端產生一接近預期大小的條帶。對照組 ZFN 或供體單獨處理在 PCR 為陰性的，表明該方法係針對於標定位之供體併入特異性地計分。所有處理係於三至六個下重複進行，且在多重重複中該意料 PCR 產物的存在 (≥ 2 於兩端) 係使用以確認標定。透過 NHEJ 的供體插入往往產生強度較低的副產物，該者係由於在該標靶及/或供體 ZFN 位點線性化端加工而產生的。此外，其係觀察的是，不同的 ZFN 引致不同程度的標定併入效率，一些 ZFN 一貫地產生高水平的供體併入，一些 ZFN 產生較不一貫的供體併入水平，而其它 ZFN 引致無併入。總體而言，對於每一試驗的玉蜀黍選定基因組基因座標靶，在該玉蜀黍代表性基因組基因座標靶內由一或多個 ZFN 的標定併入係證明，其證實這些基因座每一

者係可標定的。更進一步，該等玉蜀黍選定基因組基因座標靶每一者係適用於精確的基因轉形。這些玉蜀黍選定的基因組基因座標靶的驗證係重複多次，伴隨每一次類似的結果，從而證實該驗證過程之再現性，其中包括質體設計與構建、原生質體轉形、樣品加工與樣品分析。

[0231] 結論

[0232] 該供體質體及一設計以特異性地切開一玉蜀黍選定基因組基因座標靶的ZFN係轉染至玉米變種之Hi-II原生質體，且細胞係於24小時後收穫。藉由內-外PCR分析從對照組、ZFN處理及以供體處理原生質體的ZFN分離該基因組DNA顯示該通用供體聚核苷酸之標定插入，做為由ZFN之基因組DNA切割的結果(表8)。這些研究顯示，該通用供體聚核苷酸系統可以使用以評估assess於內源性位點之標定，並使用於篩選候選的ZFN。最後，該基於原生質體的快速標定分析與該新穎的通用供體聚核苷酸序列系統提供一改良的系統，用於篩選基因組標靶及ZFN，用於在植物中精確的基因組工程工作。該方法可以擴展，使用引入DNA雙股或單股斷裂的任何核酸酶，以在任何感興趣系統中評估基因組標靶的位點特異性切割與供體插入。

[0233] 表8.併入一通用供體聚核苷酸序列至玉米中經選擇基因組基因座標靶的結果。

名稱	ID	位置	叢集指定	ZFN (pDAB#)	供體 (pDAB#)	可標定基因座 (是/否)
OGL01	最佳_基因座_204637_G1	chr5 : 200298202..200301414	16	111879	111845	是
OGL02	最佳_基因座_204726_G1	chr5 : 200665730..200670667	03	111885	111846	是
OGL08	最佳_基因座_31710	chr1 : 194939396..194943360	23	117400	117415	是
OGL11	最佳_基因座_64542	chr2 : 72203716..72205045	14	117402	117416	是
OGL12	最佳_基因座_156393	chr4 : 154313884..154315253	10	117404	117417	是
OGL15	較佳_基因座_198387	chr5 : 164712378..164713567	25	117408	117419	是
OGL13	最佳_基因座_157315	chr4 : 158710709..158711983	30	117429	117434	是
OGL14	最佳_基因座_197372	chr5 : 158680601..158681681	26	117406	117418	是
OGL16	最佳_基因座_232228	chr6 : 144719567..144723469	28	117411	117420	是
OGL17	最佳_基因座_285621	chr8 : 118321357..118322528	06	117413	117421	是

[0234] 雖然本發明之觀點已於某些實施例描述，它們可以進一步於本揭露內容之精神及發明範圍之內修飾。因此，本申請案係意欲涵蓋使用其一般原理之本發明實施例的任何變體、用途或適應。進一步，此申請案係意欲涵蓋隨著這些實施例所屬之該技藝中知悉且慣常的實踐而落入隨附申請專利範圍之限制內的本揭露內容之此種偏離。

【符號說明】

(無)

序列表

<110> 陶氏農業科學公司

<120> 一種用於基因標定的通用供體系統

<130> 74435

<160> 219

<170> PatentIn 3.5版

<210> 1

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 1

Gln Ser Gly Asp Leu Thr Arg
1 5

<210> 2

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 2

Arg Lys Asp Gln Leu Val Ala
1 5

<210> 3

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 3

Arg Ser Asp Asp Leu Thr Arg
1 5

<210> 4

<211> 7

<212> PRT

<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 4

Thr Ser Ser Asn Arg Lys Thr
1 5

<210> 5
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 5

Arg Ser Asp Thr Leu Ser Glu
1 5

<210> 6
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 6

Ala Arg Ser Thr Arg Thr Asn
1 5

<210> 7
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 7

Arg Ser Asp Ser Leu Ser Val
1 5

<210> 8
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 8

Asp Arg Ser Asn Arg Lys Thr
1 5

<210> 9

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 9

Gln Ser Ser His Leu Thr Arg
1 5

<210> 10

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 10

Arg Ser Asp Ala Leu Ala Arg
1 5

<210> 11

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 11

Arg Ser Asp Asp Leu Thr Arg
1 5

<210> 12

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 12

Asp Pro Ser Ala Leu Arg Lys

1

5

<210> 13
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 13

Arg Ser Asp Asn Leu Ser Gln
 1 5

<210> 14
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 14

Ala Ser Asn Asp Arg Lys Lys
 1 5

<210> 15
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 15

Glu Arg Gly Thr Leu Ala Arg
 1 5

<210> 16
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 16

Arg Ser Asp His Leu Ser Arg
 1 5

<210> 17
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 17

Glu Arg Gly Thr Leu Ala Arg
 1 5

<210> 18
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 18

Gln Ser Gly His Leu Ser Arg
 1 5

<210> 19
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 19

Arg Ser Ala Asn Leu Ala Arg
 1 5

<210> 20
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 20

Asp Arg Ser Asp Leu Ser Arg
 1 5

<210> 21
 <211> 7
 <212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 21

Arg Ser Asp Thr Leu Ser Gln
1 5

<210> 22

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 22

Arg Ser Ala Asp Leu Ser Arg
1 5

<210> 23

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 23

Asp Arg Ser Asn Leu Ser Arg
1 5

<210> 24

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 24

Asn Ser Arg Asn Leu Arg Asn
1 5

<210> 25

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 25

Arg Ser Asp Ser Leu Ser Val
1 5

<210> 26

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 26

Asp Arg Ser His Leu Ala Arg
1 5

<210> 27

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 27

Asp Arg Ser Asn Leu Ser Arg
1 5

<210> 28

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 28

Arg Arg Ser Asp Leu Lys Arg
1 5

<210> 29

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 29

Arg Ser Asp Thr Leu Ser Glu
1 5

<210> 30
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋
<400> 30

Gln Asn Ala Thr Arg Ile Asn
1 5

<210> 31
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋
<400> 31

Gln Ser Gly Ser Leu Thr Arg
1 5

<210> 32
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋
<400> 32

Gln Ser Gly Asp Leu Thr Arg
1 5

<210> 33
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋
<400> 33

Arg Ser Asp Val Leu Ser Glu
1 5

<210> 34
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 34

Thr Arg Asn Gly Leu Lys Tyr
 1 5

<210> 35
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 35

Arg Ser Asp Asn Leu Ser Arg
 1 5

<210> 36
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 36

Asp Asn Ser Asn Arg Lys Thr
 1 5

<210> 37
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 37

Gln Asn Ala His Arg Lys Thr
 1 5

<210> 38

<211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 38

Gln Lys Ala Thr Arg Ile Thr
 1 5

<210> 39
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 39

Asp Arg Ser His Leu Thr Arg
 1 5

<210> 40
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 40

Arg Ser Asp Asp Arg Lys Lys
 1 5

<210> 41
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 41

Ala Ser Lys Thr Arg Thr Asn
 1 5

<210> 42
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 42

Gln Ser Gly Ser Leu Thr Arg
1 5

<210> 43
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 43

Leu Arg His His Leu Thr Arg
1 5

<210> 44
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 44

Gln Ser Ala His Leu Lys Ala
1 5

<210> 45
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 45

Gln Ser Gly Asp Leu Thr Arg
1 5

<210> 46
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 46

Ala Ser His Asn Leu Arg Thr
1 5

<210> 47

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 47

Asp Arg Ser Asn Leu Thr Arg
1 5

<210> 48

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 48

Gln Ser Ser Asp Leu Ser Arg
1 5

<210> 49

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 49

Asp Ala Gly Asn Arg Asn Lys
1 5

<210> 50

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 50

Asp Arg Ser Asp Leu Ser Arg
1 5

<210> 51
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 51

Arg Ser Asp Asn Leu Thr Arg
1 5

<210> 52
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 52

Asp Arg Ser His Leu Ser Arg
1 5

<210> 53
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 53

Thr Ser Gly Asn Leu Thr Arg
1 5

<210> 54
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 54

Gln Ser Ser Asp Leu Ser Arg
1 5

<210> 55
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 55

Gln Ser Ser Asp Leu Ser Arg
 1 5

<210> 56
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 56

His Arg Ser Thr Arg Asn Arg
 1 5

<210> 57
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 57

Arg Ser Asp Asp Leu Thr Arg
 1 5

<210> 58
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 58

Asp Arg Ser Asn Leu Lys Ala
 1 5

<210> 59
 <211> 7

<212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 59

Asp Arg Ser His Leu Thr Arg
 1 5

<210> 60
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 60

Gln Arg Ser Thr Leu Lys Ser
 1 5

<210> 61
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 61

Arg Ser Asp Ala Leu Ser Arg
 1 5

<210> 62
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 62

Arg Ser Asp Asp Leu Thr Arg
 1 5

<210> 63
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 63

Asp Arg Ser His Leu Thr Arg
1 5

<210> 64

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 64

Thr Ser Ser Asn Arg Lys Thr
1 5

<210> 65

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 65

Arg Ser Asp Thr Leu Ser Glu
1 5

<210> 66

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 66

Asp Arg Ser His Leu Ala Arg
1 5

<210> 67

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 67

Asp Arg Ser Ala Arg Thr Arg
1 5

<210> 68

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 68

Gln Ser Gly His Leu Ser Arg
1 5

<210> 69

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 69

Gln Ser Gly Asn Leu Ala Arg
1 5

<210> 70

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 70

Arg Ser Asp Val Leu Ser Thr
1 5

<210> 71

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 71

Arg Tyr Ala Tyr Leu Thr Ser

1

5

<210> 72
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 72

Arg Arg Trp Thr Leu Val Gly
 1 5

<210> 73
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 73

Arg Ser Asp Asn Leu Ser Gln
 1 5

<210> 74
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 74

Ala Ser Asn Asp Arg Lys Lys
 1 5

<210> 75
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 75

Gln Ser Gly Asp Leu Thr Arg
 1 5

<210> 76
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 76

Leu Lys Asp Thr Leu Arg Arg
 1 5

<210> 77
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 77

Gln Ser Gly Asn Leu Ala Arg
 1 5

<210> 78
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 78

Gln Ser Gly Asp Leu Thr Arg
 1 5

<210> 79
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 79

Met Gln Asn Tyr Leu Ser Arg
 1 5

<210> 80
 <211> 7
 <212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 80

Arg Ser Asp His Leu Ser Glu
1 5

<210> 81

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 81

Gln Asn Ala Asn Arg Lys Thr
1 5

<210> 82

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 82

Arg Ser Ala Asp Leu Thr Arg
1 5

<210> 83

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 83

Arg Ser Asp Asn Leu Ser Glu
1 5

<210> 84

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 84

Gln Ser Ala Asn Arg Thr Lys
1 5

<210> 85

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 85

Arg Ser Asp Ala Leu Ser Arg
1 5

<210> 86

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 86

Asp Arg Ser Ala Leu Ala Arg
1 5

<210> 87

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 87

Arg Ser Asp His Leu Ser Glu
1 5

<210> 88

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 88

Asp Ser Gln Asn Arg Ile Lys
1 5

<210> 89
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 89

Gln Ser Gly Asp Leu Thr Arg
1 5

<210> 90
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 90

Asp Lys Gly Asn Leu Thr Lys
1 5

<210> 91
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 91

Arg Ser Ala Asp Leu Thr Arg
1 5

<210> 92
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 92

Asp Arg Ser His Leu Ala Arg
1 5

<210> 93
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 93

Arg Ser Asp Thr Leu Ser Glu
 1 5

<210> 94
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 94

Asp Arg Ser Asn Arg Lys Thr
 1 5

<210> 95
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 95

Asp Arg Ser Asn Leu Ser Arg
 1 5

<210> 96
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 96

Leu Arg Gln Asp Leu Lys Arg
 1 5

<210> 97

5

<211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 97

Arg Ser Asp His Leu Ser Glu
 1 5

<210> 98
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 98

Asp Arg Ser Ala Leu Ala Arg
 1 5

<210> 99
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 99

Asp Arg Ser Ala Leu Ser Arg
 1 5

<210> 100
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
 <223> 鋅指蛋白識別螺旋

<400> 100

Asn Arg Arg Gly Arg Trp Ser
 1 5

<210> 101
 <211> 7
 <212> PRT
 <213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 101

