УКРАЇНА

(19) **UA** (11) **112643** (13) **C2**

(51) МПК
C10L 5/44 (2006.01)
C12N 15/82 (2006.01)
C12N 9/24 (2006.01)
C12N 9/42 (2006.01)
C12P 19/14 (2006.01)

ДЕРЖАВНА СЛУЖБА ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

(21) Номер заявки: **a 2013 11637**

(22) Дата подання заявки: **07.03.2012**

(24) Дата, з якої є чинними права на винахід: **10.10.2016**

(31) Номер попередньої заявки відповідно до Паризької конвенції: **61/449,769**

(32) Дата подання попередньої заявки відповідно до Паризької конвенції: **07.03.2011**

(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: **US**

(41) Публікація відомостей про заявку: **25.02.2014, Бюл.№ 4**

(46) Публікація відомостей про видачу патенту: **10.10.2016, Бюл.№ 19**

(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору PCT **PCT/US2012/028132, 07.03.2012**

(72) Винахідник(и):
Рааб Р. Майкл (US),
Чжан Дунчен (US),
Бугрі Олєг (US)

(73) Власник(и):
АГРІВІДА, ІНК.,
200 Boston Avenue, Suite 2975, Medford, Massachusetts 02155, United States of America (US)

(74) Представник:
Мошинська Ніна Миколаївна, реєстр. №115

(56) Перелік документів, взятих до уваги експертизою:
US 20100159510 A1, 24.06.2010.
US 7049485 B, 23.05.2006.
WO 2010060056 A2, 27.05.2010.
US 20090193541 A1, 330.07.2009.
WO 2008064314 A2, 29.05.2008.
Lindh et al., Production of the p24 capsid protein from HIV-1 subtype C in Arabidopsis thaliana and Daucus carota using an endoplasmic reticulum-directing SEKDEL sequence in protein expression constructs // Protein Expression and Purification. – 2009. – Vol.66. – P. 46-51.
Hood et al., Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed // Plant Biotechnology Journal.- 2007. – Vol.5. – P. 1-11.
US 20090298149 A1, 03.12.2009.
Morris et al., Cloning of the xynB Gene fromDictyoglomus thermophilum Rt46B.1 and Action of the Gene Product on Kraft Pulp // Applied and Environmental Microbiology. – 1998. – Vol.64(5). – P.1759-1765.
US 20100143967 A1, 10.06.2010.
US 5432074 A, 11.07.1995.
US 20100159494 A1, 24.06.2010.
WO 2010096510 A2, 26.08.2010.
WO 2009155601 A2, 23.12.2009.
WO 2007100897 A2, 07.09.2007.
Bae H-J et al., Production of a recombinant xylanase in plants and its potential for pulp biobleaching applications // Bioresource Technology. – 2008. - Vol.99. - № 9, - P. 3513-3519.

(54) СПОСІБ ОДЕРЖАННЯ РОЗЧИННИХ ЦУКРІВ З ГЕНЕТИЧНО МОДИФІКОВАНИХ РОСЛИН, ЩО ЕКСПРЕСУЮТЬ ФЕРМЕНТИ, ДЕГРАДУЮЧІ КЛІТИННУ СТІНКУ

UA 112643 C2

(57) Реферат:

Винахід належить до способу одержання розчинних цукрів з модифікованих способами генетичної інженерії рослин, що експресують ферменти, деградуючі клітинну стінку. А також до касет експресії і векторів для одержання зазначених трансгенних рослин.

За даною заявкою вимагається пріоритет тимчасової заявки США № 61/449769, поданої 7 березня 2011 року, яка включена в даний опис шляхом посилання в повному обсязі.

Список послідовностей, наданий в електронному вигляді разом з даною заявкою за назвою "Список послідовностей", який створений 7 березня 2012 року і має розмір файла 620037 байт, включений у даний опис шляхом посилання в повному обсязі.

ГАЛУЗЬ ВИНАХОДУ

Даний винахід стосується способів одержання розчинних цукрів з рослин, експресуючих ферменти, деградуючі клітинну стінку, трансгенних рослин, векторів експресії, нуклеїнових кислот і білків, деградуючих клітинну стінку.

