



УКРАЇНА

(19) **UA** (11) **123532** (13) **C2**
(51) МПК
C12Q 1/6811 (2018.01)
C12N 15/82 (2006.01)
C12N 9/22 (2006.01)

НАЦІОНАЛЬНИЙ ОРГАН
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ
ДЕРЖАВНЕ ПІДПРИЄМСТВО
"УКРАЇНСЬКИЙ ІНСТИТУТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ"

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

(21) Номер заявки: а 2015 09935	(72) Винахідник(и): Дешамп Стефан (US), Інгліш Джеймс (US), Лі Жонгсен (US), Ллака Віктор (US), Янг Джошуа К. (US)
(22) Дата подання заявки: 10.03.2014	(73) Володілець (володільці): Е. І. ДЮ ПОН ДЕ НЕМУР ЕНД КОМПАНІ, Chestnut Run Plaza 974 Centre Road, P. O. Box 2915 Wilmington, Delaware 19805, United States of America (US), ПАЙАНІР ХАЙ-БРЕД ІНТЕРНЕТНЛ, ІНК., 7100 N.W. 62nd Avenue, Johnston, Iowa 50131-1014, United States of America (US)
(24) Дата, з якої є чинними права інтелектуальної власності: 22.04.2021	(74) Представник: Нестеренко Наталія Олександрівна, реєстр. №413
(31) Номер попередньої заявки відповідно до Паризької конвенції: 61/777,238	(56) Перелік документів, взятих до уваги експертизою: WO 2013/006745 A2, 10.01.2013 WO 2009/006297 A2, 08.01.2009 WO 2012/129373 A2, 27.09.2012 US 2011/165679 A1, 07.07.2011
(32) Дата подання попередньої заявки відповідно до Паризької конвенції: 12.03.2013	
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: US	
(41) Публікація відомостей про заявку: 10.02.2016, Бюл.№ 3	
(46) Публікація відомостей про державну реєстрацію: 21.04.2021, Бюл.№ 16	
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ: PCT/US2014/022500, 10.03.2014	

(54) СПОСІБ ІДЕНТИФІКАЦІЇ ВАРІАНТНОГО САЙТА РОЗПІЗНАВАННЯ ДЛЯ СКОНСТРУЙОВАНОГО ЗАСОБУ, ЩО РІДКО РОЗЩЕПЛЮЄ, ДЛЯ ІНДУКЦІЇ ДВОНИТКОВОГО РОЗРИВУ

(57) Реферат:

Винахід стосується способу ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимий сайт розпізнавання, причому зазначений спосіб включає (а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатним до введення двониткового розриву в розпізнавання зазначену геному ДНК, де двонитковий розрив приводить у результаті до утворення нуклеотидного липкого кінця; (b) лігування першого адаптера із зазначеним нуклеотидним липким кінцем; (c) розрізання лігваної ДНК, отриманої на стадії (b), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив; (d) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (c), з еталонною геномною послідовністю ДНК, та (e) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу для індукції двониткового розриву. Винахід

UA 123532 C2

також стосується способу ідентифікації варіантного сайта розпізнавання з покращеною активністю розщеплення для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимий сайт розпізнавання.

Дана заявка заявляє пріоритет на підставі заявки на патент США із серійним номером 61/777238, поданої 12 березня 2013 року, яка включена в даний документ як посилання у всій своїй повноті.

ГАЛУЗЬ ВИНАХОДУ

Даний винахід відноситься до галузі молекулярної біології. Більш конкретно, даний винахід відноситься до способів ідентифікації й застосування варіантних сайтів розпізнавання для сконструйованих засобів, що рідко розщеплюють, для індукції двониткового розриву.

ПЕРЕДУМОВИ ВИНАХОДУ

Технологія рекомбінантних ДНК уможливила вбудовування чужорідних послідовностей ДНК у геном організму, таким чином, змінюючи фенотип організму. Найбільше широко використовувані способи трансформації рослин являють собою інфікування за допомогою *Agrobacterium* і бомбардування частинками в ході балістичної трансформації, при якій трансгени інтегруються в геном рослини випадковим чином і з непрогнозованим числом копій. Таким чином, мають місце спроби контролю інтеграції трансгенів у рослини.

Були розроблені способи введення або модифікації послідовності ДНК у геномі ряду організмів, і вони можуть включати методики сайт-специфічної інтеграції, які ґрунтуються на гомологічній рекомбінації (патент США № 7102055, виданий 05 вересня 2006 року) або сконструйованих ендонуклеазах, таких як мегануклеази, нуклеази "цинкові пальці" або TALEN (публікація патенту США 2009-0133152 A1, опублікованого 21 травня 2009 року).

Хоча ці системи забезпечили корисні методики для цілеспрямованої вставки послідовностей, які представляють інтерес, залишається потреба в ідентифікації більшої кількості сайтів розпізнавання для засобів, що рідко розщеплюють, для індукції двониткового розриву та в ідентифікації сайтів розпізнавання з підвищеною активністю щодо засобів, що рідко розщеплюють, для індукції двониткового розриву.

КОРОТКИЙ ОПИС ВИНАХОДУ

Передбачаються композиції та способи, у яких використовуються варіантні сайти розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання.

Передбачаються способи ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання. Один спосіб включає: а) контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатним до введення двониткового розриву у вказану геномну ДНК, де двонитковий розрив у результаті призводить до утворення нуклеотидного липкого кінця, б) лігування першого адаптера із вказаним нуклеотидним липким кінцем, с) розрізання лігваної ДНК, отриманої на стадії (b), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації і секвенування фрагментів геномної ДНК, що оточують двонитковий розрив, d) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (c), з еталонною послідовністю геномної ДНК; та е) ідентифікацію варіантного сайту розпізнавання, що містить заміну щонайменше по одному нуклеотиду (або нуклеотидній основі) в порівнянні з припустимим сайтом розпізнавання вказаного сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву. Інший спосіб включає спосіб ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому вищезазначений спосіб включає: а) контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатним до введення двониткового розриву у вказану геномну ДНК, де двонитковий розрив у результаті призводить до утворення тупого кінця; б) створення нуклеотидного липкого кінця з тупого кінця з (a); с) лігування першого адаптера з нуклеотидним липким кінцем з (b); d) розрізання лігваної ДНК, отриманої на стадії (c) і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації й секвенування фрагментів геномної ДНК, що оточують двонитковий розрив; е) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (d), з еталонною геномною послідовністю ДНК і f) ідентифікацію варіантного сайту розпізнавання, що містить заміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання вказаного сконструйованого засобу для індукції двониткового розриву.

Сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву можна вибрати із групи, що складається з мегануклеази, нуклеази "цинкові пальці", TAL-ефекторної нуклеази, транспозази, ендонуклеази Cas і сайт-специфічної рекомбінази. Нуклеотидний липкий

кінець може являти собою 3'- або 5'-нуклеотидний липкий кінець.

Крім того, передбачаються способи ідентифікації варіантного сайту розпізнавання з підвищеною активністю розщеплення для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання. Про підвищену активність сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву свідчать а) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на геномній ДНК; б) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на плазмідній ДНК; с) вища оцінка в аналізі на дріжджах для варіантного сайту розпізнавання в порівнянні з припустимим сайтом розпізнавання; або d) будь-яка комбінація з (а), (b) і (с).

Також передбачаються способи націлювання вставки полінуклеотиду, який представляє інтерес, на конкретний хромосомний сайт у геномі рослини, причому зазначений спосіб включає: а) трансформацію рослинної клітини або рослини фрагментом ДНК, що містить полінуклеотид, який представляє інтерес, де зазначений геном зазначених рослинної клітини або рослини містить щонайменше один варіантний сайт розпізнавання, вибраний із групи, що складається з SEQ ID NO: 15, 16, 17, 18, 19, 20 і 21 або SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35; і б) забезпечення наявності мегануклеази, здатної забезпечити двонитковий розрив у варіабельному сайті розпізнавання з (а); і с) відбір зазначеної рослинної клітини або рослини, що містять зазначений полінуклеотид, який представляє інтерес, інтегрований у зазначений варіантний сайт розпізнавання.

Різні композиції включають рослину, насіння або рослинну клітину, що містить у своєму геномі варіантний сайт розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання.

КОРОТКИЙ ОПИС ГРАФІЧНИХ МАТЕРІАЛІВ І ПЕРЕЛІКУ ПОСЛІДОВНОСТЕЙ

Даний винахід може бути зрозумілим повною мірою з наступного детального опису й доданих графічних матеріалів і переліку послідовностей, які утворюють частину даної заявки. Описи послідовностей і перелік послідовностей включені в якості додатка в даний документ, відповідають правилам, які регулюють розкриття нуклеотидних і амінокислотних послідовностей у патентних заявках, як викладено в §1.821-1.825 розділу 37 C.F.R. Описи послідовностей містять трьохбуквені коди амінокислот, як визначено в §§1.821-1.825 розділу 37 C.F.R., які включені в даний документ як посилання.

Фігура 1. (А) Сигнатура піка для геномного сайту розпізнавання у зразку, обробленому мегануклеазою. Нанесені на карту дані послідовності починаються із сайту розщеплення й повертаються до нього. Напрямок нанесених на карту зчитувань вказується у вигляді накопичень і показане стрілками. (В) Порожній контроль не містить сигнатури, яка відповідає збагаченню або піку, як спостерігалось для обробленого зразка.

Фігура 2. Перекриття 4 основ у сигнатурі піка сайту розпізнавання відповідає липкому кінцю, створеному мегануклеазою, і визначає послідовність геномного варіантного сайту розпізнавання. Пунктирна лінія визначає липкі кінці, отримані при розщепленні сайту розпізнавання.

Фігура 3. Відсотковий склад основ у ДНК орієнтованих геномних варіантів сайтів розпізнавання. Переважно сторонні нуклеотиди обведені, тоді як основи припустимого сайту розпізнавання заштриховані. (А) Склад основ ДНК 30 геномних варіантних сайтів розпізнавання мегануклеазою Lig3-4. Припустимий сайт розпізнавання LIG3-4 (SEQ ID NO:13) показаний внизу фігури 3А. (В) Склад основ ДНК 254 геномних варіантних сайтів розпізнавання мегануклеазою MHP14+. Припустимий сайт розпізнавання MHP14+ (SEQ ID NO:14) показаний знизу фігури 3В.

Фігура 4. А) Вирівнювання припустимого сайту розпізнавання для мегануклеази LIG3-4 і варіантних сайтів розпізнавання LIG3-4. Переважно сторонні нуклеотиди (підкреслені) вводили в припустимий сайт розпізнавання LIG3-4 окремо (що призводило до створення варіантних сайтів розпізнавання -11C, -7C, -2G, -1T, +8T, що відповідають SEQ ID NO: 15-19) і в комбінації (що призводило до створення варіантних сайтів розпізнавання, що відповідають SEQ ID NO:20-22). В) Вирівнювання припустимого сайту розпізнавання для мегануклеази MHP14+ і варіантних сайтів розпізнавання MHP14+. Переважно сторонні нуклеотиди (підкреслені) вводили в припустимий сайт розпізнавання MHP14+ окремо (що призводило до створення варіантних сайтів розпізнавання -3A, -2G, -1T, +2A, +7T, +8G, +11G, +11A, що відповідають SEQ ID NO: 23-30) і в комбінації (що призводило до створення варіантних сайтів розпізнавання, що відповідають SEQ ID NO: 31-35).

Фігура 5. Порівняння активності розщеплення плазмідної ДНК між припустимим сайтом розпізнавання та припустимими сторонніми нуклеотидами, поміщеними окремо в припустимий сайт розпізнавання для (А) мегануклеази Lig3-4 і (В) мегануклеази MHP14+. Відсоток активності розщеплення припустимого сайту розпізнавання позначений пунктирною лінією. У якості контролю також аналізували щонайменше дві основи, відмінні від відсоткового складу основ ДНК у геномних варіантних сайтах розпізнавання.

Фігура 6. Порівняння активності розщеплення плазмідної ДНК між припустимим сайтом розпізнавання й переважно сторонніми нуклеотидами, поміщеними в комбінації в припустимий сайт розпізнавання для (А) мегануклеази Lig3-4 і (В) мегануклеази MHP14+.

Фігура 7. Порівняння графіків ампліфікації ПЛР у реальному часі з бібліотек геномних варіантних сайтів розпізнавання для мегануклеази, створених або з фосфорильованими, або з нефосфорильованими біотинильованими адаптерами, що містять повністю вироджений 4-нуклеотидний 3'-липкий кінець.

На фігурі 8 показана діаграма, яка представляє систему скринінгу дріжджів, і яка застосовується для визначення мегануклеазної активності дріжджів. Фрагменти гена, що відповідають першим 1000 нуклеотидам кодуєчої послідовності Ade2 дріжджів (5'-фрагмент Ade2) і останнім 1011 нуклеотидам кодуєчої послідовності Ade2 дріжджів (3'-фрагмент Ade2), роз'єднували за допомогою фрагмента, що містить ген дріжджів *ura3* (*Ura3*) та сайти розпізнавання мегануклеазою для I-SceI.

На фігурі 9 показана числова шкала й відповідне утворення білих секторів у колоніях дріжджів, які застосовуються для кількісного визначення мегануклеазної активності. Оскільки фенотип утворення секторів є якісним показником мегануклеазної активності, було задіяно числову систему кількісних показників 0-4. Кількісний показник 0 означає, що білі сектори не спостерігалися (було відсутнє розрізання мегануклеазами); кількісний показник 4 означає повністю білі колонії (повне розрізання сайту розпізнавання); кількісні показники 1-3 означають проміжні фенотипи утворення білих секторів (і проміжні ступені розрізання сайту розпізнавання).

Фігура 10. (А) Порівняння активності розщеплення плазмідної ДНК між припустимим сайтом розпізнавання MHP14+ (SEQ ID NO: 14), варіантним сайтом розпізнавання MHP14+ (SEQ ID NO: 11), який зустрічається в природних умовах у геномі маїсу (мічений варіантний сайт розпізнавання маїсу), та варіантними сайтами розпізнавання MHP14+ (SEQ ID NO: 31-35), які не є ендегенними для генома маїсу. (В) Порівняння відносного числа копій варіантного сайту розпізнавання MHP14+ з SEQ ID NO:11 і припустимого сайту розпізнавання MHP14+ (SEQ ID NO: 14) у зрілих зародках маїсу.

Фігура 11. (А) являє собою карту плазміди PHP57712, (В) являє собою карту плазміди PHP62552.

ПОСЛІДОВНОСТІ

SEQ ID NO: 1 являє собою нуклеотидну послідовність, що кодує одноланцюговий поліпептид злиття на основі мегануклеази LIG3-4.

SEQ ID NO: 2 являє собою амінокислотну послідовність поліпептиду злиття на основі мегануклеази LIG3-4.

SEQ ID NO: 3 являє собою нуклеотидну послідовність, що кодує одноланцюгову мегануклеазу MHP14+.

SEQ ID NO: 4 являє собою амінокислотну послідовність мегануклеази MHP14+.

SEQ ID NO: 5 являє собою нуклеотидну послідовність біотинильованого дефосфорильованого адаптера, сконструйованого з повністю виродженим 4-п.о. 3'-липким кінцем.

SEQ ID NO: 6 являє собою нуклеотидну послідовність відновлюючого праймера А.

SEQ ID NO: 7 являє собою нуклеотидну послідовність відновлюючого праймера В.

SEQ ID NO: 8 являє собою нуклеотидну послідовність Illumina-сумісного адаптера.

SEQ ID NO: 9 являє собою нуклеотидну послідовність маркерної послідовності.

SEQ ID NO: 10 являє собою комплементарну нуклеотидну послідовність маркерної послідовності з SEQ ID NO: 9.

SEQ ID NO: 11 являє собою нуклеотидну послідовність 5'-3' послідовності, показаної на фігурі 2.

SEQ ID NO: 12 являє собою нуклеотидну послідовність 3'-5' послідовності, показаної на фігурі 2.

SEQ ID NO: 13 являє собою нуклеотидну послідовність припустимого сайту розпізнавання для мегануклеази LIG3-4 (також показано на фігурі 3А і фігурі 4).

SEQ ID NO: 14 являє собою нуклеотидну послідовність припустимого сайту розпізнавання для мегануклеази MHP14+ (також показано на фігурі 3В і фігурі 4).

SEQ ID NO:15-22 являють собою нуклеотидні послідовності варіантних сайтів розпізнавання для мегануклеази LIG3-4.

SEQ ID NO:23-36 являють собою нуклеотидні послідовності варіантних сайтів розпізнавання для мегануклеази MHP14+.

5 SEQ ID NO: 36 являє собою нуклеотидну послідовність гена Ade2 дріжджів.

SEQ ID NO: 37 являє собою нуклеотидну послідовність припустимого сайту розпізнавання для мегануклеази MS26.

SEQ ID NO: 38 являє собою нуклеотидну послідовність плазмиди PHP57712.

SEQ ID NO: 39 являє собою нуклеотидну послідовність плазмиди PHP62552

10 ДОКЛАДНИЙ ОПИС ВИНАХОДУ

Даний винахід тепер буде більш докладно описаний нижче в даному документі з посиланням на додані графічні матеріали, на яких показані деякі, але не всі варіанти здійснення даного винаходу. У дійсності, даний винахід можна здійснювати в багатьох різних формах і його не слід розглядати як обмежене варіантами здійснення, викладеними в даному документі;

15 скоріше дані варіанти здійснення представлені для того, щоб дане розкриття відповідало вимогам чинного законодавства. Однакові числа відносяться до однакових елементів по всьому документу.

Фахівцю в даній області, до якої відноситься даний винахід, спаде на думку безліч модифікацій та інших варіантів здійснення даного винаходу, викладеного в даному документі, при використанні ідей, представлених у вищенаведеному описі та супутніх графічних матеріалах. Таким чином, слід розуміти, що даний винахід не повинен обмежуватися конкретними розкритими варіантами здійснення, і мається на увазі, що в прикладену формулу винаходу включені модифікації та інші варіанти здійснення. Хоча в даному документі використовуються спеціальні вирази, вони використовуються тільки в загальному та описовому сенсі, а не з метою обмеження.

Всі публікації та заявки на патент, згадані в даному описі, орієнтовані на рівень фахівця в даній області, до якої відноситься даний винахід. Всі публікації та заявки на патент включені в даний документ у вигляді посилання тією самою мірою, як якби кожна окрема публікація або заявка на патент конкретно й окремо була включена у вигляді посилання.

30 Форми однини, які використовуються в даному документі та у докладній формулі винаходу, включають посилання на множину, якщо з контексту явно не випливає інше. Таким чином, наприклад, посилання на "рослину" містить у собі безліч таких рослин; посилання на "клітину" містить у собі одну або кілька клітин та їх еквіваленти, відомі фахівцям у даній області техніки, і так далі.

35 Якщо не зазначено інше, всі технічні та наукові терміни, які застосовуються в даному документі, мають таке ж значення, яке звичайно розуміється середнім фахівцем в області техніки, до якої належить даний винахід. Хоча будь-які способи й матеріали, подібні або еквівалентні описаним у даному документі, можна застосовувати на практиці для тестування даного винаходу, у даному документі описані конкретні приклади придатних матеріалів і способів.

40 У контексті даного розкриття використовується ряд термінів і скорочень. Пропонуються наступні визначення.

Як використовується в даному документі, терміни "цільовий сайт", "цільова послідовність", "геномний цільовий сайт" і "геномна цільова послідовність" використовуються в даному документі взаємозамінно й відносяться до послідовності полінуклеотиду в геномі рослинної клітини або дріжджової клітини, яка містить сайт розпізнавання для засобу для індукції двониткового розриву.

50 "Штучний цільовий сайт" являє собою цільову послідовність, яка була введена в геном організму, такого як рослина або дріжджі. Така штучна цільова послідовність може бути ідентичною за послідовністю до ендегенної або нативної цільової послідовності в геномі організму, але може розташовуватися в іншому положенні (тобто неендогенному або ненативному положенні) у геномі організму.

Терміни "ендогенна цільова послідовність" і "нативна цільова послідовність" використовуються взаємозамінно в даному документі для позначення цільової послідовності, яка є ендегенною або нативною стосовно геному хазяїна (такого як рослина або дріжджі) і перебуває в ендегенному або нативному положенні цієї цільової послідовності в геномі хазяїна (такого як рослина або дріжджі).

60 Термін "засіб для індукції двониткового розриву", який використовується у даному документі, відноситься до будь-якої нуклеази, яка робить двонитковий розрив у цільовій послідовності. Одержання двониткового розриву в цільовій послідовності або інший ДНК можна назвати в

даному документі "розрізанням" або "розщепленням" цільової послідовності або іншої ДНК.

Термін "засіб, що рідко розщеплює, для індукції двониткового розриву", який використовується у даному документі, відноситься до будь-якої нуклеази, яка робить двонитковий розрив у цільовій послідовності, але в рідких випадках ріже (на відміну від рестрикційних ферментів, наприклад) у геномі організму. Засоби, що рідко розщеплюють, для індукції двониткового розриву, включають без обмеження ендонуклеази, такі як мегануклеази (заявка на патент США 2332 і BB1990), нуклеази "цинкові пальці" (Kim, Y. G., J. Cha, et al. (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage), ендонуклеази Cas (заявка WO 2007/025097, опублікована 1 березня 2007 року) і TALEN (Christian, M., T. Cermak, et al. 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186(2): 757-61). Розщеплення ендонуклеазами, що рідко розщеплюють, звичайно призводить до утворення "липких" кінців з 3'-липкими кінцями для мегануклеаз LAGLIDADG (Chevalier, B. S. and B. L. Stoddard. 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* 29(18): 3757-74) і 5'-липкими кінцями для нуклеаз "цинкові пальці" (Smith, J., M. Bibikova, et al. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger Dna-recognition domains. *Nucleic Acids Res* 28(17): 3361-9). TALE-нуклеази на основі FokI (TALEN) характеризуються функціональною схемою, подібної до нуклеаз "цинкові пальці", причому ДНК-зв'язуючий домен "цинкові пальці" замінений на домен TALE (Li, T., S. Huang, et al. 2011. TAL nucleases (TALENs): hybrid proteins composed of TAL effectors and FokI Dna-cleavage domain. *Nucleic Acids Res* 39(1): 359-72; Christian, M., T. Cermak, et al. 2010). Розщеплення ендонуклеазами Cas, такими як ендонуклеази Cas9, може призводити до тупих кінців.

"Ендонуклеаза" відноситься до ферменту, який розщеплює фосфодієфірний зв'язок у полінуклеотидному ланцюзі.

Ендонуклеази включають рестрикційні ендонуклеази, які розщеплюють ДНК у конкретних сайтах без ушкодження основ. Рестрикційні ендонуклеази включають ендонуклеази типу I, типу II, типу III і типу IV, які додатково включають підтипи. У системах типу I і типу III ферментативні активності як метилази, так і рестриктази об'єднані в єдиний комплекс.

Рестрикційні ендонуклеази типу I і типу III розпізнають специфічні сайти розпізнавання, але, як правило, розщеплюють у положенні, відмінному від сайту розпізнавання, яке може бути розташованим на відстані сотень пар основ від сайту розпізнавання. У системах типу II рестрикційна активність не залежить від будь-якої метилазної активності, і розщеплення зазвичай відбувається в конкретних сайтах (або точках) у межах сайту розпізнавання або біля нього. Більшість ферментів типу II розрізають паліндромні послідовності, однак ферменти типу IIa розпізнають непаліндромні сайти розпізнавання й роблять розщеплення за межами сайту розпізнавання; ферменти типу IIb роблять розрізання послідовностей двічі, причому обидва сайти перебувають за межами сайту розпізнавання; і ферменти типу IIc розпізнають асиметричний сайт розпізнавання й роблять розщеплення з одного боку і на певній відстані від сайту розпізнавання, яка становить приблизно 1-20 нуклеотидів. Ферменти рестрикції типу IV цілеспрямовано діють на метильовану ДНК. Ферменти рестрикції додатково описані й класифіковані, наприклад, у базі даних REBASE (веб-сторінка на rebase.neb.com; Roberts et al., (2003) *Nucleic Acids Res* 31:418-20), Roberts et al., (2003) *Nucleic Acids Res* 31:1805-12; і Belfort et al., (2002) в *Mobile DNA II*, pp. 761-783, Eds. Craigie et al., (ASM Press, Washington, DC).

"Сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву" відноситься до будь-якого засобу, що рідко розщеплює, для індукції двониткового розриву, який є сконструйованим (модифікованим або похідним) у порівнянні з його нативною формою для специфічного розпізнавання необхідного сайту розпізнавання та індукції двониткового розриву в ньому. Таким чином, сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, можна одержати з нативної, що зустрічається в природних умовах, нуклеази, або його можна штучно створити або синтезувати. Нуклеаза може модифікуватись всього лише по одному нуклеотиду. У деяких варіантах здійснення сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву індукує двонитковий розрив у сайті розпізнавання, при цьому сайт розпізнавання не являвся послідовністю, яка повинна була бути розпізнаною нативним (несконструйованим або немодифікованим) засобом, що рідко розщеплює, для індукції двониткового розриву. Одержання двониткового розриву в сайті розпізнавання або інший ДНК можна назвати в даному документі "розрізанням" або "розщепленням" сайту розпізнавання або іншої ДНК.

"Мегануклеаза" відноситься до хомінг-ендонуклеази, яка, подібно до рестрикційних ендонуклеаз, зв'язується з конкретним сайтом розпізнавання й розрізає його, однак сайти розпізнавання для мегануклеаз звичайно є більшими, маючи розмір приблизно 18 п.о. або

більше. У деяких варіантах здійснення даного винаходу мегануклеаза була сконструйована (або модифікована) для розрізання конкретної ендегенної послідовності розпізнавання, де ендегенна цільова послідовність перед розрізанням за допомогою сконструйованого засобу для індукції двониткового розриву, не була послідовністю, яка повинна була бути розпізнаною нативною (несконструйованою або немодифікованою) ендеонуклеазою.

"Поліпептид с мегануклеазною активністю" відноситься до поліпептиду, що характеризується мегануклеазною активністю і, отже, здатний робити двонитковий розрив у розпізнаванні послідовності.

Мегануклеази були розподілені на чотири сімейства на основі консервативних мотивів послідовностей, при цьому сімейства являють собою сімейства LAGLIDADG, GIY-YIG, H-N-H і His-Cys-бокс. Ці мотиви залучені у координацію іонів металів і гідроліз фосфодіефірних зв'язків. Нееази прикметні своїми довгими сайтами розпізнавання й тим, що допускають деякі поліморфізми послідовностей у їх ДНК-субстратах. Правила номенклатури мегануклеаз схожі на правила для інших рестрикційних ендеонуклеаз. Мегануклеази також характеризують за префіксом F-, I- або PI- для ферментів, які кодуються автономними відкритими рамками зчитування, інтронами та інтеїнами, відповідно. Наприклад, мегануклеази від *Saccharomyces cerevisiae*, які кодуються інтронами, інтеїнами та автономними генами, позначаються I-SceI, PI-SceI і F-SceII, відповідно. Домени, структура та функції мегануклеаз відомі; див., наприклад, Guhan and Muniyappa (2003) Crit Rev Biochem Mol Biol 38:199-248; Lucas et al., (2001) Nucleic Acids Res 29:960-9; Jurica and Stoddard, (1999) Cell Mol Life Sci 55:1304-26; Stoddard, (2006) Q Rev Biophys 38:49-95; і Moure et al., (2002) Nat Struct Biol 9:764. У деяких прикладах застосовують, варіант, який зустрічається в природі та/або сконструйовану мегануклеазу-похідне. Відомі способи модифікації кінетичних характеристик, взаємодій з кофакторами, експресії, оптимальних умов і/або специфічності відносно сайту розпізнавання та скринінгу відносно активності, див., наприклад, Epinat et al., (2003) Nucleic Acids Res 31:2952-62; Chevalier et al., (2002) Mol Cell 10:895-905; Gimble et al., (2003) Mol Biol 334:993-1008; Seligman et al., (2002) Nucleic Acids Res 30:3870-9; Sussman et al., (2004) J Mol Biol 342:31-41; Rosen et al., (2006) Nucleic Acids Res 34:4791-800; Chames et al., (2005) Nucleic Acids Res 33:e178; Smith et al., (2006) Nucleic Acids Res 34:e149; Gruen et al., (2002) Nucleic Acids Res 30:e29; Chen and Zhao, (2005) Nucleic Acids Res 33:e154; WO 2005105989; WO 2003078619; WO 2006097854; WO 2006097853; WO 2006097784 і WO 2004031346.

У даному документі можна застосовувати будь-яку мегануклеазу, у тому числі, без обмежень, I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-CeuI, I-CeuAIP, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-TliI, I-PpoI, PI-PspI, F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-Amal, I-AnII, I-Chul, I-Cmoel, I-Cpal, I-Cpall, I-Csml, I-Cvul, I-CvuAIP, I-DdII, I-DdIII, I-Dirl, I-Dmol, I-Hmul, I-Hmull, I-HsNIP, I-Llal, I-Msol, I-Naal, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-Njal, I-Nsp236IP, I-PakI, I-PbolIP, I-PculIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PbplIP, I-SpBetaIP, I-Scal, I-SexIP, I-SnelIP, I-SpomI, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiSTe3bP, I-TdelIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPAIP, I-UarHGPA13P, I-VinIP, I-ZbilIP, PI-Mtul, PI-MtuHIP, PI-MtuHIIP, PI-Pful, PI-Pfull, PI-Pkol, PI-Pkoll, PI-Rma43812IP, PI-SpBetaIP, PI-SceI, PI-Tful, PI-Tfull, PI-Thyl, PI-TliI, PI-TliII, або будь-які їхні активні варіанти або фрагменти. У конкретному варіанті здійснення сконструйована ендеонуклеаза походить від I-Cre-I, що має послідовність, наведену під SEQ ID NO: 15, 21 або 26, або її активного варіанта або фрагмента.

TAL-ефекторні нуклеази являють собою новий клас нуклеаз, специфічних для послідовностей, які можна застосовувати для одержання двониткових розривів у конкретних цільових послідовностях у геномі рослини або іншого організму. Tal-ефекторні нуклеази створюють шляхом злиття нативного або сконструйованого ефектора, подібного активатора транскрипції (TAL), або функціональної частини такого, з каталітичним доменом ендеонуклеази, такої як, наприклад, FokI. Унікальний модульний ДНК-зв'язуючий домен TAL-ефектора створює можливість розробки білків з потенційною специфічністю розпізнавання будь-якої заданої ДНК. Таким чином, ДНК-зв'язуючі домени TAL -ефекторних нуклеаз можна конструювати для розпізнавання конкретних цільових сайтів ДНК і, таким чином, застосовувати для отримання двониткових розривів в необхідних цільових послідовностях. Див. WO 2010/079430; Morbitzer et al. (2010) PNAS 10.1073/pnas.1013133107; Scholze & Boch (2010) Virulence 1:428-432; Christian et al. Genetics (2010) 186:757-761; Li et al. (2010) Nuc. Acids Res. (2010) doi:10.1093/nar/gkq704; і Miller et al. (2011) Nature Biotechnology 29:143-148; всі з яких включені в даний документ як посилання.

Як використовується в даному документі, вираз "ген Cas" відноситься до гена, який, як правило, з'єднаний, зв'язаний або розташований поблизу або поруч із фланкуючими локусами

CRISPR.

CRISPR-локуси (короткі паліндромні повтори, регулярно розташовані групами) (також відомі як SPIDR - дисперговані повтори) складають сімейство нещодавно описаних локусів ДНК. CRISPR-локуси складаються з коротких та висококонсервативних повторів ДНК (зазвичай від 24 до 40 п.о., що повторюються від 1 до 140 разів, також позначуваних як CRISPR-повтори), які є частково паліндромними. Повторювані послідовності (зазвичай специфічні стосовно виду) відділені варіабельними послідовностями постійної довжини (зазвичай від 20 до 58, залежно від CRISPR-локусу (WO 2007/024097, опублікована 1 березня 2007 р.).

CRISPR-локуси вперше були ідентифіковані в *E. coli* (Ishino et al. (1987) J. Bacteriol. 169:5429-5433; Nakata et al. (1989) J. Bacteriol. 171:3553-3556). Подібні розсіяні короткі повторювані послідовності були ідентифіковані в *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena* і *Mycobacterium tuberculosis* (Groenen et al. (1993) Mol. Microbiol. 10:1057-1065; Hoe et al. (1999) Emerg. Infect. Dis. 5:254-263; Masepohl et al. (1996) Biochim. Biophys. Acta 1307:26-30; Mojica et al. (1995) Mol. Microbiol. 17:85-93). CRISPR-локуси відрізняються від інших SSR структурою повторів, які були названі короткими рівномірно розділеними повторами (SRSR) (Janssen et al. (2002) OMICS J. Integ. Biol. 6:23-33; Mojica et al. (2000) Mol. Microbiol. 36:244-246). Повтори являють собою короткі елементи, які зустрічаються в кластерах, які завжди рівномірно розділені варіабельними послідовностями постійної довжини (Mojica et al. (2000) Mol. Microbiol. 36:244-246).

Терміни "ген Cas", "CRISPR-асоційований (Cas) ген" використовуються в даному документі взаємозамінно. Вичерпний огляд щодо сімейства білків Cas представлений в Haft et al. (2005) Computational Biology, Plos Comput Biol 1(6): e60. doi:10.1371/journal.pcbi.0010060. Як описано в даному документі, описано 41 сімейство CRISPR-асоційованих (Cas) генів, крім чотирьох сімейств генів, відомих раніше. Показано, що CRISPR-системи належать до різних класів, з різними патернами повторів, наборами генів і сукупностями видів. Число генів Cas у певному CRISPR-локусі може різнитися між видами.

Термін "ендонуклеаза Cas", який використовується в даному документі, відноситься до білка Cas, що кодується геном Cas, причому зазначений білок Cas здатний до введення двониткового розриву в цільову послідовність ДНК. Ендонуклеаза Cas розмотує ДНК-дуплекс поблизу геномного цільового сайту й розщеплює обидві нитки ДНК після розпізнавання цільової послідовності направляючою РНК, але тільки якщо правильний прилягаючий до протоспейсера мотив (PAM) приблизно орієнтований на 3'-кінці цільової послідовності.

Термін "направляюча РНК", який використовується у даному документі відноситься до синтетичного злиття двох молекул РНК, сРНК (РНК CRISPR), що містить варіабельний направляючий домен, і tracrRNA. В одному варіанті здійснення направляюча РНК містить варіабельний направляючий домен з послідовностей від 12 до 30 нуклеотидів і фрагмент РНК, який може взаємодіяти з ендонуклеазою Cas.

Вираз "варіабельний направляючий домен" відноситься до нуклеотидної послідовності 5-прим мотиву послідовності GUUUU в направляючій РНК, яка є комплементарною одному ланцюгу двониткового цільового сайту ДНК у геномі рослинної клітини, рослини або насіння. В одному варіанті здійснення варіабельний направляючий домен має довжину від 12 до 30 нуклеотидів.

В одному варіанті здійснення направляюча РНК та ендонуклеаза Cas здатні утворювати комплекс, який дозволяє ендонуклеазі Cas вводити двонитковий розрив у цільовий сайт ДНК.

Термін "сайт розпізнавання", який використовується у даному документі, відноситься до послідовності ДНК, у якій індукується двонитковий розрив у геномі клітини за допомогою засобу, що рідко розщеплює, для індукції двониткового розриву. Терміни "сайт розпізнавання", "послідовність розпізнавання" використовуються в даному документі взаємозамінно. Сайт розпізнавання може бути ендегенним сайтом у геномі хазяїна (такого як дріжджі або рослина), або в якості альтернативи, сайт розпізнавання може бути гетерологічним відносно хазяїна (дріжджів або рослини), і, таким чином, не зустрічатися в природі в геномі, або сайт розпізнавання може перебувати в гетерологічному положенні в геномі у порівнянні з тим положенням, у якому він перебуває в природі.

Термін "ендогенний сайт розпізнавання", який застосовується у даному документі, відноситься до сайту розпізнавання засобом, що рідко розщеплює, для індукції двониткового розриву, який є ендегенним або нативним стосовно генома хазяїна (такого як рослина або дріжджі) і розташований в ендегенному або нативному положенні для такого сайту розпізнавання в геномі хазяїна (такого як рослина або дріжджі). Довжина сайту розпізнавання може варіювати, при цьому включені, наприклад, сайти розпізнавання, що мають довжину, що складає щонайменше 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,

32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70 або більше нуклеотидів. Крім того, є можливість того, що сайт розпізнавання може бути паліндромним, тобто послідовність однієї нитки читається так само у зворотному напрямку на комплементарній нитці. Сайт одониткового розриву/розщеплення може знаходитись в межах послідовності розпізнавання, або сайт одониткового розриву/розщеплення може знаходитись за межами послідовності розпізнавання. В іншому варіанті розщеплення може відбуватися в положеннях нуклеотидів безпосередньо навпроти один одного з одержанням розрізу з тупими кінцями, або, в інших випадках, розрізи можуть бути несиметрично розташованими з одержанням одониткових виступаючих кінців, також названих "липкими кінцями", які можуть бути 5'-липкими кінцями або 3'-липкими кінцями.

Термін "припустимий сайт розпізнавання", який використовується у даному документі, відноситься до послідовності розпізнавання, на яку було спрямовано сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, такий як сконструйована мегануклеаза, для специфічного розпізнавання та індукції двониткового розриву. В одному варіанті здійснення засіб, що рідко розщеплює, для індукції двониткового розриву являє собою сконструйовану мегануклеазу LIG3-4 (SEQ ID NO: 2), яка була розроблена для розпізнавання припустимої послідовності розпізнавання SEQ ID NO: 13 (публікація заявки на патент США 2009-0133152 A1, опублікована 21 травня 2009 року). В іншому варіанті здійснення засіб, що рідко розщеплює, для індукції двониткового розриву являє собою сконструйовану мегануклеазу MHP14+ (SEQ ID NO: 4), яка була розроблена для розпізнавання припустимої послідовності розпізнавання SEQ ID NO: 14 (у заявці на патент США 13/427138, поданої 22 березня 2012 року).

Термін "варіантний сайт розпізнавання", який використовується у даному документі, відноситься до варіантної нуклеотидної послідовності, яка містить зміну щонайменше по одній нуклеотидній основі у порівнянні з припустимим сайтом розпізнавання, на який було спрямовано сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, такий як мегануклеаза, для специфічного розпізнавання та індукції двониткового розриву. Така "зміна" включає, наприклад: (I) заміщення щонайменше одного нуклеотиду, (II) делецію щонайменше одного нуклеотиду, (III) вставку щонайменше одного нуклеотиду або (IV) будь-яку комбінацію з (I)-(III). Активні варіанти й фрагменти сайту розпізнавання можуть характеризуватися щонайменше 65 %, 70 %, 75 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % або більшою ідентичністю послідовності стосовно даної послідовності розпізнавання, при цьому активні варіанти зберігають біологічну активність і, отже, здатні розпізнаватись та розщеплюватись ендонуклеазою. Варіантні сайти розпізнавання можуть містити щонайменше один (1) і аж до 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 або 20 переважно сторонніх нуклеотидів. В одному варіанті здійснення варіантні сайти розпізнавання є неендогенними стосовно генома хазяїна, такі варіантні сайти розпізнавання включають без обмеження варіантні сайти розпізнавання маїсу, показані на фігурі 4 (SEQ ID NO: 15-22 і SEQ ID NO: 23-35). В іншому варіанті здійснення варіантні сайти розпізнавання присутні в геномі хазяїна (називаються геномними варіантними сайтами розпізнавання) або є ендогенними стосовно генома хазяїна, як наприклад, геноми рослин або дріжджів. У деяких варіантах здійснення варіантні сайти розпізнавання можна ввести в геном рослини за допомогою мутагенезу ендогенної геномної послідовності. Способи сайт-специфічного мутагенезу геномної ДНК відомі на технічному рівні та включають способи, описані, наприклад, у патенті США № 5565350, патенті США № 5731181 і патенті США № 6870075. Інші способи включають застосування нуклеаз "цинкові пальці", як наприклад, ті способи, які описані в публікації патентного документа США 20050208489.