Arg Pro Tyr Thr Leu Arg Leu
1 5

<210> 102

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 102

His Arg Ser Ser Leu Arg Arg
1 5

<210> 103

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 103

Arg Ser Asp Ser Leu Leu Arg
1 5

<210> 104

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 104

Trp Leu Ser Ser Leu Ser Ala
1 5

<210> 105

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 鋅指蛋白識別螺旋

<400> 105

Gln Ser Gly Asp Leu Thr Arg
1 5

<210> 106

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 106

Asp Arg Ser His Leu Ala Arg
1 5

<210> 107

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 107

Asp Arg Ser Asn Leu Ser Arg
1 5

<210> 108

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 108

Leu Lys Gln His Leu Asn Glu
1 5

<210> 109

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 109

Leu Arg His His Leu Thr Arg
1 5

<210> 110

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 110

Gln Ser Gly Asn Leu His Val
1 5

<210> 111

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 鋅指蛋白識別螺旋

<400> 111

Thr Ser Gly His Leu Ser Arg
1 5

<210> 112

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 鋅指蛋白結合序列

<400> 112

ctactccgta tgcgaggca cg

22

<210> 113

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 鋅指蛋白結合序列

<400> 113

tattcgcggt gggacacttg at

22

<210> 114

<211> 22

<212> DNA

6

<213> 人工序列	
<220>	
<223> 鋅指蛋白結合序列	
<400> 114	22
ccggagccgg ggcctccca g c	
<210> 115	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 鋅指蛋白結合序列	
<400> 115	23
atcgcgacgc gacgcgacga g a c	
<210> 116	
<211> 13	
<212> DNA	
<213> 人工序列	
<220>	
<223> 鋅指蛋白結合序列	
<400> 116	13
tgcatgcgca g t a	
<210> 117	
<211> 18	
<212> DNA	
<213> 人工序列	
<220>	
<223> 鋅指蛋白結合序列	
<400> 117	18
acaccggcgc a c g g c a c g	
<210> 118	
<211> 12	
<212> DNA	
<213> 人工序列	
<220>	
<223> 鋅指蛋白結合序列	
<400> 118	12
agaggtgtaa c c	
<210> 119	
<211> 18	

<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 119		18
tcgggcacaa gaaacgag		
<210> 120		
<211> 15		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 120		15
tacgctgaca atgca		
<210> 121		
<211> 18		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 121		18
ccagctgatg gagaggac		
<210> 122		
<211> 12		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 122		12
agagcaggcg ag		
<210> 123		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 123		16
agcaaagtga gtagtt		
<210> 124		5

<211> 15		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 124		
tgatatggaaag gaatc		15
<210> 125		
<211> 15		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 125		
gaagctacat cccag		15
<210> 126		
<211> 18		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 126		
tacgcgcaac ggaacgca		18
<210> 127		
<211> 19		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 127		
caccgggtgtc gtgttaacag		19
<210> 128		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 鋅指蛋白結合序列		
<400> 128		
cccgacgac gccgag		16

<210> 129						
<211> 20						
<212> DNA						
<213> 人工序列						
<220>						
<223> 鋅指蛋白結合序列						
<400> 129						
gacatggcac	gcgcatcgag	20				
<210> 130						
<211> 15						
<212> DNA						
<213> 人工序列						
<220>						
<223> 鋅指蛋白結合序列						
<400> 130						
gcatgtgtgg	tttttg	15				
<210> 131						
<211> 15						
<212> DNA						
<213> 人工序列						
<220>						
<223> 鋅指蛋白結合序列						
<400> 131						
ggtcaaggta	gtgac	15				
<210> 132						
<211> 2378						
<212> DNA						
<213> 人工序列						
<220>						
<223> 通用聚核苷酸供體卡匣序列						
<400> 132						
aagtttgtga	tacatgaacc	aacacgaaga	acccgccagt	cctactacca	ctctccttcg	60
tttacacttag	ctggctggtt	aagaaaactga	tgttgatacc	agccactact	ccgagctagt	120
agtaaaagat	ttttacagcc	aggggagcca	caggttagta	gagtggtagg	agcagtgttg	180
gactgatggc	tcgcttacat	aagcagttct	gtcccatgga	gccaatgtcc	tgatacggt	240
ggacacatata	cccttaatga	atcgccaaac	gcgccgggag	aggcggttg	cgtattggc	300
gctcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	360
tatcagctca	ctcaaaggcg	gtaatacgg	tatccacaga	atcagggat	aacgcaggaa	420

agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	480
cgttttcca	taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	540
ggtggcgaaa	cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	600
tgcgctctcc	tgttccgacc	ctgcccgtta	ccggataacct	gtccgcctt	ctcccttcgg	660
gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcggtc	720
gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	780
gtaactatcg	tctttagtcc	aacccggtaa	gacacgactt	atgccactg	gcagcagcca	840
ctggtaacag	gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtgg	900
ggcctaacta	cggctacact	agaagaacag	tatttgtat	ctgcgctctg	ctgaagccag	960
ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	1020
ctttgatctt	ttctacgggg	tctgacgctc	agtggAACGA	aaactcacgt	taagggattt	1140
tggcatggc	gttcggaacc	gtgctgaccc	gcaagtggca	accccccgtg	cctctgctca	1200
cctttaccgc	ctggcaactg	gcggccacct	gcagggcgat	cgcaccgagc	gcttagtggg	1260
aatttgtacc	ccttatcgaa	ccgggagcac	aggatgacgc	ctaacaattc	attcaagccg	1320
acaccgcctc	gcggcgccgc	ttaattcagg	agttaaacat	catgagggaa	gcggtgatcg	1380
ccgaagtatc	gactcaacta	tcagaggtag	ttggcgtcat	cgagcgccat	ctcgaaccga	1440
cgttgctggc	cgtacatttgc	tacggctccg	cagtggatgg	cggcctgaag	ccacacagtg	1500
atattgattt	gctggttacg	gtgaccgtaa	ggcttgatga	aacaacgcgg	cgagctttga	1560
tcaacgacct	tttgaaact	tcggcttccc	ctggagagag	cgagattctc	cgcgctgttag	1620
aagtaccat	tgttgtgcac	gacgacatca	ttccgtggcg	ttatccagct	aagcgcgaac	1680
tgcaatttgg	agaatggcag	cgcaatgaca	ttcttgagg	tatcttcgag	ccagccacga	1740
tcgacatttga	tctggctatc	ttgctgacaa	aagcaagaga	acatagcggt	gccttggtag	1800
gtccagcggc	ggaggaactc	tttgatccgg	ttccctgaaca	ggatctattt	gaggcgctaa	1860
atgaaacctt	aacgctatgg	aactcgccgc	ccgactgggc	tggcgatgag	cgaaatgtag	1920
tgcttacgtt	gtccccgcatt	tggtacagcg	cagtaaccgg	aaaaatcgcg	ccgaaggatg	1980
tcgctgccga	ctgggcaatg	gagcgcctgc	cggcccagta	tcagcccgtc	atacttgaag	2040
ctaggcaggc	ttatcttgg	caagaagatc	gcttggcctc	gcgcgagat	cagtggaaag	2100
aatttgttca	ctacgtgaaa	ggcgagatca	ccaaggtagt	cgccaaataa	tgtctaacaa	2160
ttcggtcaag	ccgacgcccgc	ttcggcggcgc	ggcttaactc	aagcgtaga	gagctgggaa	2220

agactatgcg cgatgcctg cagctacgtg cttcgata cggagtagtt tattcgcggt	2280
gggacacttg atagaaaggt taagaggaca cgcctaaaca gttgtgaata catacataca	2340
caccgataca cagatgctaa agcactaaag cccctaaa	2378

<210> 133

<211> 2424

<212> DNA

<213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 133

tttgcggat ctggatggc ttttcgctg cgcggcgta cttccctta acttctcgct	60
ctggaaaggagg cagctgcctg cttttgcttg gccgctgacc atgacacctg gccccgcctt	120
gcagcccggtg gaggactctt ggcgagccac aggttagtag agtggtagga gcagtgttgg	180
actgatggct cgcttacata agcagttctg tccatggag ccaatgtcct gatacggtag	240
gacacatatac ccttaatgaa tcggccaacg cgcggggaga ggcgggttgc gtattggcgc	300
ctcttcgcct tcctcgctca ctgactcgct gcgcctcggt gttcggctgc ggcgagcggt	360
atcagctcac tcaaaggcgg taatacggtt atccacagaa tcagggata acgcaggaaa	420
gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgtggc	480
gttttccat aggctccgccc cccctgacga gcatcacaaa aatcgacgct caagtcagag	540
gtggcgaac ccgacaggac tataaagata ccaggcggtt ccccccggaa gctccctcggt	600
gcgctctcct gttccgaccc tgccgcttac cgatcacctg tccgccttgc tcccttcggg	660
aagcgtggcg ctttctcata gtcacgctg tagtatctc agttcggtgt aggtcggtcg	720
ctccaagctg ggctgtgtgc acgaaccccc cggtcagccc gaccgctgctg ctttatccgg	780
taactatcgt cttagtcca acccgtaag acacgactta tcgcccactgg cagcagccac	840
tggtaacagg attagcagag cgaggtatgt aggccgtgct acagagtct tgaagtgggt	900
gcctaactac ggctacacta gaagaacagt atttggtatac tgctcgctgc tgaagccagt	960
taccttcgga aaaagagttt gtagctcttg atccggcaaa caaaccaccg ctggtagcgg	1020
tggttttttt gtttgcaagc agcagattac ggcgcagaaaa aaaggatctc aagaagatcc	1080
tttgcgtttt tctacgggt ctgacgctca gtggaaacgaa aactcacgtt aaggatttt	1140
ggtcatggcg ttccggaccg tgctgacccg caagtggcaa cctcccggtgc ctctgctcac	1200
ctttaccggcc tggcaactgg cggccacctg cagggcgatc gcaccgagcg ctttagtggga	1260

atttgtaccc	cttatacgaa	cgggagcaca	ggatgacgccc	taacaattca	ttcaagccga	1320
caccgcttcg	cggcgccgct	taattcagga	gttaaacatc	atgagggaag	cggtgatcgc	1380
cgaagtatcg	actcaactat	cagaggtagt	tggcgatcgc	gagcgccatc	tcgaaccgac	1440
gttgctggcc	gtacatttgt	acggctccgc	agtggatggc	ggcctgaagc	cacacagtga	1500
tattgatttg	ctgggtacgg	tgaccgtaag	gcttgatgaa	acaacgcggc	gagctttgat	1560
caacgacctt	ttggaaactt	cggcttcccc	tggagagagc	gagattctcc	gcgctgtaga	1620
agtcaaccatt	gttgcacg	acgacatcat	tccgtggcgt	tatccagcta	agcgcgaact	1680
gcaatttggaa	aatggcagc	gcaatgacat	tcttcaggt	atcttcgagc	cagccacgat	1740
cgacatttgc	ctggctatct	tgctgacaaa	agcaagagaa	catagcgttg	ccttggtagg	1800
tccagcggcg	gaggaactct	ttgatccggt	tcctgaacag	gatctatttg	aggcgctaaa	1860
tgaaacacctt	acgctatggaa	actcgccgcc	cgactggcgt	ggcgatgagc	gaaatgttagt	1920
gcttacgttg	tcccgcat	gttacagcgc	agtaaccggc	aaaatcgccgc	cgaaggatgt	1980
cgctgcccac	tgggcaatgg	agcgcctgccc	ggcccagtt	cagccgtca	tacttgaagc	2040
taggcaggct	tatcttggac	aagaagatcg	cttggcctcg	cgcgcagatc	agtttggaaaga	2100
atttgttac	tacgtgaaag	gcgagatcac	caaggtagtc	ggcaaataat	gtctaacaat	2160
tcgttcaaggc	cgacgcccgt	tcgcggcg	gcttaactca	agcgttagag	agctggggaa	2220
gactatgcgc	gatgcctgc	aggtgggtgc	ctgggaggcc	ccggctccgg	gcatcgccac	2280
gcgacgcgcac	gagacgcgt	aaagtcaccc	gtcgccgtct	cgcacgcgc	agggaggcag	2340
gcaggcagcgc	aacaaaatcg	cacgcgcgtc	gtcgactgcc	tggcctggcc	tggtccagct	2400
gaaaaccggcc	ccgggttgcc	tccg				2424

<210> 134

<211> 2225

<212> DNA

<213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 134	gcgttattggg	cgctcttcg	cttcctcgct	cactgactcg	ctgcgcgtcg	tcgttcggct	60
gcggcgagcg	gtatcagctc	actcaaaggc	gttaatacgg	ttatccacag	aatcaggggaa		120
taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc		180
cgcgttgctg	gcgttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg		240
ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg		300

aagctccctc	gtgcgcctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	360
tctcccttcg	ggaagcgtgg	cgcttctca	tagctcacgc	tgttaggtatc	tcagttcggt	420
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	480
cgccttatcc	ggtaactatac	gtcttgagtc	caacccggta	agacacgact	tatgccact	540
ggcagcagcc	actggtaaca	ggatttagcag	agcgaggtat	gtaggcggtg	ctacagagtt	600
cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtatggta	tctgcgtct	660
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	720
cgctggtagc	ggtggtttt	ttgttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	780
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	840
ttaaggatt	ttggtcatgg	cgttcggAAC	cgtgctgacc	cgcaagtggc	aaccccgt	900
gcctctgctc	acctttaccg	cctggcaact	ggcggccacc	tgcagggcga	tcgcaccgag	960
cgcttagtgg	gaatttgcac	cccttatcga	accgggagca	caggatgacg	cctaacaatt	1020
cattcaagcc	gacaccgctt	cgcggcgcgg	cttaattcag	gagttaaaca	tcatgaggga	1080
agcggtgatc	gccgaagtat	cgactcaact	atcagaggtt	gttggcgtca	tcgagcgcac	1140
tctcgaaccg	acgttgctgg	ccgtacattt	gtacggctcc	gcagtggatg	gcggcctgaa	1200
gccacacagt	gatattgatt	tgctggttac	ggtgaccgta	aggcttgcgt	aaacaacgcg	1260
gcgagcttgc	atcaacgacc	ttttggaaac	ttcggcttcc	cctggagaga	gcgagattct	1320
ccgcgcgtgt	gaagtcacca	tgttgtgca	cgacgacatc	attccgtggc	tttatccagc	1380
taagcgcgaa	ctgcaatttgc	gagaatggca	gwgcaatgac	attcttgcag	gtatttcga	1440
gccagccacg	atcgacatttgc	atctggctat	cttgcgtaca	aaagcaagag	aacatagcgt	1500
tgccttggta	ggtccagcgg	cgagggact	ctttgatccg	gttccctgaac	aggatctatt	1560
tgaggcgcata	aatgaaacct	taacgctatg	gaactcgccg	cccgaactggg	ctggcgtatga	1620
gcgaaatgt	gtgcttacgt	tgtccgcata	ttggtagcagc	gcagtaaccg	gcaaaatcgc	1680
gccgaaggat	gtcgcgtcccg	actgggcaat	ggagcgcctg	ccggcccagt	atcagcccgt	1740
catacttggaa	gctaggcagg	cttatcttgg	acaagaagat	cgcttggcct	cgcgccgaga	1800
tcagttggaa	gaatttggtc	actacgtgaa	aggcgagatc	accaaggtag	tcggcaaata	1860
atgtctaaca	attcgtaaca	gccgacgccc	ttcgcggcg	cggttaact	caagcgtag	1920
agagctgggg	aagactatgc	gwgatcgcc	gcagtcgca	gttaccttcc	ccatcatgaa	1980
cttggagacc	gaaagggtgt	gtgcctagtt	accccgaaga	aatgaaacat	ctgcattgtc	2040

agcgtagtagat cccagctgat ggagaggaca taccatgag gagccacagg ttagtagagt	2100
ggtaggagca gtgttggact gatggctgc ttacataagc agttctgtcc catggagcca	2160
atgtcctgat acggtaggac acatatccct taatgaatcg gccaacgcgc ggggagaggc	2220
ggttt	2225