РІВЕНЬ ТЕХНІКИ

Лігноцелюлозна біомаса є привабливою сировиною для виробництва біопалива, хімічних реагентів і біопродуктів. Лігноцелюлозна біомаса забезпечує множину переваг, включаючи широку доступність, потенційно низьку вартість, відновлюваність і той факт, що її звичайно не вживають люди як джерело їжі (Langeveld J.W.A. et al. 2010, Crop Sci. 50:S131-S151). Для перетворення лігноцелюлозної біомаси в відновлювану енергію і біохімічні речовини, за допомогою біопроцесів здійснюють конверсію частини лігноцелюлозної біомаси в прості цукри, які конвертують у біопалива або інші біопродукти.

Вартість одержання цукру шляхом біологічної конверсії є високою внаслідок витрат на попередню обробку і ферментативний гідроліз біомаси (Alvira P. et al. 2010, Bioresour. Technol. 101:4851; Abramson M. et al. 2010, Plant Science, 178:61; Daniel Klein-Marcuschamer et al. Biotechnol. Bioeng. 2012; 109:1083). Клітинні стінки рослин не піддаються ферментативному гідролізу, оскільки гетерогенність, хімічний склад і структурні ознаки полісахаридів клітинної стінки роблять їх недоступними для гідролітичних ферментів (Zhu L. et al. 2008, Bioresour. Technol. 99:3817). З цієї причини для ферментативного гідролізу потрібна попередня обробка, яка може забезпечити доступність клітинних стінок рослин. У технологіях попередньої обробки, що переважають у промисловості, як правило, використовуються жорсткі умови, такі як високі температури і крайні значення pH (Wyman C.E. et al. Bioresour. Technol. 96:1959; Mosier N. et al. 2005, Bioresour. Technol. 96:673). Ці умови викликають деградацію цукрів і приводять до зниження виходу й утворення токсичних сполук ферментації, що вимагають дорогих додаткових стадій детоксикації, розділення і нейтралізації, а також дорогого передового капітального обладнання.

Витрати на попередню обробку, висока вартість навантаження екзогенними ферментами, низька швидкість гідролізу й обмежені постачання ферментів також є проблемами для комерціалізації способів, що залучають лігноцелюлозну біомасу.

СУТЬ ВИНАХОДУ

В одному аспекті винахід стосується способу одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Спосіб включає попередню обробку шляхом змішування одержаного способами генної інженерії рослинного матеріалу з перетворюючим у пульпу складом з одержанням суміші. Одержаний способами генної інженерії рослинний матеріал включає першу полінуклеотидну послідовність, кодуючу перший білок, вибраний із групи, що складається з: ксиланази, ендоглюканази, екзоглюканази, ферулоїлестерази, модифікованої інтеїном ксиланази, модифікованої інтеїном ендоглюканази, модифікованої інтеїном екзоглюканази і модифікованої інтеїном ферулоїлестерази. Спосіб також включає забезпечення умов гідролізу.

В одному аспекті винахід стосується одержаної способами генної інженерії рослини. Одержана способами генної інженерії рослина включає першу полінуклеотидну послідовність, кодуючу амінокислотну послідовність з щонайменше 90 % ідентичністю з першою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10 і SEQ ID NO: 11.

В одному аспекті винахід стосується касети експресії. Касета експресії включає першу полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з першої референсної послідовності, вибраної з групи, що складається з: SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38 і SEQ ID NO: 39 [P77853:S158-30-108-35]. Касета експресії також включає другу полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з другої референсної послідовності, вибраної з групи, що складається з: SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO:

41, SEQ ID NO: 42 і SEQ ID NO: 43. SEQ ID NO, вибрана як перша референсна послідовність, відрізняється від SEQ ID NO, вибраної як друга референсна послідовність.

В одному аспекті винахід стосується касети експресії. Касета експресії включає полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з референсної послідовності, вибраної з групи послідовностей, що складається з: SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61 і SEQ ID NO: 62.

В одному аспекті винахід стосується вектора експресії. Вектор експресії включає полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з послідовності, вибраної з групи, що складається з: SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82 і SEQ ID NO: 83.

КОРОТКИЙ ОПИС КРЕСЛЕНЬ

Представлений нижче докладний опис переважних варіантів здійснення даного винаходу буде краще зрозумілим при прочитанні разом із прикладеними кресленнями. Для мети ілюстрації винаходу на кресленнях представлені варіанти здійснення, які на даний час є переважними. Однак зрозуміло, що винахід не обмежується точними представленими схемами і засобами.

На ФІГ. 1 представлена принципова технологічна схема, що ілюструє стадії об'єднаної попередньої обробки і гідролізу рослинної біомаси, експресуючої ферменти, деградуючі клітинну стінку.