"Геномний варіантний сайт розпізнавання" відноситься до варіантного сайту розпізнавання засобом, що рідко розщеплює, для індукції двониткового розриву, такого як мегануклеаза, який є ендогенним стосовно генома організму (такого як рослина або дріжджі). Одним прикладом варіантного сайту розпізнавання, який є ендогенним стосовно генома маїсу, є SEQ ID NO: 11.

Терміни "переважно сторонні нуклеотиди" або "перевага сторонніх нуклеотидів" можна використовувати взаємозамінно, і вони відносяться до нуклеотидів, які розташовані в тому ж положенні відносно нуклеотидів припустимого сайту розпізнавання, але є більш розповсюдженими в ідентифікованих геномних варіантних сайтах розпізнавання (див., наприклад, поширеність для +8T (80 %) у порівнянні з припустимим сайтом розпізнавання +8C (13 %) на фігурі 3A). У більшості випадків переважно сторонній нуклеотид, коли його поміщають у припустимий сайт розпізнавання, розщеплюється в більшому відсотку випадків, ніж припустимий сайт розпізнавання (див., наприклад, +8T (96 % розщеплення) у порівнянні з +8C (80 % розщеплення на фігурі 5A).

В одному варіанті здійснення припустима послідовність розпізнавання сконструйованою мегануклеазою LIG3-4 містить SEQ ID NO: 13, тоді як варіантний сайт розпізнавання сконструйованою мегануклеазою LIG3-4 містить SEQ ID NO: 15, 16, 17, 18, 19, 20, 21 або 22. В іншому варіанті здійснення припустима послідовність розпізнавання сконструйованою мегануклеазою MHP14+ містить SEQ ID NO: 14, тоді як варіантний сайт розпізнавання сконструйованою мегануклеазою MHP14+ містить SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 або 35.

"Локус варіантного сайту розпізнавання" являє собою положення на хромосомі, яка містить варіабельний сайт розпізнавання. Переважно, локус варіантного сайту розпізнавання перебуває в межах 0, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900 або 1000 пар основ від варіантного сайту розпізнавання.

Термін "мегануклеазна активність", який використовується у даному документі, відноситься до здатності мегануклеази робити двонитковий розрив у бажаній послідовності розпізнавання і, таким чином, до збереження активності відносно індукції двониткового розриву. Одержання двониткового розриву в послідовності розпізнавання або іншій ДНК може називатися в даному документі "розрізанням" або "розщепленням" послідовності розпізнавання або іншої ДНК.

Є відомі аналізи мегануклеазної активності, і в них, як правило, вимірюють загальну активність і специфічність мегануклеази на ДНК-субстратах що містять сайти розпізнавання. Ці ДНК-субстрати включають без обмеження геномну ДНК і плазмідну ДНК. Наприклад, мегануклеазну активність можна виміряти *in vitro*, як описано в даному документі в прикладі 3 і прикладі 9. Коротко, процедури розщеплення в динаміці можна виконувати при 37 °C, 28 °C і 23 °C (або будь-якій температурі в діапазоні між 37 °C, 36 °C, 35 °C, 34 °C, 33 °C, 32 °C, 31 °C, 30 °C, 29 °C, 28 °C, 27 °C, 26 °C, 25 °C, 24 °C і 23 °C) (а плазмідній або геномній ДНК, що містить сайт розпізнавання мегануклеазою, і % розщеплення в кожному зразку (який також називають % розщеплення або % втрати сайтів розпізнавання мегануклеазою, що є показником мегануклеазної активності) можна визначити за допомогою ПЛР у реальному часі.

Мегануклеазну активність також можна виміряти з використанням скринінг-аналізу на дріжджах, який описаний у даному документі (фігури 8 і 9 і приклад 16). Коротко, клітини дріжджів з функціональним геном Ade2 є білими, тоді як такі, у яких відсутнє функціонування Ade2, проявляють червоне забарвлення у зв'язку з накопиченням метаболіту на більш ранньому етапі біосинтезу аденіну, що в результаті призводить до утворення червоних колоній з білими секторами. Ступінь утворення білих секторів, що інколи поширюється на цілі колонії, вказує на величину активності розрізання, властивої мегануклеазам. Оскільки фенотип утворення секторів є якісним показником мегануклеазної активності, було задіяно числову систему кількісних показників 0-4. Як показано на фігурі 3, кількісний показник 0 означає, що білі сектори не спостерігалися (було відсутнє розрізання мегануклеазами); кількісний показник 4 означає повністю білі колонії (повне розрізання сайту розпізнавання); кількісні показники 1-3 означають проміжні фенотипи утворення білих секторів (та проміжні ступені розрізання сайту розпізнавання).

Додатково мегануклеазну активність можна вимірювати *in planta* шляхом визначення частоти мутацій у цільовому сайті (TS). Частота мутацій у цільовому сайті визначається в такий спосіб: (кількість трансгенних об'єктів із модифікацією в цільовому сайті/загальну кількість трансгенних об'єктів)*100 %.

Терміни "підвищена" або "збільшена" активність застосовуються в даному документі взаємозамінно. "Підвищена" або "збільшена" мегануклеазна активність включає будь-яке статистично значиме підвищення активності вихідного поліпептиду з мегануклеазною активністю, обумовлене за допомогою кожного з аналізів активності, описаних у даному документі.

Мегануклеаза може бути надана полінуклеотидом, що кодує ендонуклеазу. Такий полінуклеотид, що кодує ендонуклеазу, можна модифікувати для заміни кодонів, які мають вищу частоту використання в рослині в порівнянні з послідовністю полінуклеотида, який зустрічається в природі. Наприклад, полінуклеотид, що кодує мегануклеазу, можна модифікувати для заміни кодонів з вищою частотою використання в рослині маїсу або сої в порівнянні з послідовністю полінуклеотида, що зустрічається в природі.

Терміни "контрольна мегануклеаза" або "еталонна мегануклеаза" можна використовувати взаємозамінно, і вони відносяться до будь-якої мегануклеази, з якою порівнюють варіантну мегануклеазу. Контрольні мегануклеази можуть включати, без обмежень, вихідні або відповідні мегануклеази або будь-які мегануклеази дикого типу типу I-crei.

Нумерація полімеру на основі амінокислот або нуклеотидів, такого як будь-яка мегануклеаза по даному винаходу, відповідає нумерації обраного полімеру на основі амінокислот або

нуклеїнової кислоти, якщо положення даного мономерного компонента (амінокислотного залишку, включеного нуклеотида і т. д.) полімеру відповідає положенню того ж залишку в обраному еталонному поліпептиді або полінуклеотиді.

Як використовується в даному документі, "ділянка генома, що представляє інтерес" являє собою сегмент хромосоми в геномі рослини, введення полінуклеотиду, що представляє інтерес, або ознаки, що представляє інтерес. Ділянка генома, що представляє інтерес, може включати, наприклад, один або декілька полінуклеотидів, що представляють інтерес. Як правило, ділянка генома, що представляє інтерес, по даному винаходу містить сегмент хромосоми, розмір якого становить 0-15 сантиморганід (сМ).

Термін "полінуклеотид, що представляє інтерес", який використовується в даному документі, "у ділянці генома, що представляє інтерес", являє собою будь-яку, кодує та/або не кодує частину ділянки генома, що представляє інтерес, в тому числі, без обмеження трансгена, нативного гена, мутованого гена і генетичного маркера, такого як, наприклад, маркер одонуклеотидного поліморфізму (SNP) і маркер простої повторюваної послідовності (SSR).

Як використовується в даному документі, "фізично зчеплений", "у фізичному зчепленні" і "генетично зчеплений", використовуються для позначення будь-яких двох або більше генів, трансгенів, нативних генів, мутованих генів, змін, цільових сайтів, маркерів і подібного, які є частиною однієї молекули ДНК або хромосоми.

Як застосовується в даному документі, "виділений" полінуклеотид або поліпептид або його біологічно активна частина є по суті або практично вільними від компонентів, що звичайно супроводжують полінуклеотид або поліпептид, або взаємодіючих з ними, які існують в його природному середовищі. Таким чином, виділений або очищений полінуклеотид або поліпептид є по суті вільним від іншого клітинного матеріалу або культурального середовища при одержанні за допомогою методик рекомбінантних молекул або по суті вільними від хімічних попередників або інших хімічних речовин при хімічному синтезі. В оптимальному випадку "виділений" полінуклеотид не містить послідовностей (в оптимальному випадку послідовностей, що кодують білок), які в природних умовах фланкують полінуклеотид (тобто послідовностей, розташованих на 5'- і 3'-кінцях полінуклеотида) у геномній ДНК організму, з якого отриманий полінуклеотид. Наприклад, у різних варіантах здійснення виділений полінуклеотид може містити нуклеотидну послідовність розміром менш ніж приблизно 5 т.о., 4 т.о., 3 т.о., 2 т.о., 1 т.о., 0,5 т.о. або 0,1 т.о., яка в природних умовах фланкує полінуклеотид у геномній ДНК клітини, з якої отриманий полінуклеотид. Поліпептид, який по суті вільний від клітинного матеріалу, включає препарати поліпептидів, які мають менш ніж приблизно 30 %, 20 %, 10 %, 5 % або 1 % (по сухій вазі) забруднюючого білка. Якщо поліпептид по даному винаходу або його біологічно активна частина отримані рекомбінантним шляхом, то, в оптимальному випадку, у культуральному середовищі представлено менш ніж приблизно 30 %, 20 %, 10 %, 5 % або 1 % (по сухій вазі) хімічних попередників або хімічних речовин, які не є білком, який представляє інтерес.

Як застосовується в даному документі, полінуклеотид або поліпептид є "рекомбінантними", якщо вони є штучними або сконструйованими або отримані зі штучних або сконструйованих білка або нуклеїнової кислоти. Наприклад, полінуклеотид, що вставляється у вектор або будь-яке інше гетерологічне положення, наприклад, у геном рекомбінантного організму, так, що він не пов'язаний із нуклеотидними послідовностями, які в природних умовах фланкують полінуклеотид, як це виявляється в природі, є рекомбінантним полінуклеотидом. Поліпептид, що утворився в результаті *in vivo* або *in vitro* експресії рекомбінантного полінуклеотида, є прикладом рекомбінантного поліпептиду. Аналогічно, послідовність полінуклеотида, яка не зустрічається в природі, наприклад, варіант гена, що зустрічається в природі, є рекомбінантною.

"Підпослідовність" або "фрагмент" означає будь-яку частину повної послідовності.

Порівняння послідовностей

Для опису взаємостосунків послідовностей двох або більше полінуклеотидів або поліпептидів використовують наступні терміни: (a) "еталонна послідовність", (b) "вікно порівняння", (c) "ідентичність послідовностей" і (d) "відсоткова ідентичність послідовностей".

(a) Як застосовується в даному документі, "еталонна послідовність" є заданою послідовністю, яка застосовується в якості основи для порівняння послідовностей. Еталонна послідовність може бути частиною певної послідовності або всією такою; наприклад, у вигляді сегмента кДНК повної довжини або послідовності гена або всієї кДНК, або послідовності гена, або білкової послідовності.

(b) Як застосовується в даному документі, "вікно порівняння" відноситься до суміжного й певного сегмента послідовності поліпептиду, де послідовність поліпептиду у вікні порівняння може включати додавання або делеції (тобто гепи) у порівнянні з еталонною послідовністю (яка не містить додавань або делецій) для оптимального вирівнювання двох поліпептидів. Зазвичай

довжина вікна порівняння становить щонайменше 5, 10, 15 або 20 суміжних амінокислот, або вона може становити 30, 40, 50, 100 або більше. Фахівці в даній області розуміють, що для того, щоб уникнути високого ступеня подібності відносно до еталонної послідовності внаслідок включення гепів у послідовність поліпептиду, як правило, вводиться штраф за відкриття гепів, і його вираховують із кількості збігів.

Способи вирівнювання послідовностей для порівняння добре відомі на технічному рівні. Таким чином, визначення відсоткової ідентичності послідовності між будь-якими двома послідовностями можна виконати з використанням математичного алгоритму. Необмежувальними прикладами таких математичних алгоритмів є алгоритм по Myers and Miller (1988) CABIOS 4:11-17; алгоритм локального вирівнювання по Smith et al. (1981) Adv. Appl. Math. 2:482; алгоритм глобального вирівнювання по Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453; спосіб пошуку локального вирівнювання по Pearson and Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; алгоритм по Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 90:5873-5877. Для порівняння послідовностей можна використовувати комп'ютерні реалізації даних математичних алгоритмів, щоб визначити ідентичність послідовностей. Такі реалізації включають без обмежень CLUSTAL у програмі PC/Gene (доступна від Intelligenetics, Маунтін-В'ю, Каліфорнія); програму ALIGN (версія 2.0) і GAP, BESTFIT, BLAST, FASTA і TFASTA в GCG Wisconsin Genetics Software Package, версія 10 (доступні від Accelrys Inc., 9685 Scranton Road, Сан-Дієго, Каліфорнія, США). Вирівнювання за допомогою цих програм можуть бути виконані з використанням параметрів за замовчуванням. Програма CLUSTAL добре описана в Higgins et al. (1988) Gene 73:237-244 (1988); Higgins et al. (1989) CABIOS 5:151-153; Corpet et al. (1988) Nucleic Acids Res. 16:10881-90; Huang et al. (1992) CABIOS 8:155-65; і Pearson et al. (1994) Meth. Mol. Biol. 24:307-331. Програма ALIGN заснована на алгоритмі з Myers and Miller (1988), згаданому вище. При порівнянні амінокислотних послідовностей у програмі ALIGN можна застосовувати таблицю ваги заміни залишків PAM120, штраф за продовження гепів 12 і штраф за відкриття гепів 4. Програми BLAST по Altschul et al (1990) J. Mol. Biol. 215:403; засновані на алгоритмі Karlin and Altschul (1990), згадані вище. Пошуки нуклеотидних послідовностей в BLAST можна виконувати за допомогою програми BLASTN при параметрах вага вирівнювання = 100, довжина слова = 12 з одержанням нуклеотидних послідовностей, гомологічних нуклеотидній послідовності, що кодує білок згідно із даним винаходом. Пошуки білкових послідовностей в BLAST можна виконувати за допомогою програми BLASTX при параметрах вага вирівнювання = 50, довжина слова = 3 з одержанням амінокислотних послідовностей, гомологічних білку або поліпептиду згідно із даним винаходом. Пошуки білкових послідовностей в BLASTP можна виконувати з використанням параметрів за замовчуванням. Див. blast.ncbi.nlm.nih.gov/Blast.cgi.

Вирівнювання послідовностей і розрахунки відсоткової подібності можна виконувати за допомогою програми Megalign з комплекту біоінформаційних обчислювальних програм LASARGENE (DNASTAR Inc., Медісон, Вісконсин) або за допомогою програми AlignX з комплекту біоінформаційних обчислювальних програм Vector NTI (Invitrogen, Карлсбад, Каліфорнія). Множинне вирівнювання послідовностей виконують із використанням способу вирівнювання Clustal (Higgins and Sharp, CABIOS. 5:151-153 (1989)) з параметрами за замовчуванням (ШТРАФ ЗА ВІДКРИТТЯ ГЕПІВ=10, ШТРАФ ЗА ПРОДОВЖЕННЯ ГЕПІВ=10). Параметри за замовчуванням для парного вирівнювання й розрахунків відсотку ідентичності білкових послідовностей за допомогою програми Clustal є наступними: РОЗМІР ІДЕНТИЧНОЇ ДІЛЯНКИ=1, ШТРАФ НА ВНЕСЕННЯ ДЕЛЕЦІЇ=3, ДОВЖИНА СЕГМЕНТА=5 і ЧИСЛО БЕЗПЕРЕРВНО СПІВПАДАЮЧИХ ДІЛЯНОК=5. Для нуклеїнових кислот ці параметри є наступними: ШТРАФ ЗА ВІДКРИТТЯ ГЕПІВ=10, ШТРАФ ЗА ПРОДОВЖЕННЯ ГЕПІВ=10, KUPLE=2, ШТРАФ ЗА ВІДКРИТТЯ ГЕПІВ=5, ВІКНО=4 і ЗБЕРЕЖЕНІ ДІАГОНАЛІ=4. "Значна частина" амінокислотної або нуклеотидної послідовності включає достатню частину амінокислотної послідовності поліпептиду або нуклеотидної послідовності гена для забезпечення припустимої ідентифікації цього поліпептиду або гена шляхом оцінки послідовності фахівцем у даній області техніки вручну або шляхом автоматизованого за допомогою комп'ютера порівняння та ідентифікації послідовностей за допомогою алгоритмів, таких як BLAST (Altschul, S. F. et al., J. Mol. Biol. 215:403-410 (1993)) і BLAST з гепами (Altschul, S. F. et al., Nucleic Acids Res. 25:3389-3402 (1997)). BLASTN відноситься до програми BLAST, що порівнює запитувану нуклеотидну послідовність із нуклеотидними послідовностями з бази даних.

"Ген" відноситься до фрагменту нуклеїнової кислоти, який експресує специфічний білок, включаючи регуляторні послідовності, що передують (5'-некодуючі послідовності) та ідуть після

(3'-некодуючі послідовності) кодуєючої послідовності. Термін "нативний ген" означає ген, який існує в природі, зі своїми власними регуляторними послідовностями. "Химерний ген" або "рекомбінантна експресуюча конструкція", застосовуються взаємозамінно, відносяться до будь-якого гена, який не є нативним геном, що включає регуляторні та кодуєчі послідовності, які разом не зустрічаються в природі. Відповідно, химерний ген може містити регуляторні послідовності та кодуєчі послідовності, які отримані з різних джерел, або регуляторні послідовності та кодуєчі послідовності, які отримані з одного джерела, але розташовані в порядку, відмінному від такого, що зустрічається в природі. "Ендогенний ген" відноситься до нативного гена в його природному розташуванні в геномі організму. "Чужорідний" ген відноситься до гена, що не виявляється в нормі в організмі хазяїна, але який введений в організм хазяїна за допомогою переносу гена. Чужорідні гени можуть включати нативні гени, введені в організм, який не є нативним, або химерні гени. "Трансген" є геном, який ввели в генوم за допомогою способу трансформації.

"Кодуюча послідовність" відноситься до послідовності ДНК, яка кодує конкретну амінокислотну послідовність.

"Регуляторні послідовності" відносяться до нуклеотидних послідовностей, розташованих вище (5'-некодуючі послідовності), в межах або нижче (3'-некодуючі послідовності) кодуєючої послідовності, і таких, що впливають на транскрипцію, процесинг або стабільність РНК або трансляцію пов'язаної кодуєючої послідовності. Регуляторні послідовності можуть включати без обмеження промотори, лідерні послідовності трансляції, інтрони та послідовності розпізнавання поліаденілювання.

"Виродженість кодонів" відноситься до відхилення в генетичному коді, що дозволяє робити зміну нуклеотидної послідовності без впливу на амінокислотну послідовність поліпептиду, який кодується. Відповідно, даний винахід відноситься до будь-якого фрагмента нуклеїнової кислоти, що містить нуклеотидну послідовність, яка кодує всі амінокислотні послідовності, викладені в даному документі, або їх значну частину. Фахівцям в даній області техніки добре відомо про "зсув частоти використання кодонів", що проявляється конкретною клітиною-хазяїном відносно використання нуклеотидних кодонів, що визначають дану амінокислоту. Таким чином, при синтезі фрагменту нуклеїнової кислоти для підвищеної експресії в клітині-хазяїні бажано розробити такий фрагмент нуклеїнової кислоти, щоб його частота використання кодонів була близькою до частоти використання переважних кодонів клітини-хазяїна.

Як використовується в даному документі, "ідентичність послідовності" або "ідентичність" по відношенню до двох послідовностей полінуклеотидів або поліпептидів відноситься до залишків двох послідовностей, які є однаковими при вирівнюванні зі знаходженням максимальної відповідності в зазначеному вікні порівняння. Якщо відсоток ідентичності послідовностей застосовують відносно білків, то враховують, що положення залишків, які не є ідентичними, часто відрізняються по консервативних амінокислотних замінах, де амінокислотні залишки замінені на інші амінокислотні залишки з подібними хімічними властивостями (наприклад, зарядом або гідрофобністю). Якщо послідовності відрізняються консервативними замінами, то відсоток ідентичності послідовності можна підвищити для того, щоб внести поправку на консервативну природу заміни. Кажуть, що послідовності, які відрізняються такими консервативними замінами, характеризуються "подібністю послідовностей" або "подібністю". Засоби для здійснення такого коректування добре відомі фахівцям у даній області. Як правило, це передбачає оцінку в балах консервативної заміни скоріше в якості часткової, ніж повної розбіжності, що, таким чином, збільшує відсоток ідентичності послідовностей. Таким чином, наприклад, якщо ідентичній амінокислоті присвоюється бал 1, а неконсервативній заміні присвоюється бал нуль, то консервативній заміні присвоюється бал від нуля до 1. Оцінку в балах консервативних замін розраховують, наприклад, як реалізовано в програмі PC/GENE (Intelligenetics, Маунтін-В'ю, Каліфорнія).

Як застосовується в даному документі, "відсоток ідентичності послідовностей" означає значення, визначене шляхом порівняння двох вирівняних послідовностей у вікні порівняння, де частина послідовності полінуклеотиду у вікні порівняння може включати додавання або делеції (тобто, гепи) у порівнянні з еталонною послідовністю (яка не включає додавання або делеції) для оптимального вирівнювання двох послідовностей. Відсоткове значення розраховують шляхом визначення кількості положень, у яких ідентичний залишок нуклеїнової кислоти або амінокислотний залишок зустрічається в обох послідовностях, з одержанням кількості положень, що співпали, ділення кількості положень, що співпали, на загальну кількість положень у вікні порівняння та множення результату на 100 з одержанням відсоткової ідентичності послідовностей.

Полінуклеотидні конструкції

У даному документі передбачаються молекули полінуклеотидів або нуклеїнових кислот, які містять варіантні сайти розпізнавання для засобів, що рідко розщеплюють/розщеплюють, для індукції двониткового розриву або будь-яких їхніх активних варіантів або фрагментів. Терміни "полінуклеотид", "послідовність полінуклеотида", "послідовність нуклеїнової кислоти" і "фрагмент нуклеїнової кислоти" застосовуються в даному документі взаємозамінно. Ці вирази охоплюють нуклеотидні послідовності і т. п. Використання терміна "полінуклеотид" не має на увазі обмеження даного винаходу полінуклеотидами, що містять ДНК. Фахівці в даній області розуміють, що полінуклеотиди можуть містити рибонуклеотиди та комбінації рибонуклеотидів і дезоксирибонуклеотидів. Такі дезоксирибонуклеотиди та рибонуклеотиди містять у собі як молекули, що зустрічаються в природі, так і їх синтетичні аналоги. Полінуклеотиди по даному винаходу також охоплюють усі форми послідовностей, у тому числі без обмежень одониткові форми, двониткові форми, шпильки, структури типу "стебло-петля" і т. п.

Крім того, передбачаються рекомбінантні полінуклеотиди, що містять різні засоби, що рідко розщеплюють, для індукції двониткового розриву, такі як сконструйовані мегануклеази. Терміни "рекомбінантний полінуклеотид", "рекомбінантний нуклеотид", "рекомбінантна ДНК" і "рекомбінантна ДНК-конструкція" використовуються в даному документі взаємозамінно. Рекомбінантна конструкція містить комбінацію штучних або гетерологічних послідовностей нуклеїнових кислот, наприклад, регуляторних і кодуючих послідовностей, які не зустрічаються разом у природі. Наприклад, касета для переносу може містити сайти рестрикції та гетерологічний полінуклеотид, що представляє інтерес. В інших варіантах здійснення рекомбінантна конструкція може містити регуляторні послідовності та кодуючі послідовності, які отримані з різних джерел, або регуляторні послідовності, та кодуючі послідовності, які отримані з одного джерела, але розташовані в порядку, відмінному від такого, що зустрічається в природі. Таку конструкцію можна застосовувати саму по собі або можна застосовувати в комбінації з вектором. Якщо використовувати вектор, тоді вибір вектора залежить від способу, який будуть використовувати для трансформації клітин-хазяїв, що добре відомо фахівцям у даній області. Наприклад, можна застосовувати плазмідний вектор. Фахівцям в даній області добре відомі генетичні елементи, які повинні бути присутніми у векторі для успішної трансформації, відбору та розмноження клітин-хазяїв, що містять який-небудь із виділених фрагментів нуклеїнових кислот, які забезпечуються у даному документі. Фахівцям в даній області також буде зрозуміло, що у різних незалежних трансформантів будуть проявлятися різні рівні та патерни експресії (Jones et al., EMBO J. 4:2411-2418 (1985); De Almeida et al., Mol. Gen. Genetics 218:78-86 (1989)), і таким чином, для того, щоб одержати лінії, що проявляють бажаний рівень і патерн експресії, слід провести скринінг декількох трансгенних об'єктів. Такий скринінг, окрім іншого, можна виконувати за допомогою Саузерн-аналізу ДНК, Нозерн-аналізу експресії мРНК, аналізу експресії білка за допомогою імуноблотингу або фенотипового аналізу.

Полінуклеотиди, що кодують мегануклеази, які розкриті в даному документі, можуть бути представлені в касетах експресії для експресії в рослині, що представляє інтерес. Касета може містити в собі 5'- і 3'-регуляторні послідовності, функціонально пов'язані з полінуклеотидом, що кодує мегануклеазу або її активний варіант або фрагмент. Мається на увазі, що термін "функціонально зв'язаний" означає функціональний зв'язок між двома або більше елементами. Наприклад, функціональний зв'язок між полінуклеотидом, що представляє інтерес, і регуляторною послідовністю (тобто промотором) являє собою функціональний зв'язок, який забезпечує експресію полінуклеотида, що представляє інтерес. Функціонально зв'язані елементи можуть бути суміжними або несуміжними. При використанні для позначення з'єднання двох ділянок, що кодують білок, під функціонально зв'язаним мається на увазі, що кодуючі ділянки знаходяться в одній рамці зчитування. Касета може додатково містити щонайменше один додатковий ген, який підлягає введенню в організм шляхом котрансформації. Альтернативно, додатковий (додаткові) ген(и) можуть бути представлені в декількох касетах експресії. Така касета експресії забезпечується безліччю сайтів рестрикції та/або сайтів рекомбінації для того, щоб вставка полінуклеотиду, який кодує мегануклеазу або її активний варіант або фрагмент, зазнала регуляції транскрипції з боку регуляторних ділянок. Касета експресії може додатково містити гени маркерів, які селектуються.

У напрямку транскрипції 5'-3' касета експресії може містити ділянку ініціації транскрипції й трансляції (тобто промотор), полінуклеотид, що кодує мегануклеазу або її активний варіант або фрагмент, і ділянку термінації транскрипції й трансляції (тобто ділянка термінації), які є функціональними у рослин. Регуляторні ділянки (тобто промотори, ділянки регуляції транскрипції та ділянки термінації трансляції) та/або полінуклеотид, що кодує мегануклеазу або її активний варіант або фрагмент, можуть бути нативними/аналогічними по відношенню до клітини-хазяїна або один до одного. Альтернативно, регуляторні ділянки та/або полінуклеотид,

що кодує мегануклеазу або її активний варіант або фрагмент, можуть бути гетерологічними по відношенню до клітини-хазяїна або один до одного.

Використовувана в даному документі "гетерологічна" стосовно послідовності означає, що послідовність походить із чужорідного виду або, якщо вона походить із того ж виду, то істотно модифікована по складу та/або положенню в геномі у порівнянні з її нативною формою в результаті навмисного втручання людини. Наприклад, промотор, функціонально пов'язаний з гетерологічним полінуклеотидом, походить із виду, відмінного від виду, з якого отриманий полінуклеотид, або, якщо він походить із того ж/аналогічного виду, то один або обидва з них є значною мірою модифікованими в порівнянні з їхньою вихідною формою та/або положенням в геномі, або промотор не є нативним промотором для функціонально зв'язаного полінуклеотиду.

Незважаючи на те, що експресія послідовностей за допомогою гетерологічних промоторів може бути оптимальною, можна застосовувати нативні промоторні послідовності. Такі конструкції можуть змінювати рівні експресії полінуклеотиду, що кодує мегануклеазу, у рослині або рослинній клітині. Таким чином, фенотип рослини або рослинної клітини можна змінювати.

Ділянка термінації може бути нативною відносно ділянки ініціації транскрипції, може бути нативною відносно функціонально зв'язаного полінуклеотиду, що кодує мегануклеазу або її активний варіант або фрагмент, може бути нативною відносно рослини-хазяїна, або може бути отримана з іншого джерела (тобто чужорідного або гетерологічного відносно промотора, полінуклеотида, що кодує мегануклеазу або її активний варіант або фрагмент, рослини-хазяїна або будь-якої їхньої комбінації). Підходящі ділянки термінації доступні з Ті-плазміді *A. tumefaciens*, наприклад, ділянки термінації генів октопінсинтази й нопалінсинтази. Див. також Guerineau et al. (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon et al. (1991) *Genes Dev.* 5:141-149; Mogen et al. (1990) *Plant Cell* 2:1261-1272; Munroe et al. (1990) *Gene* 91:151-158; Ballas et al. (1989) *Nucleic Acids Res.* 17:7891-7903; i Joshi et al. (1987) *Nucleic Acids Res.* 15:9627-9639.

За необхідності полінуклеотида можна оптимізувати для підвищення експресії в трансформованій рослині. Тобто полінуклеотида можна синтезувати з використанням кодонів, кращих для рослини, для поліпшення експресії. Див., наприклад, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11; відносно розгляду використання кодонів, кращих для хазяїна. З технічного рівня доступні способи синтезу генів, кращих для рослин. Див., наприклад, патенти США №№ 5380831 і 5436391; i Murray et al. (1989) *Nucleic Acids Res.* 17:477-498, включеної в даний документ як посилання.

Відомі додаткові модифікації послідовності для посилення експресії гена в клітинного хазяїна. Вони включають виключення послідовностей, що кодують фальшиві сигнали поліаденілювання, сигнали сайту сплайсингу екзонів та інтронів, транспозоноподібні повтори та інші подібні добре вивчені послідовності, які можуть згубно впливати на експресію гена. Вміст G-C у послідовності можна коректувати до рівнів, середніх для даного клітинного хазяїна, що розраховується з урахуванням відомих генів, які експресуються у клітині-хазяїні. Якщо можливо, послідовність модифікують для того, щоб уникнути утворення прогнозованих шпилькових вторинних структур мРНК.

Касети експресії можуть додатково містити 5'-лідерні послідовності. Такі лідерні послідовності можуть сприяти посиленню трансляції. Трансляційні лідерні послідовності відомі в даній області та включають лідерні послідовності пікорнавірусів, наприклад, лідерну послідовність EMCV (5'- некодуюча ділянка геному вірусу енцефаломіокардиту) (Elroy-Stein (1989) *Proc. Natl. Acad. Sci. USA* 86:6126-6130); лідерні послідовності потівірусів, наприклад, лідерну послідовність TEV (вірусу гравіювання тютюну) (Gallie et al. (1995) *Gene* 165(2):233-238), лідерну послідовність MDMV (вірусу карликової мозаїки кукурудзи) (*Virology* 154:9-20) і мРНК білка, що зв'язує важкий ланцюг імуноглобуліну людини (BiP) (Macejak et al. (1991) *Nature* 353:90-94); нетрансльовану лідерну послідовність мРНК білка оболонки вірусу мозаїки люцерни (РНК-4 AMV) (Jobling et al. (1987) *Nature* 325:622-625); лідерну послідовність вірусу тютюнової мозаїки (TMV) (Gallie et al. (1989) в *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256) і лідерну послідовність вірусу хлоротичної плямистості маїсу (MCMV) (Lommel et al. (1991) *Virology* 81:382-385. Див. також Della-Cioppa et al. (1987) *Plant Physiol.* 84:965-968.

При отриманні касети експресії з різними фрагментами ДНК можна проводити маніпуляції так, щоб одержати послідовності ДНК у належній орієнтації та, при необхідності, у належній рамці читування. Із цією метою для з'єднання фрагментів ДНК можна використовувати адаптери або лінкери, або можна задіяти інші маніпуляції для забезпечення підходящих сайтів рестрикції, видалення надлишкової ДНК, видалення сайтів рестрикції і т. п. З цією метою можна задіяти мутагенез *in vitro*, репарацію за допомогою праймерів, рестрикцію, відпал, повторні заміни, наприклад, транзиції і трансверсії.

Для експресії різних послідовностей мегануклеаз, розкритих у даному документі, можна застосовувати ряд промоторів, у тому числі нативний промотор послідовності полінуклеотиду, що представляє інтерес. Промотори можна вибирати, виходячи з необхідного результату. Такі промотори включають, наприклад, конститутивні, кращі для тканин або інші промотори для експресії в рослинах.

Конститутивні промотори включають, наприклад, коровий промотор промотору Rsyn7 та інші конститутивні промотори, розкриті в WO 99/43838 і патенті США № 6072050; коровий промотор 35S CaMV (Odell et al. (1985) *Nature* 313:810-812); актиновий промотор рису (Mcelroy et al. (1990) *Plant Cell* 2:163-171); убівітиновий промотор (Christensen et al. (1989) *Plant Mol. Biol.* 12:619-632; i Christensen et al. (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last et al. (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten et al. (1984) *EMBO J.* 3:2723-2730); промотор ALS (патент США № 5659026) і т. п. Інші конститутивні промотори включають, наприклад, описані в патентах США № 5608149; 5608144; 5604121; 5569597; 5466785; 5399680; 5268463; 5608142 і 6177611.

Націлюючись на підвищену експресію мегануклеаз у конкретній тканині рослини можна використовувати кращі для тканин промотори. Кращі для тканини промотори включають описані в Yamamoto et al. (1997) *Plant J.* 12(2):255-265; Kawamata et al. (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen et al. (1997) *Mol. Gen Genet.* 254(3):337-343; Russell et al. (1997) *Transgenic Res.* 6(2):157-168; Rinehart et al. (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp et al. (1996) *Plant Physiol.* 112(2):525-535; Canevascini et al. (1996) *Plant Physiol.* 112(2):513-524; Yamamoto et al. (1994) *Plant Cell Physiol.* 35(5):773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco et al. (1993) *Plant Mol Biol.* 23(6):1129-1138; Matsuoka et al. (1993) *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; i Guevara-Garcia et al. (1993) *Plant J.* 4(3):495-505. За потребою, такі промотори можна модифікувати з одержанням слабкої експресії.

Промотори, які активні переважно в листках, відомі з технічного рівня. Див., наприклад, Yamamoto et al. (1997) *Plant J.* 12(2):255-265; Kwon et al. (1994) *Plant Physiol.* 105:357-67; Yamamoto et al. (1994) *Plant Cell Physiol.* 35(5):773-778; Gotor et al. (1993) *Plant J.* 3:509-18; Orozco et al. (1993) *Plant Mol. Biol.* 23(6):1129-1138; i Matsuoka et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590.

Для експресії послідовностей, що кодують мегануклеази, або їх біологічно активні варіанти і фрагменти, можна застосовувати синтетичні промотори.

Касета експресії може також містити ген маркера який селектується для відбору трансформованих клітин. Гени маркерів, які селектуються, використовуються для відбору трансформованих клітин або тканин. Маркерні гени включають гени, що кодують стійкість до антибіотиків, як, наприклад, такі, що кодують неоміцинфосфотрансферазу II (NEO) і гігromіцинфосфотрансферазу (HPT), а також гени, що надають стійкість до гербіцидних сполук, таких як гліфосат, глюфосинат амонію, бромоксиніл, сульфонілсечовини, дикамба й 2, 4-дихлорфеноксиацетат (2,4-D). Додаткові маркери, які селектуються, включають фенотипові маркери, такі як β-галактозидаза і флуоресцентні білки, такі як зелений флуоресцентний білок (GFP) (Su et al. (2004) *Biotechnol Bioeng* 85:610-9; i Fetter et al. (2004) *Plant Cell* 16:215-28), блакитний флуоресцентний білок (CYP) (Bolte et al. (2004) *J. Cell Science* 117:943-54; i Kato et al. (2002) *Plant Physiol* 129:913-42) і жовтий флуоресцентний білок (Phiyfp™ від Evrogen, див. Bolte et al. (2004) *J. Cell Science* 117:943-54). Про додаткові маркери, які селектуються, див., у цілому, в Yarranton (1992) *Curr. Opin. Biotech.* 3:506-511; Christopherson et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6314-6318; Yao et al. (1992) *Cell* 71:63-72; Reznikoff (1992) *Mol. Microbiol.* 6:2419-2422; Barkley et al. (1980) в *The Operon*, pp. 177-220; Hu et al. (1987) *Cell* 48:555-566; Brown et al. (1987) *Cell* 49:603-612; Figge et al. (1988) *Cell* 52:713-722; Deuschle et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:5400-5404; Fuerst et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle et al. (1990) *Science* 248:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow et al. (1990) *Mol. Cell. Biol.* 10:3343-3356; Zambretti et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski et al. (1991) *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb et al. (1991) *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschmidt et al. (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva et al. (1992) *Antimicrob. Agents Chemother.* 36:913-919; Hlavka et al. (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill et al. (1988) *Nature* 334:721-724. Такі розширення включені в даний документ як посилання. Наведений вище перелік генів маркерів, які селектуються, не призначений для обмеження. У даному винаході можна застосовувати будь-який ген маркера, який селектується.

Спосіб введення

Засіб, що рідко розщеплює, для індукції двониткового розриву, таке як мегануклеаза, можна вводити за допомогою будь-яких засобів, відомих з технічного рівня. Наприклад, передбачаються клітина, дріжджі або рослина, що мають припустимий або варіантний сайт розпізнавання у своєму геномі. Можна забезпечувати транзійтну експресію мегануклеази, або поліпептид сам по собі можна безпосередньо доставляти в клітину. Альтернативно, нуклеотидна послідовність, здатна до експресії мегануклеази, може бути стабільно інтегрована в геном рослини. При наявності відповідного припустимого або варіантного сайту розпізнавання і мегануклеази, донорну ДНК можна вводити в геном трансформованої рослини. Альтернативно, різні компоненти можна поєднувати шляхом статевого схрещування трансформованих рослин. Таким чином, послідовність, що кодує мегануклеазу, і/або припустимий або варіантний сайт розпізнавання можна поєднувати один з одним у ході статевого схрещування із забезпеченням присутності кожного компонента системи в одній рослині. Мегануклеаза може перебувати під контролем конститутивного промотору або промотору, що індукується. Такі промотори, які представляють інтерес, обговорюються докладніше в іншому місці в даному документі.

Для введення послідовності, яка представляє інтерес, такої як кожен із засобів, що рідко розщеплює, для індукції двониткового розриву, у рослину або частину рослини, можна застосовувати різні способи. Мається на увазі, що "введення" означає надання рослині, рослинній клітині або частині рослини полінуклеотиду або поліпептиду таким чином, що послідовність потрапляє всередину клітини рослини. Способи по даному винаходу не залежать від конкретного способу введення послідовності в рослину або частину рослини, за винятком того, що полінуклеотид або поліпептиди потрапляють всередину щонайменше однієї клітини рослини. З технічного рівня відомі способи введення полінуклеотиду та/або поліпептидів у рослину, у тому числі без обмежень способи стабільної трансформації, способи транзійтної трансформації та способи трансформації, опосередкованої вірусами.

Мається на увазі, що "стабільна трансформація" означає, що нуклеотидна конструкція, яка вводиться в рослину, інтегрується в геном рослини й може бути успадкована його нащадками. Мається на увазі, що "транзійтна трансформація" означає, що полінуклеотид вводиться в рослину й не інтегрується в геном рослини або в рослину вводиться поліпептид.