<210> 135

<211> 2234

<212> DNA

<213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 135	
gagagactggg gaagactatg cgcgatcgcc tgcagtgtcg cgaggcaggg agcagccg	60
tgtccgcctg ctgctgctgg aatggcgccg tggccgcgcg cggtgacgtc ggcaaggctg	120
gtctcgccctg ctcttggcaa gcaaagttag tagttacacg ggcgtggctg gagccacagg	180
ttagtagagt ggttaggagca gtgttggact gatggctgc ttacataagc agttctgtcc	240
catggagcca atgtcctgat acggtaggac acatatccct taatgaatcg gccaacgcgc	300
ggggagaggc ggtttgcgta ttggcgctc ttccgcctcc tcgctcactg actcgctgcg	360
ctcggtcggtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc	420
cacagaatca ggggataacg cagggaaagaa catgtgagca aaaggccagc aaaaggccag	480
gaaccgtaaa aaggccgcgt tgctggcggtt ttccatagg ctccgcctcc ctgacgagca	540
tcacaaaaat cgacgctcaa gtcagagggtg gcgaaacccg acaggactat aaagatacca	600
ggcggtttccc cctggaagct ccctcggtcg ctctcgtgtt ccgaccctgc cgcttaccgg	660
atacctgtcc gccttctcc cttcggaag cgtggcgctt tctcatagct cacgctgttag	720
gtatctcagt tcgggttagg tcgttcgctc caagctggc tgtgtgcacg aaccccccgt	780
tcagcccgac cgctgcgcct tatccggtaa ctatcgctt gagtccaacc cggtaaagaca	840
cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgttagg	900
cggtgctaca gagttcttga agtgggtggcc taactacggc tacactagaa gaacagtatt	960
tggtatctgc gctctgctga agccagttac cttcgaaaaa agagttggta gctcttgatc	1020
cgccaaacaa accaccgctg gtagcggtgg ttttttgtt tgcaagcagc agattacgcg	1080
cagaaaaaaaaa ggatctcaag aagatcctt gatctttct acgggtctg acgctcagtg	1140
gaacgaaaac tcacgttaag ggattttgtt catggcggttc ggaaccgtgc tgacccgcaa	1200
gtggcaacct cccgtgcctc tgctcacctt taccgcctgg caactggcgcc acacactgcag	1260

ggcgatcgca ccgagcgctt agtggaaatt tgtacccctt atcgaaccgg gagcacagga	1320
tgacgcctaa caattcattc aagccgacac cgcttcgcgg cgccgcttaa ttcaggagtt	1380
aaacatcatg aggaaagcgg tgatcgccga agtatcgact caactatcatg aggtagttgg	1440
cgtcatcgag cgccatctcg aaccgacgtt gctggccgta catttgcgtc gctccgcagt	1500
ggatggcggc ctgaagccac acagtgatat tgatttgcgtc gttacgggtga ccgttaaggct	1560
tgtgaaaca acgcggcgag ctttgatcaa cgaccctttg gaaacttcgg cttccctgg	1620
agagagcggag attctccgcg ctgtagaagt caccattgtt gtgcacgcg acatcattcc	1680
gtggcgttat ccagctaagc gcgaactgca atttggagaa tggcagcgca atgacattct	1740
tgcaggtatc ttcgagccag ccacgatcgatcattgtc gctatcttgc tgacaaaagc	1800
aagagaacat agcgttgccct tggttaggtcc agcggcggag gaactctttg atccggttcc	1860
tgaacaggat ctatttgagg cgctaaatga aaccttaacg ctatggaact cgccgcccga	1920
ctgggctggc gatgagcgaa atgttagtgc tacgttgcgtc cgcattttgtt acagcgcagt	1980
aaccggcaaa atcgcgccga aggatgtcgc tgccgactgg gcaatggagc gcctgcccggc	2040
ccagtatcag cccgtcatac ttgaagctag gcaggcttat cttggacaag aagatcgctt	2100
ggcctcgccgc gcagatcagt tggaagaatt tggtcactac gtgaaaggcg agatcaccaa	2160
ggtagtcggc aaataatgtc taacaattcg ttcaagccga cgccgcttcg cggcgcggct	2220
taactcaagc gtta	2234

<210> 136

<211> 2165

<212> DNA

<213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 136

gagagctggg gaagactatg cgcgatcgcc tgcagtgttgcgtc gtcagccgta cgtgccgtgc	60
gccgggttat gtctgcgtatgc gcagttacttta ttttctgaag cgagccacag gtttagtagag	120
tggttaggagc agtgttggac tgatggctcg cttacataag cagttctgc ccatggagcc	180
aatgtccctga tacggtagga cacatatccc ttaatgaatc ggccaacgcg cggggagagg	240
cggtttgcgt attgggcgtcttccgcgtcact gactcgctgc gctcggtcg	300
tcggctgcgg cgagcggtat cagctcactc aaaggcggtat atacgggtat ccacagaatc	360
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa	420

aaaggccgcg ttgctggcgt tttccatag gctccgcccc cctgacgagc atcacaaaaa	480
tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc	540
cccttggaa tccctcgtgc gctctcctgt tccgaccctg ccgcattaccg gataacctgtc	600
cgcccttctc ctttcggaa gcgtggcgct ttctcatagc tcacgctgtta ggtatctcag	660
ttcgggttag gtcgttcgt ccaagctggg ctgtgtgcac gaaccccccgttccagccccga	720
ccgctgcgcc ttatccggta actatcgct tgagtccaaac ccggtaagac acgacttatac	780
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgttag gcggtgtac	840
agagttcttg aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg	900
cgctctgctg aagccagttt ctttcggaaa aagagtttgt agctttgtat ccggcaaaaca	960
aaccaccgct ggtagcggtg gttttttgtt ttgcaagcag cagattacgc gcagaaaaaa	1020
aggatctcaa gaagatcctt tgatctttc tacgggtct gacgctcagt ggaacgaaaa	1080
ctcacgttaa gggattttgg tcatggcggtt cggaaccgtg ctgaccgcga agtggcaacc	1140
tcccggtcct ctgctcacct ttaccgcctg gcaactggcg gccacctgca gggcgatcgc	1200
accgagcgct tagtggaaat ttgttaccct tatcgaaccg ggagcacagg atgacgccta	1260
acaattcatt caagccgaca ccgcttcgac ggcggccta attcaggagt taaacatcat	1320
gagggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtatgtg gcgtcatcga	1380
gcccgcgcgc acaccgacgt tgctggccgt acattgtac ggctccgcag tggatggcg	1440
cctgaagcca cacagtgata ttgatttgct ggttacggtg accgttaaggc ttgtatgaaac	1500
aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttccctg gagagagcga	1560
gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgta	1620
tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat	1680
cttcgagcca gccacgatcg acattgatct ggctatctt ctgacaaaag caagagaaca	1740
tagcggtgcc ttggtaggtc cagcggcgga ggaactctt gatccggttc ctgaacagga	1800
tctatttgag ggcgttaatg aaaccttaac gctatggaaac tcggcccccgtt actggctgg	1860
cgatgagcga aatgtatgtc ttacgttgcc cgcattttgg tacagcgcag taaccggcaa	1920
aatcgccgc aaggatgtcg ctgcccactg ggcaatggag cgcctgcgg cccagtatca	1980
gcccgtcata ctgaaagcta ggcaggctta tcttggacaa gaagatcgct tggccctcg	2040
cgccagatcgat tggaaagaat ttgttacta cgtgaaaggc gagatcacca aggtatgtgg	2100
caaataatgt ctaacaattc gttcaagccg acgcccgttc gcggcgcggc ttaactcaag	2160
cgtta	2165

<210> 137
 <211> 2264
 <212> DNA
 <213> 人工序列

<220>
 <223> 通用聚核苷酸供體卡匣序列

<400> 137		
gagagactggg gaagactatg cgcgatcgcc tgcagtaacg atttattttc ctgtttctt	60	
gtgcccggaaa gagagaggtg taacccatcc tctataacag tgtggctttg acaacggttc	120	
aatatgtatg ttggcaaat gtagatttgc gccaattctt ggtcataatc agcgccgaca	180	
aaccggctac ccaaatttgg gagccacagg ttagtagagt ggttaggagca gtgttggact	240	
gatggctcgc ttacataaagc agttctgtcc catggagcca atgtcctgat acgttaggac	300	
acatatccct taatgaatcg gccaacgcgc ggggagaggc ggtttgcgtt ttggcgctc	360	
ttccgcttcc tcgctcactg actcgctgcg ctgggtcggtt cggctgcggc gagcgggtatc	420	
agctcactca aaggcggtaa tacggttatc cacagaatca gggataaagc cagaaagaa	480	
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggtt	540	
ttccatagg ctccgccccct ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg	600	
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaaagct ccctcggtcg	660	
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc ctgggaaag	720	
cgtggcgctt tctcatagct cacgctgttag gtatctcgtt tcgggttagg tcgttcgtc	780	
caagctggc tgggtgcacg aaccccccgt tcagccgac cgctgcgcct tatccgtaa	840	
ctatcgctt gagtccaacc cggttaagaca cgacttatcg ccactggcag cagccactgg	900	
taacaggatt agcagagcga ggtatgttagg cggtgctaca gagttcttga agtgggtggcc	960	
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac	1020	
cttcggaaaa agagttggta gctcttgatc cgccaaacaa accaccgctg gtacgggtgg	1080	
ttttttgtt tgcaaggcgc agattacgcg cagaaaaaaaaa ggatctcaag aagatcctt	1140	
gatctttct acggggctcg acgctcgtt gaaacgaaaac tcacgttaag ggatttgggt	1200	
catggcggttc ggaaccgtgc tgacccgcaa gtggcaacct cccgtgcctc tgctcacctt	1260	
tacccgcgtt caactggcgg ccacctgcag ggcgatgcgca ccgagcgtt agtggaaatt	1320	
tgtacccctt atcgaaccgg gagcacagga tgacgcctaa caattcattc aagccgacac	1380	
cgcttcggcgg cgccgttaa ttcaaggat aaacatcatg agggaaagcgg tgatcgccga	1440	

agtatcgact caactatcg aggtagttgg cgtcatcgag cgccatctcg aaccgacgtt 1500
gctggccgta catttgtacg gctccgcagt ggatggcgcc ctgaagccac acagtgata 1560
tgatttgctg gttacggtga ccgtaaggct tcatgaaaca acgcggcgag ctttgatcaa 1620
cgaccctttg gaaacttcgg cttccctgg agagagcgag attctccgcg ctgtagaagt 1680
caccattgtt gtgcacgacg acatcattcc gtggcgttat ccagctaagc gcgaaactgca 1740
atttggagaa tggcagcgca atgacattct tgcaggtatc ttcgagccag ccacgatcga 1800
cattgatctg gctatcttgc tgacaaaagc aagagaacat agcgttgcct tggtaggtcc 1860
agcggcggag gaactcttg atccggttcc tgaacaggat ctatttgggg cgctaaatga 1920
aaccttaacg ctatggaact cgccgcccga ctgggctggc gatgagcgaa atgttagtgct 1980
tagttgtcc cgcattttgtt acagcgcagt aaccggcaaa atcgccgca aggatgtcgc 2040
tgccgactgg gcaatggagc gcctgcccgc ccagtatcag cccgtcatac ttgaagctag 2100
gcaggcttat cttggacaag aagatcgctt ggcctcgccgc gcagatcagt tggaaagaatt 2160
tgttcactac gtgaaaggcg agatcaccaa ggtagtcggc aaataatgtc taacaattcg 2220
ttcaagccga cgcccgcttcg cggcgccggct taactcaagc gtta 2264