На ФІГ. 2A, 2B проілюстровані виходи глюкози (ФІГ. 2A) і ксилози (ФІГ. 2B) з попередньо обробленої трансгенної рослини, експресуючої ксиланазу A (XynA.2015.05), і трансгенної контрольної рослини, позбавленої ксиланази A (TGC.4000.12) після ферментативного гідролізу коктейлем ферментів #5 (FCt; сірий (середній стовпчик у кожнім наборі з трьох)) або коктейлем ферментів #5, позбавленим ксиланази A (Ct-Xyn; діагональні смуги (праворуч)), або без ферментів (NCt; білий (ліворуч)).

На ФІГ. 3A, 3B проілюстровані виходи глюкози (ФІГ. 3A) і ксилози (ФІГ. 3B) з попередньо обробленої трансгенної рослини, експресуючої ксиланазу B (XynB.2063.17), і попередньо обробленої контрольної рослини дикого типу (AxB) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)) або коктейлем ферментів #1, позбавленим ксиланази (Ct-Xyn; діагональні смуги (праворуч)), або без ферментів (NCt; білий (ліворуч)).

На ФІГ. 4 проілюстровані виходи глюкози з попередньо оброблених трансгенних рослин, експресуючих ендоглюканазу (EG) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)) або коктейлем ферментів #1, позбавленим ендоглюканази (Ct-EG (праворуч)), або без ферментів (NCt; білий (ліворуч)). На ФІГ. 4A проілюстрований вихід глюкози з трансгенних рослин, експресуючих ендоглюканазу A (EGA.2049.02 і EGA.2049.10), і трансгенної контрольної рослини, позбавленої ендоглюканази (TGC.4000.12). На ФІГ. 4B проілюстрований вихід глюкози з трансгенної рослини, експресуючої ендоглюканазу B (EGB.2042.03), і трансгенної контрольної рослини, позбавленої ендоглюканази (TGC.2004.8.02).

На ФІГ. 5A-5D проілюстровані результати гідролізу для трансгенних рослин, експресуючих множину білків. На ФІГ. 5A і 5C проілюстровані виходи глюкози, а на ФІГ. 5B і 5D проілюстровані виходи ксилози з досліджуваних трансгенних рослин і трансгенних контрольних рослин при обробці коктейлем #1. На ФІГ. 5A і 5B проілюстровані результати для 1) трансгенних рослин з подвійним набором XynA/AccA/B.2096.05 і XynA/AccA/B.2096.01, що експресують ксиланазу A (XynA) і допоміжні ферменти (Acc), і 2) трансгенної контрольної рослини TGC.2004.8.02 при обробці повним коктейлем ферментів (FCt; темно-сірий (середина)), повним коктейлем, позбавленим ксиланази (FCt-Xyn; стовпчики в смужку (праворуч)), і без ферментів (NCt; білі стовпчики (ліворуч)). На ФІГ. 5C і 5D проілюстровані результати для 1) трансгенної рослини EGA/XynA.2242.09, експресуючої XynA і EGA, і 2) трансгенної контрольної рослини TGC.4000.12 при обробці повним коктейлем ферментів (FCt; темно-сірий (лівий стовпчик з чотирьох для кожного зразка)), повним коктейлем, позбавленим ксиланази (Ct-Xyn; діагональні смуги (другий ліворуч)), повним коктейлем, позбавленим ендоглюканази (Ct-EG; білий (третій ліворуч)), і повним коктейлем, позбавленим ксиланази і ендоглюканази (Ct-Xyn-EG; у клітинку (четвертий ліворуч)).

На ФІГ. 6A, 6B проілюстровані виходи глюкози і ксилози, відповідно, із соломи попередньо обробленої контрольної рослини дикого типу AxB і трансгенних рослин кукурудзи XynB/EGA/CBHB.2349.56, XynB/EGA/CBHB.2349.55 і XynB/EGA/CBHA.2345.116, що експресують

потрійні набори білків. Виходи вимірювали після ферментативного гідролізу коктейлем ферментів Accelerase® 1500/XY (FCt; чорні стовпчики (праворуч)) у порівнянні з контрольною обробкою, позбавленою коктейлю ферментів (NCt; сірі стовпчики (ліворуч)).