Протоколи трансформації, а також протоколи введення поліпептидів або послідовностей полінуклеотидів у рослини можуть змінюватися залежно від типу рослини або рослинної клітини, тобто однодольної або дводольної рослини, яку цілеспрямовано трансформують. Підходящі способи введення поліпептидів і полінуклеотидів у рослинні клітини включають мікроін'єкцію (Crossway et al. (1986) *Biotechniques* 4:320-334), електропорацію (Riggs et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, *Agrobacterium*-опосередковану трансформацію (патент США № 5563055 і патент США № 5981840), прямий перенос генів (Paszowski et al. (1984) *EMBO J.* 3:2717-2722) і балістичне прискорення частинок (див., наприклад, патент США № 4945050; патент США № 5879918; патент США № 5886244 і № 5932782; Tomes et al. (1995) в *Plant Cell, Tissue and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe et al. (1988) *Biotechnology* 6:923-926); а також трансформацію за допомогою *Lec1* (WO 00/28058). Також див. Weissinger et al. (1988) *Ann. Rev. Genet.* 22:421-477; Sanford et al. (1987) *Particulate Science and Technology* 5:27-37 (цибуля); Christou et al. (1988) *Plant Physiol.* 87:671-674 (соє); McCabe et al. (1988) *Bio/Technology* 6:923-926 (соєві боби); Finan and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (соє); Singh et al. (1998) *Theor. Appl. Genet.* 96:319-324 (соє); Datta et al. (1990) *Biotechnology* 8:736-740 (рис); Klein et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (маїс); Klein et al. (1988) *Biotechnology* 6:559-563 (маїс); патенти США №№ 5240855, 5322783 і 5324646; Klein et al. (1988) *Plant Physiol.* 91:440-444 (маїс); Fromm et al. (1990) *Biotechnology* 8:833-839 (маїс); Hooykaas-Van Slogteren et al. (1984) *Nature (London)* 311:763-764; патент США № 5736369 (злаки); Bytebier et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet et al. (1985) в *The Experimental Manipulation of Ovule Tissues*, ed. Chapman et al. (Longman, New York), pp. 197-209 (пилок); Kaeppler et al. (1990) *Plant Cell Reports* 9:415-418 і Kaeppler et al. (1992) *Theor. Appl. Genet.* 84:560-566 (опосередкована ниткоподібними кристалами трансформація); D'Halluin et al. (1992) *Plant Cell* 4:1495-1505 (електропорація); Li et al. (1993) *Plant Cell Reports* 12:250-255; і Christou and Ford (1995) *Annals of Botany* 75:407-413 (рис); Osjoda et al. (1996) *Nature Biotechnology* 14:745-750 (маїс, за допомогою *Agrobacterium tumefaciens*); всі включені в даний документ у якості посилання.

У конкретних варіантах здійснення послідовність засобу, що рідко розщеплює, для індукції двониткового розриву, таку як послідовність мегануклеази, або її активний варіант або фрагменти можна доставляти в дріжджову клітину або рослину за допомогою різних способів

транзієнтної трансформації. Такі способи транзієнтної трансформації включають, без обмежень, введення білка мегануклеази або його активних варіантів і фрагментів безпосередньо в дріжджову клітину або рослину. Такі способи включають, наприклад, мікроін'єкцію або бомбардування частинками. Див., наприклад, Crossway et al. (1986) *Mol Gen. Genet.* 202:179-185; Nomura et al. (1986) *Plant Sci.* 44:53-58; Hepler et al. (1994) *Proc. Natl. Acad. Sci.* 91: 2176-2180; і Hush et al. (1994) *The Journal of Cell Science* 107:775-784, всі включені в даний документ як посилання.

Як правило, такі способи включають вбудовування нуклеотидної конструкції по даному винаходу в молекулу ДНК або РНК. Зрозуміло, що послідовність мегануклеази можна початково синтезувати як частину вірусного поліпротеїну, який потім можна піддавати процесингу шляхом протеолізу *in vivo* або *in vitro* з одержанням необхідного рекомбінантного білка. Крім того, мається на увазі, що промотори по даному винаходу також охоплюють промотори, які використовуються для транскрипції вірусними РНК-полімеразами. Способи введення в рослини полінуклеотидів, у тому числі молекул вірусної ДНК або РНК, та способи експресії білка, що ними кодується, відомі з технічного рівня. Див., наприклад, патенти США №№ 5889191, 5889190, 5866785, 5589367, 5316931; і Porta et al. (1996) *Molecular Biotechnology* 5:209-221; включені в даний документ як посилання.

З технічного рівня відомі способи цілеспрямованої вставки полінуклеотиду в конкретне положення в геномі рослини. В одному варіанті здійснення вставку полінуклеотиду в необхідне положення в геномі виконують із використанням системи сайт-специфічної рекомбінації. Див., наприклад, WO 99/25821, WO 99/25854, WO 99/25840, WO 99/25855 і WO 99/25853, всі включені в даний документ як посилання. Коротко, полінуклеотид по даному винаходу може знаходитись в касеті для переносу, фланкованої двома нерекомбіногенними сайтами рекомбінації. Касету для переносу вводять у рослину, що має у своєму геномі стабільно вбудований цільовий сайт, фланкований двома нерекомбіногенними сайтами рекомбінації, які відповідають сайтам касети для переносу. Забезпечують підходящу рекомбіназу, і касета для переносу інтегрується в цільовий сайт. Полінуклеотид, який представляє інтерес, таким чином, інтегрується в конкретне хромосомне положення в геномі рослини. Інші способи націлювання полінуклеотидів викладені в WO 2009/114321 (включеної в даний документ як посилання), у якому описані "оригінальні" мегануклеази, які одержуються для модифікації геномів рослин, зокрема, геному маїсу. Див. також Gao et al. (2010) *Plant Journal* 1:176-187.

Із клітин, які були трансформовані, можна виростити рослини відповідно до традиційних способів. Див., наприклад, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. Ці рослини можна потім вирощувати та запилювати за допомогою або тієї ж самої трансформованої лінії, або інших ліній, та ідентифікувати отриманих в результаті нащадків, які характеризуються конститутивним проявом вказаної необхідної фенотипової характеристики. Можна виростити два або більше поколінь, щоб переконатися в тому, що експресія необхідної фенотипової характеристики стабільно підтримується та успадковується, а потім зібрати насіння, щоб переконатися в тому, що була досягнута експресія необхідної фенотипової характеристики. Таким чином, у даному винаході передбачене трансформоване насіння (яке також називають "трансгенним насінням"), що має полінуклеотид по даному винаходу, наприклад, касету експресії по даному винаходу, стабільно вбудовану в його геном.

Як використовується в даному документі, "праймери" являють собою виділені полінуклеотиди, які гібридизуються з комплементарною ниткою цільової ДНК шляхом гібридизації нуклеїнових кислот з утворенням гібриду праймера та цільової нитки ДНК, які потім подовжують вздовж нитки цільової ДНК за допомогою полімерази, наприклад, ДНК-полімерази. Пари праймерів по даному винаходу відносять до їхнього використання в ампліфікації цільового полінуклеотиду, наприклад, за допомогою полімеразної ланцюгової реакції (ПЛР) або інших традиційних способів ампліфікації нуклеїнових кислот. "ПЛР" або "полімеразна ланцюгова реакція" являє собою методику, застосовувану для ампліфікації конкретних сегментів ДНК (див. патенти США №№ 4683195 і 4800159; включені в даний документ як посилання).

Зонди і праймери мають достатню довжину в нуклеотидах, щоб зв'язуватися із цільовою послідовністю ДНК і забезпечувати специфічне виявлення та/або ідентифікацію полінуклеотиду, що кодує поліпептид з мегануклеазною активністю або його активний варіант або фрагмент, описані в іншому місці в даному документі. Зрозуміло, що умови гібридизації або умови реакції для досягнення даного результату можуть визначатися оператором. Ця довжина може бути будь-якою довжиною, яка є достатньою довжиною для того, щоб бути застосованою в бажаному способі виявлення. Такі зонди і праймери можуть здійснювати специфічну гібридизацію із цільовою послідовністю в умовах гібридизації високої жорсткості. Зонди і праймери, згідно з варіантами здійснення даного винаходу, можуть мати повну ідентичність суміжних нуклеотидів у

послідовності ДНК відносно цільової послідовності, хоча за допомогою традиційних способів можна розробити зонди, що відрізняються від цільової послідовності ДНК, та зберігають здатність до забезпечення специфічного виявлення та/або ідентифікації цільової послідовності ДНК. Відповідно, зонди і праймери можуть мати приблизно 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % або більшу ідентичність послідовностей або комплементарність стосовно цільового полінуклеотиду.

Способи одержання й застосування зондів і праймерів описані, наприклад, в *Molecular Cloning: A Laboratory Manual*, 2nd ed, vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 1989 (далі, "Sambrook et al., 1989"); *Current Protocols in Molecular Biology*, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1992 (з періодичними оновленнями) (далі, "Ausubel et al., 1992"); і Innis et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press: San Diego, 1990. Пари праймерів для ПЛР можна одержувати з відомої послідовності, наприклад, шляхом застосування комп'ютерних програм, призначених для цієї мети, таких як засіб аналізу праймерів для ПЛР у Vector NTI версії 10 (Invitrogen); Primerselect (DNASTAR Inc., Медісон, Вісконсин) і Primer (версія 0.5.COPYRIGHT., 1991, інститут біомедицинських досліджень Уайтхеда, Кембрідж, Масачусетс). Додатково, послідовності можна піддавати зоровому скануванню, а праймери ідентифікувати вручну із застосуванням керівництва, відомого фахівцеві в даній області.

Дріжджі та рослини

Передбачаються дріжджі, рослини, рослинні клітини, частини та насіння рослин, а також зерно що мають варіантні послідовності розпізнавання для засобів, що рідко розщеплюють, для індукції двониткового розриву, таких як мегануклеази, які розкриті в даному документі. У конкретних варіантах здійснення дріжджі, рослини та/або частини рослин мають стабільно вбудовану щонайменше одну гетерологічну варіантну послідовність розпізнавання, розкриту в даному документі, або її активний варіант або фрагмент. Таким чином, забезпечуються дріжджі, рослини, рослинні клітини, частини та насіння рослин, що містять щонайменше одну варіантну послідовність розпізнавання з кожним з SEQ ID NO: 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 або будь-якою їх комбінацією, або їх біологічно активний фрагмент та/або варіант. У конкретних варіантах здійснення варіантні послідовності розпізнавання проявляють підвищену активність розщеплення по відношенню до засобу, що рідко розщеплює, для індукції двониткового розриву.

Використовуваний в даному документі вираз "рослина" включає рослинні клітини, протопласти рослинних клітин, тканинні культури рослинних клітин, з яких можна регенерувати рослини, рослинні калюси, скупчення рослинних клітин і рослинні клітини, які є інтактними в рослинах або частинах рослин, таких як зародки, пилки, сім'ябруньки, насіння, листки, квітки, гілки, плоди, зерна, колоски, качани, листові обгортки, стебла, коріння, кінчики корінь, пильовики й т. п. Припускається, що зерно означає зріле насіння, отримане комерційними рослинниками для цілей, відмінних від вирощування або відтворення виду. Нащадки, варіанти та мутанти регенерованих рослин також включені в обсяг даного винаходу, за умови, що ці частини містять введені полінуклеотиди.

Трансформована рослина або трансформована рослинна клітина, які забезпечуються в даному документі, являють собою такі, у яких генетична зміна, така як трансформація, торкнулася гена, який представляє інтерес, або являють собою рослину або рослинну клітину, які походять від рослини або клітини, змінених таким чином, і які містять зміну. "Трансген" є геном, який ввели в геном за допомогою способу трансформації. Відповідно до цього, "трансгенна рослина" являє собою рослину, яка містить трансген, незалежно від того, чи ввели трансген у дану конкретну рослину шляхом трансформації або схрещування; таким чином, нащадки початково трансформованої рослини охоплені визначенням. "Контроль", або "контрольна рослина", або "контрольна рослинна клітина" забезпечують точку відліку для вимірювання змін фенотипу досліджуваної рослини або рослинної клітини. Контрольна рослина або рослинна клітина можуть включати, наприклад, (а) рослину або клітину дикого типу, тобто, з тим же генотипом, що і у вихідного матеріалу для генетичної зміни, з якого в результаті одержали досліджувану рослину або клітину; (b) рослину або рослинну клітину з тим же генотипом, що і у вихідного матеріалу, але які трансформували за допомогою нульової конструкції (тобто за допомогою конструкції, яка не експресує трансген, такий як конструкція, що містить маркерний ген); (c) рослину або рослинну клітину, які являють собою нетрансформовані сегреганти серед нащадків досліджуваної рослини або рослинної клітини; (d) рослину або рослинну клітину, які є генетично ідентичними досліджуваній рослині або рослинній клітині, але які не зазнають впливу умов або подразників, які будуть індукувати експресію трансгена; або (e) досліджувану рослину або рослинну клітину як такі в умовах, в яких конструкція не

експресується.

З рослинних клітин, які трансформували для експресії мегануклеази, забезпечуваної в даної документі, можна виростити цілі рослини. Регенерація, розвиток і культивування рослин із трансформованих протопластів з однієї рослини або з різних трансформованих експлантів добре відомі в даному технічному рівні. Див., наприклад, McCormick et al. (1986) Plant Cell Reports 5:81-84; Weissbach and Weissbach, в Methods for Plant Molecular Biology, (Eds.), Academic Press, Inc. San Diego, Calif., (1988). Даний спосіб регенерації та вирощування, як правило, включає етапи відбору трансформованих клітин, культивування цих індивідуалізованих клітин протягом звичайних стадій розвитку зародка до стадії вкоріненого саджанця. Трансгенні зародки та насіння регенерують подібним чином. Отримані в результаті вкорінені трансгенні пагони згодом висаджують у відповідне середовище для вирощування рослин, таке як ґрунт. Бажано, щоб регенеровані рослини були самозапильними для одержання гомозиготних трансгенних рослин. В інших випадках пилком, отриманим із регенерованих рослин, перехресно запилюють вирощені з насіння рослини агрономічно важливих ліній. У свою чергу, пилок від рослин даних важливих ліній використовують для запилення регенерованих рослин. Можна виростити два або більше поколінь, щоб переконатися в тому, що експресія необхідної фенотипової характеристики стабільно підтримується та успадковується, а потім зібрати насіння, щоб переконатися в тому, що була досягнута експресія необхідної фенотипової характеристики. Таким чином, композиції, представлені в даному документі, забезпечують трансформоване насіння (яке також називають "трансгенним насінням"), що має полінуклеотид, забезпечуваний у даному документі, наприклад, цільовий сайт, стабільно вбудований у його геном.

Варіантні послідовності розпізнавання та їх активні варіанти і фрагменти, розкриті в даному документі, можна використовувати для трансформації будь-яких видів рослин, у тому числі, без обмежень, однодольних і дводольних. Приклади видів рослин, що представляють інтерес, включають без обмежень кукурудзу (*Zea mays*), *Brassica* sp. (наприклад, *B. napus*, *B. rapa*, *B. juncea*), зокрема, ті види *Brassica*, які є придатними в якості джерел рослинної олії, люцерну (*Medicago sativa*), рис (*Oryza sativa*), жито (*Secale cereale*), сорго (*Sorghum bicolor*, *Sorghum vulgare*), просо (наприклад, пенісетум рогозовидний (*Pennisetum glaucum*), просо культурне (*Panicum miliaceum*), щетинник італійський (*Setaria italica*), просо пальчасте (*Eleusine coracana*)), соняшник (*Helianthus annuus*), сафлор (*Carthamus tinctorius*), пшеницю (*Triticum aestivum*), сою (*Glycine max*), тютюн (*Nicotiana tabacum*), картоплю (*Solanum tuberosum*), різновиди арахісу (*Arachis hypogaea*), бавовник (*Gossypium barbadense*, *Gossypium hirsutum*), солодку картоплю (*Ipomoea batatas*), маніок (*Manihot esculenta*), кавове дерево (*Coffea* spp.), кокосову пальму (*Cocos nucifera*), ананас (*Ananas comosus*), цитрусові дерева (*Citrus* spp.), шоколадне дерево (*Theobroma cacao*), чайний кущ (*Camellia sinensis*), банан (*Musa* spp.), авокадо (*Persea americana*), інжир (*Ficus casica*), гуаву (*Psidium guajava*), манго (*Mangifera indica*), маслину (*Olea europaea*), папайю (*Carica papaya*), кеш'ю (*Anacardium occidentale*), макадамію (*Macadamia integrifolia*), мигдаль (*Prunus amygdalus*), різновиди цукрового буряка (*Beta vulgaris*), цукровий очерет (*Saccharum* spp.), різновиди вівса, ячмінь, овочі, декоративні рослини та хвойні рослини.

Овочі включають різновиди томату (*Lycopersicon esculentum*), латук (наприклад, *Lactuca sativa*), різновиди зеленої квасолі (*Phaseolus vulgaris*), різновиди лімської квасолі (*Phaseolus limensis*), різновиди гороху (*Lathyrus* spp.) і представників роду *Cucumis*, таких як огірок (*C. sativus*), канталупа (*C. cantalupensis*) і диня мускусна (*C. melo*). Декоративні рослини включають азалію (*Rhododendron* spp.), гортензію (*Macrophylla hydrangea*), гібіскус (*Hibiscus rosasanensis*), троянди (*Rosa* spp.), тюльпани (*Tulipa* spp.), нарциси (*Narcissus* spp.), петунії (*Petunia hybrida*), гвоздику (*Dianthus caryophyllus*), пуансетію (*Euphorbia pulcherrima*) і хризантему.

Хвойні рослини, які можна використовувати при практичному здійсненні даного винаходу, включають, наприклад, види сосни, такі як сосна ладанна (*Pinus taeda*), сосна Еліота (*Pinus elliotii*), сосна жовта (*Pinus ponderosa*), сосна скручена широкохвойна (*Pinus contorta*) і сосна промениста (*Pinus radiata*); псевдотсугу Мензиса (*Pseudotsuga menziesii*); тсугу канадську (*Tsuga canadensis*); ялину білу (*Picea glauca*); секвойю вічнозелену (*Sequoia sempervirens*); справжні ялиці, такі як ялиця миловидна (*Abies amabilis*) і ялиця бальзамічна (*Abies balsamea*), і туї, такі як туя гігантська (*Thuja plicata*) і каллітропсис нутканський (*Chamaecyparis nootkatensis*), а також тополя та евкаліпт. У конкретних варіантах здійснення рослини по даному винаходу є культурними рослинами (наприклад, кукурудза, люцерна, соняшник, *Brassica*, соя, бавовник, сафлор, арахіс, сорго, пшениця, просо, тютюн і т. д.). В інших варіантах здійснення рослини кукурудзи й сої є оптимальними, а ще в декількох інших варіантах здійснення оптимальними є рослини кукурудзи.

Інші рослини, що представляють інтерес, включають зернові рослини, які забезпечують

насіння, яке представляє інтерес, олійні рослини й бобові рослини. Насіння, яке представляє інтерес, включає насіння зернових культур, таких як кукурудза, пшениця, ячмінь, рис, сорго, жито і т. д. Олійні рослини включають бавовник, сою, сафлор, соняшник, Brassica, маїс, люцерну, пальму, кокосову пальму і т. д. Бобові рослини включають боби й різновиду гороху.

Боби включають гуар, ріжкове дерево, пажитник, сою, різновиди звичайної квасолі, вігну китайську, золотаву квасолю, лімську квасолю, стручкову квасолю, різновиди сочевиці, турецький горох і т. д.

Необмежувані приклади композицій і способів, розкритих у даному документі, є наступними.

1) Спосіб ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому зазначений спосіб включає:

а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання в зазначену геномну ДНК, де двонитковий розрив призводить у результаті до утворення нуклеотидного липкого кінця;

б) лігування першого адаптера із зазначеним нуклеотидним липким кінцем;

с) розрізання лігваної ДНК, отриманої на стадії (b), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив;

д) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (c), з еталонною геномною послідовністю ДНК; і

е) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву.

2) Спосіб згідно з варіантом здійснення 1, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, вибирають із групи, що складається з мегануклеази, нуклеази "цинкові пальці", Tal-ефекторної нуклеази, транспозази та сайт-специфічної рекомбінази.

3) Спосіб згідно з варіантом здійснення 1, де нуклеотидний липкий кінець являє собою 3'-нуклеотидний липкий кінець.

4) Спосіб згідно з варіантом здійснення 1, де нуклеотидний липкий кінець являє собою 5'-нуклеотидний липкий кінець.

5) Спосіб згідно з варіантом здійснення 1, де перший адаптер, лігований з нуклеотидним липким кінцем, являє собою нефосфорильований по 5' адаптер.

6) Спосіб згідно з варіантом здійснення 1, де геномну ДНК вибирають із групи, що складається із прокаріотичної ДНК, еукаріотичної ДНК і синтетичної ДНК.

7) Спосіб згідно з варіантом здійснення 6, де еукаріотичну ДНК виділяють із рослини, дріжджів або тварини.

8) Спосіб згідно з варіантом здійснення 7, де рослину вибирають із групи, що складається із сої, соняшника, бавовнику, люцерни, каноли, тютюну, картоплі, Arabidopsis, сафлору, маїсу, рису, сорго, ячменю, пшениці, проса, вівса, цукрової тростини, газонної трави й проса прутноподібного.

9) Спосіб згідно з варіантом здійснення 1, де засіб для індукції двониткового розриву одержують із I-SceI.

10) Спосіб ідентифікації варіантного сайту розпізнавання з покращеною активністю розщеплення для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому зазначений спосіб включає:

а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатним до введення рідкого двониткового розриву в припустимому сайті розпізнавання, де двонитковий розрив призводить у результаті до утворення нуклеотидного липкого кінця;

б) лігування першого адаптера із зазначеним нуклеотидним липким кінцем;

с) розрізання лігваної ДНК, отриманої на стадії (b), і лігування другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив;

д) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (c), з еталонною геномною послідовністю ДНК; і

е) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного для

сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву;

f) аналіз активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, у варіантних сайтах розпізнавання з d);

g) ідентифікацію варіантного сайту розпізнавання, який призводить у результаті до підвищеної активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, в порівнянні з активністю в припустимому сайті розпізнавання.

11) Спосіб згідно з варіантом здійснення 10, де про підвищену активність сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, свідчить:

a) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на геномній ДНК;

b) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на плазмідній ДНК;

c) вища оцінка в аналізі на дріжджах для варіантного сайту розпізнавання в порівнянні з припустимим сайтом розпізнавання; або

d) будь-яка комбінація (a), (b) і (c).

12) Спосіб введення в геном клітини варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому зазначений спосіб включає:

a) забезпечення наявності донорної ДНК, що містить варіантний сайт розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, де зазначений сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, також здатний до введення двониткового розриву в зазначений варіантний сайт розпізнавання;

b) забезпечення наявності рослинної клітини;

c) приведення в контакт рослинної клітини з донорною ДНК; і

d) ідентифікацію щонайменше однієї рослинної клітини з (c), що містить у своєму геномі зазначений варіантний сайт розпізнавання.

13) Спосіб згідно з варіантом здійснення 12, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, вибирають із групи, що складається з мегануклеази, нуклеази "цинкові пальці", Tal-ефекторної нуклеази, транспозази, ендонуклеази Cas і сайт-специфічної рекомбінази.

14) Виділений полінуклеотид що містить варіантний сайт, розпізнавання з покращеною активністю розщеплення для сконструйованої мегануклеази, здатної до введення двониткового розриву в припустимому сайті розпізнавання, де зазначений варіантний сайт розпізнавання містить нуклеотидну послідовність із заміною щонайменше 1 нуклеотидної основи в порівнянні з припустимим сайтом розпізнавання SEQ ID NO: 14.

15) Виділений полінуклеотид згідно з варіантом здійснення 1, де зазначений варіантний сайт розпізнавання містить послідовність зі змінами щонайменше 2, 3, 4, 5, 6 або 7 пар основ у порівнянні з SEQ ID NO: 14.

16) Виділений полінуклеотид згідно з варіантом здійснення 14, де зазначений варіантний сайт розпізнавання містить:

a) аденін (A) у положенні, що відповідає нуклеотидному положенню 9 в SEQ ID NO: 14;

b) гуанін (G) у положенні, що відповідає нуклеотидному положенню 10 в SEQ ID NO: 14;

c) тимін (T) у положенні, що відповідає нуклеотидному положенню 11 в SEQ ID NO: 14;

d) аденін (A) у положенні, що відповідає нуклеотидному положенню 13 в SEQ ID NO: 14;

e) тимін (T) у положенні, що відповідає нуклеотидному положенню 18 в SEQ ID NO: 14;

f) гуанін (G) у положенні, що відповідає нуклеотидному положенню 19 в SEQ ID NO: 14;

g) гуанін (G) або аденін (A) у положенні, що відповідає нуклеотидному положенню 22 в SEQ ID NO: 14; або

h) будь-яку комбінацію з a)-g).

17) Виділений полінуклеотид згідно з варіантом здійснення 14, де зазначена варіантна послідовність розпізнавання обрана із групи, що складається з SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35.

18) Виділений полінуклеотид згідно з варіантом здійснення 14, де про покращену активність розщеплення свідчить:

a) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання SEQ ID NO: 14, де сайти розпізнавання розташовані на геномній ДНК;

b) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на плазмідній ДНК;

5 c) вища оцінка в аналізі на дріжджах для варіантного сайту розпізнавання в порівнянні з припустимим сайтом розпізнавання; або

d) будь-яка комбінація (a), (b) і (c).

19) Рекombінантний фрагмент ДНК, що містить виділений полінуклеотид згідно з варіантом здійснення 14.

20) Клітина, що містить рекombінантний фрагмент ДНК згідно з варіантом здійснення 19.

10 21) Клітина згідно з варіантом здійснення 20, де клітина являє собою клітину дріжджів або рослинну клітину.

22) Трансгенна рослина або насіння, що містить рослинну клітину згідно з варіантом здійснення 21.

15 23) Трансгенна рослина згідно з варіантом здійснення 22, де зазначена рослина обрана із групи, що складається з маїсу, пшениці, рису, ячменю, цукрової тростини, сорго, жита, проса прутovidного, сої, Brassica, соняшника, бавовнику або люцерни.

24) Виділений полінуклеотид що містить варіантний сайт, розпізнавання з покращеною активністю розщеплення для сконструйованої мегануклеази, здатної до введення двониткового розриву в припустимий сайт розпізнавання, де зазначений варіантний сайт розпізнавання 20 містить нуклеотидну послідовність із заміною щонайменше 1 нуклеотидної основи в порівнянні з припустимим сайтом розпізнавання SEQ ID NO: 13.

25) Виділений полінуклеотид згідно з варіантом здійснення 24, де зазначений варіантний сайт розпізнавання містить послідовність зі змінами щонайменше 2, 3, 4 або 5 пар основ у порівнянні з SEQ ID NO: 13.

25 26) Виділений полінуклеотид згідно з варіантом здійснення 24, де зазначений варіантний сайт розпізнавання містить:

a) цитозин (C) у положенні, що відповідає нуклеотидному положенню 1 в SEQ ID NO: 13

b) цитозин (C) у положенні, що відповідає нуклеотидному положенню 5 в SEQ ID NO: 13;

c) гуанін (G) у положенні, що відповідає нуклеотидному положенню 10 в SEQ ID NO: 13;

30 d) тимін (T) аденін (A) у положенні, що відповідає нуклеотидному положенню 11 в SEQ ID NO: 13;

e) тимін (T) у положенні, що відповідає нуклеотидному положенню 19 в SEQ ID NO: 13;

f) будь-яку комбінацію з a)-e).

35 27) Виділений полінуклеотид згідно з варіантом здійснення 24, де зазначена варіантна послідовність розпізнавання обрана із групи, що складається з SEQ ID NO: 15, 16, 17, 18, 19, 20 і 21.

28) Виділений полінуклеотид згідно з варіантом здійснення 24, де про покращену активність розщеплення свідчить:

40 a) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання SEQ ID NO:13, де сайти розпізнавання розташовані на геномній ДНК;

b) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на плазмідній ДНК;

45 c) вища оцінка в аналізі на дріжджах для варіантного сайту розпізнавання в порівнянні з припустимим сайтом розпізнавання; або

d) будь-яка комбінація (a), (b) і (c).

29) Рекombінантний фрагмент ДНК, що містить виділений полінуклеотид згідно з варіантом здійснення 24.

50 30) Клітина, що містить рекombінантний фрагмент ДНК згідно з варіантом здійснення 29.

31) Клітина згідно з варіантом здійснення 30, де клітина являє собою клітину дріжджів або рослинну клітину.

32) Трансгенна рослина або насіння, що містить рослинну клітину згідно з варіантом здійснення 31.

55 33) Трансгенна рослина згідно з варіантом здійснення 32, де зазначена рослина обрана із групи, що складається з маїсу, пшениці, рису, ячменю, цукрової тростини, сорго, жита, проса прутovidного, сої, Brassica, соняшника, бавовнику або люцерни.

34) Спосіб націлювання вставки полінуклеотиду, що представляє інтерес, у специфічний хромосомний сайт у геномі рослини, причому зазначений спосіб включає:

60 a) трансформацію рослинної клітини або рослини фрагментом ДНК, що містять

полінуклеотид, що представляє інтерес, де зазначений геном зазначених рослинної клітини або рослини містить щонайменше один варіантний сайт розпізнавання, обраний із групи, яка складається з SEQ ID NO: 15, 16, 17, 18, 19, 20 і 21; і

5 б) забезпечення наявності мегануклеази, здатної робити двонитковий розрив у варіабельному сайті розпізнавання з (а); і

с) відбір зазначених рослинної клітини або рослини, що містять зазначений полінуклеотид, що представляє інтерес, інтегрований у зазначений варіантний сайт розпізнавання.

35) Спосіб згідно з варіантом здійснення 34, де забезпечення наявності зазначеної мегануклеази включає інтеграцію в геном зазначених рослинної клітини або рослини нуклеотидної послідовності, що кодує мегануклеазу SEQ ID NO: 1.

36) Рослина або рослинна клітина, отримані за допомогою способу згідно з варіантом здійснення 34.

37) Спосіб націлювання вставки полінуклеотиду, що представляє інтерес, у специфічний хромосомний сайт у геномі рослини, причому зазначений спосіб включає:

15 а) трансформацію рослинної клітини або рослини фрагментом ДНК, що містить полінуклеотид, що представляє інтерес, де зазначений геном зазначених рослинної клітини або рослини містить щонайменше один варіантний сайт розпізнавання, обраний із групи, що полягає з SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35;

20 б) забезпечення наявності мегануклеази, здатної робити двонитковий розрив у варіабельному сайті розпізнавання з (а); і

с) відбір зазначених рослинної клітини або рослини, що містять зазначений полінуклеотид, що представляє інтерес, інтегрований у зазначений варіантний сайт розпізнавання.

25 38) Спосіб згідно з варіантом здійснення 37, де забезпечення наявності зазначеної мегануклеази включає інтеграцію в геном зазначених рослинних клітин нуклеотидної послідовності, що кодує мегануклеазу SEQ ID NO: 3

39) Рослина або рослинна клітина, отримані за допомогою способу згідно з варіантом здійснення 37.

40) Спосіб ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому зазначений спосіб включає:

а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатного до введення двониткового розриву в припустимому сайті розпізнавання, де двонитковий розрив призводить у результаті до утворення тупого кінця;

б) створення нуклеотидного липкого кінця з тупого кінця з (а);

35 с) лігування першого адаптера з нуклеотидним липким кінцем з (б);

д) розрізування легованої ДНК, отриманої на стадії (с), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив;

40 е) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (д), з еталонною геномною послідовністю ДНК; й

ф) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу для індукції двониткового розриву.

41) Спосіб згідно з варіантом здійснення 40, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву, являє собою ендонуклеазу Cas.

42) Спосіб згідно з варіантом здійснення 41, де ендонуклеаза Cas здатна до утворення комплексу з направляючою РНК, де зазначений комплекс дозволяє ендонуклеазі Cas вводити двонитковий розрив у зазначену геномну ДНК.

50 43) Спосіб згідно з варіантом здійснення 40, де геномну ДНК вибирають із групи, що складається із прокаріотичної ДНК, еукаріотичної ДНК і синтетичної ДНК.

44) Спосіб згідно з варіантом здійснення 43, де еукаріотичну ДНК виділяють із рослини, дріжджів або тварини.

55 45) Спосіб ідентифікації варіантного сайту розпізнавання з покращеною активністю розщеплення для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимому сайті розпізнавання, причому зазначений спосіб включає:

а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатного до введення двониткового розриву в зазначену геномну ДНК, де двонитковий розрив призводить у результаті до утворення тупого кінця;

60 б) створення нуклеотидного липкого кінця з тупого кінця з (а);

с) лігування першого адаптера з нуклеотидним липким кінцем з (b);
 d) розрізання легованої ДНК, отриманої на стадії (с), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив;

5 е) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (d), з еталонною геномною послідовністю ДНК; й

f) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву;

10 g) аналіз активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву у варіантних сайтах розпізнавання з f); й

h) ідентифікацію варіантного сайту розпізнавання, який призводить у результаті до підвищеної активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву в порівнянні з активністю в припустимому сайті розпізнавання.

15 46) Спосіб згідно з варіантом здійснення 45, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву являє собою ендонуклеазу Cas.

47) Спосіб згідно з варіантом здійснення 46, де ендонуклеаза Cas здатна до утворення комплексу з направляючою РНК, де зазначений комплекс дозволяє ендонуклеазі Cas вводити двонитковий розрив у зазначену геномну ДНК.

20 ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА

ПРИКЛАД 1

Створення сконструйованих мегануклеаз, що рідко розщеплюють

A. Мегануклеаза LIG3-4 і припустима послідовність розпізнавання для LIG3-4

25 Ендогенний геномний цільовий сайт маїсу, що містить припустиму послідовність розпізнавання для LIG3-4 (SEQ ID NO: 13), вибирали для розробки засобу, що рідко розщеплює, для індукції двониткового розриву (SEQ ID NO: 1), як описано в публікації заявки на патент США 2009-0133152 A1 (опублікованої 21 травня 2009 року). Припустима послідовність розпізнавання для LIG3-4 являє собою полінуклеотид розміром 22 п.о., що має наступну послідовність: ATATACCTCACACGTACGCGTA (SEQ ID NO: 13).

30 B. Мегануклеази MHP14+ і сайт розпізнавання для MHP14

Ендогенний геномний цільовий сайт маїсу, що містить припустиму послідовність розпізнавання для MHP14+ (SEQ ID NO: 14), вибирали для розробки засобу, що рідко розщеплює, для індукції двониткового розриву (SEQ ID NO: 3), як описано в заявці на патент США 13/427138, поданої 22 березня 2012 року). Припустимий сайт розпізнавання для MHP14+ являє собою полінуклеотид розміром 22 п.о., розташований і , що має наступну послідовність:

(SEQ ID NO: 14) CAAACAGATTTCACGTCAGATTT.

ПРИКЛАД 2

Продуктування білка мегануклеази в E. coli

3 Метою одержання очищеного білка для in vitro аналізів активності розщеплення
 40 мегануклеазою геномної і плазмідної ДНК, фрагменти ДНК, відповідні до відкритих рамок зчитування мегануклеази Lig3-4 (SEQ ID NO: 2) і мегануклеази MHP14+ (SEQ ID NO: 4), поміщали у вектор експресії pQE80 (Qiagen), трансформували в клітини E. coli BL21-Gold (Agilent Technologies) і вирощували протягом ночі на твердому середовищі LB, що містить 100 ppm (частин на мільйон) карбеніциліну. Колонії ресуспендували в 2 мл середовища 2ХYT і 250
 45 мкл клітинної суспензії використовували для інокуляції в 50 мл культури 2ХYT, доповненої 100 ppm карбеніциліну. Культури вирощували при 37 °C протягом 1-1,5 години або до досягнення OD600 0,8, а потім експресію білка індукували додаванням 0,5 мл 100 mM IPTG. Культури охолоджували до кімнатної температури та забезпечували можливість експресії білка протягом 2 годин. Клітини осаджували центрифугуванням протягом 10 хвилин при 5000 gcf (одиниць
 50 відцентрового прискорення). Супернатант зливали, осад ресуспендували в 1 мл буфера 1 (50 mM Tris- HCl (ph 8,0), 500 mM NaCl, 10 mM імідазол) і переносили в 1,5 мл мікроцентрифужну пробірку. Клітини руйнували шляхом ультразвукової обробки у два етапи 1/8" мікронаконечником з 20 імпульсами (робочий цикл 50, потужність 4) на аналоговому обладнанні для ультразвукової обробки Branson 450 Analog Sonifer і центрифугували при 20000 gcf
 55 протягом 15 хвилин при 4 °C. Супернатант роз розводили в 4 мл буфера 1 і завантажували на одноразову колонку, що містить 0,3 мл нікель-вмісної смоли Nickel- NTA Superflow resin (Qiagen). Колонку промивали 5 мл буфера 2 (50 mM Tris- HCl (ph 8,0), 500 mM NaCl, 60 mM імідазол) і білок елюювали 0,6 мл буфера 4 (50 mM Tris- HCl (ph 8,0), 500 mM NaCl, 250 mM імідазол) у колонку Vivaspin (GE). Для концентрування зразків колонки vivaspin центрифугували
 60 при 14800 gcf протягом приблизно 6 хвилин або до того, як меніск перебував між 75 і 50. Заміну

буфера здійснювали з використанням колонки для знесолення Zeba Spin Desalting Column (Pierce), попередньо врівноваженої буфером для зберігання (25 mM Tris- HCl (pH 8,0), 100 mM NaCl, 10 mM MgCl₂, 5 mM EDTA, 50 % гліцерин). Після заміни буфера бичачий сироватковий альбумін додавали до кінцевої концентрації 100 нг/мкл й очищений білок зберігали при -20 °C до застосування.

ПРИКЛАД 3

In vitro аналізи розщеплення геномної ДНК

Для створення матеріалу для захоплення геномних варіантних сайтів розпізнавання *in vitro* аналізи проводили з 114 нМ очищеного білка мегануклеази, виділеного, як описано в прикладі 2, і 6,07 мкг очищеної геномної ДНК маїсу при 32 °C протягом 80 хвилин у кінцевому обсязі 80 мкл при наявності буфера для розщеплення (50 mM Tris- HCl (pH 7,9), 100 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 5 mM EDTA). Через 80 хвилин усю реакцію зупиняли рівним об'ємом стоп-буфера (100 mM Tris- HCl (pH 8,0), 600 mM NaCl, 2 % SDS, 100 mM EDTA, 1 мг протеїнази K на мл) та інкубували при 50 °C протягом 30-45 хвилин. Реакційну суміш із зупиненою реакцією очищали за допомогою екстракції фенолом/хлороформом й осаджували етанолом при наявності 0,2 M NaCl. Осаджену геномну ДНК двічі промивали 70 % етанолом, висушували та ресуспендували в 34 мкл води.

Концентрацію білка мегануклеази визначали візуально в гелях Nu- PAGE (Life Technologies) шляхом розрахунку інтенсивності смуги та наступного її порівняння з такою для зразків відомої концентрації в серійному розведенні, й концентрацію геномної ДНК визначали з використанням флуориметричного аналізу з барвником Хехста.

Для підтвердження розщеплення (що представляє собою % втрати сайтів розпізнавання мегануклеазою) у передбачуваному геномному сайті розпізнавання проводили ПЛР у реальному часі на 1 мкл очищеної геномної ДНК із аналізом TaqMan, що охоплює сайт розпізнавання мегануклеазою. % розщеплення або втрати сайтів розпізнавання мегануклеазою розраховували за допомогою способу $\Delta\Delta C_t$ щодо внутрішнього контролю в аналізі TaqMan з використанням порожнього контролю як каліброваного стандарту.

ПРИКЛАД 4

Уловлювання геномних варіантних сайтів розпізнавання та одержання бібліотек для глибокого секвенування Illumina

У даному способі використовується підхід з новим адаптером, спеціально підібраним для уловлювання розщеплених хомінг-ендонуклеазою I-CreI або сконструйованою I-CreI геномних варіантних сайтів розпізнавання, послідовність яких невідома й відрізняється за складом від припустимого сайту розпізнавання, який відрізняється від способів з використанням рестрикційних ферментів для здійснення секвенування зі зменшеним представленням, глибоке секвенування ДНК, пов'язане з рестрикцією (RAD-tag або RADseq), повногеномне секвенування (WGS) або генотипівування за допомогою секвенування (GBS).