<210> 138
<211> 2373
<212> DNA
<213> 人工序列

<220>
<223> 通用聚核苷酸供體卡匣序列

<400> 138 gcgtattggg cgctttccg cttccctcgct cactgactcg ctgcgctcggtcggttcggct 60
gcggcgagcg gtatcagctc actcaaaggc ggttaatacgg ttatccacag aatcagggga 120
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 180
cgcgttgctg gcgttttcc ataggctccg ccccccgtac gagcatcaca aaaatcgacg 240
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 300
aagctccctc gtgcgtctc ctgttccgac cctgcccgtt accggatacc tgcgtccctt 360
tctccctcg ggaagcgtgg cgctttctca tagctcacgc tgttaggtatc tcagttcggt 420
gtaggtcggt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 480
cgcccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 540
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 600
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgtct 660

gctgaagcca	gttacccitcg	aaaaaaagagt	tggttagctct	tgatccggca	aacaaaccac	720
cgctggtagc	ggtggttttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	780
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggAACG	aaaactcact	840
ttaaggatt	ttggtcatgg	cgttcggAAC	cgtgctgacc	cgcaagtggc	aaccccccgt	900
gcctctgctc	accttttacccg	cctggcaact	ggcgccacc	tgcagggcga	tcgcaccgag	960
cgcttagtgg	gaatttgtac	cccttatcga	accgggagca	caggatgacg	cctaacaatt	1020
cattcaagcc	gacaccgctt	cgcggcgcgg	cttaattcag	gagttaaaca	tcatgaggga	1080
agcggtgatc	gccgaagtat	cgactcaact	atcagaggta	gttggcgtca	tcgagcgcca	1140
tctcgaaccg	acgttgctgg	ccgtacattt	gtacggctcc	gcagtggatg	gcggcctgaa	1200
gccacacagt	gatattgatt	tgctggttac	ggtgaccgta	aggcttcatg	aaacaacgca	1260
gcgagctttg	atcaacgacc	ttttggaaac	ttcggcttcc	cctggagaga	gcgagattct	1320
ccgcgcgtgt	gaagtaccca	ttgttgtca	cgacgacatc	attccgtggc	gttatccagc	1380
taagcgcgaa	ctgcaatttg	gagaatggca	gchgcaatgac	attcttgcag	gtatctcga	1440
gccagccacg	atcgacattt	atctggctat	cttgcgtaca	aaagcaagag	aacatagcgt	1500
tgccttggta	ggtccagcgg	cgaggaaact	cttgcgtccg	gttcctgaaac	aggatctatt	1560
tgaggcgcta	aatgaaacct	taacgctatg	gaactcgccg	cccgactggg	ctggcgatga	1620
gcgaaatgta	gtgcttacgt	tgtccgcatt	ttggtaacgc	gcagtaaccg	gcggaaatcgc	1680
gccgaaggat	gtcgctgccc	actggcaat	ggagcgcctg	ccggcccagt	atcagcccgt	1740
catacttggaa	gctaggcagg	cttacgttgg	acaagaagat	cgcttggcct	cgcgccgaga	1800
tcagttggaa	gaatttggc	actacgtaa	aggcgagatc	accaaggtag	tcggcaaata	1860
atgtctaaca	attcgtaaca	gccgacgccc	cttcgcggcg	cggttaact	caagcgtag	1920
agagctgggg	aagactatgc	gcgatgcct	gcagataagg	aactatatac	aaaaccacac	1980
atgcacacgt	ggtcaaggta	gtgactaatac	tcgcctaata	cacggcgctg	gctgtatgcatt	2040
gcgtgacacg	gtggctagct	agctgttgc	cccgccggc	ccgtgtatgc	agcgctcgcc	2100
gtcttcaggc	ttcagaaccg	ttgtcaagg	acgtatgact	tgagagctga	atcccggtt	2160
cgtgggttgc	atctcagcgt	gtctcgctgt	cggcccggtc	ggcagcggca	gaatttcatt	2220
tcagacttgg	gccacagggtt	atgtatgtgg	taggagcgt	gttggactga	tggctcgctt	2280
acataaggcag	ttctgtccca	tggagccaa	gtcctgatac	ggtaggacac	atatccctta	2340
atgaatcgcc	caacgcgcgg	ggagaggcgg	ttt			2373

<210> 139
 <211> 2241
 <212> DNA
 <213> 人工序列

<220>
 <223> 通用聚核苷酸供體卡匣序列

<400> 139
 gagagctggg gaagactatg cgcgatcgcc tgcaggagct tgcattaact agcaaagtga 60
 ttccttccat ccatgcaaga agctacatcc cagtggtgc ggcaaaagct gtatgaaaag 120
 gttggagact tccatacaac tttgtgtgt cgagtagtagaa accaacaaca caaagtgcag 180
 ccacaggtta gtagagtggt aggagcagtggactgtatggctcgctta cataagcagt 240
 tctgtcccat ggagccaatg tcctgatacg gtaggacaca tatcccttaa tgaatcgccc 300
 aacgcgcggg gagaggcggt tgcgtatggcgccttc cgcttcctcg ctcactgact 360
 cgctgcgcgc ggtcggttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaaatac 420
 gtttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 480
 aggccaggaa ccgtaaaaag gccgcgttgc tggcggttttccataggctc cgccccctg 540
 acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaaccgcaca ggactataaa 600
 gataccaggc gttccccc ggaagctccc tcgtgcgcgc tcctgttccg accctgcccgc 660
 ttaccggata cctgtccgccc ttctccctt cggaaagcgt ggcgccttct catacgctcac 720
 gctgttaggtt tctcagttcg gtgttaggtcg ttgcgtccaa gctggctgt gtgcacgaac 780
 ccccggttca gcccggaccgc tgccgccttat ccggtaacta tcgtcttgc tccaaccgg 840
 taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 900
 atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 960
 cagtatttgg tatctgcgt ctgctgaagc cagttacctt cggaaaaaga gttggtagct 1020
 ctgtatccgg caaacaacc accgctggta gcgggggttt tttgtttgc aagcagcaga 1080
 ttacgcgcag aaaaaaagga tctcaagaag atcccttgcat ctgttctacg gggctgcacg 1140
 ctcagtgaa cggaaactca cgttaaggga tttggcat ggcgttcggaa accgtgctga 1200
 cccgcaagtgc gcaacctccc gtgcctctgc tcaccccttac cgcctggcaa ctggcggcca 1260
 cctgcaggc gatgcacccg agcgcttagt gggaaattgt accccttatc gaaccgggag 1320
 cacaggatga cgcctaacaa ttcatcaag ccgacaccgc ttgcggcgc ggcttaattc 1380
 aggagttaaa catcatgagg gaagcggtga tcgcccgaat atcgactcaa ctatcagagg 1440
 tagttggcgt catcgagcgc catctcgaac cgacgttgct ggccgtacat ttgtacggct 1500

ccgcagtgga tggccgcctg aagccacaca gtgatattga tttgctggtt acggtgaccg	1560
taaggcttga tgaaacaacg cggcgagcctt tgatcaacga ccttttgaa acttcggctt	1620
cccctggaga gagcgagatt ctccgcgtg tagaagtcac cattgttgc cacgacgaca	1680
tcattccgtg gcgttatcca gctaagcgcg aactgcaatt tggagaatgg cagcgcaatg	1740
acattcttgc aggtatcttc gagccagcca cgatcgacat tgatctggct atcttgctga	1800
caaaaagcaag agaacatagc gttgccttgg taggtccagc ggcggaggaa ctctttgatc	1860
cggttccctga acaggatcta tttgaggcgc taaatgaaac cttaacgcta tggaaactcgc	1920
cgcggactg ggctggcgat gagcgaaatg tagtgcttac gttgtccgc atttggtaca	1980
gcfgcagtaac cggcaaaatc ggcgcgaagg atgtcgctgc cgactggca atggagcgcc	2040
tgccggccca gatatcagccc gtcatacttg aagctaggca ggcttatctt ggacaagaag	2100
atcgcttggc ctcgcgcgca gatcagttgg aagaatttgt tcactacgtg aaaggcgaga	2160
tcaccaaggt agtccggcaaa taatgtctaa caattcgttc aagccgacgc cgcttcgcgg	2220
cgcggcttaa ctcaagcgta a	2241

<210> 140

<211> 2195

<212> DNA

<213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 140 gagagctggg gaagactatg cgcgatcgcc tgcagggacc ccgtcgccgt cggcacagcg	60
tcagggcggt aacatgcgtt ccgttgcgcg tacggaccac cgggtgtcgta taacaggaag	120
agctgtcagt ggagccacag gtttagtagag tggtaggagc agtgttggac tgaatggctcg	180
cttacataag cagttctgtc ccatggagcc aatgtcctga tacggtagga cacatatccc	240
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgt ctccgccttc	300
ctcgctcact gactcgctgc gctcggtcgta tcggctgcgg cgagcggtat cagctcactc	360
aaaggcggtt atacggttt ccacagaatc agggataac gcagggaaaga acatgtgagc	420
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcg tttccatag	480
gctccgcccc cctgacgagc atcacaaaaaa tcgacgctca agtcaagaggt ggcgaaaccc	540
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcggtc gctctccgt	600
tccgaccctg ccgcattaccg gatacctgtc cgccttctc ccttcggaa gcgtggcgct	660

ttctcatagc tcacgctgta ggtatctcg	720
ttcggtgtag gtcgttcgct ccaagctggg	
ctgtgtgcac gaaccccccgttca	780
ccgctgcgcc ttatccgta actatcgct	
tgagtccaaac ccggtaagac acgacttatac	840
gccactggca gcagccactg gtaacaggat	
tagcagagcg aggtatgtag gcggtgctac	900
agagttcttg aagtggggc ctaactacgg	
ctacactaga agaacagtat ttggtatctcg	960
cgctctgctg aagccagttt ccttcggaaa	
aagagtttgtt agctcttgat ccggcaaaca	1020
aaccaccgct ggttagcggtg gttttttgt	
ttgcaagcag cagattacgc gcagaaaaaa	1080
aggatctcaa gaagatcctt tgatctttc	
tacgggtct gacgctcagt ggaacgaaaaa	1140
ctcacgttaa gggatttgg tcatggcg	
cggaaccgtg ctgacccgca agtggcaacc	1200
tcccggtcct ctgctcacct ttaccgcctg	
gcaactggcg gccacctgca gggcgatcgc	1260
accgagcgct tagtggaaat ttgtaccct	
tatcgaaccg ggagcacagg atgacgccta	1320
acaattcatt caagccgaca ccgcttcg	
gcgcggctta attcaggagt taaacatcat	1380
gagggaaagcg gtgatcgccg aagtatcgac	
tcaactatca gaggttagttg gcgtcatcg	1440
gcgccatctc gaaccgacgt tgctggccgt	
acattttgtac ggctccgcag tggatggcg	1500
cctgaagcca cacagtata ttgatttgct	
ggttacggtg accgtaaggc ttgatgaaac	1560
aacgcggcga gctttgatca acgacc	
tttttggaaacttcg gcttccctg gagagagcga	1620
gattctccgc gctgtagaag tcaccattgt	
tgtgcacgac gacatcattc cgtggcgta	1680
tccagctaag cgcaactgc aatttggaga	
atggcagcgc aatgacattc ttgcaggtat	1740
cttcgagcca gccacgatcg acattgatct	
ggctatctg ctgacaaaag caagagaaca	1800
tagcgttgcc ttggtaggtc cagcggcg	
ggaactctt gatccggttc ctgaacagga	1860
tctatttgag ggcctaataatg aaaccta	
gctatggaaac tcgcccggc actgggctgg	1920
cgatgagcga aatgttagtgc ttacgttg	
ccgcatttgg tacagcgcag taaccggcaa	1980
aatcgcgcg aaggatgtcg ctgcccactg	
ggcaatggag cgccctggccc cccagtatca	2040
gcccgtcata cttgaagcta ggcaggctt	
tcttgacaa gaagatcgct tggcctcg	2100
cgcaactgcg ttcgatcg tggaaagaat ttgttca	
cgtgaaaggc gagatcacca aggtagtcgg	2160
caaataatgt ctaacaattc gttcaagccg	
acgcccgttc gcggcgccgc ttaactcaag	2195
cgtta	

<210> 141
 <211> 2222
 <212> DNA
 <213> 人工序列

<220>

<223> 通用聚核苷酸供體卡匣序列

<400> 141	
gagagactggg gaagactatg cgcgatcgcc tgcaggctcg tcgctgatca ccagtatcta	60
ctcgtacagt actccatgga tgcgtacgac cccggaaact cctcggcgctc gtccgggctg	120
accgacatgg cacgcgcac gaggatgtag atgcacgtga gccacagggtt agtagagtg	180
taggagcagt gttggactga tggctcgctt acataaggcag ttctgtccca tggagccaat	240
gtcctgatac ggttaggacac atatccctta atgaatcgcc caacgcgcgg ggagaggcgg	300
tttgcgtatt gggcgcttt ccgcgttcctc gctcactgac tcgctgcgct cggcggttcg	360
gctgcggcga gcggtatcag ctcactcaaa ggcggttaata cggttatcca cagaatcagg	420
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa	480
ggccgcgttg ctggcggttt tccataggct ccgcggccct gacgagcatc acaaaaatcg	540
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgttcccc	600
tgaaagctcc ctcgtgcgct ctcctgttcc gaccctgccc cttaccggat acctgtccgc	660
ctttctccct tcggaaagcg tggcgcttc tcatacgctca cgctgttaggt atctcagttc	720
ggtaggttc gttcgctcca agctggctg tgtgcacgaa ccccccgttc agcccgaccg	780
ctgcgcctta tccggtaact atcgtcttga gtccaaacccg gtaagacacg acttatcgcc	840
actggcagca gccactggta acaggattag cagagcgagg tatgttaggcg gtgtacaga	900
gttcttgaag tggtggccta actacggcta cactagaaga acagtatttgcgtatcgcc	960
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgcgtatccg gcaaaacaaac	1020
caccgctggt agcggtggtt ttttgggtt caagcagcag attacgcgca gaaaaaaagg	1080
atctcaagaa gatcccttga tctttctac ggggtctgac gctcagtggaa acgaaaactc	1140
acgttaaggg atttggta tcggcggtt aaccgtgctg acccgcaagt ggcaacactcc	1200
cgtgcctctg ctcacctta ccgcgtggca actggcgcc accgtcaggcgatcgacc	1260
gagcgcttag tggaaatttgcgtatccctt cgaaccggaa gcacaggatg acgcctaaca	1320
attcattcaa gccgacacccg cttcgccggcg cggcttaatt caggagttaa acatcatgag	1380
ggaagcggtt atcgccgaag tatcgactca actatcagag gtatgtggcg tcacgtcg	1440
ccatctcgaa ccgacgttgc tggccgtaca tttgtacggc tccgcagtgg atggccgcct	1500
gaagccacac agtgtatatttgcgtatccctt caggatggaa acatcgatggatgttggag	1560
gcggcgagct ttgtatcgatggaa aacttcggct tccctggag agagcgagat	1620
tctccgcgtt gtatcgatggaa aacttcggct tccctggag agagcgagat	1680

agctaaggcgc	gaactgcaat	ttggagaatg	gcagcgcaat	gacattcttg	caggtatctt	1740
cgagccagcc	acgatcgaca	ttgatctggc	tatcttgcgt	acaaaagcaa	gagaacatag	1800
cgttgccttg	gtaggccag	cggcggagga	actctttat	ccggttccctg	aacaggatct	1860
atttgaggcg	ctaaatgaaa	ccttaacgct	atggaactcg	ccgcccgaact	gggctggcga	1920
tgagcgaaat	gtagtgctt	cgttgtcccg	cattgggtac	agcgcagtaa	ccggcaaaat	1980
cgcgccgaag	gatgtcgctg	ccgactgggc	aatggagcgc	ctgcccggccc	agtatcagcc	2040
cgtcatactt	gaagcttaggc	aggcttatact	tggacaagaa	gatcgcttgg	cctcgccgcgc	2100
agatcagttg	gaagaatttg	ttcactacgt	gaaaggcgag	atcaccaagg	tagtcggcaa	2160
ataatgtctt	acaattcgtt	caagccgacg	ccgcctcgcg	gcgccggcttia	actcaagcgt	2220
ta						2222