На ФІГ. 7A, 7B проілюстровані виходи глюкози і ксилози, відповідно, із трансгенних рослин. На ФІГ. 7A показані виходи глюкози з попередньо оброблених трансгенних рослин проса, експресуючих ксиланазу A (XynA.pv2015.3c, XynA.pv2015.4c), і попередньо обробленої рослини проса (Alamo) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)); коктейлем #1, позбавленим ксиланази (Ct-Xyn; діагональні смуги (праворуч)); і контрольної обробки, позбавленої коктейлю ферментів (NCt; білий (ліворуч)). На ФІГ. 7B показані результати виходу ксилози з попередньо обробленої трансгенної рослини першого покоління, експресуючої ксиланазу A (XynA.2015.05.T0), трансгенної рослини другого покоління, експресуючої ксиланазу A (XynA.2015.05.T1), і попередньо обробленої трансгенної рослини, позбавленої ксиланази (TGC.4000.11) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)); коктейлем #1, позбавленим ксиланази (Ct-Xyn; діагональні смуги (праворуч)); і контрольної обробки, позбавленої коктейлю ферментів (NCt; білий (ліворуч)).

На ФІГ. 8A, 8B проілюстровані виходи глюкози і ксилози, відповідно, із двох попередньо оброблених трансгенних рослин проса, експресуючих ксиланазу A (XynA.pv2015.3c і XynA.pv2015.4c), у порівнянні з контрольною нетрансгенною рослиною проса (Alamo) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)); коктейлем ферментів #1, позбавленим ксиланази (Ct-Xyn; діагональні смуги (праворуч)), і контрольної обробки, позбавленої коктейлю ферментів (NCt; білий (ліворуч)).

На ФІГ. 9A, 9B проілюстровані виходи глюкози і ксилози, відповідно, з попередньо обробленого трансгенної рослини (iXynA.2229.110), експресуючої модифіковану інтеїном XynA (iXynA) і попередньо обробленої контрольної рослини дикого типу (AxB) після ферментативного гідролізу коктейлем ферментів #1 (FCt; сірий (середина)); коктейлем ферментів #1, позбавленим ксиланази (Ct-Xyn; діагональні смуги (праворуч)), і контрольної обробки без коктейлю ферментів (NCt; білий (праворуч)).

На ФІГ. 10 проілюстрований вихід глюкози з перебігом часу після ферментативного гідролізу попередньо обробленої трансгенної рослини, експресуючої ксиланазу A (XynA.2015.5T1; зафарбований трикутник), і попередньо обробленої трансгенної рослини, експресуючої ксиланазу B (XynB.2063.17; зафарбоване коло), проти трансгенної контрольної рослини (TGC.2004.8.02; незафарбований квадрат) з використанням коктейлю ферментів #1.

На ФІГ. 11 проілюстрований вихід глюкози з перебігом часу після ферментативного гідролізу попередньо обробленої трансгенної рослини (EGA.2049.10) і попередньо обробленого трансгенного контролю (TGC.4000.11) з використанням коктейлю ферментів #1 (EGA.2049.10.FCt, зафарбований квадрат; TGC.4000.11.FCt, незафарбований ромб) і коктейлю ферментів #1, позбавленого ендоглюканази (EGA.2049.10.Ct-EG, зафарбований трикутник; TGC.4000.11.Ct-EG, незафарбоване коло).

На ФІГ. 12A, 12B проілюстрований вихід глюкози з перебігом часу після ферментативного гідролізу попередньо обробленої трансгенної рослини (EGA/XynA.2242.09) і попередньо обробленої трансгенної контрольної рослини (TGC.4000.11) з використанням повного коктейлю ферментів (EGA/XynA.2242.09.FCt, зафарбований квадрат; TGC.4000.11.FCt, незафарбований ромб) у порівнянні з обробкою з використанням повного коктейлю ферментів, позбавленого ендоглюканази (EGA/XynA.2242.09.Ct-EG, зафарбоване коло; TGC.4000.11.Ct-EG, незафарбоване коло на ФІГ. 12A), і повного коктейлю ферментів, позбавленого ксиланази (EGA/XynA.2242.09.Ct-Xyn, зафарбоване коло; TGC.4000.11.Ct-Xyn, незафарбоване коло).

На ФІГ. 13A, 13B проілюстровані виходи глюкози і ксилози, відповідно, з перебігом часу з попередньо обробленої трансгенної рослини, експресуючої ксиланазу A і ферулоїлестеразу B (XynA/AccB.2092.103), і попередньо обробленої трансгенної контрольної рослини (TGC.4000.11) після ферментативного гідролізу повним коктейлем ферментів (XynA/AccB.2092.103.FCt, зафарбований квадрат; TGC.4000.11.FCt, незафарбований ромб) і повним коктейлем ферментів, позбавленим ксиланази (XynA/AccB.2092.103.Ct-Xyn, зафарбований трикутник; TGC.4000.11.Ct-Xyn, незафарбоване коло).