Оскільки хомінг-нуклеаза I- CreI утворює 4-нуклеотидний 3'-липкий кінець в центрі свого сайту розпізнавання довжиною 22 п.о. (+2, +1, -1, -2) при розщепленні (Thompson et al. (1992) Gene 119:247-51; i Durrenberger et al. (1993) Mol. Gen. Genet. 236:409-14), і, як було продемонстровано, розщеплює свій сайт розпізнавання в контексті різних комбінацій центральних 4 пар основ (+2, +1, -1, -2) (Molina et al. (2012) Nucleic Acids Res. 40:6936-45), створювали адаптери, що містять 4-нуклеотидний 3'-липкий кінець, що містить усі можливі комбінації нуклеотидів ДНК (G, T, A або C) липкого кінця в еквімолярному розподілі. Таким чином, забезпечення ефективного лігування і повної комплементарності всім можливим липким кінцям, створеним при розщепленні сайту розпізнавання в геномі, який піддавався розщепленню хомінг-ендонуклеазою I-CreI або сконструйованою I-CreI.

Подібні стратегії можна використовувати для інших засобів, що рідко розщеплюють, для індукції двониткового розриву, таких як нуклеази "цинкові пальці" і нуклеази TALEN, які розщеплюють ДНК неспецифічним каталітичним доменом FokI, створюючи липкі кінці з варіабельними довжиною та нуклеотидним складом у спейсерній ділянці між ними (Smith et al. (2000) Nucleic Acids Res. 28:3361-69; i Li et al. (2011) Nucleic Acids Res. 39:359-72). Для вловлювання геномних варіантних сайтів розпізнавання нефосфорильовані біотинильовані адаптери синтезували та очищали за допомогою HPLC (Integrated DNA Technologies, Inc.), що містить повністю вироджений 4-нуклеотидний 3'-нуклеотидний липкий кінець, комплементарний 4-нуклеотидному 3'-липкому кінцю, створеному при розщепленні сайту розпізнавання мегануклеазою (SEQ ID NO: 5), і лігували приблизно з 2 мкг розщепленої мегануклеазою геномної ДНК (отриманої, як описано в прикладі 3) в 100 мкл реакційної суміші з лігазою T4 (NEB) (представляючи перший адаптер у способі для ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового

розриву). Зразки, що містять ліговану ДНК, потім завантажували в мікропробірки для ультразвукової обробки та довільним чином розрізали до середнього розміру піка 300 п.о. шляхом ультразвукової обробки в системі Covaris E220. Налаштування були наступними: 10% робочий цикл, 140 пікова потужність падаючої хвилі й 200 циклів на розрив. Фрагменти з довжиною в діапазоні від 200 до 500 п.о. фракціонували за допомогою електрофорезу в агарозному гелі з подальшою екстракцією з гелю з використанням набору для екстракції з гелю Qiagen Gel Extraction Kit відповідно до рекомендацій виробника.

Небіотинільовані кінці піддавали репарації з використанням набору для репарації End-It End repair kit (Epicentre) в 75 мкл реакційної суміші та очищали на колонці (Qiagen). Подовження 3'-липкого кінця на один А здійснювали шляхом інкубування репарованої ДНК при 37 °С протягом 30 хвилин в 50 мкл реакційної суміші, що містить АТР, 1× буфер Кленова (NEBnext) і 15 одиниць фрагменту Кленова (без екзонуклеазної активності). Зразки потім очищували за допомогою колонки (Qiagen) і лігували в індексовані TruSeq-сумісні адаптери Illumina (які представляють собою другий набір адаптерів у способі для ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву) в 50 мкл реакційної суміші, що містить 0,3 мМ індексований адаптер, 1× буфер для лігування Quick ligation buffer і 5 одиниць ДНК лігази T4 (NEB) при кімнатній температурі. Після лігування зразки інкубували при 65 °С протягом 15 хвилин і об'єм доводили до 100 мкл. Магнітне уловлювання за допомогою стрептавідину здійснювали з використанням стрептавідинових гранул Dynabeads M-280 (Invitrogen). Усього 100 мкл ресуспендованих Streptavidin-Dynabeads (M-280) двічі промивали в ТЕ і ресуспендували в 100 мкл 2X B&W буфера (10 мМ Tris-HCl, 1 мМ EDTA, 100 мкл 0,5 М EDTA, 2 М NaCl). Зразки інкубували при 30 °С протягом 30 хвилин, супернатант видаляли і гранули промивали 4 рази 1 мл 1X B&W буфера. Кінцевий збагачений зразок ресуспендували в 30 мкл буфера EB.

Фрагменти вилучали із гранул за допомогою 12-циклової ПЛР із використанням майстер-міксу Phusion (NEB) в 50 мкл реакційної суміші при наявності 0,4 пмоль праймера А для вилучення (5'GTTGACATGCTGGATTGAGACTTC; SEQ ID NO: 6) і праймера В для вилучення (5'CAAGCAGAAGACGGCATACGA; SEQ ID NO: 7) відповідно до інструкцій виробника, за винятком того, що використовували температуру відпалу 66 °С і час подовження 30 секунд. Вилучену ДНК розщеплювали SbfI (NEB) і очищували двічі з гранулами Agencourt AMPure XP Beads (SPRI) відповідно до інструкцій виробника. Зразок лігували з Illumina-сумісним адаптером з SbfI-сумісним липким кінцем (SEQ ID NO: 8). Супернатант очищували двічі з використанням гранул Agencourt AMPure XP Beads, спочатку з використанням співвідношення зразка до гранул 1:1,8, а потім 1:1. Кінцеві зразки ресуспендували в 50 мкл буфера EB. Проводили другу ампліфікацію зі стандартним коктейлем ПЛР праймерів TruSeq PCR (Illumina) з використанням температури відпалу 60 °С, з наступним дворазовим очищенням з використанням гранул Agencourt AMPure XP Beads при співвідношенні зразка до гранул 1:1,8, а потім 1:1. Кінцевий зразок ресуспендували в 20 мкл. Зразки оцінювали на біоаналізаторі, піддавали відносному кількісному аналізу з використанням qPCR із праймерами для qPCR Illumina і збирали в пул. Перед секвенуванням пули піддавали селекції по розміру з використанням лабораторного чіпа для кристалографії Xchip (Caliper) відповідно до інструкцій виробника.

ПРИКЛАД 5

Нефосфорильовані адаптери підсилюють збагачення розщепленими мегануклеазою геномними сайтами розпізнавання

Для оцінки ефекту, який фосфорильовання здійснює на здатність першого адаптера (який описано в прикладі 4) до вловлювання та збагачення відносно розщеплених мегануклеазою геномних сайтів розпізнавання, були отримані бібліотеки як з фосфорильованими, так і нефосфорильованими адаптерами. Після нормалізації концентрації ДНК бібліотек їх оцінювали відносно збагачення розщепленим припустимим сайтом розпізнавання за допомогою ПЛР у реальному часі з аналізом TaqMan, що безпосередньо межує з припустимим сайтом розпізнавання. Як показано на фігурі 7, обидві бібліотеки демонстрували збагачення в порівнянні з порожнім контролем, але на графіках ампліфікації з бібліотеки, отриманої з нефосфорильованим адаптером, логарифмічна ампліфікація досягалася на набагато більш ранньому циклі ампліфікації, ніж з бібліотеки, отриманої з фосфорильованим адаптером. При використанні стандартних кривих, отриманих з геномної ДНК маїсу, бібліотеку, отриману з нефосфорильованим адаптером, оцінювали як приблизно в 900 разів більш збагачену відносно лівої половини припустимого сайту розпізнавання, ніж бібліотеку, отриману з фосфорильованим адаптером.

ПРИКЛАД 6

Глибоке секвенування Illumina і підрізання після обробки

Після вловлювання й збагачення відносно геномних варіантних сайтів розпізнавання та одержання ДНК, як описано в прикладі 4, створення кластерів і секвенування рідів спареного кінця здійснювали на приладі Illumina cBot і геномному аналізаторі Genome Analyzer IIx відповідно, згідно з інструкціями виробника. Приблизно 30 % (про./про.) контролю phiX ДНК (Illumina) додавали до розчину бібліотеки ДНК перед виділенням кластерів. Очікувалося, що довільний склад основ фрагментів phiX ДНК компенсує будь-яке відхилення в складі основ поблизу сайту для мегануклеази. Набори команд на 100 циклів для спарених кінців використовували на геномному аналізаторі Illumina Genome Analyzer. Послідовності та оцінки якості одержували за допомогою програмного забезпечення Illumina pipeline версії 2.9 для аналізу зображення та розпізнавання основ. У ході розпізнавання основ дані контролю phiX використовували для одержання оцінок погрішності та повторного калібрування неопрацьованих оцінок якості для інших зразків. Після первинного розпізнавання основ здійснювали додаткову фільтрацію за допомогою програмного забезпечення Illumina, при якій ріді виключали, якщо оцінка шуму перевищує граничні значення, визначені при розпізнаванні основ за допомогою Illumina pipeline. За перетворенням результатів розпізнавання основ у формат FASTQ (з використанням програмного забезпечення Illumina CASSAVA) слідувала додаткова фільтрація, при якій ріді підрізали та фільтрували відповідно до відповідної оцінки якості для кожної основи (де основи з якісною оцінкою нижче 10 підрізали з 3'-кінця рідів).

ПРИКЛАД 7

In silico збагачення та нанесення на фізичну карту даних про послідовність відносно геномного стандарту

Для додаткового збагачення стосовно розщеплених мегануклеазою і легованих з адаптером фрагментів геномної ДНК набір рідів з експерименту по секвенуванню фільтрували з використанням оригінального сценарію, який шукає пари рідів або одиночні об'єкти, де щонайменше один член з пари (або одиночний об'єкт) містить маркерну послідовність, GCAGGACGT (SEQ ID NO: 9), на початку рідів або її комплементарну послідовність, ACGTCCTGC (SEQ ID NO: 10), в кінці рідів. Пари або одиночні об'єкти рідів, що збігаються із цією послідовністю, записували в новий файл для використання на фазі нанесення на карту, тоді як інші відкидали.

Для возз'єднання обох половин розщепленого геномного сайту розпізнавання набір збагачених рідів наносили на фізичну карту відносно цільового стандартного геному з використанням програми bowtie версії 0.12.7 (Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.) для ідентифікації положень у геномі, гомологічних рідів. Налаштування вирівнювання для специфічності коректували індивідуально в кожному конкретному випадку залежно від характеристик цільового геному та від подібності цього геному з даними вихідного генотипу. Отримані в результаті вирівнювання потім використовували для виявлення піка та ідентифікації варіантних послідовностей розпізнавання, які присутні в геномі.

ПРИКЛАД 8

Виявлення піка та ідентифікація геномного варіантного сайту розпізнавання та композиція з використанням геномних вирівнювань, отриманих у прикладі 7, виявлення піка виконували з використанням алгоритму MAC (модельний аналіз Chip-Seq) для виявлення піка в Genedata Expressionist Refiner 7.5 (Genedata) з параметрами, перерахованими в таблиці 1.

Таблиця 1

Налаштування в алгоритмі MAC (модельний аналіз Chip-Seq)
для виявлення піка в Genedata Expressionist Refiner 7.5 (Genedata)

Експериментальна група	Зразки для ChiP з контролем
Спосіб	MACS
Мінімальна кратність збагачення	10
Максимальна кратність збагачення	1,00E+09
Ширина полоси	200 п.о.
Порогове значення P	1,00E-05

Хромосомні ділянки зі значним збагаченням у порівнянні з порожнім контролем експортували в excel. Ділянкам з найбільшою відмінністю між обробленими зразками та порожнім контролем віддавали перевагу, і їх підтверджували як збагачені в порівнянні з порожнім контролем, і такі, що мають сигнатуру піка, яка є результатом розщеплення геномної

ДНК, у функціональному компоненті Genome Browser в Genedata Expressionist Refiner 7.5. Як показано на фігурі 1, сигнатури піка для сайту розпізнавання містять дані про послідовність, що походять і розходяться із сайтом розщеплення, причому центр, що перекривається, відповідає липкому кінцю, створеному засобом для індукції двониткового розриву. Виходячи з липкого кінця, визначеного за сигнатурою піка, можна ідентифікувати точну послідовність варіантного сайту розпізнавання, який присутній в геномній ДНК, як показано на фігурі 2. Для визначення правильної орієнтації послідовностей геномного варіантного сайту розпізнавання, їх вирівнювали з припустимим цільовим сайтом як у змістовій, так і в антизмістовій орієнтації, при цьому найбільш відповідну орієнтацію використовували як послідовність сайту розпізнавання. Невелика частка геномних варіантних сайтів розпізнавання, що однаково добре відповідають обом орієнтаціям. Їх залишали в змістовій орієнтації. Орієнтовані геномні варіантні сайти розпізнавання вирівнювали та відсотковий склад нуклеотидів ДНК розраховували для кожного окремого положення сайту розпізнавання та порівнювали з припустимим сайтом розпізнавання. Як показано на фігурі 3, деякі положення в сайтах розпізнавання проявляли перевагу до сторонніх нуклеотидів; перевага по відношенню до нуклеотиду, відмінного від того, на який відбувається націлювання в припустимому сайті розпізнавання.

ПРИКЛАД 9

In vitro аналізи розщеплення плазмідної ДНК

Активність розщеплення мегануклеазою *in vitro* також можна аналізувати з використанням плазмідної ДНК. Для порівняння активності розщеплення мегануклеазою у припустимих і варіантних сайтах розпізнавання гібридизовані олігонуклеотиди (синтезовані Integrated DNA Technologies, Inc.), які містять припустимий або варіантний сайт розпізнавання з липкими кінцями EcoRI HindIII, клонували в сайти рестрикційних ендонуклеаз HindIII і EcoRI у плазміді pBluescript SK+ (Stratagene, зараз компанія Agilent Technologies) і аналізували *in vitro* активність розщеплення ДНК, як описано в прикладі 3, з наступними модифікаціями. Хронометровані розщеплення здійснювали з 0,25 нМ субстрату-лінеаризованою плазмідною, що містить один припустимий або варіантний сайт розпізнавання, з 25 нМ очищеного білка мегануклеази. *In vitro* аналізи проводили при 37 °C, 28 °C і 23 °C для найкращої оцінки активності розщеплення в заданому варіантному сайті розпізнавання. Реакційні суміші після зупинки реакцій очищали за допомогою колонки для очищення ПЛР продуктів Qiagen PCR purification column відповідно до інструкцій виробника та очищену ДНК розводили в 200 разів перед кількісним визначенням активності розщеплення або % втрати сайтів розпізнавання за допомогою qPCR.

ПРИКЛАД 10

Ідентифікація варіантних сайтів розпізнавання та вплив переваг до сторонніх нуклеотидів на активність розщеплення мегануклеазою

Для оцінки впливу, який переваги до сторонніх нуклеотидів здійснюють на активність розщеплення мегануклеазою, переваги до сторонніх нуклеотидів вводили в припустимий сайт розпізнавання окремо та у комбінації (див. фігуру 4), у такий спосіб створюючи варіантні сайти розпізнавання. Приклади таких варіантних сайтів розпізнавання показано в таблиці 2 для мегануклеази LIG3-4 і в таблиці 3 для мегануклеази MHP14+.

Таблиця 2. Перелік варіантних сайтів розпізнавання мегануклеази LIG3-4. Нуклеотиди, поміщені в клітини з білою заливкою, вказують на модифікацію в порівнянні з нуклеотидом з відповідним розташуванням в припустимому сайті розпізнавання (SEQ ID NO:13).

5

		нуклеотидне положення еталонного нуклеотиду (SEQ ID NO: 13)																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
SEQ ID NO:	Назва варіабельного сайту розпізнавання																						
13	передбачуваний LIG3-4	A	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
15	-11C	C	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
16	-7C	A	T	A	T	C	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
17	-2G	A	T	A	T	A	C	C	T	C	G	C	A	C	G	T	A	C	G	C	G	T	A
18	-1T	A	T	A	T	A	C	C	T	C	A	T	A	C	G	T	A	C	G	C	G	T	A
19	+8T	A	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	T	G	T	A
20	-7C, +8T	A	T	A	T	C	C	C	T	C	A	C	A	C	G	T	A	C	G	T	G	T	A
21	-11C, -7C, -2G, -1T, +8T	C	T	A	T	C	C	C	T	C	G	T	A	C	G	T	A	C	G	T	G	T	A
22	-11C, -7C, -1T, +8T	C	T	A	T	C	C	C	T	C	A	T	A	C	G	T	A	C	G	T	G	T	A

Таблиця 3. Перелік варіантних сайтів розпізнавання мегануклеази МНР14+. Нуклеотиди, поміщені в клітини з білою заливкою, вказують на модифікацію в порівнянні з нуклеотидом з відповідним розташуванням в припустимому сайті розпізнавання (SEQ ID NO:14).

SEQ ID NO:	назва варіабельного сайту розпізнавання	нуклеотидне положення еталонного нуклеотиду (SEQ ID NO:13 або SEQ ID NO:14)																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
14	Передбачуваний МНР14+	С	А	А	А	С	А	Г	А	Т	Т	С	А	С	Г	Т	С	А	Г	А	Т	Т	Т
23	-3A	С	А	А	А	С	А	Г	А	А	Т	С	А	С	Г	Т	С	А	Г	А	Т	Т	Т
24	-2G	С	А	А	А	С	А	Г	А	Т	Г	С	А	С	Г	Т	С	А	Г	А	Т	Т	Т
25	-1T	С	А	А	А	С	А	Г	А	Т	Т	Т	А	С	Г	Т	С	А	Г	А	Т	Т	Т
26	+2A	С	А	А	А	С	А	Г	А	Т	Т	С	А	А	Г	Т	С	А	Г	А	Т	Т	Т
27	+7T	С	А	А	А	С	А	Г	А	Т	Т	С	А	С	Г	Т	С	А	Т	А	Т	Т	Т
28	+8G	С	А	А	А	С	А	Г	А	Т	Т	С	А	С	Г	Т	С	А	Г	Г	Т	Т	Т
29	+11G	С	А	А	А	С	А	Г	А	Т	Т	С	А	С	Г	Т	С	А	Г	А	Т	Т	Г
30	+11A	С	А	А	А	С	А	Г	А	Т	Т	С	А	С	Г	Т	С	А	Г	А	Т	Т	А
31	-3A, -2G, -1T, +2A, +7T, +8G, +11G	С	А	А	А	С	А	Г	А	А	Г	Т	А	А	Г	Т	С	А	Т	Г	Т	Т	Г
32	-3A, -2G, -1T, +2A, +7T, +8G, +11A	С	А	А	А	С	А	Г	А	А	Г	Т	А	А	Г	Т	С	А	Т	Г	Т	Т	А
33	-3A, -2G, -1T, +7T, +8G, +11G	С	А	А	А	С	А	Г	А	А	Г	Т	А	С	Г	Т	С	А	Т	Г	Т	Т	Г
34	-2G, -1T, +2A, +7T, +8G, +11G	С	А	А	А	С	А	Г	А	Т	Г	Т	А	А	Г	Т	С	А	Т	Г	Т	Т	Г
35	-2G, -1T, +7T, +8G, +11G	С	А	А	А	С	А	Г	А	Т	Г	Т	А	С	Г	Т	С	А	Т	Г	Т	Т	Г

5 сайтів розпізнавання на субстратах-плазмідній ДНК, як описано в прикладі

9.

Варіантні сайти розпізнавання потім клонували в сайти для рестрикційних ендонуклеаз HindIII і EcoRI у векторі pBluescript SK+ і мегануклеазну активність аналізували шляхом визначення % розщеплення або % втрати сайтів розпізнавання на субстратах-плазмідній ДНК, як описано в прикладі 9.

При введенні окремо в їхні відповідні припустимі сайти розпізнавання Lig3-4 і МНР14+ переваги до сторонніх нуклеотидів надавали активність розщеплення плазмідної ДНК рівну або більшу, ніж припустимий сайт розпізнавання (див. фігуру 5А і 5В). Хоча нуклеотиди, що не переважають, не розщеплювалися добре (див. фігуру 5А і 5В). Цікаво, що навіть переваги до сторонніх нуклеотидів у межах центральних 4 основ (+2, +1, -1, -2), як повідомлялося, не були безпосередньо пов'язані з підвищеною активністю розщеплення хомінг-ендонуклеазою I-CreI (Jurica et al. (1998) Mol. Cell 2:469-76; Grizot et al. (2011) Nucleic Acids Res. 39:6124-36, і Ulge et al. (2011) Nucleic Acids Res. 39:4330-9)). Одним виключенням із цього була слабка активність розщеплення для аденіну в положенні +2 для мегануклеази МНР14+. Проте, якщо всі переважні нуклеотиди для МНР14+ оцінювали в комбінації, як показано на фігурі 6В (сайти розпізнавання -3A, -2G, -1T, +2A, +7T, +8G, +11G і -3A, -2G, -1T, +2A, +7T, +8G, +11A), спостерігали активність розщеплення, значно вищу, ніж така у припустимому сайті розпізнавання. Також, якщо +2A

видаляли, як у сайті розпізнавання -3A, -2G, -1T, +7T, +8G, +11G, було присутнє невелике підвищення активності розщеплення (див. фігуру 6B), вказуючи на те, що розщеплення по аденіну в положенні +2 є контекст-специфічним і залежить від розпізнавання або конформації нуклеотидних основ, що прилягають до нього.

У комбінації, переваги до сторонніх нуклеотидів виявляють адитивний вплив на ефективність розщеплення, причому найкращі комбінації розщеплюються приблизно в 5-6 разів ефективніше, ніж припустимий сайт розпізнавання (фігура 6). Багато з варіантних сайтів розпізнавання можна навіть розщеплювати при температурах до 23 °C у випадку, коли тільки невелика активність розщеплення спостерігалася в припустимому сайті розпізнавання (фігура 6A і B). Деякі комбінації переважних сторонніх нуклеотидів в 4 центральних основах (+2,+1,-1,-2) і в безпосередньо прилягаючих положеннях (+3,-3) впливали на величину набутої активності розщеплення в порівнянні з припустимим сайтом розпізнавання.

Узяті в сукупності, дані авторів даного винаходу вказують на те, що способи, описані в даному документі, можна застосовувати для виведення переважних контактів основ ДНК, утворених мегануклеазою в окремих положеннях уздовж її ДНК-зв'язувальної поверхні контакту, забезпечуючи ретельну оцінку специфічності розщеплення, забезпечуючи новий підхід до вивчення специфічності мегануклеази в контексті геномної ДНК. Способи авторів даного винаходу також дозволяють ідентифікацію варіантних сайтів розпізнавання, які розщеплюються більш ефективно, ніж припустимий сайт розпізнавання.

ПРИКЛАД 11

Застосування варіантних сайтів розпізнавання у рослини або тварин

Варіантні сайти розпізнавання, ідентифіковані в прикладі 10, з покращеною активністю розщеплення в порівнянні з припустимим сайтом розпізнавання (-7C, 8T (Lig3-4); -11C, -7C, -2G, -1T, +8T (Lig3-4); -11C, -7C, -1T, +8T (Lig3-4); -3A, -2G, -1T, +2A, +7T, +8G, +11G (MHP14+); -3A, -2G, -1T, +2A, +7T, +8G, +11A (MHP14+); -3A, -2G, -1T, +7T, +8G, +11G (MHP14+); -2G, -1T, +2A, +7T, +8G, +11G (MHP14+); -2G, -1T, +7T, +8G, +11G (MHP14+), що відповідають SEQ ID NO: 13-35), або будь-який інший варіантний сайт розпізнавання, ідентифіковані за допомогою способу, описаного в даному документі, можна трансформувати в геном будь-якої рослини або тварини та піддати цілеспрямованому впливу для мутагенезу або вставки гена. Оскільки активність розщеплення в цих сайтах розпізнавання є підвищеною в порівнянні з припустимим сайтом розпізнавання, також можна підвищити і частоти модифікації сайту, в тому числі делеції, вставки або будь-якої комбінації цих двох модифікацій. Варіантні сайти розпізнавання також можна розташувати окремо або в комбінації на трансгенних касетах експресії, забезпечуючи можливість зміни, вирізання або вставки трансгенних частин.

ПРИКЛАД 12

Застосування до інших реактивів, що рідко розщеплюють, для індукції двониткового розриву

Оскільки олігонуклеотиди як з 5'-, так і з 3'-виродженими кінцями можна синтезувати в широкому діапазоні специфічних для користувача конфігурацій і значень довжини (Integrated DNA Technologies, Inc) і гібридизувати з утворенням двониткових ДНК адаптерів з 5'-, або з 3'-виродженим липким кінцем, встановлені в даному документі способи можуть бути застосовні до будь-якого реактиву, що рідко розщеплює, для індукції двониткового розриву, який створює липкий кінець із основ ДНК при розщепленні. Вони будуть включати, але не обмежуються іншими хомінг-ендонуклеазами, нуклеазами "цинкові пальці" і TALEN.

ПРИКЛАД 13

Трансформація незрілих зародків маїсу

Трансформація може бути здійснена за допомогою різних способів, про які відомо, що вони є ефективними в рослинах, включаючи опосередковану частинками доставку, опосередковану *Agrobacterium* трансформацію, опосередковану PEG доставку та електропорацію.

а. Опосередкована частинками доставка

Трансформацію незрілих зародків маїсу за допомогою опосередкованої частинками доставки проводили в такий спосіб. Склади середовищ представлені нижче.

Качани очищували від листової обгортки та піддавали поверхневій стерилізації в 30 % відбілювачі Сіорох разом з 0,5 % миючого засобу Мікро протягом 20 хвилин, та ополіскували двічі стерильною водою. Незрілі зародки виділяли й поміщали рубчиком униз (щитком нагору), 25 зародків на планшет, у середовище 560Y на 4 години і потім вирівнювали в межах 2,5-см зони-мішені в препараті для бомбардування. Альтернативно, виділені зародки поміщали в 560L (середовище для ініціювання) і поміщали в темряву при температурах, що варіюють у діапазоні від 26 °C до 37 °C, на 8-24 години перед поміщенням в 560Y на 4 години при 26 °C перед бомбардуванням, як описано вище.

Плазміди, що містять ген засобу, що індукує двонитковий розрив, і донорну ДНК,

конструювали за допомогою стандартних методик молекулярної біології та піддавали спільному бомбардуванню із плазмідами, що містять контролюючі розвиток гени ODP2 (фактор транскрипції ODP2 з доменами AP2 (відповідальний за розвиток сім'язародку білок 2); US20090328252 A1) і Wushel (US2011/0167516).

Плазміди і ДНК, що представляє інтерес, осаджували на гранулах золота розміром 0,6 мкм (середній діаметр) за допомогою водорозчинного катіонного ліпиду Tfx™-50 (№ по кат. E1811, Promega, Медісон, Вісконсин, США) у такий спосіб. Розчин ДНК готували на льоду із застосуванням 1 мкг плазмідної ДНК і необов'язкових інших конструкцій для спільного бомбардування, таких як 50 нг (0,5 мкл) кожної плазміди, що містить контролюючі розвиток гени ODP2 (фактор транскрипції ODP2 з доменами AP2 (відповідальний за розвиток сім'язародку білок 2); US20090328252 A1) і Wushel. До попередньо змішаної ДНК додавали 20 мкл підготовлених часток золота (15 мг/мл) і 5 мкл TFX-50 у воді та обережно перемішували. Частки золота осаджували в мікроцентрифузі при 10000 об./хв. протягом 1 хв. і супернатант видаляли. Отриманий у результаті осад обережно ополіскували за допомогою 100 мл 100 % EtOH без ресуспендування осаду й промивний EtOH обережно видаляли. Додавали 105 мкл 100 % EtOH і частки ресуспендували шляхом нетривалої ультразвукової обробки. Потім 10 мкл наносили на центральну частину кожного макроносія і їм дозволяли висохнути приблизно за 2 хвилини до бомбардування.

Альтернативно, плазміди і ДНК, що представляє інтерес, осаджували на гранулах вольфраму розміром 1,1 мкм (середній діаметр) за допомогою процедури осадження хлоридом кальцію (CaCl₂) шляхом змішування 100 мкл підготовлених часток вольфраму у воді, 10 мкл (1 мкг) ДНК у буфері Tris- EDTA (1 мкг загальної ДНК), 100 мкл 2,5 М CaCl₂ і 10 мкл 0,1 М спермідину. Кожен реактив додавали послідовно до суспензії часток вольфраму з перемішуванням. Кінцеву суміш піддавали нетривалій ультразвуковій обробці та забезпечували інкубацію при постійному перемішуванні на вихровій мішалці протягом 10 хвилин. Після періоду осадження пробірки центрифугували протягом нетривалого часу, рідину видаляли, а частки промивали 500 мл 100 % етанолу з наступним центрифугуванням протягом 30 секунд. Рідину знову видаляли і до кінцевого осаду часток вольфраму додавали 105 мкл 100 % етанолу. Для бомбардування генною гарматою частки вольфраму із ДНК піддавали нетривалій ультразвуковій обробці. 10 мкл часток вольфраму із ДНК наносили на центральну частину кожного макроносія, після чого нанесеним часткам дозволяли висохнути приблизно за 2 хвилини до бомбардування.

Планшети зі зразками бомбардували на рівні № 4 за допомогою гелієвої генної гармати Biorad. Усі зразки одержували однократний постріл при 450 фунт/кв. дюйм, при цьому загальна кількість аліквот, взятих з кожної пробірки з підготовленими частками із ДНК, становила десять.

Після бомбардування зародки інкубували в 560P (підтримуючому середовищі) протягом 12-48 годин при температурах, що варіюють у діапазоні від 26 °C до 37 °C, а потім поміщали в умови температури 26 °C. Через 5-7 днів зародки переносили в селективне середовище 560R, що містить 3 мг/літр біолафосу, і субкультивували кожні 2 тижні при 26 °C. Через приблизно 10 тижнів відбору калюсні клони, стійкість яких була виявлена в результаті відбору, переносили на середовище 288J для ініціювання регенерації рослин. Після дозрівання соматичних зародків (2-4 тижні) добре розвинені соматичні зародки переносили в середовище для проростання та переносили в освітлювану кімнату для культивування. Через приблизно 7-10 днів саджанці, що розвиваються, переносили в безгормональне середовище 272V у пробірки на 7-10 днів, поки саджанці добре вкореняться. Рослини потім переносили у вставки в лотках (еквівалентні горщику на 2,5 дюйма), що містять ґрунтову горщикову суміш, і вирощували протягом 1 тижня у вегетаційній камері, згодом вирощуючи протягом додаткових 1-2 тижнів у теплиці, переносючи потім у горщики Classic 600 (1,6 галона) і вирощуючи до зрілості. За рослинами стежили та оцінювали в балах по ефективності трансформації та/або модифікації регенеративних здібностей.

Середовище для ініціювання (560L) містить 4,0 г/л основних солей N6 (SIGMA C-1416), 1,0 мл/л вітамінної суміші Еріксона (1000X SIGMA-1511), 0,5 мг/л тіаміну-HCl, 20,0 г/л сахарози, 1,0 мг/л 2,4-D і 2,88 г/л L-проліну (доведених до об'єму за допомогою D-I H₂O після доведення pH до 5,8 за допомогою KOH); 2,0 г/л Gelrite (доданого після доведення до об'єму за допомогою D-I H₂O) і 8,5 мг/л нітрату срібла (доданого після стерилізації середовища та охолодження до кімнатної температури).

Підтримуюче середовище (560P) містить 4,0 г/л основних солей N6 (SIGMA C-1416), 1,0 мл/л вітамінної суміші Еріксона (1000X SIGMA-1511), 0,5 мг/л тіаміну-HCl, 30,0 г/л сахарози, 2,0 мг/л 2,4-D і 0,69 г/л L-проліну (доведених до об'єму за допомогою D-I H₂O після доведення pH до 5,8 за допомогою KOH); 3,0 г/л Gelrite (доданого після доведення до об'єму за допомогою D-I

H₂O) і 0,85 мг/л нітрату срібла (доданого після стерилізації середовища та охолодження до кімнатної температури).

Середовище для бомбардування (560Y) містить 4,0 г/л основних солей N6 (SIGMA C-1416), 1,0 мл/л вітамінної суміші Еріксона (1000X SIGMA-1511), 0,5 мг/л тіаміну-HCl, 120,0 г/л сахарози, 1,0 мг/л 2,4-D і 2,88 г/л L-проліну (доведених до об'єму за допомогою D-I H₂O після доведення pH до 5,8 за допомогою KOH); 2,0 г/л Gelrite (доданого після доведення до об'єму за допомогою D-I H₂O) і 8,5 мг/л нітрату срібла (доданого після стерилізації середовища та охолодження до кімнатної температури).

Селективне середовище (560R) містить 4,0 г/л основних солей N6 (SIGMA C-1416), 1,0 мл/л вітамінної суміші Еріксона (1000X SIGMA-1511), 0,5 мг/л тіаміну-HCl, 30,0 г/л сахарози й 2,0 мг/л 2,4-D (доведених до об'єму за допомогою D-I H₂O після доведення pH до 5,8 за допомогою KOH); 3,0 г/л Gelrite (доданого після доведення до об'єму за допомогою D-I H₂O) і 0,85 мг/л нітрату срібла та 3,0 мг/л біалафосу (обидва з яких додані після стерилізації середовища та охолодження до кімнатної температури).

Середовище для регенерації рослин (288J) містить 4,3 г/л солей по MS (GIBCO 11117-074), 5,0 мл/л вихідного розчину вітамінів по MS (0,100 г нікотинової кислоти, 0,02 г/л тіаміну-HCl, 0,10 г/л піридоксину-HCl і 0,40 г/л гліцину, доведених до об'єму за допомогою очищеної D-I H₂O) (Murashige and Skoog (1962) *Physiol. Plant.* 15:473), 100 мг/л міоінозиту, 0,5 мг/л зеатину, 60 г/л сахарози й 1,0 мл/л 0,1 мМ абсцизової кислоти (доведених до об'єму очищеною D-I H₂O після доведення pH до 5,6); 3,0 г/л Gelrite (доданого після доведення до об'єму за допомогою D-I H₂O), а також 1,0 мг/л індолілоцтової кислоти й 3,0 мг/л біалафосу (доданих після стерилізації середовища та охолодження до 60 °C).

Безгормональне середовище (272V) містить 4,3 г/л солей по MS (GIBCO 11117-074), 5,0 мл/л вихідного розчину вітамінів по MS (0,100 г/л нікотинової кислоти, 0,02 г/л тіаміну-HCl, 0,10 г/л піридоксину-HCl і 0,40 г/л гліцину, доведених до об'єму за допомогою очищеної D-I H₂ PRO), 0,1 г/л міоінозит і 40,0 г/л сахарози (доведених до об'єму за допомогою очищеної D-I H₂ після доведення pH до 5,6); і 6 г/л бактоагару (доданого після доведення до об'єму за допомогою очищеної D-I H₂O), її стерилізували та охолоджували до 60 °C.

b. Опосередкована *Agrobacterium* трансформація

Опосередковану *Agrobacterium* трансформацію проводили по суті відповідно до описаного в Djukanovic et al. (2006) *Plant Biotech J* 4:345-57. Коротко, незрілі зародки у віці 10-12 днів (розміром 0,8-2,5 мм) вирізали зі стерилізованих зерен і поміщали в рідке середовище (4,0 г/л основних солей N6 (Sigma C-1416), 1,0 мл/л вітамінної суміші Еріксона (Sigma E-1511), 1,0 мг/л тіаміну-HCl, 1,5 мг/л 2,4-D, 0,690 г/л L-проліну, 68,5 г/л сахарози, 36,0 г/л глюкози, pH 5,2). Після збору зародків середовище заміняли на 1 мл *Agrobacterium* у концентрації 0,35-0,45 OD₅₅₀. Зародки маїсу інкубували з *Agrobacterium* протягом 5 хв. при кімнатній температурі, потім суміш виливали на планшети із середовищем, що містить 4,0 г/л основних солей N6 (Sigma C-1416), 1,0 мл/л вітамінної суміші Еріксона (Sigma E-1511), 1,0 мг/л тіаміну-HCl, 1,5 мг/л 2,4-D, 0,690 г/л L-проліну, 30,0 г/л сахарози, 0,85 мг/л нітрату срібла, 0,1 мМ ацетосірингону і 3,0 г/л Gelrite, pH 5,8. Зародки інкубували рубчиком униз у темряві протягом 3 днів при 20 °C, потім інкубували протягом 4 днів у темряві при 28 °C, потім переносили на нові планшети із середовищем, що містить 4,0 г/л основних солей N6 (Sigma C-1416), 1,0 мл/л вітамінної суміші Еріксона (Sigma E-1511), 1,0 мг/л тіаміну-HCl, 1,5 мг/л 2,4-D, 0,69 г/л L-проліну, 30,0 г/л сахарози, 0,5 г/л буфера MES, 0,85 мг/л нітрату срібла, 3,0 мг/л біалафосу, 100 мг/л карбеніциліну й 6,0 г/л агару, pH 5,8. Зародки піддавали субкультивуванню кожні три тижні до ідентифікації трансгенних об'єктів. Соматичний ембріогенез індукували шляхом переносу невеликої кількості тканини в середовище для регенерації (4,3 г/л солей по MS (Gibco 11117), 5,0 мл/л вихідного розчину вітамінів по MS, 100 мг/л міоінозиту, 0,1 мкМ ABA, 1 мг/л IAA, 0,5 мг/л зеатину, 60,0 г/л сахарози, 1,5 мг/л біалафосу, 100 мг/л карбеніциліну, 3,0 г/л Gelrite, pH 5,6) та інкубації в темряві протягом двох тижнів при 28 °C. Весь матеріал з видимими пагонами та коріннями переносили в середовище, що містить 4,3 г/л солей по MS (Gibco 11117), 5,0 мл/л вихідного розчину вітамінів по MS, 100 мг/л міоінозиту, 40,0 г/л сахарози, 1,5 г/л Gelrite, pH 5,6, та інкубували під штучним освітленням при 28 °C. Через тиждень сіянці переміщали в скляні пробірки, що містили те ж саме середовище, і вирощували до відбору зразків з них та/або пересадження їх у ґрунт.

ПРИКЛАД 14

Транзйентна експресія ВВМ підсилює трансформацію

Параметри протоколу трансформації можуть бути модифіковані для підтвердження того, що активність ВВМ є транзйентною. Один такий спосіб включає осадження плазмід, що містить ВВМ, способом, який дозволяє проходити транскрипції та експресії, але запобігає наступному вивільненню ДНК, наприклад, за допомогою застосування хімічної сполуки PEI.

В одному прикладі плазміді з BBM осаджували на частках золота за допомогою PEI, тоді як трансгенну касету експресії (UBI:moPAT~GFPm:PinII; moPAT являє собою оптимізований для експресії в маїсі ген PAT), яка підлягає інтеграції, осаджували на частинках золота за допомогою стандартного способу із застосуванням хлориду кальцію.