<210> 142
 <211> 110
 <212> DNA
 <213> 人工序列

<220>
 <223> 解析域

<400> 142	60				
gagccacagg	ttagtagagt	ggtaggagca	gtgttggact	gatggctcgc	ttacataagg
agttctgtcc	catggagcca	atgtcctgat	acggtaggac	acatatccct	110

<210> 143
 <211> 118
 <212> DNA
 <213> 人工序列

<220>
 <223> 解析域

<400> 143	60				
actagttttc	atagggatat	gtgaggacta	accttggcca	aaggagctgg	aactgcctgc
agttatgtaa	gggccttagt	ccaaattgct	ccaccctctg	ggaagctaat	ggactagt
					118

<210> 144
 <211> 21
 <212> DNA
 <213> 人工序列

<220>
 <223> 引子

<400> 144	21				
cggcacaaat	ctgaaccagc	a			

<210> 145
 <211> 24
 <212> DNA
 <213> 人工序列

<220>
 <223> 引子

<400> 145
 ccacgatcga cattgatctg gcta

24

<210> 146
 <211> 21
 <212> DNA
 <213> 人工序列

<220>
 <223> 引子

<400> 146
 gcgacatatac aggccaaacag g

21

<210> 147
 <211> 25
 <212> DNA
 <213> 人工序列

<220>
 <223> 引子

<400> 147
 gggatatgtg tcctaccgta tcagg

25

<210> 148
 <211> 21
 <212> DNA
 <213> 人工序列

<220>
 <223> 引子

<400> 148
 ccagcataaca gtttagggccc a

21

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

<220>
 <223> 引子

<400> 149

149

gttgccttgg taggtccagc	20
<210> 150	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 150	
cgaaaaactca gcatgcggga a	21
<210> 151	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 151	
gagccatcag tccaaacactg c	21
<210> 152	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 152	
acaggcgtac agcaacacca	20
<210> 153	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 153	
gaccctatgg ttttggatcc ca	22
<210> 154	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	

<400> 154	21
cgggagctag gcaacaaatc g	
<210> 155	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 155	23
tctgactaaa cgggtggatg ctg	
<210> 156	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 156	26
cggatcagg t gattcgctca ctttca	
<210> 157	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 157	21
gccgaaaagc agcaactgga a	
<210> 158	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	
<400> 158	21
gattgctacg cagaccgcct a	
<210> 159	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引子	

<400> 159
cactattcct ccggcatgca g

21

<210> 160
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 160
tgacctatgg atcggtcggc tc

22

<210> 161
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 161
tgccttgaat ctcagggatg ca

22

<210> 162
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 162
gccgaagcta actagcggac a

21

<210> 163
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 163
catggagtag cagctgtgct g

21

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

<220>

<223> 引子

<400> 164
gaaaagcagt caccggctct g

21

<210> 165
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 165
ccatggacat gaattcggca cg

22

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

<220>
<223> 引子

<400> 166
cttttgacc acggagcaga c

21

<210> 167
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 167
gctagcaaaa ctttgaagct cgctc

25

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

<220>
<223> 引子

<400> 168
gagggtccctt acgggtcatc g

21

<210> 169
<211> 22
<212> DNA
<213> 人工序列

§

<220>
<223> 引子

<400> 169
accagggtcta tcttgcgca g ac

22

<210> 170
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 170
aatagcgtgg tcgggtccta g

21

<210> 171
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 171
acgaacgatc caagggtgcag t

21

<210> 172
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 172
tagagacgag gactctgggc t

21

<210> 173
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 173
aagtccaaaca tgggcacaac c

21

<210> 174
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 174
cctcgtaag ggtgcagg t g

21

<210> 175
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 175
ccaagtcagc ttctaagcca tcaaac

26

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

<220>
<223> 引子

<400> 176
aaccctagac ttctgcctgg tg

22

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

<220>
<223> 引子

<400> 177
gctcacttac gagcagatcc ca

22

<210> 178
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 178
ggtgcacgca tggctctcatg t

21

<210> 179
<211> 21
<212> DNA

5

<213> 人工序列

<220>

<223> 引子

<400> 179

tgttaccgc agccatgctt g

21

<210> 180

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 引子

<400> 180

gttgtatacg gcatccatcc gct

23

<210> 181

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 引子

<400> 181

gaatgaaact ggtggctgc tcc

23

<210> 182

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 引子

<400> 182

ccgacgaggt acaagtagca gg

22

<210> 183

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 引子

<400> 183

cccgtagtcc agattcttgt ggt

23

<210> 184

<211> 21

<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 184
gtcgtttgtt cggaagggga g

21

<210> 185
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 185
cgtatgtc cggcatgtcc t

21

<210> 186
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 186
tgtatccctt cggtgagcac g

21

<210> 187
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引子

<400> 187
tgaatcgact cgctgacagg tg

22

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

<220>
<223> 引子

<400> 188
ccacgatcga cattgatctg gcta

24

<210> 189

9

<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 189		
gggatatgtg tcctaccgta tcagg		25
<210> 190		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 190		
tttgccttgg taggtccagc		20
<210> 191		
<211> 21		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 191		
gagccatcag tccaaacactg c		21
<210> 192		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 192		
tggcactaat ctcaccggct		20
<210> 193		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 193		
agtcttagaa gtacgctacc gt		22

<210> 194		
<211> 19		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 194		
tac tggctt cgg cggcga		19
<210> 195		
<211> 21		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 195		
gggtgacttt tacgcgtctc g		21
<210> 196		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 196		
gtcacgacg catggcctaa		20
<210> 197		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 197		
aggatgcatg gatcaccgtc		20
<210> 198		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 198		
gctctgttgt gcagccgtac		20

<210> 199		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 199		
cgttgcagat accacagtgt ac	22	
<210> 200		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 200		
gcttagtagct gtttacacgg cgtct	25	
<210> 201		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 201		
aggtcgagac aaccaagtag ag	22	
<210> 202		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 202		
acaggacatc gagcttgcatt	20	
<210> 203		
<211> 23		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 203		
cagaagaaag gcatcaactc atg	23	

<210> 204		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 204		
ctctttcacc tctactttta cttag	26	
<210> 205		
<211> 21		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 205		
attgaaccgt tgtcaaagcc a	21	
<210> 206		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 206		
cacagcgtca gggcggtaac	20	
<210> 207		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 207		
ggcacgcacc tgtcactgac	20	
<210> 208		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引子		
<400> 208		

gtacgcgccc	gggaactcct	20				
<210>	209					
<211>	20					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	引子					
<400>	209					
cctgcggccc	acgtgcacatct	20				
<210>	210					
<211>	3213					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	最佳基因座_ 204637					
<400>	210					
gaaaaacgaga	gtgggataaaa	caagcgtaaa	aatgaacgg	gaacgagagt	attaagcgga	60
aacttataat	ttaatacaaaa	acacacgtga	tattgatggc	gagctagcag	tacaacaaca	120
tggaacaaac	agatacagag	agacaattaa	tactacattg	ttatttgctt	gtttgtggtg	180
tatatatata	gctacgacca	gcgttgttta	agtctttagg	acatttgctca	ttaaggggag	240
gcaattaata	ctacattgtt	atttgcttgt	tttgttgtta	catatagcta	ggtagaataaa	300
ccttttatgc	tataatttgc	atataataatg	cttagctaa	tgcattaatg	catcttgggc	360
cgatcccgaa	tttctaaaat	atataataaa	taccggttca	acctggctaa	aaacgagtgg	420
gatagaccct	atcatgtttt	agaatttttt	cacaataaca	aaaacagacg	gatcaaataat	480
agaaaaatga	tattggtcgg	gacgtcctct	ttcaacctta	catgcgggtc	ggtcgctcgag	540
ccaattttca	gcttgcggg	tgaagtaaaa	gaagcgaagc	caagtagcaa	gcacgcatgg	600
gcatggccgc	atggccgtg	cggcgtgca	tggcacatgg	cagcgcgcag	cacatcgatg	660
gatcctcggc	tccagccggc	gtcaggcgcac	acagatgtct	tccgtcaaa	cggccgctga	720
aagagacgag	acgaccgcgt	tcttgtggaa	gaaaagaagc	gagatcgagg	ctcggcgggg	780
cgcggcccttgc	ggcagcaggc	aaggcatcggt	tacggatcca	attccatgcgt	tggaaatctc	840
ccggagctag	gccggaggga	agggagccctg	gggagagccg	aagggaaac	agtgcgggtcg	900
gcagcgggtcc	gcggtaggaa	gggtgatgac	aggagcggga	gtaaatgcct	gtagcccgta	960
ggggcacggt	cgtctccgt	caaccggctg	gccgccgcgc	ccgcctcggt	cgccatgcct	1020
cctgcagcag	cagcagcatc	gagcccctgc	atgcattgca	tgccatgcgt	gcagcagcag	1080

cagcaggcca	ggtcaggcca	ggttagcgtag	cgcccgtagc	gcgggctccg	tccgiccatg	1140
tccctgtccc	tggccgcagc	tactaggagt	actatagtgt	gtagtaactgc	tcagcagccg	1200
tgcagtgcgg	tgcgcggcgg	aggggcgggc	gagctggcaa	agtggcaact	ggcttggcta	1260
ggctagctt	ccgttggcgt	ggttggcctt	tgcaggacag	ctggtgacag	tgacggggcg	1320
ggagggggca	tgcttcccta	tatcagccac	agctaccgcc	gtactagccc	ccgcaccta	1380
gcttagcta	gggccagccg	gccgcccgtc	tccggctgc	gttgtggttg	tggtgagtg	1440
gatggagat	gagatggatg	atgccattgt	cacatcacgc	acctgcccgt	gccgcgcgcg	1500
caaggcgcac	ttgcctgccc	tggccctcc	tcccaccggg	ggtctcgct	ggcaccgtca	1560
cttcatgggg	aggggaagca	tggcatggc	tgcatggtgc	gtgcaccacg	cttcattcgt	1620
atatacagct	gttagccggg	atggatcgac	tgtatctc	aactgaagt	atcgataaaa	1680
agcggtgct	tggtatttagc	tggcctcggt	ccgtccggcg	ggacgcgttg	gaagctttt	1740
cgactccggg	gagactttt	tttagactt	gcaatccact	ctcactgatc	cactgcttt	1800
ctgctcggt	gctaaactgc	taactcggca	atatacgatc	actcattgc	ggtgcaactg	1860
cagtgcgcag	cctacagatc	caaagctacg	tactgtactc	taaattcctt	gctgtagtagc	1920
gtgaaacgtc	atgcgtatcc	ggccggctcc	cgcgagacag	gaagagacag	cgactcgcag	1980
acttgcagct	gcccagcagg	cagtagcttg	cccttttaa	tcgcagcact	tgggttgac	2040
agtcacacac	acactggtcg	cggcgccgca	gatgaggcgc	cacaaatctg	aaccagcata	2100
cagttagggc	ccatgatttc	ccccctttct	ctatggttt	tatataaaaa	aaaactgatt	2160
tatataa	ttttgtatga	ttcctcctaa	attcctgcat	ttcaaataag	acatcgagct	2220
agtgtaa	taggcctgca	aagctgcagc	tgcctgttcc	ctgtactact	tggcactaat	2280
ctcaccggct	acgtgccttc	gcatacggag	tagtttattc	gccccggac	acttgataga	2340
aaggctacgg	tagcgtactt	ctaagactat	ggtcatttaa	aacactattt	tatactatag	2400
tctatactgt	tactacacta	gagttaaaat	ataaaagatga	atatagataa	attactgaag	2460
atagtctaaa	catctttcc	attttttaag	aggacacgccc	taaacagttt	tgaatacata	2520
catacacacc	gatacacaga	tgctaaagca	ctaaagcccc	taaaaagttt	gtgatacatg	2580
aaccaacacg	aaatcgatca	cctttgaacc	cgccagtcct	actaccactc	tccttcgttt	2640
gacctagctg	gctggtaag	aaaaaaagtt	cgtttccgt	aaaagcgatc	caagtgacaa	2700
cgacgagacg	tgattttccg	tacctgactg	atgttgatac	cagccactac	tccgagctag	2760
tagtaaaaga	tttttacagc	caggggagag	acggcggtcg	tctgctagtg	gctagaccgc	2820

gcgtgttgcg cgacgccggg ccggggcgcg cggtcgcca cggactggca gccgggtgcga	2880
ggcacgcccgc gtggtccagc tagatcgccg accgggcacg gcagacgacg gccgggtccag	2940
cggctggaca cgggcacacg agagtacgag actgcgcgcg ttttaggttgg taccactacc	3000
accatgcatg tcgcgactcg ctcgtccgat tatggagcaa aaacacctgt cgtcatgcga	3060
atcattgtac ggcataloga taagaaaaaa aaacttaccg cttctagcat gcccgctgta	3120
tgcgacgacg aagttgagaa atcatgccca agcaagcttc ccgcatalogt agtttcgct	3180
gaccctgtg gcctgatatg tcgctgaccc tga	3213