На ФІГ. 14 проілюстровані виходи глюкози після ферментативного гідролізу попередньо обробленої трансгенної рослини, експресуючої наступні білки: ксиланаза B (XynB.2063.17), ендоглюканаза (EGA.2049.10), ксиланаза A і ферулоїлестераза B (XynA/Acc.B.2092.103), ксиланаза A і ендоглюканаза (EGA/XynA.2242.09), модифікована інтеїном ксиланаза A (iXynA.2229.110), у порівнянні з нетрансгенною контрольною рослиною (AxB) і трансгенною контрольною рослиною, позбавленою ферментів (TGC.4000.11). Попередню обробку проводили при температурах 65 °C (PT_65) і 75 °C (PT_75). Навантаження ферментами включає 0,2 мл

Accellerase® 1500 або 0,1 мл Accellerase® XY на грам біомаси і 0,05 мкМ β-глюкозидазу (BGL). Стовпчики над кожною з груп попередньої обробки PT_65 і PT_75 відповідають даним для трансгенних і контрольних рослин наступним чином зліва направо: AxB; TGC.4000.11; XynA.2015.05T1; XynB.2063.17; XynA/AccB.2092.103; iXynA.2229.110; EGA.2049.10 і XynA/EGA.2242.09.

На ФІГ. 15 проілюстровані виходи глюкози після ферментативного гідролізу попередньо обробленого трансгенного проса (EGC.2253.4b, зафарбоване коло) і проса дикого типу (Alamo, незафарбований квадрат) коктейлем ферментів #5. Температури попередньої обробки: 65 °C, 75 °C і 95 °C.

На ФІГ. 16A, 16B проілюстрований ефект температури попередньої обробки і часу на вихід глюкози (ФІГ. 16A) і ксилози (ФІГ. 16B) після ферментативного гідролізу попередньо обробленої трансгенної рослини, експресуючої ендоглюканазу і ксиланазу A (EGA/XynA.2342.105; чорні стовпчики (праворуч)), і попередньо обробленої контрольної рослини (TGC.2342.01; сірі стовпчики (ліворуч)).

На ФІГ. 17A, 17B проілюстрований ефект температури попередньої обробки на вихід глюкози (ФІГ. 17A) і ксилози (ФІГ. 17B) з попередньо оброблених трансгенних рослин EGA/XynA.2242.09.01 і EGA/XynA.2242.09.07, експресуючих ендоглюканазу A і ксиланазу A, і контрольних рослин: AxB дикого типу і трансгенна TGC.4000.11.

На ФІГ. 18 проілюстрований ефект зниження навантаження зовнішніх ферментів на виходи глюкози з попередньо оброблених трансгенних рослин XynA.2015.05T1 (зафарбоване коло), XynB.2063.17 (зафарбований трикутник) і контрольної рослини TGC.2004.8.04 (зафарбований квадрат) після гідролізу при знижуваному навантаженні ферментом: повний коктейль #1 (FCt), 75 % повний коктейль #1 (0,75 FCt), 50 % повний коктейль #1 (0,50 FCt), 25 % повний коктейль #1 (0,25 FCt), 10 % повний коктейль #1 (0,10 FCt), і без ферментів (0 FCt).

На ФІГ. 19 проілюстрований ефект зниження навантаження зовнішніх ферментів на вихід глюкози з трансгенних рослин XynE/EGC/CBHA.2339.03, XynE/EGC/CBHA.2339.04 і XynE/EGC/CBHA.2339.05 і контрольної рослини BxA. Попередню обробку проводили з використанням 0,17M бісульфіту амонію і карбонату амонію (pH 8,1) при 75 °C, при співвідношенні рідини і твердої речовини, що дорівнює 10, протягом 16 годин. Ферментативний гідроліз проводили при вмісті твердих речовин приблизно 2 % без ферментів (NCt; білий (ліворуч)), з 20 % повним коктейлем (0,2 FCt; сірий (середина)) і з повним коктейлем Accellerase® 1500/XY при 0,2 мл/0,1 мл на грам соломи (FCt; чорний (праворуч)) при 50 °C і pH 5,0 протягом 3 діб.