Коротко, частинки золота покривали за допомогою PEI у такий спосіб. Спочатку частинки золота промивали. Тридцять п'ять мг частинок золота із середнім діаметром 1,0 (A.S.I. № 162-0010) відважували в мікроцентрифужну пробірку, і додавали 1,2 мл абсолютного EtOH, і перемішували вихровою мішалкою протягом однієї хвилини. Пробірку інкубували протягом 15 хвилин при кімнатній температурі, а потім центрифугували з високою швидкістю, використовуючи мікроцентрифугу, протягом 15 хвилин при 4 °C. Надосадову рідину зливали, і додавали свіжу аліквоту етанолу (EtOH) на 1,2 мл, перемішували вихровою мішалкою протягом однієї хвилини, центрифугували протягом однієї хвилини, і надосадову рідину знову зливали (це повторювали двічі). Додавали свіжу аліквоту EtOH на 1,2 мл, і цю суспензію (частинки золота в EtOH) зберігали при -20 °C протягом декількох тижнів. Для покриття частинок поліетиленіміном (PEI; Sigma №P3143) 250 мкл суміші промитих частинок золота/EtOH центрифугували, і EtOH зливали. Частинки промивали один раз в 100 мкл ddH₂O для видалення залишкового етанолу, додавали 250 мкл 0,25 мМ PEI з наступною імпульсною ультразвуковою обробкою для суспендування частинок, і потім пробірку занурювали у ванну із сухим льодом/EtOH для миттєвої заморозки суспензії, яку потім ліофілізували протягом ночі. На даній стадії сухі покриті частинки можуть зберігатися при -80 °C протягом щонайменше 3 тижнів. Перед використанням частинки ополіскували 3 рази за допомогою аліквот 2,5 мМ буфера HEPES на 250 мкл, pH 7,1, з 1х імпульсною ультразвуковою обробкою та наступним швидким перемішуванням вихровою мішалкою перед кожним центрифугуванням. Потім частки суспендували в кінцевому об'ємі 250 мкл буфера HEPES. Аліквоту частинок на 25 мкл додавали в невикористані пробірки перед прикріпленням ДНК. Для прикріплення непокритої ДНК частинки піддавали імпульсній ультразвуковій обробці, потім додавали 1 мкг ДНК (в 5 мкл води) з наступним перемішуванням шляхом втягування й випуску піпеткою кілька разів за допомогою Pipetteman та інкубували протягом 10 хвилин. Частки центрифугували протягом нетривалого часу (тобто 10 секунд), надосадову рідину видаляли, і додавали 60 мкл EtOH. Частки осадженої за допомогою PEI ДНК-1 промивали двічі в 60 мкл EtOH. Частинки центрифугували, надосадову рідину зливали й частки ресуспендували в 45 мкл води. Для прикріплення другої ДНК (ДНК-2) застосовували осадження за допомогою TFX-50. 45 мкл суспензії частинки/ДНК-1 піддавали нетривалій ультразвуковій обробці, і потім додавали 5 мкл 100 нг/мкл ДНК-2 і 2,5 мкл TFX-50. Розчин поміщали на ротаційний струшувач на 10 хвилин, центрифугували при 10000 g протягом 1 хвилини. Супернатант видаляли й частинки ресуспендували в 60 мкл EtOH. Розчин наносили на макроносії, і частинки золота, до яких були послідовно прикріплені ДНК-1 і ДНК-2, доставляли в клітини щитка незрілих зародків Hi-II на стадії 10 DAP, використовуючи стандартний протокол для PDS-1000. У цьому експерименті плазміді із ДНК-1 містила касету експресії UBI:RFP:pinII, а плазміді із ДНК-2 містила касету експресії UBI:CFP:pinII. Через два дні після бомбардування спостерігали транзйентну експресію флуоресцентних маркерів як CFP, так і RFP у вигляді чисельних червоних і синіх клітин на поверхні незрілого зародка. Зародки потім поміщали в неселективне культуральне середовище, і їм дозволяли рости протягом 3 тижнів перед оцінюванням у балах стабільних колоній. Після цього 3-тижневого періоду спостерігали 10 багатоклітинних, стабільно експресуючих синіх колоній у порівнянні з лише однією червоною колонією. Це показало, що осадження за допомогою PEI можна застосовувати для ефективного введення ДНК для транзйентної експресії, при цьому суттєво знижуючи ступінь інтеграції, осадженої за допомогою PEI ДНК, що вводиться і, таким чином, ослаблюючи регенерацію трансгенних об'єктів, експресуючих RFP. Таким чином, осадження за допомогою PEI можна застосовувати для забезпечення транзйентної експресії BBM і/або WUS2.

Наприклад, частинки спочатку покривали UBI:BBM:pinII за допомогою PEI, потім покривали UBI: moPAT ~TFP за допомогою TFX-50, а потім вводили в клітини щитка на поверхні незрілих зародків за допомогою бомбардування. Осадження, опосередковане PEI, обумовлює високу частоту зустрічі клітин із транзйентною експресією на поверхні незрілого зародка та надзвичайно низькі частоти регенерації стабільних трансформантів (у порівнянні зі способом з використанням TFX-50). Таким чином, припускається, що осаджена за допомогою PEI касета з BBM експресується транзйентно й стимулює сплеск ембріогенного росту на поверхні тканини, яка бомбардується (тобто поверхні щитка), але ця плазміді не буде інтегруватися. Припускається, що плазміді з PAT~GFP, звільнена від частинок золота, на яких вона була осаджена за допомогою Ca⁺⁺, інтегрується та експресує маркер, який селектується, із частотою, яка обумовлює суттєво покращену регенерацію трансгенних об'єктів. У якості

контрольної обробки осаджені за допомогою PEI частки, що містять UBI:GUS:pinII (замість BBM), змішували із частками з PAT~GFP, осадженими за допомогою Ca⁺⁺. Незрілі зародки після обох обробок переміщали в культуральне середовище, що містить 3 мг/л біолафосу. Припускають, що GFP+ калюси, стійкі до біолафосу, будуть спостерігатися через 6-8 тижнів з набагато вищою частотою у випадку обробки PEI/BBM у порівнянні з контрольною обробкою (PEI/GUS).

У якості альтернативного способу плазмиду з BBM осаджували на частинках золота за допомогою PEI і потім вводили в клітини щитка на поверхні незрілих зародків, і наступна транзйентна експресія гена BBM викликала швидку проліферацію в ході росту ембріогенного калюсу. Протягом даного періоду індукованого росту експланти обробляли *Agrobacterium*, використовуючи стандартні способи для маїсу (див. приклад 1), з доставкою T-ДНК у клітину за допомогою введення трансгенної касети експресії, такої як UBI: moPAT ~GFPm:pinII. Після спільного культивування експлантам дозволяли регенерувати в нормальному культуральному середовищі, і їх потім переміщали в культуральне середовище, що містить 3 мг/л біолафосу. Припускають, що GFP+ калюси, стійкі до біолафосу, будуть спостерігатися через 6-8 тижнів з набагато вищою частотою у випадку обробки PEI/BBM у порівнянні з контрольною обробкою (PEI/GUS).

Може бути необхідним "активізувати" ріст калюсу шляхом транзйентної експресії BBM і/або полінуклеотидних продуктів WUS2. Це може бути виконано шляхом доставки 5'-кепованої поліаденільованої РНК BBM і WUS2, касет експресії, що містять ДНК BBM і WUS2, або білків BBM і/або WUS2. Всі ці молекули можуть бути доставлені за допомогою біолістичної генної гармати. Наприклад, 5'-кепована поліаденільована BBM і/або WUS2 РНК може бути легко створена *in vitro*, застосовуючи набір Ambion's mMessage mMachine. РНК піддавали спільній доставці разом із ДНК, що містить полінуклеотид, який представляє інтерес, і маркер, який використовується для відбору/скринінгу, як наприклад, Ubi: moPAT ~GFPm:PinII. Припускається, що клітини, які отримують РНК, негайно почнуть ділитися швидше, і більша частина цих клітин буде мати інтегрований агрономічно важливий ген. Може бути додатково підтверджено, що ці трансгенні об'єкти являють собою трансгенні клональні колонії, тому що вони будуть також експресувати білок злиття PAT~GFP (і, таким чином, будуть проявляти зелену флуоресценцію при відповідному освітленні). Рослини, регенеровані з даних зародків, можуть потім зазнати скринінгу щодо наявності полінуклеотиду, який представляє інтерес.

ПРИКЛАД 15

Одержання та трансформація в модельній системі культур соматичних зародків сої векторами експресії в сої і регенерація рослин

Умови культивування

Ембріогенні суспензійні культури сої (сорт Jack) підтримували в 35 мл рідкого середовища SB196 (нижче) на ротаційному струшувачі при 150 об./хв. і 26 °C з люмінесцентними лампами холодного білого світла з фотоперіодом 16:8 годин день/ніч при інтенсивності світла 60-85 мкЕрг/м²/с. Культури пересівали від одного разу на 7 днів до одного разу на два тижні шляхом інокуляції приблизно 35 мг тканини в 35 мл свіжого рідкого SB196 (кращий інтервал між пересіваннями становив 7 днів).

Ембріогенні суспензійні культури сої трансформували за допомогою плазмід експресії для сої за допомогою способу бомбардування частинками (Klein et al., *Nature*, 327:70 (1987)) із застосуванням обладнання DuPont Biolistic PDS1000/HE (гелієва модифікована установка) для всіх трансформацій.

Ініціація ембріогенних суспензійних культур сої

Здійснювали ініціацію культур сої двічі на місяць із інтервалом 5-7 днів між кожною ініціацією. Стручки з незрілими насіннями від доступних рослин сої збирали через 45-55 днів після посадки. Насіння видаляли зі стручків і поміщали в стерилізований контейнер Magenta. Насіння сої стерилізували шляхом струшування протягом 15 хв. в 5 % розчині Clorox з 1 краплею мила "Ivory" (тобто 95 мл автоклавованої дистильованої води плюс 5 мл Clorox і 1 краплею мила, ретельно змішане). Насіння промивали з використанням 2 1-літрових флаконів зі стерильною дистильованою водою і насіння розміром менше 4 мм поміщали на окремі предметні скельця мікроскопу. Зрізали гострий кінець насіння та видавлювали сім'ядолі з насінної оболонки. Коли культури були підготовлені для проведення трансформації, сім'ядолі переносили на плашки, що містять середовище SB1 (25-30 сім'ядоль на плашку). Планшети обертали волокнистою стрічкою й підтримували при 26 °C з освітленням флуоресцентними лампами холодного білого з фотоперіодом день/ніч 16:8 годин при інтенсивності світла 60-80 мкЕрг/м²/с протягом восьми тижнів із заміною середовища кожні 4 тижні. Коли культури були підготовлені для експериментів у модельній системі, сім'ядолі переносили на плашки, що

містять середовище SB199 (25-30 сім'ядоль на плашку), на 2 тижні, а потім переносили на SB1 на 2-4 тижні. Умови освітлення і температури були такими ж, як описано вище. Після інкубування на середовищі SB1 вторинні зародки вирізали й поміщали в рідке середовище SB196 на 7 днів.

5 Одержання ДНК для бомбардування

Для бомбардування використовували або інтактну плазмиду, або фрагмент ДНК плазмиди, що містить гени, які представляють інтерес, і маркерний ген, що селектується. Фрагменти з плазмід експресії в сої одержували шляхом виділення розщеплених плазмід у гелі. У кожному випадку 100 мкг плазмідної ДНК використовували в 0,5 мл специфічної суміші ферментів, описаної нижче. Плазмиди розщеплювали Ascl (100 одиниць) у буфері NEBuffer 4 (20 мМ Tris-ацетат, 10 мМ ацетат магнію, 50 мМ ацетат калію, 1 мМ дітіотриетол, рН 7,9), 100 мкг/мл BSA і 5 мМ бета-меркаптоетанолі при 37 °С протягом 1,5 години. Отримані фрагменти ДНК розділяли за допомогою електрофорезу в гелі на 1 % агарозі SeaPlaque GTG (BioWhittaker Molecular Applications) і фрагменти ДНК, що містять касети з генами, вирізали з агарозного гелю. ДНК очищали від агарози з використанням ферменту GELase, який розщеплює, згідно із протоколом виробника.

50 мкл аліквоту, що містить 3 мг золотих частинок (3 мг золота), стерильної дистильованої води додавали до 30 мкл розчину ДНК у концентрації 10 нг/мкл (або інтальної плазмиди, або фрагменту ДНК, отриманого, як описано в даному документі), 25 мкл 5 М CaCl₂ і 20 мкл 0,1 М спермидину. Суміш струшували 3 хвилини на струшувачі Vortex при рівні 3 і центрифугували протягом 10 секунд на настільній мікроцентрифузі. Супернатант видаляли з наступним промиванням 400 мкл 100 % етанолу та ще одного короткого центрифугування. 400 мкл етанолу видаляли і осад ресуспендували в 40 мкл 100 % етанолу. П'ять мкл суспензії ДНК розподіляли на кожен диск-носій в обладнанні для біолістичного бомбардування Biolistic PDS1000/HE. Кожна 5 мкл аліквота містила приблизно 0,375 мг золота на бомбардування (наприклад, на диск).

Для трансформацій у модельній системі протокол був ідентичним за винятком декількох незначних змін (тобто 1 мг золотих частинок додавали до 5 мкл 1 мкг/мкл розчину ДНК, використовували 50 мкл 2,5 М CaCl₂ і осад в остаточному підсумку ресуспендували в 85 мкл 100 % етанолу, у такий спосіб забезпечуючи 0,058 мг золотих частинок на бомбардування).

30 Одержання тканин і бомбардування ДНК

Приблизно 150-200 мг семиденних ембріогенних суспензійних культур поміщали в порожню стерильну 60 x 15 мм чашку Петрі, чашку накривали пластмасовою сіткою. Забезпечували розрідження в камері до зниженого тиску 27-28 дюймів ртутного стовпа й тканину бомбардували одним або двома пострілами на плашку, причому тиск розривам мембрани встановлювали на 1100 фунт/дюйм². Тканину поміщали приблизно в 3,5 дюймах від затримуючого/обмежувального екрану. Умови трансформації в модельній системі були ідентичними за винятком використання 100-150 мг ембріогенної тканини, тиск розриву встановлювали на 650 фунт/дюйм² і тканину поміщали приблизно в 2,5 дюймах від затримуючого екрану.

40 Відбір трансформованих зародків

Відбір трансформованих зародків проводили або з використанням гігromіцину (коли ген гігromіцин-В-фосфотрансферази (HPT) використовували в якості маркера, який селектується) або хлорсульфурона (коли ген ацетолактат-синтази (ALS) використовували в якості маркера, який селектується).

Після бомбардування тканину поміщали у свіже середовище SB196 і культивували, як описано вище. Через шість-вісім днів після бомбардування SB196 заміняли свіжим SB196, що містить або 30 мг/л гігromіцину, або 100 нг/мл хлорсульфурону залежно від використовуваного маркера, який селектується. Селективне середовище оновлювали щотижня. Через чотири-шість тижнів після відбору спостерігали ріст зеленої трансформованої тканини з нетрансформованих некротичних ембріогенних кластерів.

50 Дозрівання зародків

Для отримання трансформантів виділену зелену тканину видаляли та інокулювали в багатолункові планшети для утворення за допомогою клонального розмноження нових трансформованих ембріогенних суспензійних культур. Трансформовані ембріогенні кластери культивували протягом чотирьох-шести тижнів у багатолункових планшетах при 26 °С в SB196 під флуоресцентними лампами холодного білого світла (Phillips cool white Econowatt F40/CW/RS/EW) і лампами Agro (Phillips F40 Agro) (40 Вт) з фотоперіодом 16:8 год. при інтенсивності освітлення 90-120 мкЕрг/м²с. Після цього періоду часу зародкові кластери видаляли у твердому агаровому середовищі, SB166, протягом одного-двох тижнів, а потім

субкультивували в середовище SB103 на 3-4 тижні для дозрівання зародків. Після дозрівання на планшетах в SB103 окремі зародки видаляли із кластерів, сушили та піддавали скринінгу відносно змін у складі їх жирних кислот, як описано в прикладі 7.

- Для трансформацій у модельній системі дозрівання зародків забезпечували в рідкому середовищі для тканинного диференціювання та дозрівання сої (рідке середовище SHaM; Schmidt et al., Cell Biology and Morphogenesis 24:393 (2005)) з використанням модифікованої процедури. Коротко, після 4 тижнів відбору в SB196, як описано вище, зародкові кластери видаляли в 35 мл SB228 (рідкого середовища SHaM) в 250 мл колбу Ерленмеєра. Тканину підтримували в рідкому середовищі SHaM на ротаційному струшувачі при 130 об./хв. і при 26 °C з освітленням флуоресцентними лампами холодного білого з фотоперіодом день/ніч 16:8 годин при інтенсивності світла 60-85 мкЕрг/м²/с протягом 2 тижнів до дозрівання зародків. Зародки вирощували протягом 2 тижнів у рідкому середовищі SHaM були еквівалентними за розміром та вмістом жирних кислот зародкам, культив

Прописи середовищ

- 15 SB 196 - рідке середовище для проліферації FN Lite (на літр)
MS FeEDTA - 100x матковий розчин 1 10 мл
MS сульфат - 100x матковий розчин 2 10 мл
Галогеніди FN Lite - 100x матковий розчин 3 10 мл
P, B, Mo FN Lite - 100x матковий розчин 4 10 мл
20 Вітаміни B5 (1 мл/л) 1,0 мл
2,4-D (кінцева концентрація 10 мг/л) 1,0 мл
KNO₃ 2,83 г
(NH₄)₂SO₄ 0,463 г
Аспарагін 1,0 г
25 Сахароза (1 %) 10 г
pH 5,8

Маткові розчини FN Lite

- | Номер маткового розчину | 1000 мл | 500 мл |
|--------------------------------------|---------|---------|
| 1 MS Fe EDTA, 100x матковий розчин | | |
| 30 Na ₂ EDTA* | 3,724 г | 1,862 г |
| FeSO ₄ -7H ₂ O | 2,784 г | 1,392 г |

*Додають у першу чергу, потім розчиняють у флаконі з темного скла при перемішуванні

- 2 MS сульфат, 100x матковий розчин
- | | | |
|---|----------|-----------|
| 35 MgSO ₄ -7H ₂ O | 37,0 г | 18,5 г |
| MnSO ₄ -H ₂ O | 1,69 г | 0,845 г |
| ZnSO ₄ -7H ₂ O | 0,86 г | 0,43 г |
| CuSO ₄ -5H ₂ O | 0,0025 г | 0,00125 г |
- 3 Галогеніди FN Lite, 100x матковий розчин
- | | | |
|---|----------|-----------|
| 40 CaCl ₂ -2H ₂ O | 30,0 г | 15,0 г |
| KI | 0,083 г | 0,0715 г |
| CoCl ₂ - 6H ₂ O | 0,0025 г | 0,00125 г |
- 4 P, B, Mo FN Lite, 100x матковий розчин
- | | | |
|---|---------|----------|
| 45 KH ₂ PO ₄ | 18,5 г | 9,25 г |
| H ₃ BO ₃ | 0,62 г | 0,31 г |
| Na ₂ MoO ₄ -2H ₂ O | 0,025 г | 0,0125 г |

Щільне середовище SB1 (на літр)

- 1 упаковка солей MS (Gibco/ BRL - № у каталозі 11117-066)
1 мл вітамінів B5, 1000X матковий розчин
50 31,5 г глюкози
2 мл 2,4-D (кінцева концентрація 20 мг/л)
pH 5,7
8 г агару TC

Щільне середовище SB199 (на літр)

- 55 1 упаковка солей MS (Gibco/ BRL - № у каталозі 11117-066)
1 мл вітамінів B5, 1000X матковий розчин
30 г сахарози
4 мл 2,4-D (кінцева концентрація 40 мг/л)
pH 7,0
60 2 г Gelrite

- Щільне середовище SB 166 (на літр)
 1 упаковка солей MS (Gibco/ BRL - № у каталозі 11117-066)
 1 мл вітамінів B5, 1000X матковий розчин
 60 г мальтози
 5 750 мг гексагідрату $MgCl_2$
 5 г активованого вугілля
 pH 5,7
 2 г gelrite
- Щільне середовище SB 103 (на літр)
 10 1 упаковка солей MS (Gibco/ BRL - № у каталозі 11117-066)
 1 мл вітамінів B5, 1000X матковий розчин
 60 г мальтози
 750 мг гексагідрату $MgCl_2$
 pH 5,7
 15 2 г gelrite
- Щільне середовище SB 71-4 (на літр)
 1 флакон солей Gamborg's B5 із сахарозою (Gibco/BRL - № у каталозі 21153-036)
 pH 5,7
 5 г агару TC
 20 Матковий розчин 2,4-D
 Одержують попередньо приготованим від Phytotech № у каталозі D 295 - концентрація 1 мг/мл.
- Матковий розчин вітамінів B5 (на 100 мл)
 Аліквоти зберігають при -20 °C
 25 10 г міоїнозиту
 100 мг нікотинової кислоти
 100 мг піридоксину-HCl
 1 г тіаміну
 Якщо розчин не розчиняється досить швидко, застосовують низький рівень нагрівання за допомогою гарячої плити для нагрівання.
 30 SB 228 - середовище для тканинної диференціації та дозрівання сої (SHaM) (на літр)
 DDI H₂O 600 мл
 Макросолі FN-Lite для SHaM, 10X 100 мл
 Мікросолі MS, 1000x 1 мл
 35 MS FeEDTA, 100x 10 мл
 CaCl₂, 100x 6,82 мл
 Вітаміни B5, 1000x 1 мл
 L-метіонін 0,149 г
 Сахароза 30 г
 40 Сорбіт 30 г
 Доводять об'єм до 900 мл
 pH 5,8
 Автоклавують
 Додають в охолоджене середовище (<30 °C):
 45 *глутамін (кінцева концентрація 30 mM) 4 % 110 мл
 *Примітка: Кінцевий об'єм буде становити 1010 мл після додавання глутаміну.
 Оскільки глутамін розкладається відносно швидко, його краще додавати безпосередньо перед застосуванням середовища. Закінчення строку придатності відбувається через 2 тижні після додавання глутаміну; базове середовище можна зберігати довше без глутаміну.
- 50 Макросолі FN-Lite для SHaM, 10X - матковий розчин № 1 (на літр)
 (NH₄)₂SO₄ (сульфат амонію) 4,63 г
 KNO₃ (нітрат калію) 28,3 г
 MgSO₄*7H₂O (гептагідрат сульфату магнію) 3,7 г
 KH₂PO₄ (фосфат калію, одноосновний) 1,85 г
 55 Доводять до об'єму
 Автоклавують
 Мікросолі MS, 1000X - матковий розчин № 2 (на 1 літр)
 H₃BO₃ (борна кислота) 6,2 г
 MnSO₄*H₂O (моногідрат сульфату марганцю) 16,9 г
 60 ZnSO₄*7H₂O (гептагідрат сульфату цинку) 8,6 г

- $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ (дигідрат молібдату натрію) 0,25 г
 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (пентагідрат сульфату міді) 0,025 г
 $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ (гексагідрат хлориду кобальту) 0,025 г
 KI (йодид калію) 0,8300 г
- 5 Доводять до об'єму
Автоклавують
- Fe EDTA , 100X - матковий розчин № 3 (на літр)
- Na_2EDTA^* (EDTA натрію) 3,73 г
 $\text{Fe SO}_4 \cdot 7\text{H}_2\text{O}$ (гептагідрат сульфату заліза) 2,78 г
- 10 *EDTA повинен повністю розчинитися перед додаванням заліза.
Доводять до об'єму
Розчин є світлочутливим. Флакон(и) повинні бути загорнуті у фольгу щоб уникнути світла.
Автоклавують
- Ca, 100X - матковий розчин № 4 (на літр)
- 15 $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (дигідрат хлориду кальцію) 44 г
Доводять до об'єму
Автоклавують
- Вітаміни B5, 1000X - матковий розчин № 5 (на літр)
- Тіамін*HCl 10 г
 Нікотинова кислота 1 г
 Піридоксин*HCl 1 г
 Міоїнозит 100 г
- 20 Доводять до об'єму
Зберігають замороженим
- 25 4 % Глутамін - матковий розчин № 6 (на літр)
DDI вода, підігріта до 30 °C 900 мл
L-глутамін 40 г
Поступово додають при перемішуванні та застосуванні слабкого нагрівання.
Не перевищують температуру 35 °C.
- 30 Доводять до об'єму
Стерилізують фільтруванням
Зберігають замороженим*
- *Примітка: підігрівають розморожений матковий розчин на бані температурою 31 °C для повного розчинення кристалів.
- 35 Регенерація соматичних зародків сої в рослини
- Для отримання цілих рослин з ембріогенних суспензійних культур тканину потрібно регенерувати. Забезпечували дозрівання зародків, як описано вище. Після субкультивування на середовищі SB103 протягом 3 тижнів окремі зародки можна вилучити з кластерів і піддати скринінгу щодо змін у складі їх жирних кислот, як описано в прикладі 7. Слід зазначити, що на цій стадії можна піддати скринінгу будь-який фенотип, який виявляється, і який є результатом експресії генів, які представляють інтерес. Він включає без обмеження зміни в профілі жирних кислот, профіль білків і їх склад, склад вуглеводів, швидкість росту, життєздатність або здатність до нормального розвитку в рослину сої.
- 40 Зрілі окремі зародки висушували шляхом поміщення їх у порожню невелику чашку Петрі (35 x 10 мм) приблизно на 4-7 днів. Планшети обертали волокнистою стрічкою (створюючи невелику камеру з регульованою вологістю). Висушені зародки висаджували в середовище SB71-4, де їх залишили проростати за тих самих умов культивування, що описані вище. Пророщені саджанці видаляли із середовища для проростання та ретельно промивали водою, а потім висаджували в середовище Redi-Earth в 24-лунковий пакувальний підніс, накривали прозорим пластиковим куполом. Через 2 тижні купол видаляли і рослини піддавали загартуванню протягом додаткового тижня. Якщо саджанці виглядали загартованими, їх пересаджували в 10" горщик із середовищем Redi-Earth у кількості до 3 саджанців на горщик. Після 10-16 тижнів зріле насіння збирали, лушили та аналізували відносно жирних кислот.
- 50 ПРИКЛАД 16
- 55 Система для скринінгу по відношенню до мегануклеазної активності на основі дріжджів
- Створювали штами дріжджів для скринінгу в якості хазяїв для скринінгу по відношенню до мегануклеазної активності. Ген дріжджів Ade2 (Genetika 1987 Jul-23(7):1141-8) (SEQ ID NO: 36) використовували як маркерний ген з видимим проявом, а також як маркер, який селектується, на схемі, зображеній на фігурі 7. Фрагменти гена, що відповідають першим 1000 нуклеотидам кодувальної послідовності, Ade2 (5'-фрагмент Ade2) і останнім 1011 нуклеотидам кодувальної
- 60

послідовності, Ade2 (3'-фрагмент Ade2), руйнували за допомогою

Між 5'-фрагментом Ade2 і 3'-фрагментом Ade2 знаходились 305 нуклеотидів дуплікованих послідовностей. Отримані в результаті конструкції використовували для заміщення гена Ade2 (положення нуклеотидів 566193-564480 у хромосомі 15) штаму дріжджів BY4247. Отримані в результаті штами дріжджів для скринінгу VER8145, VER8189 і HD1327 можна охарактеризувати як BY4742 MATa his3delta1 leu2delta0 lys2delta0 ura3delta0 Gal2+. Якщо розрізання мегануклеазою відбувається між дуплікованими послідовностями, може відбуватися гомологічна рекомбінація, у результаті чого утворюється функціональний ген Ade2.

Утворення функціонального гена Ade2 можна застосовувати в якості критерію відбору: якщо дріжджові клітини ростуть на середовищі з нестачею аденіну, то до росту здатні тільки ті з них, які мають функціональний ген Ade2.

Утворення функціонального гена Ade2 також можна використовувати як критерій скринінгу. Клітини дріжджів з функціональним геном Ade2 були білими, тоді як такі, у яких відсутнє функціонування Ade2, проявляли червоне забарвлення у зв'язку з більш раннім нагромадженням метаболіту в шляху біосинтезу аденіну, що в результаті призводить до утворення червоних колоній з білими секторами, показаному на фігурах 8 і 9. Ступінь утворення білих секторів, що іноді поширюється на цілі колонії, вказує на величину активності розрізання, властивій мегануклеазам. Оскільки фенотип утворення секторів є якісним показником мегануклеазної активності, задіяли числову систему кількісних показників 0-4. Як показано на фігурі 9, кількісний показник 0 означає, що білі сектори не спостерігалися (було відсутнє розрізання мегануклеазами); кількісний показник 4 означає повністю білі колонії (повне розрізання сайту розпізнавання); кількісні показники 1-3 означають проміжні фенотипи утворення білих секторів (і проміжні ступені розрізання сайту розпізнавання).

ПРИКЛАД 17

Аналіз варіантних послідовностей розпізнавання для мегануклеази LIG3-4 і MHP14+ у маїсі Для демонстрації активності розщеплення ідентифікованих у даному документі варіантних сайтів розпізнавання in planta варіантний сайт розпізнавання, який зустрічається в природі, (SEQ ID NO: 11) для мегануклеази MHP14+ ідентифікували в маїсі та активність його розщеплення in planta (вимірювана по частоті мутагенезу сайту розпізнавання) і активність розщеплення in vitro порівнювали з активністю розщеплення припустимого сайту розпізнавання для MHP14+.

Для визначення активності розщеплення сайтів розпізнавання in planta плазмідну ДНК, що містить касету експресії для мегануклеази MHP14+, доставляли в зародки маїсу за допомогою бомбардування частинками для забезпечення можливості виникнення двониткового розриву з наступною ПЛР у реальному часі, яку здійснювали з аналізами TaqMan, які охоплюють сайти розпізнавання. Відносне число копій розраховували за допомогою способу $\Delta\Delta Ct$ відносно внутрішнього контролю в аналізі TaqMan з використанням нетрансформованих зародків в якості калібруючого стандарту. Зародки з відносним числом копій менше 0,8 вважали розщепленими та/або мутуваними. Активність розщеплення мегануклеазою in vitro аналізували, як описано в прикладі 9 на плазмідній ДНК.

Активність розщеплення плазмідної ДНК in vitro у варіантному сайті розпізнавання SEQ ID NO:11 була приблизно в 3 рази ефективнішою, ніж у припустимому сайті розпізнавання з SEQ ID NO: 14, залежно від температури реакційної суміші (див. фігуру 10A). Подібну тенденцію спостерігали in planta, причому 35 % зрілих демонстрували розщеплення в ендегенному варіантному сайті розпізнавання, що підтверджувалося мутагенезом, хоча тільки 15 % демонстрували розщеплення (мутагенез) у припустимому сайті розпізнавання (див. фігуру 10B).

Подібний аналіз можна здійснювати для варіантних сайтів розпізнавання для LIG3-4, описаних у даному документі або для будь-яких інших ідентифікованих варіантних сайтів розпізнавання. Очікується, що частоти модифікації сайту, у тому числі делеції, вставки або будь-які комбінації цих двох модифікацій, будуть підвищуватися для інших варіантних сайтів розпізнавання з покращеною активністю розщеплення, ідентифікованих у прикладі 10 (-7C, 8T (Lig3-4); -11C, -7C, -2G, -1T, +8T (Lig3-4); -11C, -7C, -1T, +8T (Lig3-4); -3A, -2G, -1T, +2A, +7T, +8G, +11G (MHP14+); -3A, -2G, -1T, +2A, +7T, +8G, +11A (MHP14+); -3A, -2G, -1T, +7T, +8G, +11G (MHP14+); -2G, -1T, +2A, +7T, +8G, +11G (MHP14+); -2G, -1T, +7T, +8G, +11G (MHP14+), що відповідають SEQ ID NO: 13-35) або будь-якого іншого варіантного сайту розпізнавання, ідентифікованого за допомогою способу, описаного в даному документі, які будуть штучно вбудовувати в геном.

ПРИКЛАД 18

Аналіз варіантних послідовностей розпізнавання для мегануклеази LIG3-4 і MHP14+ у сої Для тестування припустимих і варіантних сайтів розпізнавання для мегануклеаз LIG3-4 і

MHP14+ у дводольній рослині, такий як соя, послідовності сайтів розпізнавання маїсу можна клонувати в ДНК конструкції для трансформації та вводити в сою за допомогою біолістичної трансформації, як описано в прикладі 15.

Порівняння декількох сайтів розпізнавання, розташованих в одному тому ж локусі

Для порівняння активності розщеплення (ефективностей розрізання) у різних послідовностей розпізнавання можна розташувати кілька послідовностей розпізнавання в одній ДНК конструкції та ввести різні конструкції, що містять кілька сайтів розпізнавання переважно в тому самому геномному локусі для виключення ефектів розташування. Трансформаційна система на основі FLP/ FRT-опосередкованої сайт-специфічної інтеграції є цінним інструментом для досягнення вищевикладеної мети шляхом розташування різних донорних ДНК конструкцій у раніше охарактеризованих цільових сайтах (Plant Physiology, Li et al., 2009; заявка на патент США №12/634775). Після того, як сайти розпізнавання інтегрувалися в геном сої, трансгенні трансформанти з однією копією можна ідентифікувати, характеризувати та вибирати в якості нових матеріалів для наступної трансформації відповідними мегануклеазами для оцінки активності розщеплення (ефективності розрізання) кожного сайту розпізнавання відповідною йому мегануклеазою. Оскільки кілька сайтів розпізнавання введені в той самий геномний сайт, ефективність розрізання відповідними мегануклеазами можна порівняти.

Із цією метою припустимий сайт розпізнавання для LIG34 (SEQ ID NO:13), а також припустимі сайти розпізнавання для MHP14+ (SEQ ID NO:14) і MS26 (SEQ ID NO: 37) клонували в донорну конструкцію для SSI PHP57712 (SEQ ID NO: 38) між геном маркера, який селекується, GM-ALS і касет з генами ознак надекспресії DGAT2 (діацилгліцерол-ацилтрансферази) для надекспресії для високого вмісту масла, штучної мікроРНК для косупресії FAD3 (ω-3-десатурази) для високого вмісту ненасичених жирних кислот і косупресії GAS (галактинол-синтази) за допомогою шпильки для високого рівня доступної енергії (фігура 11 А). Трансгенні трансформанти з донорною ДНК PHP57712, інтегрованою в декілька раніше охарактеризованих геномних сайтів, одержували за допомогою біолістичної SSI трансформації сої частинками, як описано вище. Трансгенні трансформанти із чистими вставками генів ознак і послідовностей розпізнавання для мегануклеази відбирали, і їх будуть використовувати для наступного раунду трансформації мегануклеазами LIG34, MS26 і MHP14+ для дослідження активності розщеплення трьох припустимих сайтів розпізнавання.

Порівняння припустимих сайтів розпізнавання у порівнянні та варіантних сайтів розпізнавання, розташованих у тому самому локусі

Варіантний сайт розпізнавання для LIG3-4 (SEQ ID NO:22) і варіантний сайт розпізнавання для MHP14+ (SEQ ID NO:35), а також припустимий сайт розпізнавання для MS26 (SEQ ID NO: 37) клонували в іншу донорну конструкцію SSI PHP62252 (SEQ ID NO: 39) (фігура 11 В) і трансформували в деякі з тих же геномних сайтів сої за допомогою SSI трансформації. Трансгенні трансформанти із чистими вставками генів ознак і послідовностей розпізнавання для мегануклеази будуть відбирати і будуть використовувати для іншого раунду трансформації мегануклеазами MS26, MHP14+ і LIG34 для дослідження активностей розщеплення сайтів розпізнавання. Оскільки всі сайти розпізнавання введені в ті самі геномні сайти, ефективності розрізання різних сайтів розпізнавання можна порівняти більш змістовно. Активність розщеплення варіантних сайтів розпізнавання для LIG3-4 і MHP14+ (SEQ ID NO: 22 і 35) можна порівняти з активністю розщеплення їх припустимим сайтам розпізнавання (SEQ ID NO: 13 і 14), як описано в прикладі 17.

ПРИКЛАД 19

Спосіб націлювання вставки полінуклеотиду, який представляє інтерес, у специфічний хромосомний сайт у геномі рослини

Нуклеотидну послідовність, що містить варіабельну послідовність розпізнавання засобом для індукції двониткового розриву, вводили в геном цільового організму, досягаючи цільового сайту (який містить варіабельну послідовність розпізнавання) для вставки нуклеотидних послідовностей, які представляють інтерес. Можна потім отримати бібліотеку стабільних рослин або культивованих тканин, яка містить варіабельний сайт розпізнавання в різних положеннях по всьому геному рослини.

Одним прикладом таких варіантних сайтів розпізнавання є SEQ ID NO: 15, 16, 17, 18, 19, 20 і 21, які можуть розщеплюватися мегануклеазою LIG3-4, яка кодується SEQ ID NO: 1. Іншим прикладом варіантних сайтів розпізнавання є SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35, які можуть розщеплюватися мегануклеазою MHP14+, яка кодується SEQ ID NO: 3. В одному варіанті здійснення SEQ ID NO: 15, 16, 17, 18, 19, 20 і 21 і SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35 не є ендегенними відносно геному маїсу або рослини.

Після того, як були отримані стабільна рослина або культивована тканина, фрагмент ДНК,

який містить полінуклеотид, який представляє інтерес, вводили в стабільно трансформовану рослину або тканини з наявністю білка, що індукує двонитковий розрив, такого як білок мегануклеази. Цей процес призводить в результаті до вставки полінуклеотиду, який представляє інтерес, у варіабельну послідовність розпізнавання.

Вважається, що трансформована рослина може містити кілька цільових сайтів, наприклад, без обмеження кілька сайтів розпізнавання, які можуть розщеплюватися засобом для індукції двониткового розриву, а також сайти рекомбінації, такі як сайти FRT або сайти LOX. Приклади сайтів розпізнавання відомі в даній технічній галузі та включають сайти FRT (див., наприклад, Schlake and Bode (1994) *Biochemistry* 33:12746-12751; Huang et al. (1991) *Nucleic Acids Research* 19:443-448).

ПРИКЛАД 20

Вловлювання геномних варіантних сайтів розпізнавання ендонуклеазою Cas і створення бібліотек для глибокого секвенування Illumina

Для вловлювання в геномній ДНК варіантних сайтів розпізнавання для засобів, що рідко розщеплюють ДНК, для індукції двониткового розриву, у яких більша частина розщеплених продуктів дасть в результаті тупі кінці, як наприклад, для ендонуклеаз Cas (Gasiunas et al. (2012) *Proc. Natl. Acad. Sci. USA* 109:E2579-86, Jinek et al. (2012) *Science* 337:816-21), можна використовувати приєднання аденіну до 3'-кінців розщепленого(их) варіантного(их) сайту(ів) розпізнавання в геномній ДНК. Адаптери, що містять комплементарний 3'-тиміновий липкий кінець можна потім використовувати для селективного лігування з тупими кінцями, які утворюються в результаті розщеплення ендонуклеазою Cas, і збагачення ними.

Для створення матеріалу для вловлювання геномних варіантних сайтів розпізнавання будуть проводити *in vitro* аналізи розщеплення, та очищення будуть проводити, по суті, як описано в прикладі 3, за винятком того, що замість білка мегануклеази будуть використовувати очищений білок ендонуклеази Cas і компонент(и)-нуклеїнові кислоти, необхідні для утворення функціонального комплексу ендонуклеази Cas, здатного до розщеплення цільового сайту ДНК. Реакції *in vitro* можна проводити в різних буферах, при різних температурах або з різними значеннями довжини(довжин) інкубації для одержання ідеальних умов розщеплення ендонуклеазою Cas.

3'-липкий кінець з одного аденіну будуть потім приєднувати до розщеплених ендонуклеазою Cas тупих кінців при інкубуванні з розщепленою *in vitro* геномною ДНК при 37 °C протягом 30 хвилин в 50 мкл реакційної суміші, що містить АТР, 1× буфер Кленова (NEBnext) і 15 фрагмента Кленова (без екзонуклеазної активності) і очищати. Нефосфорильовані або фосфорильовані біотинільовані адаптери, синтезовані та очищені за допомогою HPLC, що містять 3'-тиміновий нуклеотидний липкий кінець, комплементарний 3'-аденіновому нуклеотидному липкому кінцю, можна потім лігувати із приблизно 2 мкг розщепленої ендонуклеазою Cas подовженою аденіном 3' геномною ДНК в 100 мкл реакційної суміші з лігазою T4 (NEB). Отримані в результаті ліговані з адаптером, розщеплені ендонуклеазою Cas припустимі та варіантні сайти розпізнавання можна потім збагачувати, секвенувати й ідентифікувати, подібно описаним у прикладах 4, 6, 7 і 8.