<210> 211

<211> 4938

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_204726

<400> 211

ctcaatcgcc aaagtagaga atgatgttt tcagcctaca ggccccaggt agcagctgtg	60
tgttgcaaaag caacttgtat aatacaattg gacaagttcg agcttcaaata aactcatata	120
ccatgcaaaag caaatggat ggtttcagc ccacgtaatt aaatggaaat caagtttcc	180
aacctgtcaa gcactacatg attaattgaa accttcagtt ctgcatalogt gccaatctgg	240
ccaatagaag tggatttgct tatgcagttt gcacatgatt taaatagatg aaaacaaata	300
acaaagtaga gtttagtgctt gcagtcagct tgtccactc ccataacttcc tttgcacttt	360
tcctcccccagc aaatttagtac cagttttac gtgtcctcat cataacaagt attgaccatg	420
aatcaaaata cagacatcac tcacagctca cagttccagg ccatggttac aacctagcac	480
taagttgacc aaggcttcta aattttagaac tttaatagc agaagtgccct tgaagatgt	540
aaggctgcaa atatacagga gggacaatga tcagattctc caccaacata actcactagt	600
gtttgggtt aacaataatc atataaagga accaagttat gctgacagaa aattatttgg	660
aataatacca gcaatcaaataaaaagaacca agttgtgctg acaaaagcaa cataaagacc	720
aagaaactat atgacaaaaaa ttttgtgaat acatcattct tttgtggcg aagagtatac	780
gtagatatta gtttatttaaaaattttttaa aattttttaa aatttttagtt aaagctcatt	840
ttttgcctga gtcaaggtaa tacaacacat cagttatcata tatccaaatc tgtggattca	900
ccagctcaaa ggaagatagc aaagataaca acaggttagc gttgctgaat tgtaaagtac	960
agaacttagct agtgttgctg aattgtaaag tccacaagca aattgttgca gctcgcttc	1020
atgctaagaa attgacatgc tgaattgtg agcatgcacg tgaaacacac ttagataaac	1080

taaggatata atacatgcag cctaattaaa gtactagaac tactaattac ttggtaatgc 1140
actacatatg tgtagaatgt tcagaggaac agctagcggt gttccaacat gtgcataac 1200
ttttaactga aagggcatgt tttatcaatc tcttagacc tcagtagagt ttaggcagtg 1260
gaaagctact aactaggaga gatgtcattt attttttaac tatcacttaa caggcagaat 1320
tagccaaaca aaaaaaggac catgcttta catagacaac ctgtaaagtc catgtgacag 1380
aaaatatatt ggtatgcaaac tgacttagga actaaaaaaag cttgaatctt ttgtattta 1440
gatccaaaga acccctgacc attttcaaa atggctgtta agaaaacctt gaaaattctt 1500
gtccattata tcctattgtt acttgatttc taaaaccgat gtataaataa tcagagcaaa 1560
ctttactgaa aagtggtgac acatcgagga ggagaggaag caaccaggaa aggacacggca 1620
taggaacgtg cgctgtgcac atgcggcattt attgtcttag cttagaaaagt ggtgtgagtc 1680
aggtcaggc aggtcagagg gggatcgatg gaccgatccg attcggcagg cggaacaagt 1740
gaaggttcag tgcacggcga ggcgagtc taaaaccgat gtataaataa tcagagcaaa 1800
aaaagggaca gaggatccc agaacggggc gggggtgcgg acgacggaaa gcaaaggcga 1860
cgacacggag gggtggcactt cattagatta ctggtatctg gacggccgac caggtacagg 1920
cgtacagcaa caccagccac acacaccgac ccagccagca aaaccgctcc tccgccttc 1980
ggcttccagc caggtttcg tttcgctgg tttccctc cctccgcgc gcctactaaa 2040
tcaatcactc acccgccgtc gtcgggtgg ctggcacgcg tgcgtccag gggtggctg 2100
ggtgcgtca cccgcgcgtt taagctagct ctggcgcata gcccgcgc tcaccggag 2160
ctaggcaaca aatcgaggc ctctcggcgcg cgccgcgtc tgcgtccgc agccgcagcc 2220
tgctgtactg tacttggctt cggcggcgcac ccccgctcc aagtggttgc ctgggaggcc 2280
ccggctccgg gcatcgac gcaacgcgcac gagacgcgtaa aagtccaccc gtcgcggct 2340
cgcatgcgcg aggaggcag gcaggcagcg aacaaaatcg cacgcgcgtc gtcgactgcc 2400
tggcctggcc tggccatgt gaaaaccgc cccgtttgc tccgttgcg gatctggat 2460
gggctttt cgcgcgtggcg ggggtcaatg caaaacggc gatgtgcgt gtgggtgcg 2520
ccgcgcggcg tcaacttccc ttaacttctc ggcgtggaaag aggcaactgc ctgcctttgc 2580
cgcaactgaa aaggaggagg aagaaaatga tcggccggat tggatggcga cggcggcgg 2640
aagcggaaacc ttcttgctt gccgcgtgacc atgacacccgt gccccctt gcagccgcgt 2700
gaggactttt ggcagctttt acgcgtggcg cccgtgtgcg tgcgtgtct aggcaaggca 2760
cgcaaggcag agcacgcgcag ggcggatgag gaattttgac actgcaaaaa tcccaaattcc 2820

gcaagtatgc acactcaccc gcacgaccag tagcgtgact gcgtgagta	2880
gtcacccata ggagtagcgt tctatata	2940
tagtcgttcc gaatttcaaa aaaattggaa atcagtcga cggcctacta	3000
ccttggtagt cggcacgacg ggtatatacg ggatggaacc gagccggtgc	3060
tggacgtgac attgattgga gtggcatac agactgctgg gactggagtc	3120
ttgcatacaa acatccaaga cagcatccac ccgttagtc agagtcagac	3180
tcgagta ccaaga gccaaatcca tttggatcc aacaccatag ggtcattt	3240
catggaccg gatcatgcgt atatactatg tgtctattat tagcatgccg	3300
acgcgcgcata gataggccct ctgacagtaa ttaatata	3360
aggaggagag gaaatactt ccacctattt tttctcgct atagaggatc	3420
gttgcataa aaacaaccctt tttggttagg tttgaccc tcgttatcca	3480
attcaattt agtgaattaa tttttagtga caaatgatac atcaataaga	3540
catttatgat aactcaatt ttcaagat	3600
atgaacctt ccagttggat acctagataa aggttagatg tacttcatt	3660
aaggatggac atgtgttagg tgtggaccata aataataa tgtat	3720
ttgaa aagtacttaa agtaggaact tgtatcctaa tatgtat	3780
tttgcacgc gtgacttggt ggcctcgccg accgagcgcg gaggcaaggc	3840
atgtggcacc gatcgaagca ggcacatata actgtcgccg cgccgtatag	3900
agtcggcact gaccacccca cccctgcgtc ctagctagct agttgctact	3960
agctccgatc gcatcggtt ctggtagctt ctggctgccg gccggctgct	4020
acttgtacgg ccaataatgc gtggtgagtgcgtgc atcggtctct	4080
aggcgaggcg tctgatcatg cagcatatgc gtggtgcccc ggccgccacc	4140
cgcgtacgtt gttgcgaagc cgggaggaca gcacattgtc cgtgcacccgg	4200
ccccatacta cttatggacg acagaatatacgtggcagca gaacccatgc	4260
acgagggatcgg ccgcgacgtg acgagggaaac aaccacccca acccaaccaa	4320
cgcgcgtagg ctcagcccg cggccggcccc tgcccaggag ccggccggc	4380
tccacgtaca tcgatcgtt ggattcgatc actgtgcaga tccacgtacc	4440
atgccccatc tccatgcgtg tcttgcctc ttgcttgcta tcaacctgac	4500
atccaagtgg tgcacagcgt gaggagatcg gagcggagatcggtgt	4560
tctgtgcgcg gcatgtccgt acgacgtgcg tacgcccgcag tcctgcct	4620
gcacgcgtgc acggacggg actgagccaa ctcgatctga ggatgattgt	
ccgcgtcgccgcgtccgtgt ttgttcgacg tggtttgtg catctacggc	
ggatcgatcgta	

tagtagcatg	tctgaaatct	aatgcctgg	tgatagctca	gctgcctgcc	tgcccccgtc	4680
ggcattgcag	ttgcagccct	gttgattctc	gatttggaa	caccaagaca	ggcgagatg	4740
aattgcgt	gtccgtaata	tgcctgcact	gcaccctgca	taggttgcag	taacaaacta	4800
acaatccctg	gacaacgtac	gcgtgacaac	gtcttgacca	ggagaagagc	aacgcgcgcc	4860
acaggcatcc	ggccagagat	gccgcaaggc	ctgactagcc	cagcggccag	cgccactgca	4920
tgcatgtcat	ccaggacc					4938

<210> 212
 <211> 1370
 <212> DNA
 <213> 人工序列

<220>
 <223> 最佳_基因座_156393

acggaagtgt	gcccgcgag	agcaggacgt	ggtttggacc	agaccagatc	agagcgtcgg	60
tcggttccctg	ttcctgggtgg	tatggcaccc	atggaggccg	ataaaactaac	tagtcaaggc	120
attggatccc	gcgggcctgg	aaaagcagtc	accggctctg	gacctgggac	tggaggcaaa	180
gccctccggc	ttttgcacca	cgaggcagac	acgggctcca	ggaatcgatt	cagtgcagcg	240
gcacgcccgg	gcgcctcaca	tgattcctca	gcacatgcct	gcctgcactg	cacagccggg	300
aagtctgttc	acttgtactg	cgtcgccct	tggcgcgagg	atgatagcta	aggctcgctt	360
cccgcgacac	acgtacccac	gcacgcacgg	aaccagaacg	ttttgccgac	ttcgtgtgat	420
caaggaccgt	tgatcacgccc	gctgctgctac	taaaggatata	ataaaaaaaa	tgacagcgct	480
cggcgtcttc	aggcttcaga	accgttgcac	aaggacgatg	agcttgagag	ctgaatccgg	540
acggatcgga	ctgcgtcgac	gacgataaac	agcagtgcct	gcatgctctg	ttgtgcagcc	600
gtacgtgccg	tgcgccgggt	tatgtctgca	tgcgcagttac	ttatgttctg	aagcaggttat	660
atacaatata	ctagctagta	cactgtggta	tctgcaacgt	agtactgatc	tggattctcc	720
atgttggcgt	gtatactctc	ttttcttcac	aggatcaagt	tcatgcactg	agcaccagat	780
cagattcagc	caccgatcga	gggagagggg	gtgccggagc	cgaggcagga	cagggggcgg	840
cgacgtccca	cgtgcacggc	gatctcaatt	atattctcat	gcatgcagcg	gacaccggat	900
cgaggaagcg	cggcaaagcc	tctgtatttc	gatggacgc	cccgtcgca	tgcgtcgca	960
tcgcgtcgca	tcgcgtcgat	gctgcaaggc	agaaccgcag	ggcatggcgc	ccgaaacggag	1020
cccatcagcg	cccagtgcgt	cggtcaagccc	agagcgtac	tcaaagttt	gctagcttcg	1080

tgccgaattc atgtccatgg atatcggaat gtgcattgcag cagttgtgc cgggcaattg	1140
atgcagacaa ctctttttt ttttcaaattt atccgggtttt atttatatat atacgagata	1200
caatagatag ctttggccac aagcgattaa atgacgctgg tcggttaggaa agcttctgg	1260
ggagaccaga cttgagaaaaaaa aaaaataaacc caagacctat aaaggcaaga caatcatg	1320
tgctctgtcc tccgggaaga taaatggctg catgaagcag ttgaataaaaa	1370

<210> 213

<211> 1190

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_198387

<400> 213	
gaacatcctg catgcataata gcctggctgg tggcagcattt ctttttttca tgcgagccag	60
acccttttta tcagggaaacc aatcaaacc tagacttctg cctgggtgcac gcatgttctc	120
atgtcattta agaaatgcac ataattttgc cggttctgttca cgattaggac ttttcat	180
attcgtaca ttttaccaa aagtttagtca gattagttca aaataaataaa aaaactagtt	240
tttccccaaa aatcataaga ccatttgg taaagcatta ttatatgata acaaaagaaaa	300
actaaggctt tttttgggtt accttgttgc ttgtgaagct gctttgtttt tggtaaaaac	360
caaaagtcaa taaagaggag aaaagaaaata gctacttttgc ctctagtaca taatttatga	420
ctctttcacc tctactttta cttagttgtt gaaaaataac gatttattttt cctcgtttct	480
tgtccccgaa agagagaggt gtaacccatc ctctataaca gtgtggctttt gacaacggtt	540
caaatatgtat gtttggatgc aggcccaagg tcgcttggcc aaatattggc cttgacaaac	600
caatttggcc catgttttagt tggaggccaa atgttggcca tgaccaggaa attttggctg	660
tgc当地atgtt gatttggcc aattcttggat cataatcagc gcgacaaac cggcttacc	720
aatttggctt ggtcgatattt catttttaacc ataaacccaa catccaaact tgaacccaaat	780
tttggcttga tccaaacttttgc ggcagctacc taatttttagc atttcttgc tccacaacac	840
aacaggccctt agtgcgagcg tggccctggc acgcttaccctt gccttgcgcgtt gttgtgaagcc	900
tttagtggccatc cctcaaccgc aggtgtggat ccggccggccgg tcatggcgatc tgcggccctg	960
acaatggccatc tggcgaggc gcttccaaaca tggcgatggc gtggacacgc atgcggccct	1020
gtccccagcg catgtttgc cccaaacccaa gcatggctgc ggtaaacaac tttggcgaga	1080
ggaagatgga aatgtgagtc taacaagtgg gtctgttat aagtgtgtga gagagaaagt	1140
tgagtggatc acatgtggatc tctgctcgatc agtggatc aataaacac	1190