На ФІГ. 20A, 20B проілюстрований вихід глюкози і ксилози, відповідно, з попередньо оброблених рослин, експресуючих ксиланазу A і ендоглюканазу (XynA/EGA.2309.54 і XynA/EGA2309.107), у порівнянні з попередньо обробленою нетрансгенною контрольною рослиною (BxA) після ферментативного гідролізу при повному навантаженні коктейлем ферментів Accelerase®1500/XY (FCt; чорний (праворуч)), 20 % навантаженні коктейлем (0,2 FCt; сірий (середина на ФІГ. 20A і праворуч на ФІГ. 20B)) і без ферментів (NCt; білий (ліворуч на ФІГ. 20A)).

На ФІГ. 21A, 21B проілюстрований ефект зниження навантаження зовнішніми ферментами на вихід глюкози (ФІГ. 21A) і ксилози (ФІГ. 21B) із трансгенних рослин EGA/XynA.2242.09.16, CBHA.2069.01.03 і контрольної рослини TGC.4000.11 після ферментативного гідролізу повним коктейлем при 0,2 мл Accellerase® 1500 на грам соломи + 0,1 мл Accellerase® XY на грам соломи (FCt;), 80 % повним коктейлем: 0,16 мл Accellerase® 1500 на грам соломи + 0,08 мл Accellerase® XY на грам соломи (0,8 FCt), 60 % повним коктейлем: 0,12 мл Accellerase® 1500 на грам соломи + 0,06 мл Accellerase® XY на грам соломи (0,6 FCt), 40 % повним коктейлем: 0,08 мл Accellerase® 1500 на грам соломи + 0,04 мл Accellerase® XY на грам соломи (0,4 FCt), 20 % повним коктейлем: 0,04 мл Accellerase® 1500 на грам соломи + 0,02 мл Accellerase® XY на грам соломи (0,2 FCt) і без ферментів (OFCt).

На ФІГ. 22 проілюстроване продукування етанолу при одночасному оцукрюванні і ферментації (SSF) попередньо оброблених трансгенних рослин EGA.2049.10 і EGA/XynA.2242.09 проти контрольних рослин з використанням 1) коктейлів ферментів Accellerase® 1500 і Accellerase® XY; і 2) штами дріжджів Saccharomyces cerevisiae D5A.

На ФІГ. 23 проілюстроване розчинення біомаси в розрахунку на зниження маси в трансгенній рослині, експресуючій екзоглюканазу CBHA (CBHA.2069.3.17; білий), і контрольній рослині дикого типу (AxB; сірий).

На ФІГ. 24A, 24B проілюстрований вихід оцтової кислоти (HAc) з нетрансгенної контрольної рослини (AxB; ФІГ. 24A) і трансгенної рослини, експресуючої екзоглюканазу CBHA

(CBHA.2063.3.17; ФІГ. 24B). Обробку проводили при 75 °C протягом 16 годин (білий (ліворуч)); 85 °C протягом 7 годин (стовпчики в смужку (у середині)) і 95 °C (у клітинку (праворуч)).

На ФІГ. 25A, 25B проілюстрований вихід продуктів деградації цукрів гідроксиметилфурфуралю (HMF) і фурфуралю з нетрансгенної контрольної рослини (AxB; ФІГ. 25A) і трансгенної рослини, експресуючої екзоглюканазу CBHA (CBHA.2063.3.17). Обробка, зазначена для білих стовпчиків, стовпчиків у смужку або у клітинку, була наступною зліва направо: білий, HMF_75 °C протягом 16 годин; у смужку, HMF_85 °C протягом 7 годин; у клітинку, HMF_95 °C протягом 16 годин; у смужку, фурфураль_95 °C протягом 16 годин).

На ФІГ. 26 проілюстрований вихід ксилози з трансгенної рослини, експресуючої ксиланазу B окремо (XynB.2063.15), трансгенної рослини, експресуючої ксиланазу A і два допоміжних ферменти A і B (XynA/AccA/B.2096.1), і нетрансгенної контрольної рослини (WT AxB) після попередньої обробки 0,17M бісульфітом амонію і карбонатом амонію (BSC; pH 8,1) і аутогідролізу. Вихід ксилози оцінювали для ксилози як мономера (чорний (праворуч)) і ксилози як олігосахариду (сірий (ліворуч)).