РСТ_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ
ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<110> Е. І. ДЮ ПОН ДЕ НЕМУР ЕНД КОМПАНІ
ПАЙАНІР ХАЙ-БРЕД ІНТЕРНЕТНЛ, ІНК
ДЕШАМП, Стефан
ІНГЛІШ, Джеймс
ЛІ, Жонгсен
ллака, Віктор
янг, Джошуа

<120> СПОСОБИ ІДЕНТИФІКАЦІЇ ВАРІАНТНИХ САЙТІВ РОЗПІЗНАВАННЯ ДЛЯ
СКОНСТРУЙОВАНИХ ЗАСОБІВ, ЩО РІДКО РОЗЩЕПЛЮЮТЬ, ДЛЯ ІНДУКЦІЇ
ДВОНИТКОВОГО РОЗРИВУ, КОМПОЗИЦІЇ З НИМИ ТА ЇХ ЗАСТОСУВАННЯ

<130> ВВ2223 РСТ

<150> US 61/777238

<151> 2013-03-12

<160> 39

<170> PatentIn версія 3.5

<210> 1

<211> 1053

<212> ДНК

<213> штучна послідовність

<220>

<223> Мегануклеаза LIG3-4

<400> 1

atgaacaccca agtacaacaa ggagttcctg ctctacctgg ccggcttcgt ggacggcgac	60
ggctccatca aggcgcagat caagccgaac cagtcctgca agttcaagca ccagctctcc	120
ctgaccttcc aggtgaccca gaagacgcag aggcgctggt tcctcgacaa gctggctgac	180
gagatcgggg tgggctacgt ctacgaccgc gggctcgggt cgcactacga gctctcccag	240
atcaagcccc tgcacaactt cctcaccag ctccagccgt tcctcaagct gaagcagaag	300
caggcgaacc tcgtcctgaa gatcatcgag cagctcccct cggccaagga gtccccggac	360
aagttcctgg aggtgtgcac gtgggtcgac cagatcgcg cctcaacga cagcaagacc	420
cgcaagacga cctcggagac ggtgcgggcg gtcctggact ccctcccagg atccgtggga	480
ggtctatcgc catctcaggc atccagcgcc gcatcctcgg ctctcctcaag ccggggttca	540
gggatctccg aagcactcag agctggagca actaagtcca aggaattcct gctctacctg	600
gccggcttcg tggacggcga cgggtccatc atcgcgtcca tcaagccgcg ccagtgtac	660
aagttcaagc acgagctccg cctggagttc accgtgaccc agaagacgca gaggcgctgg	720
ttcctcgaca agctggctga cgagatcggg gtgggttacg tctacgaccg cgggtcgggtg	780
tccgactacc gcctctccca gatcaagccc ctgcacaact tcctcaccsa gctccagccg	840
ttcctcaagc tgaagcagaa gcaggcgaac ctctcctga agatcatcga gcagctcccc	900
tcggccaagg agtccccgga caagttcctg gaggtgtgca cgtgggtcga ccagatcgcg	960
gccctcaacg acagcaagac ccgcaagacg acctcggaga cgggtcggggc ggtcctggac	1020
tcctcagcg agaagaagaa gtcgtccccc tga	1053

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<210> 2
 <211> 350
 <212> БІЛОК
 <213> штучна послідовність

<220>
 <223> Мегануклеаза LIG3-4

<400> 2

Met Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe
 1 5 10 15

Val Asp Gly Asp Gly Ser Ile Lys Ala Gln Ile Lys Pro Asn Gln Ser
 20 25 30

Cys Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys
 35 40 45

Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val
 50 55 60

Gly Tyr Val Tyr Asp Arg Gly Ser Val Ser Asp Tyr Glu Leu Ser Gln
 65 70 75 80

Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys
 85 90 95

Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu
 100 105 110

Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp
 115 120 125

Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr
 130 135 140

Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Pro Gly Ser Val Gly
 145 150 155 160

Gly Leu Ser Pro Ser Gln Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser
 165 170 175

Ser Pro Gly Ser Gly Ile Ser Glu Ala Leu Arg Ala Gly Ala Thr Lys
 180 185 190

Ser Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly
 195 200 205

Ser Ile Ile Ala Ser Ile Lys Pro Arg Gln Cys Tyr Lys Phe Lys His
 210 215 220

Glu Leu Arg Leu Glu Phe Thr Val Thr Gln Lys Thr Gln Arg Arg Trp

225 PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ 230 235 240

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp
245 250 255

Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Gln Ile Lys Pro Leu His
260 265 270

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
275 280 285

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

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
305 310 315 320

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
325 330 335

Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
340 345 350

<210> 3

<211> 1053

<212> ДНК

<213> штучна послідовність

<220>

<223> мегануклеаза МНР14+

<400> 3

atgaacacca agtacaaca ggagttcctg ctctacctgg ccggcttcgt ggacggcgac	60
ggctccatca tcgcgcagat caagccgaac cagtcctaca agttcaagca ccagctcatg	120
ctgaccttca ccgtgaccca gaagacgcag aggcgctggg ttctcgacaa gctggctcgac	180
gagatcgggg tgggcaaggc ccgcgaccgc gggtcggtgt ccgactacat cctctcccag	240
atcaagcccc tgcacaactt cctcaccag ctccagccgt tcctcaagct gaagcagaag	300
caggcgaacc tcgtcctgaa gatcatcgag cagctcccct cggccaagga gtccccggac	360
aagttcctgg aggtgtgcac gtgggtcgac cagatcgcgg ccctcaacga cagcaagacc	420
cgcaagacga cctcgagac ggtgcgggcg gtcctggact ccctcccagg atccgtggga	480
ggtctatcgc catctcaggc atccagcgcc gcacacctcg cttcctcaag cccgggttca	540
gggatctccg aagcactcag agctggagca actaagtcca aggaattcct gctctacctg	600
gccggcttcg tggacggcga cggctccatc atcgcggcga tcaagccgaa ccagtcctac	660
aagttcaagc accagctctc cctgaccttc accgtgacct agaagacgca gaggcgctgg	720
ttcttcgaca agctggctga cgagatcggg gtgggctacg tccgcgacca ggggtcggtg	780
tccactacc agctctccca gatcaagccc ctgcacaact tcctcaccca gctccagccg	840

РСТ_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ
 ttccctcaagc tgaagcagaa gcaggcgaac ctcgtcctga agatcatcga gcagctcccc 900
 tcggccaagg agtccccgga caagttcctg gaggtgtgca cgtgggtcga ccagatcgcg 960
 gccctcaacg acagcaagac ccgcaagacg acctcggaga cgggtcgggc ggttctagac 1020
 tccttcagcg agaagaaga gtcgtccccc tga 1053

<210> 4
 <211> 350
 <212> БІЛОК
 <213> штучна послідовність

<220>
 <223> Мегануклеаза MHP14+

<400> 4

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

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

195 200 205

Ser Ile Ile Ala Ala Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His
210 215 220

Gln Leu Ser Leu Thr Phe Thr Val Thr Gln Lys Thr Gln Arg Arg Trp
225 230 235 240

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp
245 250 255

Gln Gly Ser Val Ser His Tyr Gln Leu Ser Gln Ile Lys Pro Leu His
260 265 270

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
275 280 285

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

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
305 310 315 320

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
325 330 335

Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
340 345 350

<210> 5
<211> 41
<212> ДНК
<213> штучна послідовність

<220>
<223> дефосфорильований адаптер

<220>
<221> misc_feature
<222> (38)..(41)
<223> н являє собою а, с, г або т

<400> 5
ggttgacatg ctggattgag acttcctgc aggacgtnnn n

41

<210> 6
<211> 24
<212> ДНК
<213> штучна послідовність

<220>
<223> Праймер А

<400> 6
gttgacatgc tggattgaga cttc

24

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<210>	7	
<211>	21	
<212>	ДНК	
<213>	штучна послідовність	
<220>		
<223>	Праймер В	
<400>	7	
	саагсагааг асggсатасg а	21
<210>	8	
<211>	61	
<212>	ДНК	
<213>	штучна послідовність	
<220>		
<223>	Illumina-сумісний адаптер	
<400>	8	
	aatgatacgg сgaccaccga gatctacact ctttcctac acgacgctct tccgatctgc	60
	а	61
<210>	9	
<211>	9	
<212>	ДНК	
<213>	штучна послідовність	
<220>		
<223>	маркерна послідовність	
<400>	9	
	gcaggacgt	9
<210>	10	
<211>	9	
<212>	ДНК	
<213>	штучна послідовність	
<220>		
<223>	комплементарна послідовність маркерної послідовності	
<400>	10	
	acgtcctgc	9
<210>	11	
<211>	22	
<212>	ДНК	
<213>	штучна послідовність	
<220>		
<223>	5'-3' послідовність з фігури 2	
<400>	11	
	саагаагааг сасgtсagct та	22
<210>	12	
<211>	22	
<212>	ДНК	
<213>	штучна послідовність	

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<220>
 <223> 3'-5' послідовність з фігури 2

<400> 12
 gtctctcttc gtgcagtcga at 22

<210> 13
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> передбачувана послідовність розпізнавання для LIG3-4

<400> 13
 atatacctca cacgtacgsg ta 22

<210> 14
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> передбачувана послідовність розпізнавання для MHP14+

<400> 14
 caaacagatt cacgtacgat tt 22

<210> 15
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> варіантна послідовність розпізнавання для LIG3-4, -11 C

<400> 15
 ctatacctca cacgtacgsg ta 22

<210> 16
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> варіантна послідовність розпізнавання для LIG3-4, -7C

<400> 16
 atatccctca cacgtacgsg ta 22

<210> 17
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> варіантна послідовність розпізнавання для LIG3-4, -2G

<400> 17
 atatacctcg cacgtacgsg ta 22

<210> 18

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для LIG3-4, -1T
 <400> 18
 atatacctca tacgtacgsg ta 22

<210> 19
 <211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для LIG3-4, +8T
 <400> 19
 atatacctca cacgtacgtg ta 22

<210> 20
 <211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для LIG3-4, -7C, +8T
 <400> 20
 atatccctca cacgtacgtg ta 22

<210> 21
 <211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для LIG3-4, -11C, -7C, -2G, -1T, +8T
 <400> 21
 ctatccctcg tacgtacgtg ta 22

<210> 22
 <211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для LIG3-4, -11C, -7C, -1T, +8T
 <400> 22
 ctatccctca tacgtacgtg ta 22

<210> 23
 <211> 22
 <212> ДНК
 <213> штучна послідовність
 <220>
 <223> Варіантна послідовність розпізнавання для MNP14+, -3A

РСТ_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<400> 23 сааасагаат сасгтсгагт tt	22
<210> 24 <211> 22 <212> днк <213> штучна послідовність	
<220> <223> варіантна послідовність розпізнавання для МНР14+, -2G	
<400> 24 сааасагагг сасгтсгагт tt	22
<210> 25 <211> 22 <212> днк <213> штучна послідовність	
<220> <223> варіантна послідовність розпізнавання для МНР14+, -1T	
<400> 25 сааасагагг тасгтсгагт tt	22
<210> 26 <211> 22 <212> днк <213> штучна послідовність	
<220> <223> варіантна послідовність розпізнавання для МНР14+, +2A	
<400> 26 сааасагагг саагтсгагт tt	22
<210> 27 <211> 22 <212> днк <213> штучна послідовність	
<220> <223> варіантна послідовність розпізнавання для МНР14+, +7T	
<400> 27 сааасагагг сасгтсгагт tt	22
<210> 28 <211> 22 <212> днк <213> штучна послідовність	
<220> <223> варіантна послідовність розпізнавання для МНР14+, +8G	
<400> 28 сааасагагг сасгтсгаггт tt	22
<210> 29 <211> 22 <212> днк	

РСТ_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, +11G

<400> 29
сааасаgatt сасgтсagat tg 22

<210> 30
<211> 22
<212> ДНК
<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, +11A

<400> 30
сааасаgatt сасgтсagat та 22

<210> 31
<211> 22
<212> ДНК
<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, -3A, -2G, -1T, +2A, +7T, +8G, +11G

<400> 31
сааасаgааg тааgтсatgt tg 22

<210> 32
<211> 22
<212> ДНК
<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, -3A, -2G, -1T, +2A, +7T, +8G, +11A

<400> 32
сааасаgааg тааgтсatgt та 22

<210> 33
<211> 22
<212> ДНК
<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, 3A, -2G, -1T, +7T, +8G, +11G

<400> 33
сааасаgааg тасgтсatgt tg 22

<210> 34
<211> 22
<212> ДНК
<213> штучна послідовність

<220>

<223> Варіантна послідовність розпізнавання для МНР14+, -2G, -1T, +2A, +7T, +8G, +11G

РСТ_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

[illegible]

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

tattcaaaaa	gagcatttaa	tcaaaaatgg	tatagcagtt	acccaaagt	ttcctgtgga	1380
acaagccagt	gagacgtccc	tattgaatgt	tggaagagat	ttgggttttc	cattcgtctt	1440
gaagtcgagg	actttggcat	acgatggaag	aggtaacctc	gttgtaaaga	ataaggaaat	1500
gattccgga	gctttggaag	tactgaagga	tcgtcctttg	tacgccgaaa	aatgggcacc	1560
atttactaaa	gaattagcag	tcattgattgt	gagatctgtt	aacggtttag	tgttttctta	1620
cccaattgta	gagactatcc	acaaggacaa	tatttgtgac	ttatgttatg	cgcctgctag	1680
agttccggac	tccgttcaac	ttaaggcgaa	gttgttgga	gaaaatgcaa	tcaaatcttt	1740
tcccggttgt	ggtatatattg	gtgtggaaat	gttctattta	gaaacagggg	aattgcttat	1800
taacgaaatt	gccccaggc	ctcacaactc	tggacattat	accattgatg	cttgcgtcac	1860
ttctcaattt	gaagctcatt	tgagatcaat	attggatttg	ccaatgccaa	agaatttcac	1920
atctttctcc	accattacaa	cgaacgccat	tatgctaaat	gttcttgag	acaacatac	1980
aaaagataaa	gagctagaaa	cttgcgaaag	agcattggcg	actccagggt	cctcagtgt	2040
cttatatgga	aaagagtcta	gacctaacag	aaaagtaggt	cacataaata	ttattgcctc	2100
cagtatggcg	gaatgtgaac	aaaggctgaa	ctacattaca	ggtagaactg	atattccaat	2160
caaaatctct	gtcgtcaaaa	agttggactt	ggaagcaatg	gtcaaaccat	tggttggaa	2220
catcatggga	tcagactctg	acttgccggt	aatgtctgcc	gcatgtgcgg	ttttaaaga	2280
ttttggcgtt	ccatttgaag	tgacaatagt	ctctgctcat	agaactccac	ataggatgtc	2340
agcatatgct	atttccgcaa	gcaagcgtgg	aattaaaaca	attatcgctg	gagctggtgg	2400
ggctgctcac	ttgccaggta	tggtggctgc	aatgacacca	cttcctgtca	tcggtgtgcc	2460
cgtaaaagg	tcttgtctag	atggagtaga	ttctttacat	tcaattgtgc	aaatgcctag	2520
agggtgtcca	gtagctaccg	tcgctattaa	taatagtacg	aacgctgcgc	tggtggctgt	2580
cagactgctt	ggcgcttatg	attcaagtta	tacaacgaaa	atggaacagt	ttttattaaa	2640
gcaagaagaa	gaagttcttg	tcaaagcaca	aaagttagaa	actgtcgggt	acgaagctta	2700
tctagaaaac	aagtaatata	taagtttatt	gatatacttg	tacagcaaat	aattataaaa	2760
tgatatacct	attttttagg	ctttgttatg	attacatcaa	atgtggactt	catacataga	2820
aatcaacgct	tacaggtgtc	cttttttaag	aatttcatac	ataagatcac	ttattataca	2880
tacatacata	tccagtaaca	agaagcaagg	aataattacc	tgcttaagtc	tgcgattaaa	2940
aaaataacgt	ttcgatacag	ttcatataag	gcggctcaat	gcagaaccga	ggatagcgct	3000
acgtcaggat	atctttgtag	ttcccaaata	taaatgcgac	aatatagttt	ctttctttca	3060
tatcaataat	atccttttct	ccactgaaat	cacgaatcaa	acctggagca	aaaactaaag	3120
ccaaattata	aagcgtcatt	cgattccagt	gactgtaccg	tgtaaccttt	tctatatgtt	3180
cactcagtac	tcttaacacc	ctataatgtt	cccttggaag	gtcttccaat	atgtttttta	3240
aagcgtctct	gctcgacata	taagtgtccg	aattctttgc	ttctaaggac	aactttcctc	3300
caacaaatgg	caagttttcc	atcatttttt	tagatttaac	taacctcatc	aacggctcgt	3360

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ
 atatttgaaa ggtaaagata gggttgggga gctttcttaa gtatcgcttc aacacaccag 3420
 taacaacggt gagatcttgt tccgttaaaa tatttggcgt ttcggtatgt tgttgacttt 3480
 tccatgcaga aaattgcttt tctatttctt ctatgactag ctgggaacct gattttctat 3540
 aaatgccctc cgatctcata ttttcttcgt ctgattcaat aaaatctatg cagacagata 3600
 gtatcatcgg tatttcattg ttttcataat tgcacctagc aacgagactt gaaccataca 3660
 aattgcttcc atccaaatat tcttctccat ctttactttg tcccatatgt gcatcg 3716

<210> 37
 <211> 22
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> передбачувана послідовність розпізнавання для MS26

<400> 37
 gatggtgacg tacgtgccct ac 22

<210> 38
 <211> 16824
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> плазміда РНР57712

<220>
 <221> misc_feature
 <222> (12027)..(12027)
 <223> н являє собою а, с, г або т

<400> 38
 ggagatcсаа gcttggcgcg ccggcctctg cctgcgttct gctgtggaag ttcctattcc 60
 gaagttccta ttctccagaa agtataggaa cttcacatgc tgcctcgtgc aagtcacgat 120
 ctcgagttct atagtgtcac ctaaactgta tgtgtatgat acataagggt atgtattaat 180
 tgtagccgcg ttctaacgac aatatgtcca tatgggtgcac tctcagtaca atctgtctgt 240
 atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg ccctgacggg 300
 cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg agctgcatgt 360
 gtcagagggt ttaccgctca tcaccgaaac gcgcgagacg aaagggcctc gtgatacgcc 420
 tatttttata ggtaaatgtc atgaccaaaa tcccttaacg tgagttttcg ttccactgag 480
 cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa 540
 tctgtgtgct gcaaacaaaa aaaccaccgc taccagcggg gggttggttg ccggatcaag 600
 agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg 660
 tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat 720
 acctcgtct gctaactctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta 780
 ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg 840

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc	900
gtgagcattg agaaagcgcc acgcttcccc aagggagaaa ggcggacagg tatccggtaa	960
gcggcagggc cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggatc	1020
tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt	1080
caggggggag gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct	1140
tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct gtggataacc	1200
gtattaccgc ctttgagtga gctgataccg ctccgcccag ccgaacgacc gagcgagcg	1260
agtcagttag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt	1320
ggccgattca ttaatgcagg ttgatcagat ctcgatcccc cgaaattaat acgactcact	1380
ataggagagc cacaacgggt tccctctaga aataattttg tttaacttta agaaggagat	1440
atacccatgg aaaagcctga actcaccgag acgtctgtcg agaagtttct gatcgaaaag	1500
ttcgacagcg tctccgacct gatgcagctc tcggaggggc aagaatctcg tgctttcagc	1560
ttcgatgtag gagggcggtg atatgtcctg cgggtaaata gctgcgccga tggtttctac	1620
aaagatcggt atgtttatcg gcactttgca tcggccgcgc tcccgattcc ggaagtgtt	1680
gacattgggg aattcagcga gagcctgacc tattgcatct cccgcctgac acaggggtgc	1740
acgttgcaag acctgcctga aaccgaactg cccgctgttc tgcagccggc cgcgagggt	1800
atggatgcga tcgtgcggc cgatcttagc cagacgagcg ggttcggccc attcggaccg	1860
caaggaatcg gtcaatacac tacatggcgt gatttcatat gcgcgattgc tgatcccat	1920
gtgtatcact ggcaactgt gatggacgac accgtcagtg cgtccgtcgc gcaggctctc	1980
gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt gcacgcggat	2040
ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcggatcat tgactggagc	2100
gaggcgatgt tcggggattc ccaatacgag gtcgccaaca tcttcttctg gagggcgtgg	2160
ttggcttgta tggagcagca gacgcgctac ttcgagcgga ggcatccgga gcttgagga	2220
tcgccgcggc tccgggagta tatgtccgc attggtcttg accaactcta tcagagcttg	2280
gttgacggca atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc aatcgtccga	2340
tccggagccg ggactgtcgg gcgtacacaa atcgcccga gaagcgggc cgtctggacc	2400
gatggctgtg tagaagtact cgccgatagt ggaaaccgac gcccagcac tcgtccgagg	2460
gcaaaggaat agtgaggtac agcttgatc gatccggtc ctaacaaagc ccgaaaggaa	2520
gctgagttgg ctgctgccac cgctgagcaa taactagcat aacccttggt ggcctctaaa	2580
cgggtcttga ggggtttttt gctgaaagga ggaactatat ccggatgatc gtcgagcct	2640
cacgtgttaa cagaagttcc tattccgaag ttcctattct ctgaaagta taggaacttc	2700
caccacacaa cacaatggcg gccaccgctt ccagaaccac ccgattctct tcttctctt	2760
cacacccac cttcccaaaa cgcattacta gatccaccct ccctctctct catcaaacc	2820
tcaccaaac caaccacgct ctcaaatca aatgttccat ctccaaacc cccacggcgg	2880

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ						
cgcccttcac	caaggaagcg	ccgaccacgg	agcccttcgt	gtcacgggtc	gcctccggcg	2940
aacctcgcaa	gggcgcggac	atccttgttg	aggcgctgga	gaggcagggc	gtgacgacgg	3000
tgttcgcgta	ccccggcggg	gcgtcgatgg	agatccacca	ggcgctcacg	cgctccgccg	3060
ccatccgcaa	cgtgctccc	cgccacgagc	agggcggcgt	cttcgccgcc	gaaggctacg	3120
cgcgttcctc	cggcctcccc	ggcgctctgca	ttgccacctc	cggccccggc	gccaccaacc	3180
tcgtgagcgg	cctcgccgac	gctttaatgg	acagcgctcc	agtcgtcgcc	atcaccggcc	3240
aggtcagccg	ccggatgata	ggcaccgacg	ccttccaaga	aaccccgatc	gtggagggtga	3300
gcagatccat	cacgaagcac	aactacctca	tcctcgacgt	cgacgacatc	ccccgcgtcg	3360
tcgccgaggc	tttcttcgtc	gccacctccg	gccgccccgg	tccggctctc	atcgacattc	3420
ccaaagacgt	tcagcagcaa	ctcgccgtgc	ctaattggga	cgagcccgtt	aacctccccg	3480
gttacctcgc	caggctgccc	aggccccccg	ccgaggccca	attggaacac	attgtcagac	3540
tcacatgga	ggcccaaaag	cccgttctct	acgtcggcgg	tggcagtttg	aattccagtg	3600
ctgaattgag	gcgctttgtt	gaactcactg	gtattcccgt	tgctagcact	ttaatgggtc	3660
ttggaacttt	tcctattggt	gatgaatatt	cccttcagat	gctgggtatg	catggtactg	3720
tttatgctaa	ctatgctgtt	gacaatagtg	atttgttgct	tgcccttggg	gtaaggtttg	3780
atgaccgtgt	tactgggaag	cttgaggcct	ttgctagtag	ggctaagatt	gttcacattg	3840
atattgattc	tgccgagatt	gggaagaaca	agcaggcgca	cgtgtcgggt	tgcgcggatt	3900
tgaagttggc	cttgaaggga	attaatatga	ttttggagga	gaaaggagtg	gagggtaagt	3960
ttgatcttgg	aggttggaga	gaagagatta	atgtgcagaa	acacaagttt	ccattgggtt	4020
acaagacatt	ccaggacgcg	atttctccgc	agcatgctat	cgaggttcct	gatgagttga	4080
ctaattggaga	tgctattgtt	agtactgggg	ttgggcagca	tcaaattgtg	gctgcgcagt	4140
tttacaagta	caagagaccg	aggcagtggt	tgacctcagg	gggtcttgga	gccatgggtt	4200
ttggattgcc	tgccgctatt	ggtgctgctg	ttgctaacc	tggggctgtt	gtggttgaca	4260
ttgatgggga	tggtagtttc	atcatgaatg	ttcaggagtt	ggccactata	agagtggaga	4320
atctcccagt	taagatattg	ttgttgaaca	atcagcattt	gggtatggtg	gttcagtggg	4380
aggatagggt	ctacaagtcc	aatagagctc	acacctatct	tggagatccg	tctagcgaga	4440
gcgagatatt	cccaaacatg	ctcaagtttg	ctgatgcttg	tgggataccg	gcagcgcgag	4500
tgacgaagaa	ggaagagcct	agagcggcaa	ttcagagaat	gttggacacc	cctggcccct	4560
accttcttga	tgctattgtg	ccccatcagg	agcatgtgtt	gccgatgatt	cccagtaatg	4620
gacccctcaa	ggatgtgata	actgagggtg	atggtagaac	gaggtactga	ttgcctagac	4680
caaattgtcc	ttgatgcttg	ttttgtacaa	tatatataag	ataatgctgt	cctagttgca	4740
ggatttggtc	tggtgtgagc	atcatagtct	gtagtagttt	tggtagcaag	acattttatt	4800
ttccttttat	ttaacttact	acatgcagta	gcatctatct	atctctgtag	tctgatatct	4860
cctgttgtct	gtattgtgcc	gttggatttt	ttgctgtagt	gagactgaaa	atgatgtgct	4920

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

agtaataata tttctgttag aaatctaagt agagaatctg ttgaagaagt caaaagctaa	4980
tggaatcagg ttacatatct aatgtttttc ttttttttagc gggttggtaga cgtgtagatt	5040
caacttctct tggagctcac ctaggcaatc agtaaaatgc atattccttt tttaacttgc	5100
catttattta ctttttagtg aaattgtgac caatttggtc atgtagaacg gatttggacc	5160
attgcgtcca caaaacgtct cttttgctcg atcttcacaa agcgataccg aaatccagag	5220
atagttttca aaagtcagaa atggcaaagt tataaatagt aaaacagaat agatgctgta	5280
atcgacttca ataacaagtg gcatcacgtt tctagttcta gacccatcag gggcagatct	5340
aggcgcgcgc catatacctc acacgtacgc gtagatggtg acgtacgtgc cctaccaaac	5400
agattcacgt cagatttgaa gttcctattc cgaagttcct attctacata gagtatagga	5460
acttccgata tctactgagt ggccggcggc gcgccgtcga cggatccgta cgatccatgc	5520
ccttcatttg ccgcttatta attaatgttg taacagtccg tactaatcag ttacttatcc	5580
ttccccatc ataattaatc ttggtagtct cgaatgccac aacctgact agtctcttgg	5640
atcataagaa aaagccaagg aacaaaagaa gacaaaacac aatgagagta tcctttgcat	5700
agcaatgtct aagttcataa aattcaaaca aaaacgcaat cacacacagt ggacatcact	5760
tatccactag ctgatcagga tcgccgcgtc aagaaaaaaa aactggaccc caaaagccat	5820
gcacaacaac acgtactcac aaaggtgtca atcgagcagc ccaaaacatt caccaactca	5880
acccatcatg agccctcaca ttgtttgttt ctaacccaac ctcaaactcg tattctcttc	5940
cgccacctca tttttgttta tttcaacacc cgtcaaactg catgccacc cgtggccaaa	6000
tgtccatgca tgtaacaag acctatgact ataaatagct gcaatctcgg ccaggtttt	6060
catcatcaag aaccagttca atatcctagt acaccgtatt aaagaattta agatatactg	6120
cgcccgcatg actatcgact cacaatacta caagtcgcga gacaaaaacg acacggcacc	6180
caaaatcgcg ggaatccgat atgccccgct atcgacacca ttactcaacc gatgtgagac	6240
cttctctctg gtctggcaca ttttcagcat tcccactttc ctcaaatctt tcatgctatg	6300
ctgcgcaatt ccaactgctt ggccatttgt gattgcgtat gtagtgtacg ctgttaaaga	6360
cgactccccg tccaacggag gagggtcaa gcgatactcg cctatttcaa gaaacttctt	6420
catctggaag ctctttggcc gctacttccc cataactctg cacaagacgg tggatctgga	6480
gcccacgcac acatactacc ctctggacgt ccaggagtat cacctgattg ctgagagata	6540
ctggccgcag aacaagtacc tccgagcaat catcaccacc atcgagtact ttctgcccgc	6600
cttcatgaaa cggctctctt ctatcaacga gcaggagcag cctgccgagc gagatcctct	6660
cctgtctccc gtttctccca gctctccggg ttctcaacct gacaagtgga ttaaccacga	6720
cagcagatat agccgtggag aatcatctgg ctccaacggc cagcctcgg gctccgaact	6780
taacggcaac ggcaacaacg gcaccactaa ccgacgacct ttgtcgtccg cctctgctgg	6840
ctccactgca tctgattcca cgcttcttaa cgggtccctc aactcctacg ccaaccagat	6900
cattggcgaa aacgaccac agctgtcgc caaaaactc aagccactg gcagaaaata	6960

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

catcttcggc taccaccccc acggcattat cggcatggga gcctttgggtg gaattgccac	7020
cgagggagct ggatgggtcca agctctttcc gggcatccct gtttctctta tgactctcac	7080
caacaacttc cgagtgcctc tctacagaga gtacctcatg agtctgggag tcgcttctgt	7140
ctccaagaag tcctgcaagg ccctcctcaa gcgaaaccag tctatctgca ttgtcgttgg	7200
tggagcacag gaaagtcttc tggccagacc cgggtgcatg gacctgggtc tactcaagcg	7260
aaagggtttt gttcgacttg gtatggaggt cggaatgtc gcccttggtc ccatcatggc	7320
ctttgggtgag aacgacctct atgaccaggt tagcaacgac aagtcgtcca agctgtaccg	7380
attccagcag tttgtcaaga acttccttgg attcaccctt cctttgatgc atgcccgagg	7440
cgtcttcaac tacgatgtcg gtcttgtccc ctacaggcga cccgtcaaca ttgtggttgg	7500
ttccccatt gacttgcctt atctcccaca cccaccgac gaagaagtgt ccgaatacca	7560
cgaccgatac atcgccgagc tgcagcgaat ctacaacgag cacaaggatg aatatttcat	7620
cgattggacc gaggagggca aaggagcccc agagtccga atgattgagt aagcggccgc	7680
aagtatgaac taaaatgcat gtaggtgtaa gagctcatgg agagcatgga atattgtatc	7740
cgaccatgta acagtataat aactgagctc catctcactt cttctatgaa taaacaaagg	7800
atgttatgat atattaacac tctatctatg caccttattg ttctatgata aatttcctct	7860
tattattata aatcatctga atcgtgacgg cttatggaat gcttcaaata gtacaaaaac	7920
aaatgtgtac tataagactt tctaaacaat tctaacctta gcattgtgaa cgagacataa	7980
gtgttaagaa gacataacaa ttataatgga agaagtttgt ctccatttat atattatata	8040
ttaccactt atgtattata ttaggatgtt aaggagacat aacaattata aagagagaag	8100
tttgatcca tttatatatt atatactacc ctttatata ttatacttat ccacttattt	8160
aatgtcttta taaggtttga tccatgatat ttctaataatt ttagttgata tgtatatgaa	8220
aaggactat ttgaactctc ttactctgta taaaggttgg atcatcctta aagtgggtct	8280
atttaatttt attgcttctt acagataaaa aaaaaattat gagttggttt gataaaatat	8340
tgaaggattt aaaataataa taaataacat ataatatatg tatataaatt tattataata	8400
taacatttat ctataaaaaa gttaaatttg tcataaatct atacaatcgt ttagccttgc	8460
tggaacgaat ctcaattatt taaacgagag taaacatatt tgactttttg gttatttaac	8520
aaattattat ttaacactat atgaaatttt tttttttatc agcaaagaat aaaattaaat	8580
taagaaggac aatggtgtcc caatccttat acaaccaact tccacaagaa agtcaagtca	8640
gagacaacaa aaaaacaagc aaaggaaatt ttttaatttg agttgtcttg tttgctgcat	8700
aatttatgca gtaaaacact acacataacc cttttagcag tagagcaatg gttgaccgtg	8760
tgcttagctt cttttatttt atttttttat cagcaaagaa taaataaaat aaaatgagac	8820
acttcaggga tgtttcaacg tactttctag acgtacgtct ttccacaata cataactatt	8880
aattaatctt aaataaataa aggataaaat attttttttt cttcataaag ttaaaatatg	8940
ttattttttg tttagatgta tattcgaata aatctaaata tatgataatg attttttata	9000

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

ttgattaaac atataatcaa tattaatat gatatttttt tatatagggt gtacacataa	9060
ttttataagg ataaaaata tgataaaaat aaatttttaa tatttttata tttacgagaa	9120
aaaaaaatat ttttagccata aataaatgac cagcatattt tacaacctta gtaattcata	9180
aattcctata tgtatatttg aaattaaaaa cagataatcg ttaagggaag gaatcctacg	9240
tcatctcttg ccatttgttt ttcattgcaa cagaaagga cgaaaaacca cctcaccatg	9300
aatcactctt cacaccattt ttactagcaa acaagtctca acaactgaag ccagctctct	9360
ttccgtttct ttttacaaca ctttctttga aatagtagta tttttttca catgatttat	9420
taacgtgcc aagatgctt attgaataga gtgcacattt gtaatgtact actaattaga	9480
acatgaaaaa gcattgttct aacacgataa tcctgtgaag gcgttaactc caaagatcca	9540
atttcactat ataaattgtg acgaaagcaa aatgaattca catagctgag agagaaagga	9600
aagggttaact aagaagcaat acttcagcgg ccgcttctag ctagctaggg tttgggtagt	9660
gagtgtataa aagttgcaa gtttttggtt aggttacgtt ttgaccttat tattatagtt	9720
caaagggaaa cattaattaa aggggattat gaagtgggt ctcttgattc ttggatgagg	9780
atcttactgg gtgaattgag ctgcttagct atggatccca cagttctacc catcaataag	9840
tgcttttggt gtagtcttgt ggcttcata tctggggagc ttcatttgcc tttatagtat	9900
taaccttctc caagaacaaa gagagccac acccttctct tcttttctct cataataatt	9960
taaatttggt atagactcta aactttaaat gtttttttg aagttttcc gtttttctct	10020
tttgccatga tcccgttctt gctgtggagt aacctgtcc gaggtatgtg catgattaga	10080
tccatactta atttgtgtgc atcacgaagg tgaggttgaa atgaacttg ctttttgac	10140
cttttaggaa agttcttttg ttgcagtaat caattttaat tagttttaat tgacactatt	10200
acttttattg tcatctttgt tagttttatt gttgaattga gtgcataatt cctaggaaat	10260
tctcttacct aacatttttt atacagatct atgctcttgg ctcttgccct tactcttggc	10320
cttgtgttgg ttatttgtct acatatttat tgactggtcg atgagacatg tcacaattct	10380
tgggcttatt tgttggtcta ataaaaggag tgcttattga aagatcaaga cggagattcg	10440
gttttatata aataaactaa agatgacata ttagtgtgtt gatgtctctt caggataatt	10500
tttgtttgaa ataatatggt aatgtcttgt ctaaaattgt gtacataatt ctactgatt	10560
ttttggattg ttggattttt ataaacaaat ctgcggccgc atgagccgta aaggttcaat	10620
acaacgagtg cttgttttct tagggacaag cattgtactt atgtatgatt ctgtgtaacc	10680
atgagtcttc cacgttgtac taatgtgaag ggcaaaaata aaacacagaa caagttcgtt	10740
tttctcaa atgtgaagg tagaaaatgg aacctgcct cctctcttgc atgtgattta	10800
aaatattagc agatggtacg tcgagtcgac ctgcaggtcg actcgacgta cgtcctcgaa	10860
gagaaggggt aataacacat tttttaacat ttttaacaca aatttttagtt atttaaaaat	10920
ttattaaaaa atttaaaaata agaagaggaa ctcttttaaat aaatctaact taaaaaattt	10980
atgattttta ataagttttc accaataaaa aatgtcataa aaatatgtta aaaagtatat	11040

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

tatcaatatt ctctttatga taaataaaaa gaaaaaaaaa ataaaagtta agtgaaaatg	11100
agattgaagt gacttttaggt gtgtataaat atatcaaccc cgccaacaat ttattttaatc	11160
caaatatatt gaagtatatt attccatagc ctttatttat ttatatattt attatataaa	11220
agctttattt gttctagggt gttcatgaaa ttttttttg gttttatctc cgttgtaaga	11280
aaatcatgtg ctttgtgtcg ccactcacta ttgcagcttt ttcattgcatt ggtcagattg	11340
acggttgatt gtatttttgt tttttatggt tttgtgttat gacttaagtc ttcattctctt	11400
tatctcttca tcagggttga tgggttaccta atatgggtcca tgggtacatg catggttaaa	11460
ttaggtggcc aactttgttg tgaacgatag aatttttttt atattaagta aactattttt	11520
atattatgaa ataataataa aaaaaatatt ttatcattat taacaaaatc atattagtta	11580
atttgttaac tctataataa aagaaatact gtaacattca cattacatgg taacatcttt	11640
ccaccctttc atttgttttt tgtttgatga ctttttttct tgtttaaatt tttttccctt	11700
cttttaaatt tggaatacat tatcatcata tataaactaa aatactaaaa acaggattac	11760
acaaatgata aataataaca caaatattta taaatctagc tgcaatatat ttaaactagc	11820
tatatcgata ttgtaaaata aaactagctg cattgatact gataaaaaaa tatcatgtgc	11880
tttctggact gatgatgcag tatacttttg acattgcctt tttttattt ttcagaaaag	11940
ctttcttagt tctgggttct tcattatttg tttcccatct ccattgtgaa ttgaatcatt	12000
tgcttcgtgt cacaaataca atttagntag gtacatgcat tggtcagatt cacggtttat	12060
tatgtcatga ctttaagttca tggtagtaca ttacctgcca cgcatgcatt atattggtta	12120
gatttgatag gcaaatttgg ttgtcaacaa tataaatata aataatgttt ttatattacg	12180
aaataacagt gatcaaaaca aacagtttta tctttattaa caagattttg tttttgttg	12240
atgacgtttt ttaatgttta cgctttcccc cttcttttga atttagaaca ctttatcatc	12300
ataaaatcaa atactaaaaa aattacatat ttcataaata ataacacaaa tttttttaa	12360
aaatctgaaa taataatgaa caatattaca tattatcacg aaaattcatt aataaaaaa	12420
ttatataaat aaaatgtaat agtagttata tgtaggaaaa aagtactgca cgcataatat	12480
atacaaaaag attaaaatga actattataa ataataacac taaattaatg gtgaatcata	12540
tcaaaataat gaaaaagtaa ataaaatttg taattaaactt ctatatgtat tacacacaca	12600
aataataaat aatagtaaaa aaaattatga taaatattta ccatctcata agatatttaa	12660
aataatgata aaaatataga ttatttttta tgcaactagc tagccaaaaa gagaacacgg	12720
gtatatataa aaagagtacc tttaaattct actgtacttc ctttattcct gacgttttta	12780
tatcaagtgg acatacgtga agattttaat tatcagtcta aatatttcat tagcacttaa	12840
tacttttctg ttttattcct atcctataag tagtcccgat tctcccaaca ttgcttattc	12900
acacaactaa ctaagaaagt cttccatagc cccccaagcg gccgctagtc gactaagtca	12960
tcaactattc caagctacgt atttgggagt ttgtggagta cagcaagatg atatacctag	13020
acggtgatat ccaagttttt gacaacattg accacttggt tgacttgcct gataactact	13080