<210> 214
<211> 3965
<212> DNA
<213> 人工序列

<220>
<223> 最佳_基因座_31710

aagtaaaactc	gtaaaagcggc	accatatata	ggcattgtg	agaacggtcg	cagtaggtat	1500
tgggtcttag	acatgtgctt	ttggcgtgag	attcagaaga	tctaaggatt	gctacgcaga	1560
ccgcctaatac	atataaaatta	tgaaccaatg	catggattaa	ctaataatat	cgatagatgt	1620
gtctttttt	tctctctatt	ctgtgggatc	ggatcagttg	attcgctcac	tttcactcac	1680
ttttcttttg	cggcgtatgt	tccgcagttt	ccttcccat	catgaacttg	gagaccgaaa	1740
ggtgtggtgc	ctagttaccc	cgtactgcaa	tttaccaatg	gcgctgggt	ggaaataaag	1800
ggcaggcgca	caacacagaa	aggaccagca	gcagcaccac	caccaccggg	cactagctag	1860
ctagctctt	ggtcaatgga	gcatgcagct	tagctggat	tcagacactg	ttgctggctt	1920
gctgctgctg	caattgttg	ttcaaggcga	acacaagaga	gatccgaagc	atggttgagt	1980
ctctggcgt	agaaaacagt	gccggtttcc	atcggtttaga	caagaacgca	agagagagat	2040
cgtccatacg	acgtgcatga	ggaccagatg	gagattcaga	gacgacccca	tcttctttgc	2100
ataaaagaaac	tagatgaagc	ttgtgtcgtc	gcctataaga	aatgaaacat	ctgcattgtc	2160
agcgttagtat	cccagctgat	ggagaggaca	tacccatgag	gcaataaatac	aaacgaccctt	2220
ttaattttta	cagccaatgc	attgcatgct	gcattatcaa	actggcgaaa	agtgttttt	2280
ttaagggggg	ggaaggctac	catgggttgc	attgaccgag	tgcggctct	cgtagggcct	2340
atgccacatt	actttcttga	aaggcaacca	caggctgatc	tggaaagtcca	ttccacggtg	2400
gcatgcatgg	gcggcgccac	acgtcagtc	tccgcgtcag	ggcaagcttc	ggtcgatcga	2460
tgacgctgcc	cggctgcatg	ccggaggaat	agtggagaag	ggtgcagagc	agagggatgg	2520
tcagatcatt	ccagttgttg	ctttcggca	cttgcataatg	tgatatatgg	agcaaacagc	2580
accgctggct	ggatggatcg	gatgcatac	accatatatg	atcgaagatg	agcctagcca	2640
tgcattttgc	aatatggact	actatgtaca	tctgtcgac	gtactagctt	gatattaatt	2700
cagtgtacat	ccaattacac	acatgcata	gcagtaactg	gacgtttatt	tgcacacttg	2760
cactatatac	acatcgatca	gacgtcgctc	cagagtacat	tcagcactgg	gtttttcttt	2820
tctttgggac	tgactgctat	gctagcaagt	cactaactca	ctatagctag	ctagctagct	2880
ttgtcttgg	tttttttcc	ctgtggatgg	atcgccctgg	cagtggagga	gattgttggt	2940
gggggtgggg	agctgagaatg	cagcattatt	tgaaagagat	gcgcgtcaaa	tgattaagtc	3000
agggacgcct	ttccacat	tcgtgattca	gttggcata	ttgcactgca	ggcaggctag	3060
ctagttgacc	agaaaagtga	gttgggttg	cattgcattg	catgcata	atccaagttac	3120
gtgttaacct	gggttaggatg	ttcctaaca	tgcacaca	cttcccatc	atcatatcac	3180
tagtaagcta	gcccttgctg	gaatgtacag	gacccaacaa	aaagaccatg	catgcata	3240

tacaagtccg	caaggctggt	ctgcctgct	cttgc当地	aaagttagta	gttacacggc	720
gctggctgat	gcatgcgtga	cacgggtggct	agctagctgt	tgtatccc当地	cggcccggtgg	780
acgaaccctgt	acgcgtacgt	gtgtttcggt	tgc当地ccaa	acaagtgc当地	ggttgc当地tt	840
gccgtggacg	atcgatcgat	cggcccacgc	atcacatcat	cattcattcc	ggccggatcg	900
atcgggcact	gtggaagcaa	cgaaaccgcg	cgc当地ccgccc	actgtggctt	ggcc当地tggtg	960
gcagcaggag	gacgc当地ctg	cggccgggtgg	ggc当地ggggcg	gaggcggagg	cggaggcgc当地	1020
gagcaaggcag	cacagctgct	actccatgca	tgc当地ccctg	agattcaagg	catcaagcaa	1080
gcgtctgct	ctctgagact	ctgactctct	gtagcgacta	gtggtggcg	agggtcgag	1140
ggtcgacgca	gtaagc当地ag	gtcaggccca	ctggccagc	atcgat	ttta tggcaagaca	1200
tgggc当地cg	cgagatgggc	acccacagtc	gactgccc当地	ctgc当地gtgg	ccctggcacc	1260
ccgc当地ctgc当地	gcctacacaa	cacaacagcg	ggagc当地gtcg	gctgc当地agagg	atccc当地gtccc	1320
cagaaggccca						1330

<210> 216

<211> 1081

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_197372

<400> 216

agaacgggtc	ggtatcgagg	agcggcgaga	taatagaggc	aatgtgacaa	tgtgtaaaca	60
cttagggtcc	tagcatcctt	gaagtatgtg	tatgtgctaa	agaaatagca	cttggctc当地	120
tggaaagggtcc	tagagacgag	gactctggc	tcctcgtaa	gggtgcaggt	tgaagtggag	180
acccagagct	actagcaatt	acaggccatg	aggagatgt	atccctcatc	caacattgag	240
tatccatggt	ataggtactt	atttataacc	aactgctctt	gttgc当地ttct	ttgttgtatg	300
aactatagta	actcaagcat	atatactagat	cagtagttagt	ggtgaaatcc	ctagtggaag	360
tcgatgccac	acgttaaattc	tttgc当地tgc	tactcaatag	aagaagattt	atatacccac	420
accccaagaa	acttgc当地cc	cagccatcca	tcaatccaaa	gataacccaa	attgttagcac	480
tagatgttga	aacgagtc当地	gttgc当地ag	gaacaatgtc	tacactgaag	atggtggc当地	540
gacaggacat	cgagcttgca	ttaactagca	aagtgattcc	ttccatccat	gcaagaagct	600
acatccc当地	gggtgc当地ggca	aaagccacac	taacacttca	tgagttgatg	ccttcttct	660
gaatgttcaa	ttcaaatactt	gcatgggc当地	gggagctct	gtctaagctg	tatgaaaagg	720
ttggagactt	ccataacaact	gttgc当地gtgc	gagtagttaga	aaccaacaac	aaagtcat	780

tatggtactt agagtagata gggacaacat gctattgtgt	840
gccaaatgc aagaaacaac	
aatggagcca ctatatgcta ctttagattag ttggactagc	900
tttagaccata gttttccaaat	
ggcttagagc tgaattggaa gatctggtca tgtctgagtt	960
ggacttgcac tcttgtcaaa	
ctaggtttgg accatggttt gttgatggct tagaagctga	1020
cttggaaatac ttggttgtgc	
ccatgttggaa cttgtactct tagttgaatt gggtttggaa	1080
catagttgtc ccttagctta	
g	1081

<210> 217

<211> 3903

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_232228

<220>

<221> 不能歸類的特性

<222> (2818)..(2917)

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

<400> 217

ccgtgtgtca gcatcacgtg cgccggcgaga ggaaggatga cgatggcttg gttgctcgta

60

gtttagaga gatagaatga aagcttcttt cagtttttg tatatgtaaa ttcttgctaa

120

taatttttttg ttttaaatat acatatatat taacgttttc gtggatggtt attgtttag

180

tacttaacat gagaatgagt gcatggccgg aatggatgg catgatgaac catticaatt

240

tagtatcagc cttcggtggg aagcaaacca taagcgcgcg tacacataacc acaccacacc

300

ggcatgagtg caaggcatta acgactcgca atggatgtcc tgtcgaagcg agagggacgg

360

gggtgggtggc ctgggtccgg cattatccct ttgccgagct cgagcagatc aggcgaaaaaa

420

aaatcgatgtt aggctgctcc ggcgcggagt cggccaccga agcgagtgca ggcgcaggca

480

gccagccgtc gtggatcccg cgatccggc cccggatct ccgctgcccc catggcccg

540

tggccgcccag ggtgggtggcc gcacggagag gcgtccggat cgggcaccgc ggctgcccgt

600

gggaggacgt gggacccacc tctggacatc aggacgtcag ggcgcatttc tcgttaggcca

660

cgcgctgaca cgatcaactg ttgcggaaaca aaaatggcgc tgcgtcacgt ccatcatcca

720

tgccgtcagt ataagcaagt actactaatc gagcgatcg cacgctgatt ggcgaccgag

780

gtggtgctaa ttattgggtgc ctgtactacc taaatggcta ctgcattggaa cgaaatcacc

840

ggaggcaggc aggcaaggcag cgatgacgag tcttaggcca gagtcgcgtc gggcacctgt

900



cgcctcacgg	960
gcccgtggtt	
gccgtcacgg	
gctgtgcctg	
ctcggcttcg	
gcttcggctt	
cggctccgcc	1020
tgccgccccg	
cgcgcgctcg	
gatcggttgg	
tttggctgct	
cctatccgga	
acgggatgga	1080
ttccgatagg	
tggatgagta	
aaactaactg	
gagcgagccg	
ggaccggcag	
gcagaggcgc	1140
cgcgcgtcc	
acgtcgacca	
accagctcct	
ccagtcaccc	
tagccgggga	
cggggaggac	1200
agcatccgat	
gtggccacac	
ccgcttcacc	
agccggccgg	
tgaccacttc	
gaccagcgag	1260
cgagccgttg	
tgtggctgct	
ggttggacag	
ccgacgcgcg	
cgccgggggg	
gccccgttg	1320
ttgtatacgg	
catccatccg	
ctgcccgtgc	
ggcgggcggg	
cggcggggtg	
ccgacgaggt	1380
acaagtagca	
ggacaggagt	
gcaggatggc	
tgcgccgcta	
cctgaccgtg	
acccgcgacg	1440
gggacggggg	
ggaccccg	
gccgtcgca	
cagcgtcagg	
gcggtaacat	
gcgttccgtt	1500
gcgcgtacgg	
accacgggt	
tcgtctaaca	
ggaagagctg	
tcagtgacag	
gtgcgtgccg	1560
ccgagtgacg	
gccgcgcaac	
ccatccgcgc	
gctcctgc	
cggctagcct	
atccagtgcg	1620
gttcgtggtg	
tccctctcag	
cgtgtctcg	
cgtcggccgg	
gtcggcagcg	
gcagaatttc	1680
atttcagact	
gttgcctcgc	
cgaggattgt	
gcagattttt	
ttttgtgtgc	
cggccgatgg	1740
agagtttccc	
gtacaaaagt	
tgtggatgca	
ggttacacca	
tatccagtga	
atactcattc	1800
ttagccttgt	
ttacgtacta	
cggtccaaa	
gataggctaa	
cttaatactt	
ctgcatgaac	1860
aaccctcc	
ccacacacca	
tctcagctca	
ctatccaaac	
ttcaatacgt	
ttagcaaatg	1920
aaaaacaaag	
tacaaacaaa	
cgaagctaag	
aataaagatg	
ttgggacgat	
taagttataa	1980
tcatgcattt	
acggtttcc	
acgttagacag	
gggtatatga	
tgaaaaagaa	
cattaattat	2040
gccaaaactt	
tagtggtact	
gaacttagtat	
aggtattgtat	
ccactgaatg	
ataaaaaaa	2100
aatcaacatc	
cattgtgcct	
cacaaagaga	
taggcaaaac	
ttcacggcaa	
agtaatatga	2160
tccagggcaa	
atatgtatgg	
cacattaaat	
gataactcca	
agtggatctt	
tcatgttctt	2220
tatcttcagt	
gctaatacca	
caagaatctg	
gactacggga	
gcagaccacc	
agtttcattt	2280
atataatttt	
tttatgaaaa	
aaaattatag	
tacgtatcag	
gggtgggtct	
gtgggtgggg	2340
ctacttgctg	
cactcaaaac	
gtcaagtagt	
tgtctagatg	
gcacatagcc	
tatcaactat	2400
caagttcatt	
acaaaaat	
tatacaaatg	
caacaggtca	
atgaaaggat	
catatthaag	2460
ctgggttcag	
gagcagtagt	
gaacaataca	
cactttcag	
tccactagcc	
taacacttca	2520
gtgcagcagt	
tgacctatac	
acaatagcaa	
tgttttagtt	
caaatttacc	
tgaaacacag	2580
cataccactt	
gaactcccag	
actccagatt	
ggtggtctac	
cacatgtcaa	
ctgagttgtg	2640
ccaggggttga	
catataaagc	
agtacacaaa	
atgtttcagt	
cataaatgtg	
gaaaaacaca	2700
tcatgcaact	
actaggattt	
tactaccaga	
tagatagtgg	
acacaaacat	

caccgagata	cttccctgaa	aaggaaagaa	attgcaacat	cacggtacaa	gtaaaatgtg	2760
aaggacaaca	acacgggcgt	agaacatgga	gagacgtctg	tgtcgagcaa	agtgtan	2820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnna	atatagtac	actgcaaag	2940
atatggacaa	gttcaagtgt	caactgatgt	taaatggtaa	ttaaacattc	atattaatag	3000
gagagcaatg	ttctacgatg	tgggggtggg	gggggggtgc	tgagctagac	agtcaaagaa	3060
agccctcgac	aacctaaaag	aactatgcgg	tttttggatg	acaactagca	tgactaggca	3120
gccatcctgc	aatggagatt	gacgttgtt	cctcaaaaaa	aaaatgtagg	ccgggtttgc	3180
atgcatggaa	tttttttg	aactgccaaa	tgcccttct	gcctaaaagt	ttaggcttta	3240
agcacaactg	gttaagctaa	ttaggactac	taaaatgtag	acttaaacaa	tttatgcag	3300
tacctgacag	atgcagcttc	gtgctcagca	gctagcgtta	gtaaagaagc	atggataactc	3360
tttaaatata	gctatttata	gcatctaccg	tctaggtgag	taggcagaca	ttgttcagca	3420
tttggccct	tggaatttgc	tagaaaaatg	tcattttat	aagaattgca	aaaggttcca	3480
cccatgcagc	catgctgcac	tagttcgaaa	gagaaatgg	gtgtgtttt	tttaaggaaa	3540
cattttccag	tagtcttcta	aacaatccct	cgaagctggt	agccgaccag	ctacatgccc	3600
aatcccaa	ttgtttttc	taacaaaatt	tcaaaaatgt	aactgaaaac	aagatctagc	3660
agcaaata	actaaaaaac	ctgctaagca	gtaacaatag	gactttaagc	tctgggtgca	3720
gtataaaat	aatcaaaatg	cggtggttat	gtgatctgta	gttcagattt	caaagatttt	3780
caagctagga	ataacgcctt	gcatcatcag	attcccttgg	ttcaaacata	tgataaacct	3840
gagataaaat	ttaaactatt	cttgttcttc	ccacatagca	ttctggtaa	gccattaaac	3900
ttc						3903