На ФІГ. 27 проілюстровані виходи ксилози з двох трансгенних рослин, експресуючих ксиланазу B, ендоглюканазу і CBHB (XynB/EGA/CBHB.2349.56 і XynB/EGA/CBHB.2349.55), трансгенної рослини, експресуючої ксиланазу B, ендоглюканазу і CBHA (XynB/EGA/CBHA.2345.116), і нетрансгенної контрольної рослини (AxB) після попередньої обробки 0,17M бісульфітом амонію і 0,165M карбонатом амонію (BSC; pH 8,1) і аутогідролізу.

На ФІГ. 28A, 28B проілюстровані виходи глюкози і ксилози, відповідно, з попередньо оброблених трансгенних рослин з використанням Accellerase® XY. EGA/XynA.2242.09T1 (зафарбоване коло) одночасно експресувала ендоглюканазу і ксиланазу A, і трансгенна контрольна рослина являла собою TGC.4000 (зафарбований квадрат).

На ФІГ. 28C, 28D представлені, відповідно, виходи глюкози і ксилози з XynB/EGA/CBHB.2349.55 (незафарбоване коло), XynB/EGA/CBHB.2349.229 (зафарбований трикутник, вершиною нагору) і XynB/EGA/CBHB.2349.56 (зафарбований трикутник, вершиною вниз), кожна з яких одночасно експресувала ендоглюканазу A, ксиланазу B і целобіогідролазу B, і iXynA.2329.14, що експресувала модифіковану інтеїном ксиланазу A, і контрольної рослини дикого типу AxB.

ДОКЛАДНИЙ ОПИС ПЕРЕВАЖНИХ ВАРІАНТІВ ЗДІЙСНЕННЯ

Визначену термінологію використовують в подальшому описі тільки для зручності і не для обмеження. Слова "правий", "лівий", "верх" і "низ" позначають напрямки на кресленнях або у конкретних прикладах, на які наводиться посилання. Форму однини, як використовують у формулі винаходу й у відповідних частинах опису, визначають як таку, що включає один або декілька зі згадуваних об'єктів, якщо немає інших конкретних указань. Вираз "щонайменше один" з наступним переліком двох або більше об'єктів, таким як "A, B або C", означає будь-який окремий з A, B або C, а також будь-яку їх комбінацію.

Варіанти здійснення в рамках даного винаходу передбачають технології для експресії набору білків, деградуючих клітинну стінку (CWD), у рослині. Білки CWD можуть являти собою ферменти CWD або модифіковані форми ферментів CWD. Модифіковані форми можуть являти собою модифіковані інтеїном білки CWD. Рослина може являти собою кукурудзу, сорго, просо або іншу рослину. Варіанти здійснення в рамках даного винаходу передбачають збирання рослинної біомаси з білками CWD у рослині для застосування як сировини для одержання цукру. При експресії ферментів у рослині рослина використовується як "фабрика" замість мікробної ферментації для одержання промислових ферментів CWD. Ця стратегія має перевагу, яка полягає в доставці білків безпосередньо в сировину біомаси для продукування ферментованих цукрів. Для біомаси трансгенної рослини з гідролітичними ознаками може не вимагатися жорстка попередня обробка з метою підвищення доступності целюлози клітинної стінки для екзогенних ферментів. Експресія різних класів білків CWD в одній рослині може забезпечити низьковитратну основу цукрів для виробництва біопалива біохімічних речовин. Варіанти здійснення в рамках даного винаходу передбачають способи продукування розчинних цукрів з використанням м'якої хімічної попередньої обробки лігноцелюлозної біомаси, що походить з рослин, модифікованих способами генетичної інженерії так, щоб вони включали один або декілька типів білка CWD.

Один варіант здійснення передбачає спосіб одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Спосіб може включати попередню обробку одержаного способами генної інженерії рослинного матеріалу шляхом змішування з перетворюючим у пульпу складом, формуючи суміш. Одержаний способами генної інженерії рослинний матеріал може включати першу полінуклеотидну послідовність, кодуючу перший білок. Перший білок може являти собою фермент CWD. Перший білок може являти собою

модифікований інтеїном фермент CWD. Перший білок може являти собою ксиланазу, ендоглюканазу, екзоглюканазу, ферулоїлестеразу, модифіковану інтеїном ксиланазу, модифіковану інтеїном ендоглюканазу, модифіковану інтеїном екзоглюканазу або модифіковану інтеїном ферулоїлестеразу. Перший білок може бути здатний здійснювати гідроліз компонента одержаного способами генної інженерії рослинного матеріалу. Є здатним гідролізувати компонент означає, що перший білок каталізує гідроліз компонента в умовах гідролізу. У випадку модифікованого інтеїном першого білка, наявність здатності гідролізувати компонент означає, що після видалення сплайсингом інтеїну з пептиду білок може гідролізувати компонент в умовах гідролізу. Крім того, спосіб може включати надання умов гідролізу. Умови гідролізу можуть бути придатні для гідролізу компонента.