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

tctatgcggt gatggactgt ttctgtgagc caacttgggg ccacactaaa caatatcaga	13140
tcggttactg ccagcagtgc ccccataagg ttcagtggcc cactcacttt gggcccaaac	13200
ctcctctcta tttcaatgct ggcattgttg tgtatgagcc caatttggct acttaccgtg	13260
acctccttca aacagtcca gtcaccagc ccacttcctt tgctgaacag gattttttga	13320
acatgtactt caaggacaaa tataggccaa ttcctaattgt ctacaatctt gtgctggcca	13380
tgctgtggcg tcaccctgag aacgttgagc ttgacaaagt taaagtgggt cactactgtg	13440
ctgctgggtc taagccttgg aggtacactg ggaagtgact cgaggtcatc aattactcca	13500
agctacgtat ttgggagttc gtggagtaca agaagacgat atacctagac ggtgacatcc	13560
aagtattttg aaacatagac cacttgtttg atctgcctga taattatttc tatgcggtga	13620
tggattgttt ctgcgagaag acttggagcc acaccctca gttccagatt gggactgcc	13680
aacagtgtcc tgataagggt caatggccct ctacttttg ttccaaacct cctctatatt	13740
tcaatgctgg catgtttgtt tatgagccta atctcgacac ctaccgtgat cttctcaaaa	13800
ctgtccaact caccaagccc acttcttttg ctgagcagga ctttctcaac atgtacttca	13860
aggacaagta caagccaata ccgaacatgt acaaccttgt gctggccatg ttgtggcgtc	13920
accctgaaaa tgttgaactt gataaagtcc aagtgttca ttactgtgct gctgggtcta	13980
agccttgag gttcactggg aagtaactgc aggtcatcaa ctactccaag ctccgtatat	14040
gggagtttgt ggagtacagc aagatgatat acttggacgg agacattgag gtatatgaga	14100
acatagacca cctatttgac ctacctgatg gtaactttta cgctgtgatg gattgtttct	14160
gcgagaagac atggagtcac acccctcagt acaaggtagg ttactgccag caatgcccgg	14220
agaagggtcg gtggcccacc gaattgggtc agcccccttc tctttacttc aacgctggca	14280
tgttcgtgtt cgaacccaac atcgccacct atcatgacct attgaaaacg gtgcaagtca	14340
ccactccac ctcgttcgct gaacaagatt tcttgaacat gtacttcaag gacatttaca	14400
agccaatccc tttaaattac aatcttgtcc tcgccatgct gtggcgccac ccggaaaacg	14460
ttaaattaga ccaagtcaag gttgttcaat attgcgagc ggggtccaag ccatggagat	14520
atacggggaa gtagcctagg cgtacgcagg taagtctctg cttctacctt tgatatatat	14580
ataataatta tcattaatta gtagtaatat aatatttcaa atattttttt caaaataaaa	14640
gaatgtagta tatagcaatt gcttttctgt agtttataag tgtgtatatt ttaatttata	14700
acttttctaa tatatgacca aaacatgggt atgtgcagg cctaggctac ttccccgtat	14760
atctccatgg cttggacccc gctgcgcaat agtgaacaac cttgacttgg tctaatttaa	14820
cgttttccgg gtggcgccac agcatggcga ggacaagatt gtaattttaa gggattggct	14880
tgtaaagtgc cttgaagtac atgttcaaga aatcttgttc agcgaacgag gtgggagtgg	14940
tgacttgac cgttttcaat aggtcatgat aggtggcgat gttgggttcg aacacgaaca	15000
tgccagcgtt gaagtaaaga gaagggggct gaccaattc ggtgggccac cgacacctct	15060
ccgggcattg ctggcagtaa cccaccttgt actgaggggt gtgactccat gtcttctcgc	15120

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

agaaacaatc catcacagcg taaaagttac catcaggtag gtcaaatagg tggcttatgt	15180
tctcatatac ctcaatgtct ccgtccaagt atatcatctt gctgtactcc acaaactccc	15240
atatacggag cttggagtag ttgatgacct gcagttactt cccagtgaac ctccaaggct	15300
tagaccagc agcacagtaa tgaaccactt gaactttatc aagttcaaca ttttcagggt	15360
gacgccacaa catggccagc acaaggttgt acatgttcgg tattggcttg tacttgcct	15420
tgaagtacat gttgagaaag tcctgctcag caaaagaagt gggcttggtg agttggacag	15480
tttggagaag atcacggtag gtgtcgagat taggctcata aacaacatg ccagcattga	15540
aatatagagg aggtttggaa ccaagttag agggccattg aaccttatca gggcactgtt	15600
ggcagtaccc aatctggaac tgaggggtgt ggctccaagt cttctcgag aaacaatcca	15660
tcaccgcata gaaataatta tcaggcagat caaacaagt gtctatgttt ccaaactctt	15720
ggatgtcacc gtctaggtat atcgtcttct tgtactccac gaactccaa atacgtagct	15780
tggagtaatt gatgacctcg agtcacttcc cagtgtacct ccaaggctta gaccagcag	15840
cacagtagtg aaccacttta actttgtcaa gctcaacgtt ctcagggtga cgccacagca	15900
tggccagcac aagattgtag acattaggaa ttggcctata tttgtccttg aagtacatgt	15960
tcaaaaaatc ctgttcagca aaggaagtgg gctgggtgac ttggactgtt tgaaggaggt	16020
cacggttaagt agccaaattg ggctcataca caaacatgcc agcattgaaa tagagaggag	16080
gtttgggccc aaagttagtg ggccactgaa ccttatgggg gactgctgg cagtaaccga	16140
tctgatattg tttagtgtgg cccaagtgg gctcacagaa acagtccatc accgcataga	16200
agtagttatc aggcaagtca aacaagtgg caatgttgtc aaaaacttgg atatcaccgt	16260
ctaggatatat catcttgctg tactccacaa actcccaaat acgtagcttg gaatagttga	16320
tgacttagtc gactagcggc cgcgacacaa gtgtgagagt actaaataaa tgctttgggt	16380
gtacgaaatc attacactaa ataaaataat caaagcttat atatgccttc cgctaaggcc	16440
gaatgcaaag aaattggttc tttctcgta tcttttgcca cttttactag tacgtattaa	16500
ttactactta atcatctttg ttacggctc attatatccg gtctaggcca aggccgcgaa	16560
gttaaaagca atgttgctac ttgtacgtac taacacatga tgtgatagtt tatgctagct	16620
agctataaca taagctgtct ctgagtgtgt tgtatattaa taaagatcat cactgggtga	16680
tggtgatcgt gtacgtaccc tacttagtag gcaatggaag cacttagagt gtgctttgtg	16740
catggccttg cctctgtttt gagacttttg taatgttttc gagtttaaat ctttgccttt	16800
gcgtacgtgg gcggatcccc tgca	16824

<210> 39
 <211> 16822
 <212> ДНК
 <213> штучна послідовність

<220>
 <223> Плазміда РНР62552

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<220>

<221> misc_feature

<222> (9373)..(9373)

<223> н являє собою а, с, г або т

<400> 39

gaagttccta ttccgaagtt cctattctct agaaagtata ggaacttcca ccacacaaca	60
caatggcggc caccgcttcc agaaccaccc gattctcttc ttctcttca caccacacct	120
tccccaaacg cactactaga tccaccctcc ctctcttca tcaaaccctc accaaaccca	180
accacgtct caaaatcaaa tggtccatct ccaaaccccc caccggcgcg cccttcacca	240
aggaagcgcc gaccacggag cccttcgtgt caccggttcgc ctccggcgaa cctcgcaagg	300
gcgcggacat ccttgtggag gcgctggaga ggcagggcgt gacgacggtg ttgcggtacc	360
ccggcggcgc gtcgatggag atccaccagg cgctcacgcg ctccgccgcc atccgcaacg	420
tgctcccgcg ccacgagcag ggcggcgtct tcgccgccga aggtacgcg cgttctctcg	480
gcctccccgg cgtctgcatt gccacctccg gccccggcgc caccaacctc gtgagcggcc	540
tcgccgacgc tttaatggac agcgtcccag tcgtcgccat caccggccag gtcagccgcc	600
ggatgatcgg caccgacgcc ttccaagaaa ccccgatcgt ggaggtgagc agatccatca	660
cgaagcaca ctacctatc ctgcagctcg acgacatccc ccgcgtcgtc gccgaggctt	720
tcttcgtcgc cacctccggc cgcgccggc cggtcctcat cgacattccc aaagacgttc	780
agcagcaact cgccgtgcct aattgggacg agcccgttaa cctccccggt tacctcgcca	840
ggctgcccag gcccccgcc gaggccaat tggaacacat tgcagactc atcatggagg	900
cccaaagcc cgttctctac gtcggcgggt gcagtttgaa ttccagtgtc gaattgaggc	960
gctttgttga actcactggt attcccgttg ctagcacttt aatgggtctt ggaacttttc	1020
ctattggtga tgaatattcc cttcagatgc tgggtatgca tgggtactgt tatgctaact	1080
atgctgttga caatagtgat ttgttgcttg ctttgggggt aagggttgat gaccgtgtta	1140
ctgggaagct tgaggctttt gctagtaggg ctaagattgt tcacattgat attgattctg	1200
ccgagattgg gaagaacaag caggcgcacg tgtcggtttg cgcgatttg aagttggcct	1260
tgaagggaa taatatgatt ttggaggaga aaggagtggg gggtaagttt gatcttgag	1320
gttgagaga agagattaat gtgcagaaac acaagtttcc attgggttac aagacattcc	1380
aggacgcgat ttctccgcag catgctatcg aggttcttga tgagttgact aatggagatg	1440
ctattgttag tactggggtt gggcagcatc aaatgtgggc tgcgcagttt tacaagtaca	1500
agagaccgag gcagtgggtg acctcagggg gtcttgagc catgggtttt ggattgcctg	1560
cggctattgg tgctgctgtt gctaaccctg gggctgttgt ggttgacatt gatggggatg	1620
gtagtttcat catgaatgtt caggagttag ccactataag agtggagaat ctcccagtta	1680
agatattgtt gttgaacaat cagcatttgg gtatggtggt tcagtgggag gataggttct	1740
acaagtcaa tagagctcac acctatcttg gagatccgtc tagcgagagc gagatattcc	1800
caaacatgct caagtttgct gatgcttggt ggataccggc agcgcgagtg acgaagaagg	1860

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

aagagcttag agcggcaatt cagagaatgt tggacacccc tggccsstac cttcttgatg	1920
tcattgtgcc ccatcaggag catgtgttgc cgatgattcc cagtaatgga tccttcaagg	1980
atgtgataac tgaggggtgat ggtagaacga ggtactgatt gcctagacca aatgttcctt	2040
gatgcttggt ttgtacaata tatataagat aatgctgtcc tagttgcagg atttggcctg	2100
tggtgagcat catagtctgt agtagttttg gttagcaagac attttatttt ccttttattt	2160
aacttactac atgcagtagc atctatctat ctctgtagtc tgatatctcc tgttgtctgt	2220
attgtgccgt tggatttttt gctgtagtga gactgaaaat gatgtgctag taataatatt	2280
tctgttagaa atctaagtag agaatctgtt gaagaagtca aaagctaatt gaatcagggt	2340
acatattcaa tgtttttctt tttttagcgg ttggtagacg tgtagattca acttctcttg	2400
gagctcacct aggcaatcag taaaatgcat attccttttt taacttgcca tttatttact	2460
tttagtgaa attgtgacca atttgttcat gttagaacgga tttggaccat tgcgtccaca	2520
aaacgtctct tttgctcgat cttcacaag cgataccgaa atccagagat agttttcaaa	2580
agtcagaaat ggcaaagtta taaatagtaa aacagaatag atgctgtaat cgacttcaat	2640
aacaagtggc atcacgtttc tagttctaga cccatcaggg gcagatctag gcgcgccgat	2700
ggtagacgtac gtgccctacc aaacagatgt acgtcatgtt gctatccctc atacgtacgt	2760
gtagaagttc ctattccgaa gttcctattc tacatagagt ataggaactt ccgatatcac	2820
tgcagtggcc ggcggcgcgc cgtcgacgga tccgtacgat ccatgccctt catttgccgc	2880
ttattaatta atttggtaac agtccgtact aatcagttac ttatccttcc cccatcataa	2940
ttaatcttgg tagtctcgaa tgccacaaca ctgactagtc tcttggatca taagaaaaag	3000
ccaaggaaca aaagaagaca aaacacaatg agagtatcct ttgcatagca atgtctaagt	3060
tcataaaatt caaacaacaa cgcaatcaca cacagtggac atcacttatc cactagctga	3120
tcaggatcgc cgcgtcaaga aaaaaaact ggaccccaaa agccatgcac aacaacacgt	3180
actcacaag gtgtcaatcg agcagcccaa aacattcacc aactcaacc atcatgagcc	3240
ctcacatttg ttgtttctaa cccaacctca aactcgtatt ctcttccgcc acctcatttt	3300
tgtttatttc aacacccgtc aaactgcatg ccaccccggtg gccaaatgtc catgcatgtt	3360
aacaagacct atgactataa atagctgcaa tctcggccca ggttttcatc atcaagaacc	3420
agttcaatat cctagtagac cgtattaaag aatttaagat atactgcggc cgcatgacta	3480
tcgactcaca atactacaag tcgcgagaca aaaacgacac ggcacccaaa atcgcgggaa	3540
tccgatatgc cccgctatcg acaccattac tcaaccgatg tgagaccttc tctctggtct	3600
ggcacatttt cagcattccc actttctca caattttcat gctatgctgc gcaattccac	3660
tgctctggcc atttgtgatt gcgtatgtag tgtacgtgt taaagacgac tccccgtcca	3720
acggaggagt ggtcaagcga tactcgccca tttcaagaaa cttcttcac tggaaactct	3780
ttggccgcta cttccccata actctgcaca agacgggtgga tctggagccc acgcacacat	3840
actaccctct ggacgtccag gagtatcacc tgattgtgta gagatactgg ccgcagaaca	3900

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

agtacctccg agcaatcatc accaccatcg agtactttct gcccgcttc atgaaacggg	3960
ctctttctat caacgagcag gagcagcctg ccgagcgaga tcctctcctg tctcccgttt	4020
ctcccagctc tccgggttct caacctgaca agtggattaa ccacgacagc agatatagcc	4080
gtggagaatc atctggctcc aacggccacg cctcgggctc cgaacttaac ggcaacggca	4140
acaacggcac cactaaccga cgacctttgt cgtccgcctc tgctggctcc actgcatctg	4200
attccacgct tcttaacggg tccctcaact cctacgcaa ccagatcatt ggcgaaaacg	4260
accacagct gtcgcccaca aaactcaagc ccactggcag aaaatacatc ttcggctacc	4320
acccccacgg cattatcggc atgggagcct ttggtggaat tgccaccgag ggagctggat	4380
ggtccaagct ctttccgggc atccctgttt ctcttatgac tctcaccaac aacttccgag	4440
tgctctcta cagagagtac ctcatgagtc tgggagtcgc ttctgtctcc aagaagtcct	4500
gcaaggccct cctcaagcga aaccagtcta tctgcattgt cgttgggtgga gcacaggaaa	4560
gtcttctggc cagaccgggt gtcattggacc tgggtgctact caagcgaag ggttttgttc	4620
gacttggat ggaggctgga aatgtcgccc ttgttcccat catggccttt ggtgagaacg	4680
acctctatga ccaggttagc aacgacaagt cgtccaagct gtaccgattc cagcagtttg	4740
tcaagaactt ccttggattc acccttcctt tgatgcatgc ccgaggcgtc ttcaactacg	4800
atgtcgggtct tgtcccctac aggcgacccg tcaacattgt ggttgggttcc cccattgact	4860
tgctttatct cccacacccc accgacgaag aagtgtccga ataccacgac cgatacatcg	4920
ccgagctgca gcgaatctac aacgagcaca aggatgaata ttcatcgat tggaccgagg	4980
agggcaaagg agccccagag ttccgaatga ttgagtaagc ggccgcaagt atgaactaaa	5040
atgcatgtag gtgtaagagc tcatggagag catggaatat tgtatccgac catgtaacag	5100
tataataact gagctccatc tcacttcttc tatgaataaa caaaggatgt tatgatatat	5160
taacactcta tctatgcacc ttattgttct atgataaatt tcctcttatt attataaatc	5220
atctgaatcg tgacggctta tggaatgctt caaatagtac aaaaacaaat gtgtactata	5280
agactttcta aacaattcta accttagcat tgtgaacgag acataagtgt taagaagaca	5340
taacaattat aatggaagaa gtttgtctcc atttatatat tatatattac ccacttatgt	5400
attatattag gatgttaagg agacataaca attataaaga gagaagtttg tatccattta	5460
tatattatat actaccatt tatatattat acttatccac ttatttaatg tctttataag	5520
gtttgatcca tgatatttct aatatttttag ttgatatgta tatgaaaagg tactatttga	5580
actctcttac tctgtataaa ggttggatca tccttaaggt gggctctattt aattttattg	5640
cttcttacag ataaaaaaaa aattatgagt tggtttgata aaatattgaa ggatttaaaa	5700
taataataaa taacatataa tatatgtata taaatttatt ataataaac atttatctat	5760
aaaaaagtaa atattgtcat aaatctatac aatcgtttag ccttgctgga acgaatctca	5820
attatttaaa cgagagtaaa catatttgac tttttggtta ttttaacaaat tattatttaa	5880
cactatatga aatttttttt ttatcagca aagaataaaa ttaaattaag aaggacaatg	5940

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

gtgtcccaat ccttatacaa ccaacttcca caagaaagtc aagtcagaga саасаааааа	6000
acaagcaaag gaaatTTTT aatttgagtt gtcttgtttg ctgcataatt tatgcagtaa	6060
aacactacac ataacccttt tagcagtaga gcaatggttg accgtgtgct tagcttcttt	6120
tattttattt tttatcagc aaagaataaa taaaataaaa tgagacactt cagggatggt	6180
tcaacgtact ttctagacgt acgtctttcc acaatacata actattaatt aatcttaaat	6240
aaataaagga taaaatattt tttttcttc ataaagttaa aatatgttat tttttgttta	6300
gatgtatatt cgaataaatc taaatatatg ataatgattt tttatattga ttaaacadat	6360
aatcaatatt aaatatgata tttttttata taggttgtac acataatttt ataaggataa	6420
aaaatatgat aaaaataaat tttaaatatt tttatattta cgagaaaaaa aaatatTTTA	6480
gccataaata aatgaccagc atattttaca accttagtaa ttcataaatt cctatatgta	6540
tatttgaaat taaaaacaga taatcgtaa gggaaaggaat cctacgtcat ctcttgccat	6600
ttgtttttca tgcaaacaga aaggagcga aaaccacctc accatgaatc actcttcaca	6660
ccatttttac tagcaacaa gtctcaacaa ctgaagccag ctctctttcc gttctttttt	6720
acaacacttt ctttgaaata gtagtatttt tttcacatg atttattaac gtgcaaaaag	6780
atgcttattg aatagagtgc acatttgtaa tgtactacta attagaacat gaaaagcat	6840
tgttctaaca cgataatcct gtgaaggcgt taactccaaa gatccaattt cactatataa	6900
attgtgacga aagcaaatg aattcacata gctgagagag aaaggaaagg ttaactaaga	6960
agcaatactt cagcgccgc ttctagctag ctagggtttg ggtagttagt gtaataaagt	7020
tgcaaagttt ttggttaggt tacgttttga cttattatt atagttcaaa gggaaacatt	7080
aattaaagg gattatgaag tgggctctct tgattcttg atgaggatct tactgggtga	7140
attgagctgc ttagctatgg atcccacagt tctaccatc aataagtgc tttgtggtag	7200
tcttggtgct tccatatctg gggagcttca tttgccttta tagtattaac cttctccaag	7260
aacaaagaga gccacaccc ttctctctt ttctctcata ataatttaa tttgttatag	7320
actctaaact ttaaatgtt tttttgaagt tttccgttt ttctcttttg ccatgatccc	7380
gttcttgctg tggagtaacc ttgtccgagg tatgtgcatg attagatcca tacttaattt	7440
gtgtgcatca cgaaggtag gttgaaatga actttgcttt ttgacctt taggaaagtt	7500
cttttggtgc agtaatcaat tttaattagt tttaattgac actattactt ttattgtcat	7560
ctttgttagt ttattgttg aattgagtgc atatttccta ggaaattctc ttacctaa	7620
ttttttatac agatctatgc tcttggtct tgcccttact cttggccttg tgttggttat	7680
ttgtctacat atttattgac tggtcgatga gacatgtcac aattcttggg cttatttggt	7740
ggtctaataa aaggagtgc tattgaaaga tcaagacgga gattcgggtt tatataaata	7800
aactaaagat gacatattag tgtgttgatg tctcttcagg ataatttttg tttgaaataa	7860
tatggtaatg tcttgcttaa atttgtgtac ataattctta ctgatttttt ggattgttgg	7920
atTTTTataa acaaatctgc ggccgcatga gccgtaaagg ttcaatacaa cgagtgcctg	7980

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

ttttcttagg gacaagcatt gtacttatgt atgattctgt gtaaccatga gtcctccacg	8040
ttgtactaat gtgaaggcca aaaataaaac acagaacaag ttcgtttttc tcaaataatg	8100
tgaaggtaga aaatggaacc atgcctcctc tcttgcattgt gatttaaaat attagcagat	8160
ggtacgtcga gtcgacctgc aggtcgactc gacgtacgtc ctcgaagaga agggttaata	8220
acacattttt taacattttt aacacaaaatt ttagttattt aaaaatttat taaaaaattt	8280
aaaataagaa gaggaactct ttaaataaat ctaacttaca aaatttatga tttttaataa	8340
gttttcacca ataaaaaatg tcataaaaat atgttaaaaa gtatatattc aatattctct	8400
ttatgataaa taaaaagaaa aaaaaataa aagttaagtg aaaaatgagat tgaagtgact	8460
ttaggtgtgt ataaatatat caaccccgcc aacaatttat ttaatccaaa tatattgaag	8520
tatatatttc catagccttt atttatttat atatttatta tataaaagct ttatttggtc	8580
taggttggtc atgaaatatt tttttggttt tatctccgtt gtaagaaaat catgtgcttt	8640
gtgtcgccac tcactattgc agctttttca tgcattgggc agattgacgg ttgattgtat	8700
ttttgttttt tatggttttg tgttatgact taagtcttca tctctttatc tcttcattcag	8760
gtttgatggt tacctaatat ggtccatggg tacatgcattg gttaaattag gtggccaact	8820
ttgttgatgaa cgatagaatt ttttttatat taagtaaact atttttatat tatgaaataa	8880
taataaaaaa aatattttat cattattaac aaaatcatat tagttaattt gtttaactcta	8940
taataaaaga aatactgtaa cattcacatt acatggtaac atctttccac cttttcattt	9000
gttttttggt tgatgacttt ttttcttgtt taaatttatt tcccttcttt taaatttgga	9060
atacattatc atcatatata aactaaaata ctaaaaacag gattacacaa atgataaata	9120
ataacacaaa tattttataaa tctagctgca atatatttta actagctata tcgatattgt	9180
aaaataaaac tagctgcatt gatactgata aaaaaatatc atgtgctttc tggactgatg	9240
atgcagtata cttttgacat tgcctttatt ttatttttca gaaaagcttt cttagtcttg	9300
ggttcttcat tatttgtttc ccatctccat tgtgaattga atcatttgct tcgtgtcaca	9360
aatacaattt agntaggtac atgcattggg cagattcacg gtttattatg tcatgactta	9420
agttcatggt agtacattac ctgccacgca tgcattatat tggttagatt tgataggcaa	9480
atttggttgt caacaatata aatataaata atgtttttat attacgaaat aacagtgatc	9540
aaaacaaaca gttttatctt tattaacaag attttgtttt tgtttgatga cgttttttaa	9600
tgtttacgct ttcccccttc ttttgaattt agaacacttt atcatcataa aatcaaatac	9660
taaaaaaatt acatatttca taaataataa cacaaatatt tttaaaaaat ctgaaataat	9720
aatgaacaat attacatatt atcacgaaa ttcattaata aaaatattat ataaataaaa	9780
tgtaatagta gttatatgta ggaaaaaagt actgcacgca taatatatac aaaaagatta	9840
aatgaacta ttataaataa taacactaaa ttaatgggtga atcatatcaa aataatgaaa	9900
aagtaataa aatttgtaat taacttctat atgtattaca cacacaaata ataaataata	9960
gtaaaaaaa ttatgataaa tatttaccat ctcataagat atttaaaata atgataaaaa	10020

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

tatagattat	tttttatgca	actagctagc	caaaaagaga	acacgggtat	atataaaaag	10080
agtaccttta	aattctactg	tacttccttt	attcctgacg	tttttatatc	aagtggacat	10140
acgtgaagat	tttaattatc	agtctaaata	tttcattagc	acttaatact	tttctgtttt	10200
attcctatcc	tataagtagt	cccgattctc	ccaacattgc	ttattcacac	aactaactaa	10260
gaaagtcttc	catagcccc	caagcggccg	ctagtcgact	aagtcatcaa	ctattccaag	10320
ctacgtatth	gggagtttgt	ggagtacagc	aagatgatat	acctagacgg	tgatatccaa	10380
gtttttgaca	acattgacca	cttgtttgac	ttgcctgata	actacttcta	tgcggtgatg	10440
gactgtttct	gtgagccaac	ttggggccac	actaaacaat	atcagatcgg	ttactgccag	10500
cagtgccttc	ataaggttca	gtggccctac	cactttgggc	ccaaacctcc	tctctatttc	10560
aatgctggca	tgtttgtgta	tgagcccaat	ttggctactt	accgtgacct	ccttcaaaca	10620
gtccaagtca	cccagcccac	ttcctttgct	gaacaggatt	ttttgaacat	gtacttcaag	10680
gacaaatata	ggccaattcc	taatgtctac	aatcttgggc	tgcccatgct	gtggcgctac	10740
cctgagaacg	ttgagcttga	caaagttaaa	gtgggttact	actgtgctgc	tggtgtctaag	10800
ccttgagggt	acactgggaa	gtgactcgag	gtcatcaatt	actccaagct	acgtatthgg	10860
gagttcgtgg	agtacaagaa	gacgatatac	ctagacgggtg	acatccaagt	atttggaaac	10920
atagaccact	tgthtgatct	gcctgataat	tatttctatg	cggtgatgga	ttgtthctgc	10980
gagaagactt	ggagccacac	ccctcagttc	cagattgggt	actgccaaca	gtgccctgat	11040
aaggttcaat	ggccctctca	ctttgggtcc	aaacctctc	tatatttcaa	tgctggcatg	11100
tttgthtatg	agcctaattc	cgacacctac	cgtgatcttc	tccaaactgt	ccaactcacc	11160
aagcccactt	ctthtgctga	gcaggacttt	ctcaacatgt	acttcaagga	caagtacaag	11220
ccaataccga	acatgtacaa	ccttggtgctg	gccatgttgt	ggcgctaccc	tgaaaatgth	11280
gaacttgata	aagttcaagt	ggttcattac	tgtgctgctg	ggtctaagcc	ttggaggthc	11340
actgggaagt	aactgcaggt	catcaactac	tccaagctcc	gtatatggga	gtthtgaggag	11400
tacagcaaga	tgatatactt	ggacggagac	attgaggtat	atgagaacat	agaccaccta	11460
tttgacctac	ctgatggtaa	cttttacgct	gtgatggatt	gtthctgcga	gaagacatgg	11520
agtcacaccc	ctcagtacaa	ggtgggttac	tgccagcaat	gcccggagaa	ggtgcggtgg	11580
cccaccgaat	tggttcagcc	cccttctctt	tacttcaacg	ctggcatgth	cgtgttcgaa	11640
cccaacatcg	ccacctatca	tgacctattg	aaaacgggtc	aagtcaccac	tcccacctcg	11700
ttcgctgaac	aagatttctt	gaacatgtac	ttcaaggaca	tttacaagcc	aatcccttta	11760
aattacaatc	ttgtcctcgc	catgctgtgg	cgccaccggg	aaaacgttaa	attagaccaa	11820
gtcaaggthg	ttcactattg	cgcagcgggg	tccaagccat	ggagatatac	ggggaagtag	11880
cctaggcgta	cgcaggtaag	tttctgcttc	tacctthgat	atatatataa	taattatcat	11940
taattagtag	taatataata	tttcaaatac	ttttttcaaa	ataaaagaat	gtagtatata	12000
gcaattgctt	ttctgtagth	tataagtgth	tatattttta	tttataactt	ttctaataata	12060

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ		
tgacaaaaac atggtgatgt gcaggtccta ggctacttcc cgtatatct ccatggcttg	12120	
gaccccgctg cgcaatagtg aacaaccttg acttggctcta atttaacggt ttccgggtgg	12180	
cgccacagca tggcgaggac aagattgtaa tttaaagga ttggcttgta aatgtccttg	12240	
aagtacatgt tcaagaaatc ttgttcagcg aacgaggtgg gagtggtgac ttgcaccgtt	12300	
ttcaataggt catgataggt ggcgatgttg ggctcgaaca cgaacatgcc agcgttgaag	12360	
taaagagaag ggggctgacc caattcggtg ggccaccgca cttctccgg gcattgctgg	12420	
cagtaaccca cttgtactg aggggtgtga ctccatgtct tctcgagaa acaatccatc	12480	
acagcgtaaa agttaccatc aggtaggtca aataggtggg ctatgttctc atatacctca	12540	
atgtctccgt ccaagtatat catcttgctg tactccacaa actcccatat acggagcttg	12600	
gagtagttga tgacctgacg ttacttccca gtgaacctcc aaggcttaga cccagcagca	12660	
cagtaatgaa ccacttgaac tttatcaagt tcaacatttt cagggtgacg ccacaacatg	12720	
gccagcaciaa ggttgtagat gttcggtatt ggcttgtagt tgccttgaa gtacatgttg	12780	
agaaagtcct gctcagcaaa agaagtgggc ttggtgagtt ggacagtttg gagaagatca	12840	
cggtaggtgt cgagattagg ctcataaaca aacatgccag cattgaaata tagaggaggt	12900	
ttggaaccaa agtgagaggg ccattgaacc ttatcagggc actgttggca gtacccaatc	12960	
tggaactgag ggggtgtggc ccaagtcttc tcgcagaaac aatccatcac cgcatagaaa	13020	
taattatcag gcagatcaaa caagtggctt atgtttccaa atacttggat gtcaccgtct	13080	
aggtatatcg tcttcttgta ctccacgaac tcccaaatac gtagcttggg gtaattgatg	13140	
acctcgagtc acttcccagt gtacctcaa ggcttagacc cagcagcaca gtagtgaacc	13200	
actttaactt tgtcaagctc aacgttctca ggggtgacgc acagcatggc cagcacaaga	13260	
ttgtagacat taggaattgg cctatatattg tccttgaagt acatgttcaa aaaatcctgt	13320	
tcagcaaagg aagtgggctg ggtgacttg actgtttgaa ggaggtcacg gtaagtagcc	13380	
aaattgggct catacaciaa catgccagca ttgaaataga gaggaggttt gggcccaaag	13440	
tgagtgggccc actgaacctt atgggggcac tgctggcagt aaccgatctg atattgttta	13500	
gtgtggcccc aagtgggctc acagaaacag tccatcacg catagaagta gttatcaggc	13560	
aagtcaaaca agtgggtcaat gttgtcaaaa acttggatat caccgtctag gtatatcatc	13620	
ttgctgtact ccacaaactc ccaaatagct agcttggagt agttgatgac ttagtcgact	13680	
agcggccgag acacaagtgt gagagtacta aataaatgct ttggttgtag gaaatcatta	13740	
cactaaataa aataatcaaa gcttatatat gccttccgct aaggccgaat gcaaagaaat	13800	
tggttctttc tcgttatctt ttgccacttt tactagtacg tattaattac tacttaatca	13860	
tctttgttta cggctcatta tatccggtct aggccaaagg cgcgaaagta aaagcaatgt	13920	
tgtcacttgt acgtactaac acatgatgtg atagtttatg ctagctagct ataacataag	13980	
ctgtctctga gtgtgttgta tattaataaa gatcatcact ggtgaatggg gatcgtgtac	14040	
gtaccctact tagtaggcaa tggaagcact tagagtgtgc tttgtgcatg gccttgccctc	14100	

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

tgttttgaga cttttgtaat gttttcgagt ttaaatcttt gcctttgctg acgtgggagg	14160
atcccctgca ggagatccaa gcttggcgcg cgggcctctg cctgcgttct gctgtggaag	14220
ttcctattcc gaagttccta ttctccagaa agtataggaa cttcacatgc tgcctcgtgc	14280
aagtcacgat ctcgagttct atagtgtcac ctaaatcgta tgtgtatgat acataagggt	14340
atgtattaat tgtagccgcg ttctaacgac aatatgtcca tatgggtgcac tctcagtaca	14400
atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg	14460
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg	14520
agctgcatgt gtcagagggt ttaccgtca tcaccgaaac gcgcgagacg aaagggcctc	14580
gtgatacgcc tttttttata ggttaatgtc atgacaaaaa tcccttaacg tgagttttcg	14640
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt	14700
ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg gggttggttg	14760
ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	14820
ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca	14880
ccgcctacat acctcgctct gctaatcctg ttaccagtggt ctgctgccag tggcgataag	14940
tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	15000
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	15060
tacctacagc gtgagcattg agaaagcgcc acgcttcccg aaggggagaaa ggcggacagg	15120
tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac	15180
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg	15240
tgatgctcgt cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg	15300
ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct	15360
gtggataacc gtattaccgc ctttgagtga gctgataccg ctcgccgcag ccgaacgacc	15420
gagcgcagcg agtcagttag cgaggaagcg gaagagcgcc caatacgaa accgcctctc	15480
cccgcgcgtt ggccgattca ttaatgcagg ttgatcagat ctcgatcccg cgaattaat	15540
acgactcact atagggagac cacaacggtt tccctctaga aataattttg tttaacttta	15600
agaaggagat ataccatgg aaaagcctga actcaccgcg acgtctgtcg agaagtttct	15660
gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggagggcg aagaatctcg	15720
tgctttcagc ttcgatgtag gagggcggtg atatgtcctg cgggtaaata gctgcgccga	15780
tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc tcccgattcc	15840
ggaagtgtct gacattgggg aattcagcga gagcctgacc tattgcatct cccgccgtgc	15900
acaggggtgt acgttgcaag acctgcctga aaccgaactg cccgctgttc tgcagccggt	15960
cgcgagggtc atggatgcga tcgctgcggc cgatcttagc cagacgagcg ggttcggccc	16020
attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat gcgcgattgc	16080
tgatcccat gtgtatcact ggcaactgt gatggacgac accgtcagtg cgtccgtcgc	16140

PCT_US2014_022500_ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ	
gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt	16200
gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcgggtcat	16260
tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaaca tcttcttctg	16320
gaggccgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga ggcattccgga	16380
gcttgacagga tcgccgcggc tccgggagga tatgctccgc attgggtcttg accaactcta	16440
tcagagcttg gttgacggca atttcgatga tgacgcttg ggcgaggggc gatgcgacgc	16500
aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc	16560
cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac gcccagcac	16620
tcgtccgagg gcaaaggaat agtgaggtag agcttgatc gatccggctg ctaacaaagc	16680
ccgaaaggaa gctgagttgg ctgctgccac cgctgagcaa taactagcat aacccttgg	16740
ggcctctaaa cgggtcttga ggggtttttt gctgaaagga ggaactatat ccggatgatc	16800
gtcgaggcct cacgtgttaa ca	16822

ФОРМУЛА ВИНАХОДУ

- 5 1. Спосіб ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимий сайт розпізнавання, причому зазначений спосіб включає
 - а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатним до введення двониткового розриву в розпізнавання
- 10 зазначену геному ДНК, де двонитковий розрив приводить у результаті до утворення нуклеотидного липкого кінця;
 - б) лігування першого адаптера із зазначеним нуклеотидним липким кінцем;
 - с) розрізання лігваної ДНК, отриманої на стадії (б), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування
- 15 фрагментів геномної ДНК, що оточують двонитковий розрив;
 - д) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (с), з еталонною геномною послідовністю ДНК; та
 - е) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного
- 20 сконструйованого засобу для індукції двониткового розриву.
 2. Спосіб за п. 1, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву вибирають із групи, що складається з мегануклеази, нуклеази "цинкові пальці", TAL-ефекторної нуклеази, транспозази і сайт-специфічної рекомбінази.
 3. Спосіб за п. 1, де нуклеотидний липкий кінець являє собою 3'-нуклеотидний липкий кінець.
- 25 4. Спосіб за п. 1, де нуклеотидний липкий кінець являє собою 5'-нуклеотидний липкий кінець.
 5. Спосіб за п. 1, де перший адаптер, лігований з нуклеотидним липким кінцем, являє собою нефосфорильований по 5' адаптер.
 6. Спосіб за п. 1, де геномну ДНК вибирають із групи, що складається із прокаріотичної ДНК, еукаріотичної ДНК і синтетичної ДНК.
- 30 7. Спосіб за п. 6, де еукаріотичну ДНК виділяють із рослини, дріжджів або тварини.
 8. Спосіб за п. 7, де рослину вибирають із групи, що складається із сої, соняшнику, бавовнику, люцерни, канолі, тютюну, картоплі, *Arabidopsis*, сафлору, маїсу, рису, сорго, ячменя, пшениці, проса, вівса, цукрової тростини, газонної трави та проса прутноподібного.
 9. Спосіб за п. 1, де засіб для індукції двониткового розриву отримують із I-CreI.
- 35 10. Спосіб ідентифікації варіантного сайту розпізнавання з покращеною активністю розщеплення для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимий сайт розпізнавання, причому зазначений спосіб включає
 - а) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для

індукції двониткового розриву, здатним до введення двониткового розриву в розпізнавання зазначену геному ДНК, де двонитковий розрив приводить у результаті до утворення нуклеотидного липкого кінця;

b) лігування першого адаптера із зазначеним нуклеотидним липким кінцем;

5 c) розрізання лігваної ДНК, отриманої на стадії (b), і лігування другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, що оточують двонитковий розрив;

d) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (c), з еталонною геномною послідовністю ДНК; та

10 e) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву;

f) аналіз активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву у варіантних сайтах розпізнавання з (d); та

15 g) ідентифікацію варіантного сайту розпізнавання, який приводить в результаті до підвищеної активності сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву в порівнянні з активністю в припустимому сайті розпізнавання.

11. Спосіб за п. 10, де про підвищену активність сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву свідчить

20 a) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на геномній ДНК;

b) вищий відсоток (%) розщеплення варіантного сайту розпізнавання в порівнянні з відсотком (%) розщеплення припустимого сайту розпізнавання, де сайти розпізнавання розташовані на плазмідній ДНК;

25 c) вища оцінка в аналізі на дріжджах для варіантного сайту розпізнавання в порівнянні з припустимим сайтом розпізнавання; або

d) будь-яка комбінація (a), (b) і (c).

30 12. Спосіб націлювання вставки полінуклеотиду, який представляє інтерес, у специфічний хромосомний сайт у геномі рослини, причому зазначений спосіб включає

a) трансформацію рослинної клітини або рослини фрагментом ДНК, що містить полінуклеотид, який представляє інтерес, де зазначений геном зазначених рослинної клітини або рослини містить щонайменше один варіантний сайт розпізнавання, вибраний із групи, що складається з SEQ ID NO: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 і 35;

35 b) забезпечення мегануклеази, здатної утворювати двонитковий розрив у варіабельному сайті розпізнавання з (a); і

c) відбір зазначених рослинної клітини або рослини, що містить зазначений полінуклеотид, який представляє інтерес, інтегрований у зазначений варіантний сайт розпізнавання.

40 13. Спосіб за п. 12, де забезпечення зазначеної мегануклеази включає інтеграцію в геном зазначених рослинних клітин нуклеотидної послідовності, що кодує мегануклеазу SEQ ID NO: 3.

14. Рослина або рослинна клітина, отримана за допомогою способу за п. 12.