<210> 218

<211> 1172

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_285621

<400> 218

tacaacaaac cgaattttaa tttacaagaa aataagaatt tttgaacaaa tatattctg

60

aacttatgag aatctctagt ttttttata tttttcatg ctttgggtttt gaattcgaaa

120

ttccaaacga aaaaaaatca aaaaaaagac ggtagaattt gtttctgtcg tttgttcgga

180

©

aggggagaaa acggcgaaac ccggcgaaatc ccgagcgga tttgtatccc ttcggtgagc	240
acgcggccca cgcacacggc agcgatcgta tccacggtgt tatccggcag cacaacactg	300
ctgcgcaga acttgcgcatt ccggcgctcc atgcgcaggt attggcgctc cgccgcctc	360
tggttgccgc tcgtcgctga tcaccagttt ctactcgta agtactccat ggtatgcgtac	420
gcgcggcgggaa actccctcgcc gtcgtccggg ctgaccgaca tggcacgcgc atcgaggatg	480
tagatgcacg tggccgcag gccaggcttg gccgtctggg gccgtgacgc cgggagccgc	540
gccgcgtgtca aaccgcgtca gctcaatctt agggtccgtt ttctcaattt ccaacgcgt	600
agacgaaagg ggtgggtgg gcctgagagg tagagggccc cggccggggc tactcgatcc	660
ggccatcatc ggctcgccgg agctcaccgt gtgcgtgtcc tacaccacgt gaccactcgc	720
ctccatgaat catcatcaat tccacccctt tgccaaggca caatggccaa ccatgtgcgt	780
gctgttaatt aatttaattt aacggAACCT tgagggtttc tttttgaaa aaaataattg	840
accagaataa tccgatcgca tcaccacttc accggcagtc ctgcagtcca acactggaag	900
cgaagcgacc tctctctgtc ttgcgtgcaa cccggccggag cgccgtcagc gtccatctgc	960
cgcttcgttc tggcttggcg caaatgagca tcacatgatt ggccgacacc agcaggcatt	1020
gcagatcgca caagcggggg atgtgatgtg acgtgaggca tacacctgtc agcgagtcga	1080
ttcagggcga cacaacgaat ggagtcgagt agaacgagag gacatgcggg acaactacgc	1140
ctcgccgagc ctgcgagctg taccgaccgg at	1172

<210> 219

<211> 1275

<212> DNA

<213> 人工序列

<220>

<223> 最佳_基因座_157315

<400> 219

ggaccctata cactgaagat ctgccttgac tatggctaag gctgaggccc accgagcagg	60
ggttgctcatt aaggctcatt ggagaggaga cagttccaca ccacgagttt gatagaacag	120
tctagaacaa cccggggccc tataggcagc tccaaccata agcgcttcta gatttagttc	180
tcatggacgg ccgaggtccc ttacgggtca tcggatctt aataagggtcc tcagtttat	240
tagagaatag cgtggtcggg tcctagctag ggcctctccg gacttaaggc cggccctcctg	300
atgagaagat aaatgagttc ctgatggaa gcaccagcat atcagagaac ataacttatt	360
acgatttttt ttgtttgtg cgaggagagc ttattattt gatggctagg tgttagtcattc	420
aaggttagaaa acgatcttcc attagcaca acaccttact gtaaataata agtagtcgtg	480

tagaaaacat ttgaccttca aaggctgcaa cgagtaaatt aaaagaaca tgcatagtc	540
ttattcaatg tttcaacgaa aacgtggaac ttgatttta ttttgcgtga gggagcttaa	600
ttatttctgt tgatgtgtaa acacgttgaa aatggggccc ccgaaccaca cgtcatactt	660
tttttaatta ttttgttagg ttgctatgct gagctagtag ctgtttacac ggcgtctgtt	720
attctgataa ggaactataat acaaaaccac acatgcacac gtggtcaagg tagtgactaa	780
tctcgcttaa taccactgac aaatcatctc tacttggttgc tctcgaccta tagttgctaa	840
aatccatgtg gctaaacttt agttttctct gagttgacga ttaattatta catatacttt	900
atacttgaac tatagttacc tgcatagttgt aaaattatct aaaaattcac gtaggtcaaa	960
tgtggaagaa gtgccataga aaacaacagt cccatattgt gcatcgaa tgaataatac	1020
cacgtgctac cgaacacggt gaggattact ggtcggtcg cccgtgatca agctggtagg	1080
tggtatataat gcagcttta tgaagttttt gagaagattt acgcagtgtat ctatcgtaa	1140
actgaaacgg agacgggcat cactgcacct tggatcggttgc tggcggttcc atttccattt	1200
ctctgctgcc cgcagccagt gtgcagccgt cgtctgctgca agatagacct ggtcatccgc	1260
tagagcgagc cctgt	1275

I669395

發明摘要

※ 申請案號 : 103138081

C12N 15/63 (2006.01)

※ 申請日 : 103年11月3日

※ I P C 分類 : C12N 15/66 (2006.01)

C12N 15/82 (2006.01)

【發明名稱】(中文/英文)

C07H 21/04 (2006.01)

一種用於基因標定的通用供體系統

A UNIVERSAL DONOR SYSTEM FOR GENE TARGETING

【中文】

一種通用的供體聚核苷酸係描述的，該者可以在植物基因組中於一標定位置插入，以促進一供體分子在一特定基因組位置中快速及高通量的併入。

【英文】

A universal donor polynucleotide is described that can be inserted at targeted locations in plant genomes to facilitate rapid and high throughput integration of a donor molecule within a specific genomic location.

【代表圖】

【本案指定代表圖】：第()圖。(無)

【本代表圖之符號簡單說明】：

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

(無)

tgactaaagt ccataatgta acactgacag acagaggata tagatagaac agccgcgcgg	3300
gcctgaagaa acatcagcag gtcataacc acttagtggc tgcatgctcc tactagtc	3360
ggcaggccat ttctttctag aggacaaaaa cgcttgcag ggtactctag cttagacggg	3420
cagagatcga gatgggtagc tagtggagg tgtggtaatg ctgtgccat ttgctggagg	3480
ctagctgctt gttggagagc tagctaaggc gtgggagagg ccagggggg ataagagatg	3540
atgctcgggc aaggcaactt gggtggtcgg gccgagacct gcaaaaggca agtagtagct	3600
gagctggac gacggcattt gaggcctt tttccgatc ccgcatttacc acttcttct	3660
atcgcatctc atcgatcccc caatcccagc cgatcctcgc tcatacggtc cccataatgc	3720
gtgctgcccct ctagcttagc tgccattgcc attggaagca ccagtccata gcagcgcgtt	3780
gctgcatccg gcggctgcct gttggggagg agacgatgca tgcacgcgtc caatccgtt	3840
gttcgttctc aataatatca cccggcccg gcccgtattt agagatattc gcccgagcgc	3900
agtccagcta gcacagcaac gcagtcgcag tgcatgcaat gcaagtttga gtcgtcagc	3960
gttac	3965

<210> 215
 <211> 1330
 <212> DNA
 <213> 人工序列

<220>
 <223> 最佳_基因座_64542

tatataatatc cggttggcgc ttaaaaaaaaa tccaaatctac gcggtgtgtg tcatacgatcg	60
gttggcgtt ggagagcatg gaaaggctgg gcatcactgg gactggtaga gcattgcatt	120
gggtccctac tcgactagcc gactccttcc ttgattgctc gccggcgcgc cgccggcgc	180
gcgggcccatt gcaacatgca tgcatgatga atggggcccg ccgggacagg gcgcattgc	240
tggggcgtt gacatttgcg tccggcgcgc ggcgcgcgc attccaccac	300
caccgtcgcc ggtcacccgc cgaagctaac tagcggacac ccggcgcgcgc gctccggcgc	360
ttttccccca tggcccatc cctctccaa agtggaaaag gggggaaaag ggataagaag	420
ggacaaaaacc acgaaaggta cgtgaaacgt gggtgagggt ctggaaagcca agatagggtc	480
acgacgcattt gcctaaaaag tgcgcgagg cagggagcag ccgcgtgtcc gcctgtgt	540
gctggaaatgg cgccgtggcc ggcgcgcgtg acgggtatcc atgcattctc ctgtcctatg	600
tacaacctca aagctactac gactactact gtggtttta tactccgtat atatgagaac	660

申請專利範圍

1. 一種聚核苷酸供體卡匣，其包含一位點特異性核酸酶結合域、一解析域(analytical domain)及一質體域，其中該解析域係對序列辨識編號：142具至少99%序列一致性的一聚核苷酸序列。
2. 一種轉基因細胞，其包含如請求項1之該聚核苷酸供體卡匣。
3. 如請求項2之轉基因細胞，其中該細胞係一轉基因植物細胞。
4. 如請求項1之聚核苷酸供體卡匣，其中該聚核苷酸供體卡匣包含選自由下列所組成之群組中之一序列：對序列辨識編號：132具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：133具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：134具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：135具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：136具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：137具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：138具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：139具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：140具至少90%序列一致性的一聚核苷酸序列及對序列辨識編號：141具至少90%序列一致性的一聚核苷酸序列。

5. 如請求項1之聚核苷酸供體卡匣，其中該聚核苷酸供體卡匣係選自由下列所組成之群組：對序列辨識編號：132具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：133具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：134具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：135具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：136具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：137具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：138具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：139具至少90%序列一致性的一聚核苷酸序列、對序列辨識編號：140具至少90%序列一致性的一聚核苷酸序列及對序列辨識編號：141具至少90%序列一致性的一聚核苷酸序列。
6. 如請求項1之聚核苷酸供體卡匣，其中該聚核苷酸供體卡匣包含一或多個同源臂序列，其中該一或多個同源臂序列與選自由下列所組成之群組的一最佳基因組基因座共享至少80%之序列一致性：最佳_基因座_204637 (序列辨識編號：210)、最佳_基因座_204726 (序列辨識編號：211)、最佳_基因座_156393 (序列辨識編號：212)、最佳_基因座_198387 (序列辨識編號：213)、最佳_基因座_31710 (序列辨識編號：214)、最佳_基因座_64542 (序列辨識編號：215)、最佳_基因座_197372 (序列辨識編號：216)、最佳_基因座_232228 (序列辨識編號：217)、

最佳_基因座_285621 (序列辨識編號：218)及最佳_基因座_157315 (序列辨識編號：219)。

7. 一種用於在一植物細胞之基因組內標定併入(targeted integration)一聚核苷酸供體卡匣的方法，該方法包含：
 - a. 表現一位點特異性DNA結合核酸酶，其包含至少一DNA結合域及至少一核酸酶域，其中該至少一DNA結合域結合至該植物細胞基因組內的一標靶位點；
 - b. 將該植物細胞與如請求項1之聚核苷酸供體卡匣接觸；
 - c. 以該位點特異性DNA結合核酸酶切開在該植物細胞基因組內的該標靶位點；且
 - d. 將如請求項1之聚核苷酸供體卡匣併入至該植物細胞基因組內的該標靶位點中。
8. 如請求項7之方法，其中該至少一DNA-結合域係選自由下列所組成之群組：一鋅指結合域、一巨核酸酶結合域、一TALEN結合域、一巨核酸酶及一CRISPR結合域。
9. 如請求項7之方法，其中該核酸酶域係源自一IIS型限制性核酸內切酶。
10. 如請求項9之方法，其中該IIS型限制性核酸內切酶係選自由FokI及StsI所組成之群組。
11. 如請求項7之方法，其中該聚核苷酸供體卡匣表現一多勝肽，或其中該聚核苷酸供體卡匣包含一非編碼的核酸序列。
12. 如請求項7之方法，其中該聚核苷酸供體卡匣包含一或

多個鋅指結合序列，或其中該聚核苷酸供體卡匣包含一或多個引子結合序列。

13. 如請求項7之方法，其中該併入包含一同源性導向的修復機制，或其中該併入包含一非同源性末端接合導向的修復機制。
14. 如請求項7之方法，其中該方法符合下列之一者：
 - i) 該植物細胞係為一單子葉植物細胞，
 - ii) 該植物細胞係為一雙子葉植物細胞。
15. 如請求項14之方法，其中該單子葉植物細胞係選自由玉米植物細胞、小麥植物細胞及水稻植物細胞(a rice plant cell)所組成之群組；或其中該雙子葉植物細胞係選自由大豆植物細胞、棉花植物細胞及油菜植物細胞(a canola plant cell)所組成之群組。
16. 如請求項7之方法，其中該標靶位點包含一最佳基因組基因座，其中該最佳基因組基因座係為包含下列特徵的一低甲基化非基因序列：
 - a. 該非基因序列之甲基化程度為1%或更低；
 - b. 該非基因序列與在該基因組中所含有的任意其它序列共享小於40%的序列一致性；
 - c. 該非基因序列係位於一已知或預測的表現編碼序列之40 Kb區域之內；及
 - d. 該非基因序列在該基因組之內展現大於0.00041 cM/Mb之一重組頻率。