15. Спосіб ідентифікації варіантного сайту розпізнавання для сконструйованого засобу, що рідко розщеплює, для індукції двониткового розриву, здатного до введення рідкого двониткового розриву в припустимий сайт розпізнавання, причому зазначений спосіб включає

45 a) приведення в контакт геномної ДНК із сконструйованим засобом, що рідко розщеплює, для індукції двониткового розриву, здатного до введення двониткового розриву в зазначену геномну ДНК, де двонитковий розрив приводить у результаті до утворення тупого кінця;

b) створення нуклеотидного липкого кінця з тупого кінця з (a);

c) лігування першого адаптера з нуклеотидним липким кінцем з (b);

50 d) розрізання лігваної ДНК, отриманої на стадії (c), і лігування щонайменше одного другого адаптера з розрізаним нуклеотидним кінцем для забезпечення ампліфікації та секвенування фрагментів геномної ДНК, які оточують двонитковий розрив;

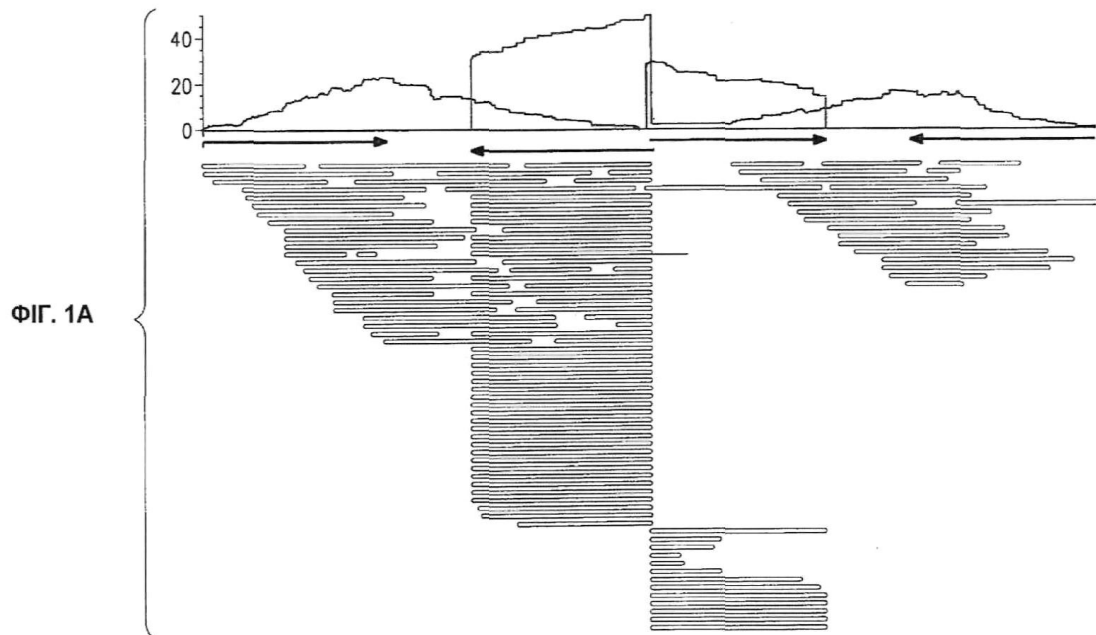
e) вирівнювання нуклеотидних послідовностей фрагментів ДНК, отриманих в (d), з еталонною геномною послідовністю ДНК; та

55 f) ідентифікацію варіантного сайту розпізнавання, що містить зміну щонайменше по одній нуклеотидній основі в порівнянні з припустимим сайтом розпізнавання зазначеного сконструйованого засобу для індукції двониткового розриву.

16. Спосіб за п. 1 або п. 15, де сконструйований засіб, що рідко розщеплює, для індукції двониткового розриву являє собою ендонуклеазу Cas.

60 17. Спосіб за п. 16, де ендонуклеаза Cas здатна до утворення комплексу з направляючою РНК,

де зазначений комплекс дозволяє ендонуклеазі Cas вводити двонитковий розрив у зазначену геномну ДНК.



ФІГ. 2

% склад основ ДНК	22 п.о. сайт розпізнавання																					
	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
ХС	10%	3%	17%	23%	7%	7%	0%	0%	0%	50%	0%	23%	0%	87%	3%	37%	7%	90%	3%	97%	3%	27%
ХС	50%	7%	30%	27%	67%	40%	93%	0%	100%	0%	43%	3%	67%	10%	20%	10%	63%	3%	13%	0%	10%	0%
ХТ	17%	63%	17%	23%	7%	47%	7%	87%	0%	37%	57%	10%	23%	0%	70%	0%	20%	0%	80%	3%	87%	7%
ХА	23%	27%	37%	27%	20%	7%	0%	13%	0%	13%	0%	63%	10%	3%	7%	53%	10%	7%	3%	0%	0%	67%
Передбачуваний сайт розпізнавання	A	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A

ФІГ. 3А

% склад основ ДНК	22 п.о. сайт розпізнавання																					
	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
ХС	5%	11%	9%	39%	2%	15%	79%	2%	7%	75%	4%	30%	5%	96%	2%	2%	12%	39%	51%	15%	4%	36%
ХС	59%	2%	1%	9%	51%	2%	6%	1%	9%	3%	26%	2%	31%	0%	4%	90%	4%	5%	8%	7%	12%	16%
ХА	4%	86%	89%	36%	43%	77%	1%	71%	46%	7%	31%	67%	48%	3%	37%	8%	82%	7%	29%	10%	9%	36%
ХТ	32%	1%	2%	17%	4%	6%	14%	26%	38%	15%	39%	2%	17%	1%	57%	0%	2%	49%	12%	68%	75%	12%
Передбачуваний сайт розпізнавання	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	G	A	T	T	T

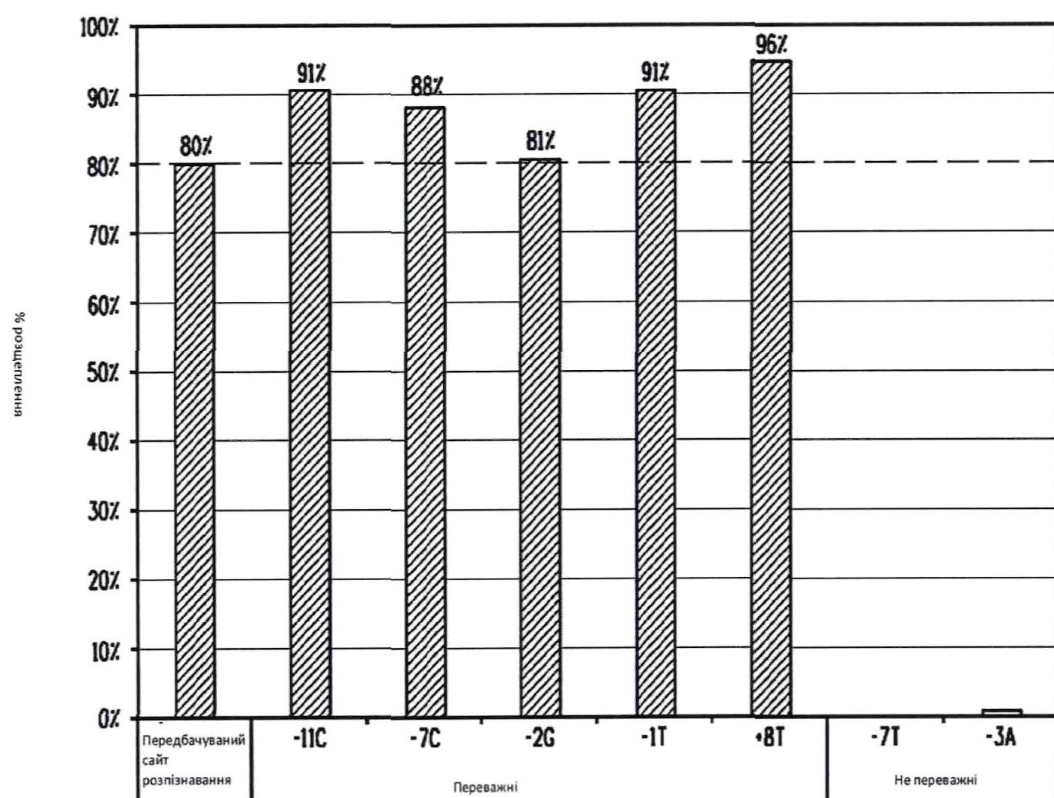
ФІГ. 3В

Мегануклеаза Uig3-4	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
Передбачуваний сайт розпізнавання	A	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
-11C	C	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
-7C	A	T	A	T	C	C	C	T	C	A	C	A	C	G	T	A	C	G	C	G	T	A
-2G	A	T	A	T	A	C	C	T	C	G	C	A	C	G	T	A	C	G	C	G	T	A
-1T	A	T	A	T	A	C	C	T	C	A	T	A	C	G	T	A	C	G	C	G	T	A
-8T	A	T	A	T	A	C	C	T	C	A	C	A	C	G	T	A	C	G	T	G	T	A
-7C, -8T	A	T	A	T	C	C	C	T	C	A	C	A	C	G	T	A	C	G	T	G	T	A
-11C, -7C, -2G, -1T, -8T	C	T	A	T	C	C	C	T	C	G	T	A	C	G	T	A	C	G	T	G	T	A
-11C, -7C, -1T, -8T	C	T	A	T	C	C	C	T	C	A	T	A	C	G	T	A	C	G	T	G	T	A

ФІГ. 4А

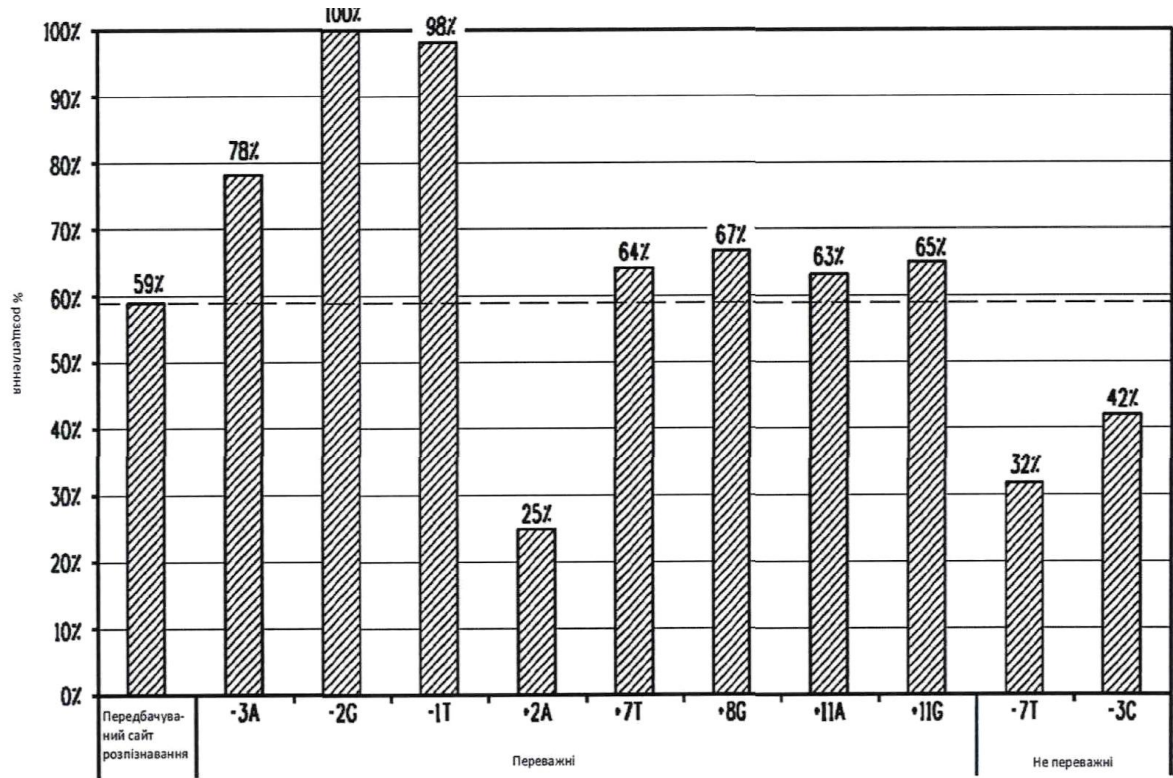
Мегануклеаза MPH14+	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
Передбачуваний сайт розпізнавання	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	G	A	T	T	T
-3A	C	A	A	A	C	A	G	A	A	T	C	A	C	G	T	C	A	G	A	T	T	T
-2G	C	A	A	A	C	A	G	A	T	G	C	A	C	G	T	C	A	G	A	T	T	T
-1T	C	A	A	A	C	A	G	A	T	T	T	A	C	G	T	C	A	G	A	T	T	T
+2A	C	A	A	A	C	A	G	A	T	T	C	A	A	G	T	C	A	G	A	T	T	T
-7T	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	T	A	T	T	T
+8G	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	G	G	T	T	T
+11G	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	G	A	T	T	G
+11A	C	A	A	A	C	A	G	A	T	T	C	A	C	G	T	C	A	G	A	T	T	A
-3A,-2G,-1T,+2A, +7T,+8G,+11G	C	A	A	A	C	A	G	A	A	G	T	A	A	G	T	C	A	T	G	T	T	G
-3A,-2G,-1T,+2A, +7T,+8G,+11A	C	A	A	A	C	A	G	A	A	G	T	A	A	G	T	C	A	T	G	T	T	A
-3A,-2G,-1T, +7T,+8G,+11G	C	A	A	A	C	A	G	A	A	G	T	A	C	G	T	C	A	T	G	T	T	G
-2G,-1T,+2A, +7T,+8G,+11G	C	A	A	A	C	A	G	A	T	G	T	A	A	G	T	C	A	T	G	T	T	G
-2G,-1T,+7T, +8G,+11G	C	A	A	A	C	A	G	A	T	G	T	A	C	G	T	C	A	T	G	T	T	G

ФІГ. 4B

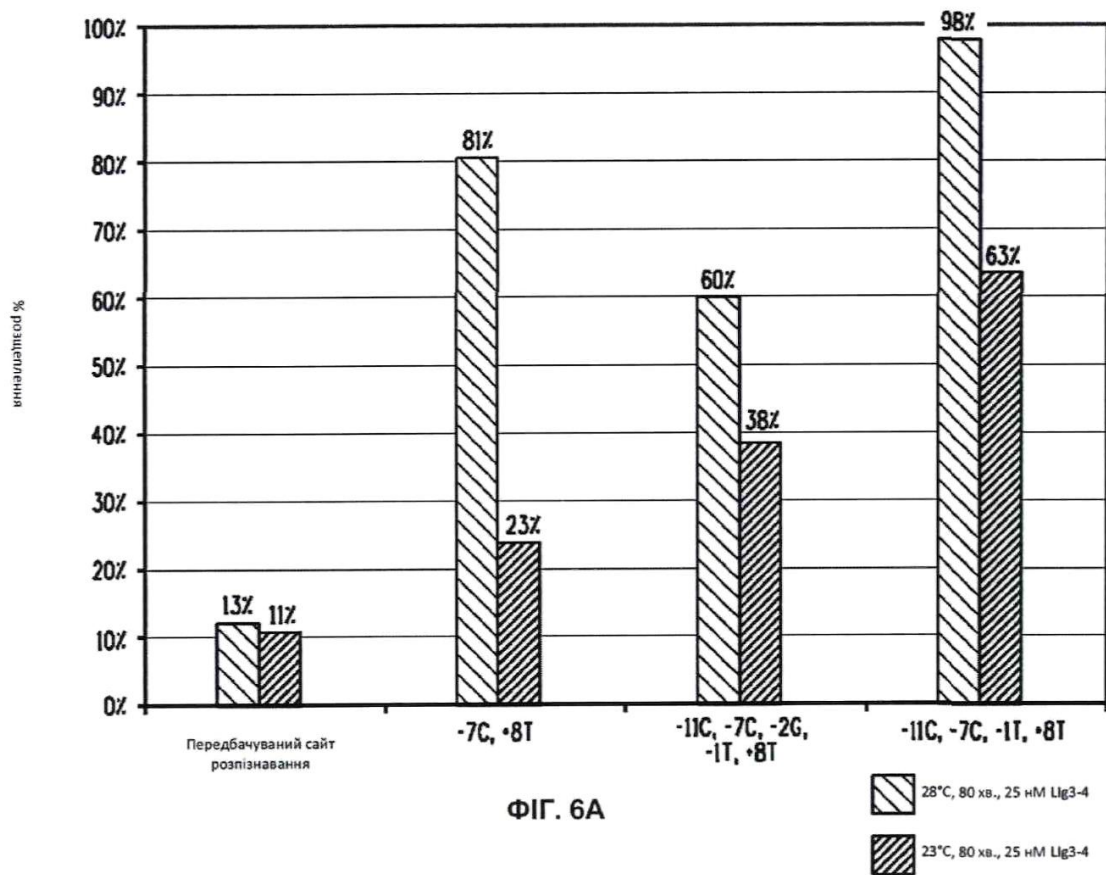


ФІГ. 5A

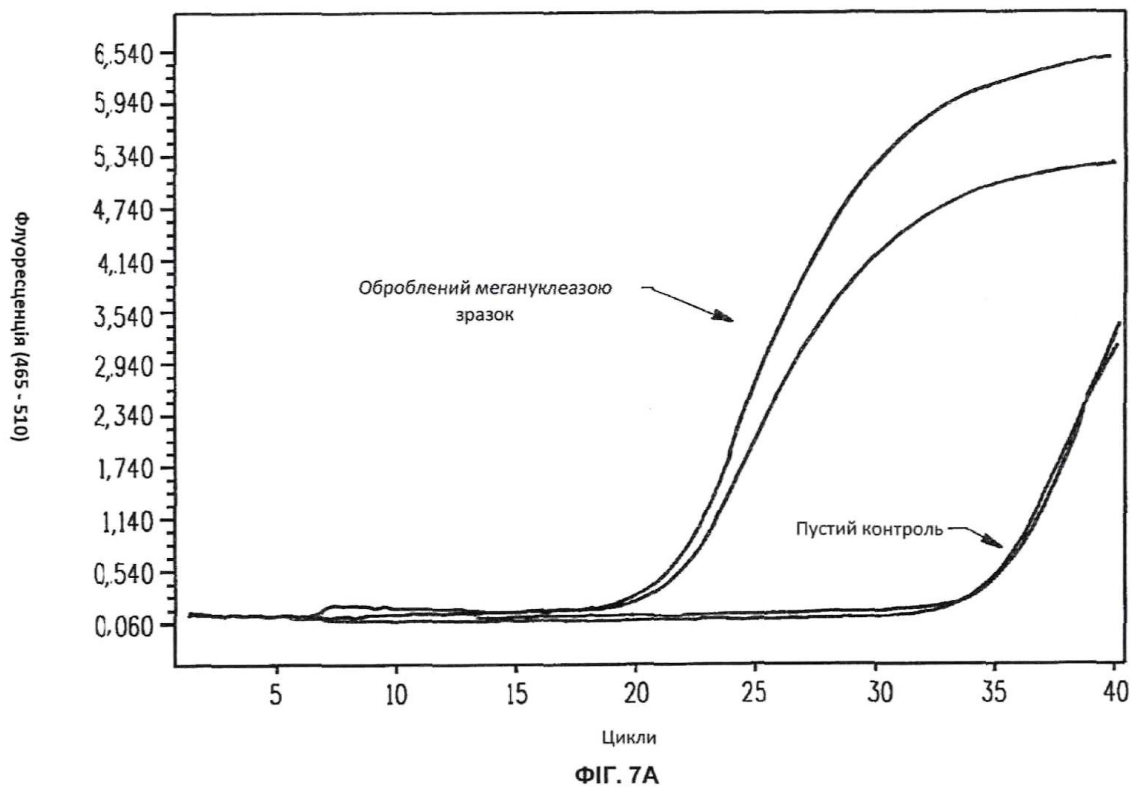
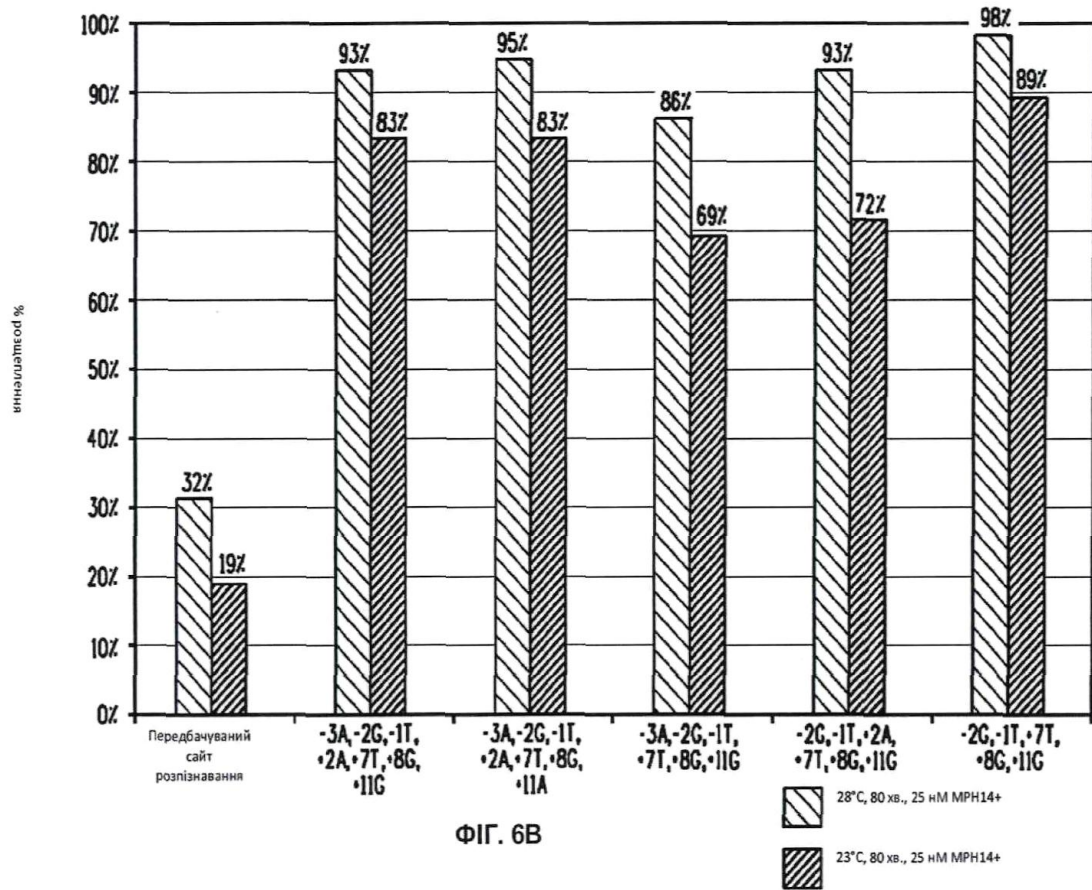
37 градусів Цельсія, 80 хвилин, 25 нМ Llg3-4

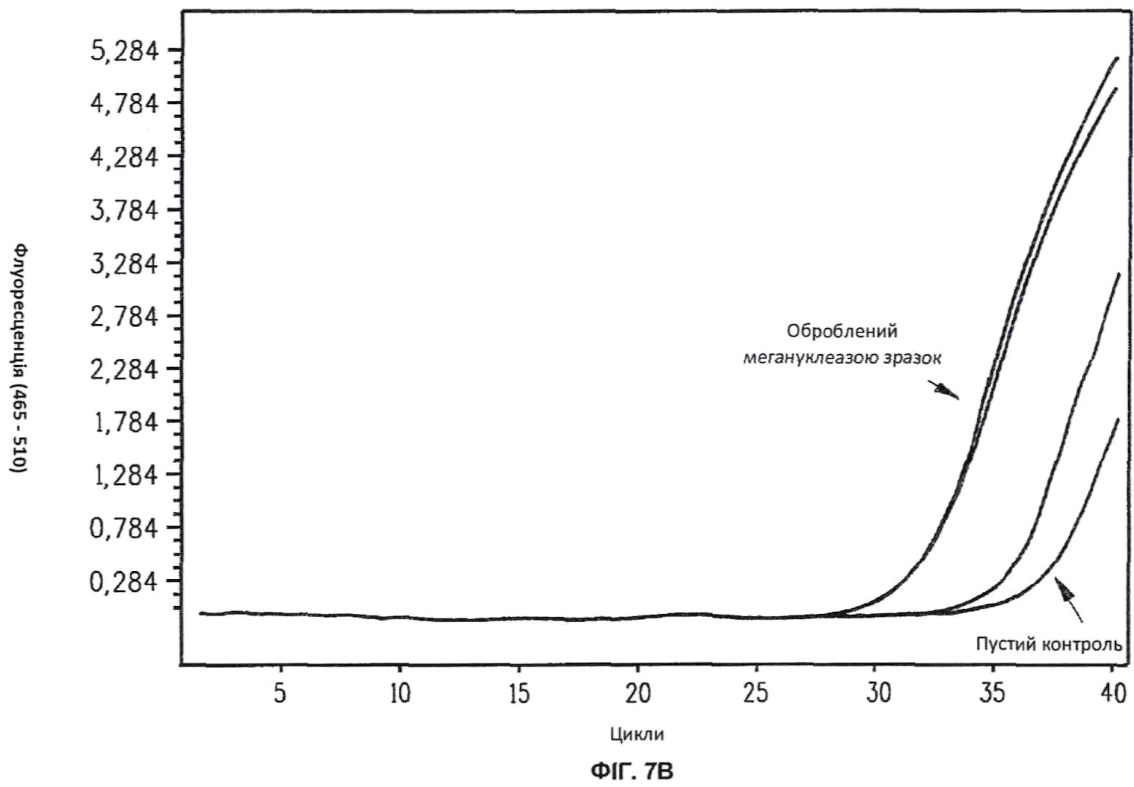


ФІГ. 5В

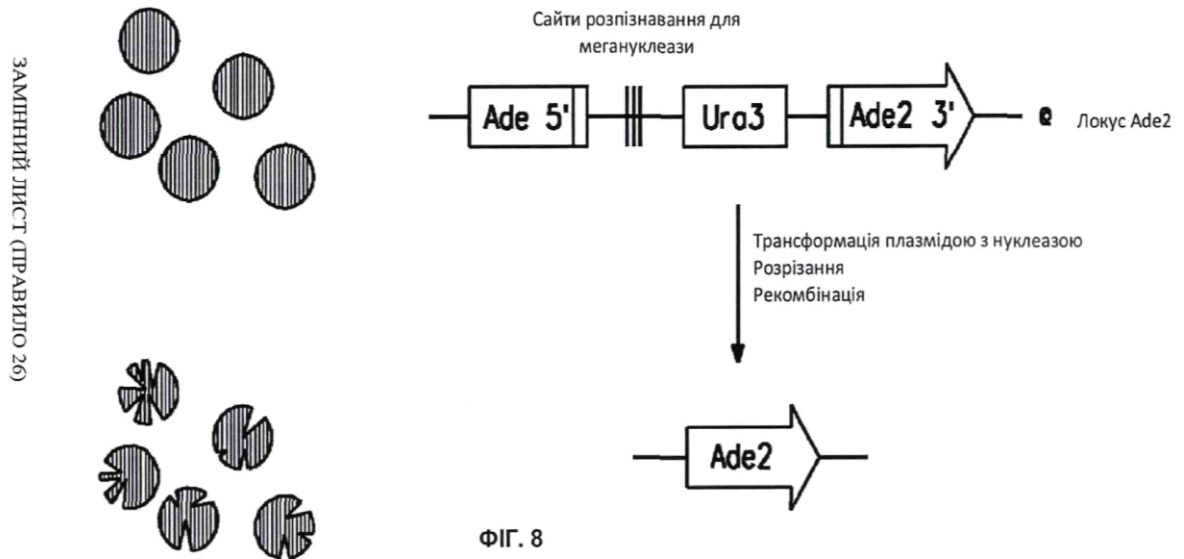


ФІГ. 6А

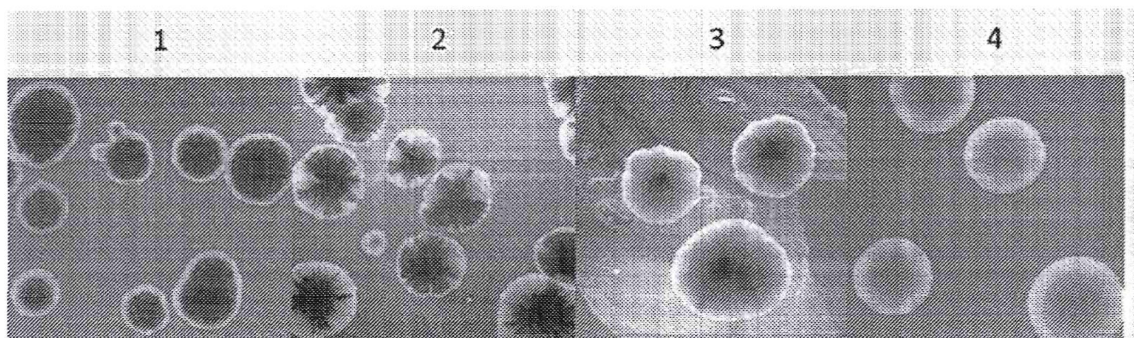




Система для скринінгу на основі дріжджів

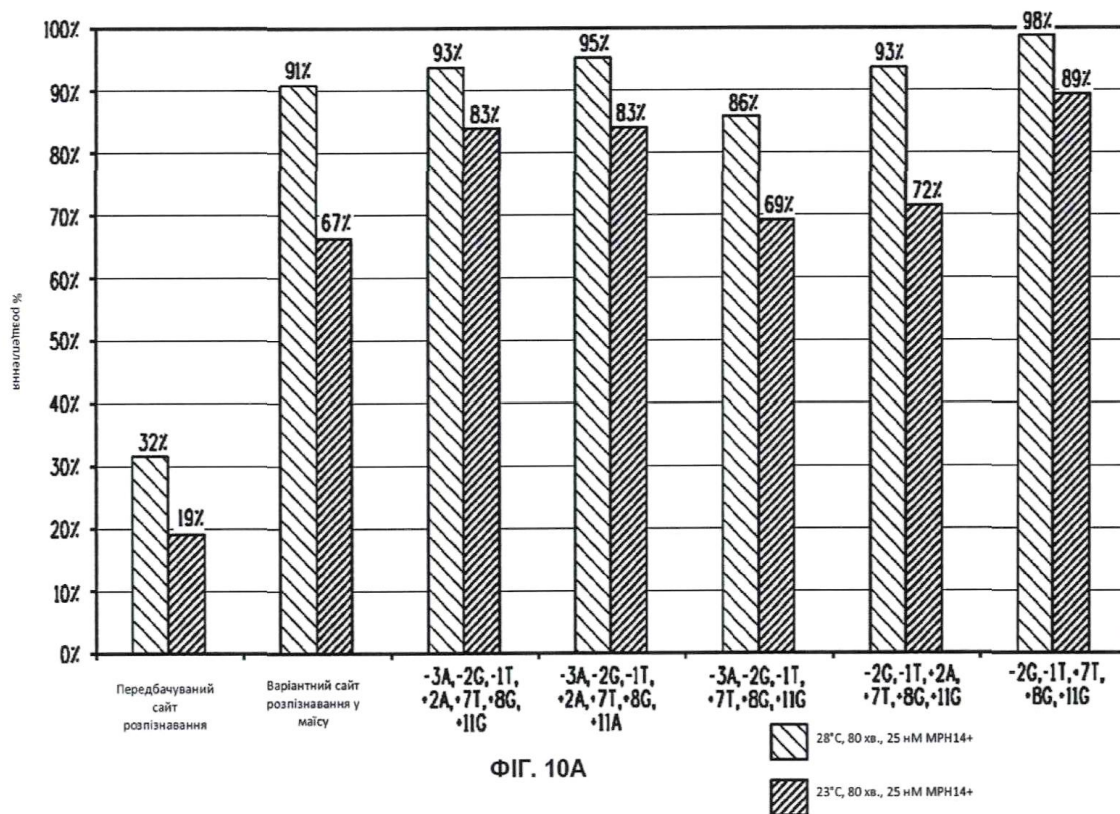


Числова шкала для скринінгу з дріжджами

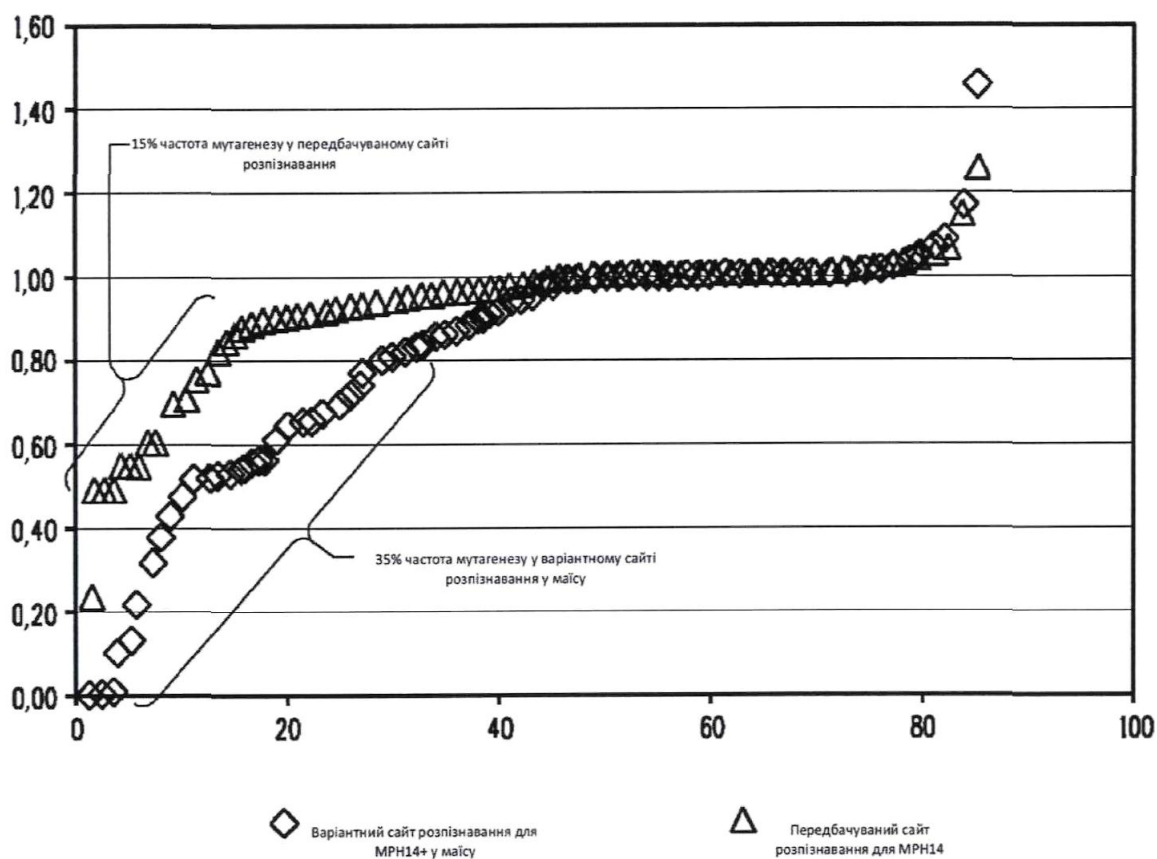


LIG34 30 град.
2% галактоза

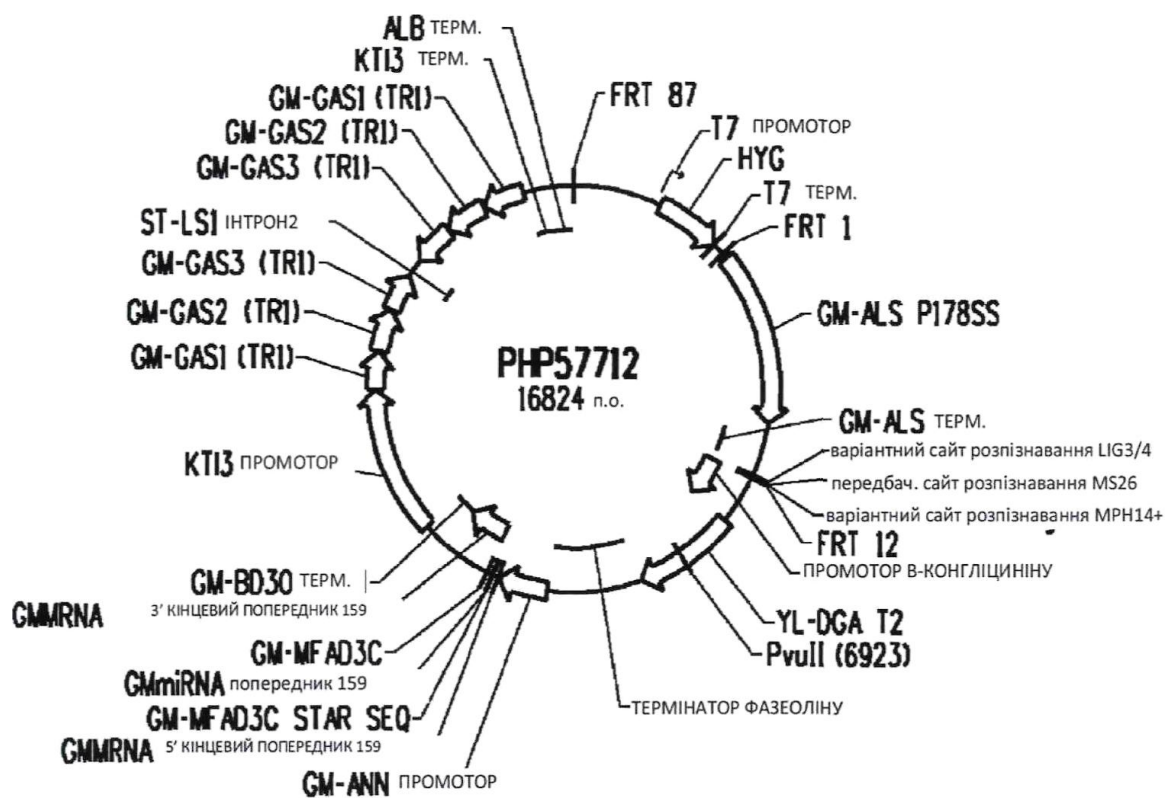
ФІГ. 9



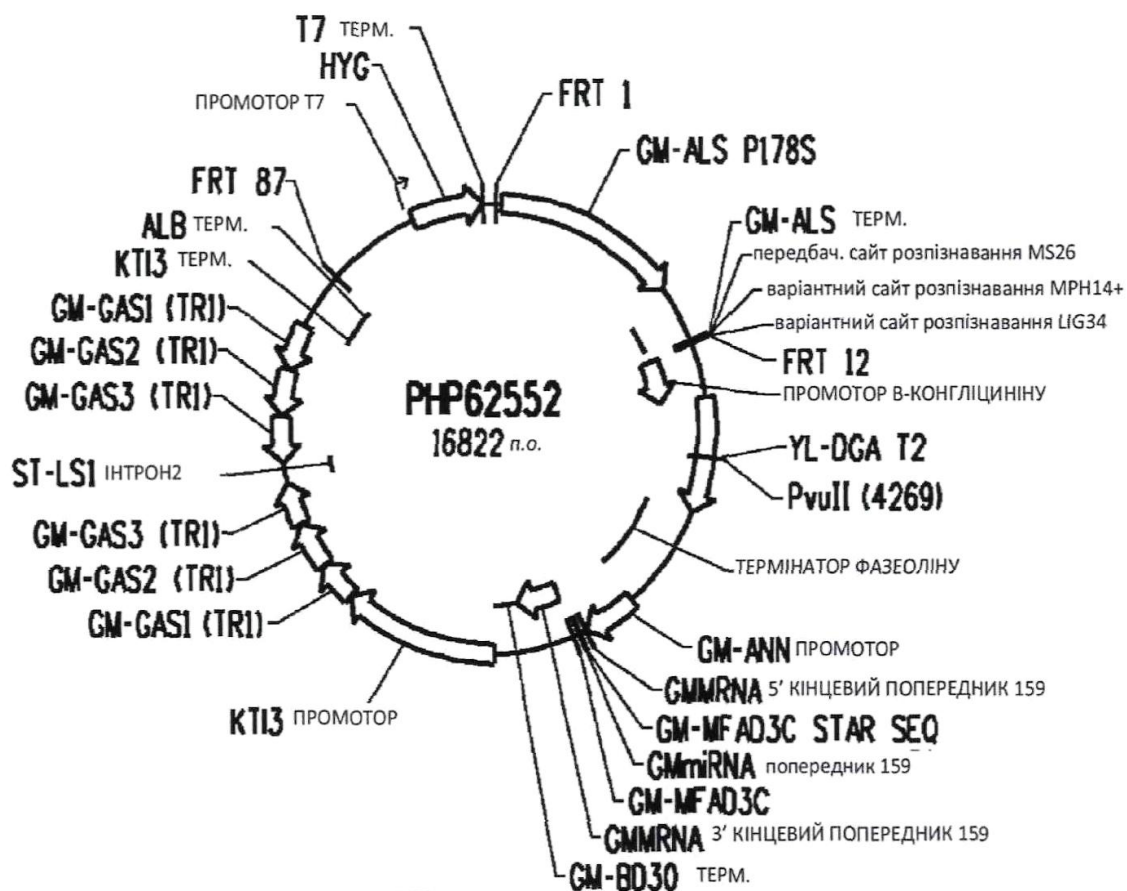
ФІГ. 10A



ФІГ. 10В



ФІГ. 11А



ФІГ. 11В