Одержаний способами генної інженерії рослинний матеріал, крім того, може включати другу полінуклеотидну послідовність, кодуючу другий білок. Другий білок може являти собою білок CWD. Другий білок може являти собою модифікований інтеїном білок CWD. Другий білок може являти собою ксиланазу, ендоглюканазу, екзоглюканазу, ферулоїлестеразу, модифіковану інтеїном ксиланазу, модифіковану інтеїном ендоглюканазу, модифіковану інтеїном екзоглюканазу або модифіковану інтеїном ферулоїлестеразу. Білок, вибраний як другий білок, може являти собою білок, відмінний від білка, вибраного як перший білок. Другий білок може бути здатний здійснювати гідроліз компонента одержаного способами генної інженерії рослинного матеріалу. Є здатним гідролізувати компонент означає, що другий білок каталізує гідроліз компонента в умовах гідролізу. У випадку модифікованого інтеїном другого білка, наявність здатності гідролізувати компонент означає, що, після видалення інтеїну сплайсингом з пептиду, білок може гідролізувати компонент в умовах гідролізу.

Одержаний способами генної інженерії рослинний матеріал, крім того, може включати третю полінуклеотидну послідовність, кодуючу третій білок. Третій білок може являти собою білок CWD. Третій білок може являти собою модифікований інтеїном білок CWD. Третій білок може являти собою ксиланазу, ендоглюканазу, екзоглюканазу, ферулоїлестеразу, модифіковану інтеїном ксиланазу, модифіковану інтеїном ендоглюканазу, модифіковану інтеїном екзоглюканазу або модифіковану інтеїном ферулоїлестеразу. Білок, вибраний як третій білок, може відрізнятися від білка, вибраного як перший білок. Білок, вибраний як третій білок, може відрізнятися від білка, вибраного як другий білок. Третій білок може бути здатний гідролізувати компонент одержаного способами генної інженерії рослинного матеріалу. Є здатним гідролізувати компонент означає, що третій білок каталізує гідроліз компонента в умовах гідролізу. У випадку модифікованого інтеїном третього білка, наявність здатності гідролізувати компонент означає, що, після видалення інтеїну сплайсингом з пептиду, білок може гідролізувати компонент в умовах гідролізу.

Одержаний способами генної інженерії рослинний матеріал належить до трансгенної рослини, потомства трансгенної рослини, низхідного родича трансгенної рослини або частини будь-якого з вищевказаних. Одержаний способами генної інженерії рослинний матеріал може включати фермент, деградуючий клітинну стінку, що не зустрічається в природі в рослині, або ген, кодуючий його. Одержаний способами генної інженерії рослинний матеріал може являти собою трансгенну рослину, експресуючу білок CWD, або будь-яку частину трансгенної рослини. Одержаний способами генної інженерії рослинний матеріал може являти собою будь-яку трансгенну рослину, експресуючу модифіковану форму білка CWD, або будь-яку частину трансгенної рослини. Трансгенна рослина може являти собою рослину будь-якого типу. Тип трансгенної рослини може являти собою, але не обмежуватися ними, кукурудзу, цукровий буряк, цукрову тростину, сорго, просо, міскантус, евкаліпт, вербу або тополю. Одержаний способами генної інженерії рослинний матеріал може являти собою цілу трансгенну рослину або частини рослини. Частини можуть являти собою, але не обмежуватися ними, листи, стебла, квітки, бруньки, пелюстки, зав'язі, плоди або насіння. Одержаний способами генної інженерії рослинний матеріал може являти собою калюс із трансгенної рослини. Одержаний способами генної інженерії рослинний матеріал може бути таким, що регенерував з частин трансгенної рослини або рослин. Одержаний способами генної інженерії рослинний матеріал може являти собою продукт статевого схрещування першої трансгенної рослини і другої трансгенної рослини або нетрансгенної рослини, де рослина-продукт зберігає полінуклеотидну послідовність, введену в першу трансгенну рослину. Трансгенна рослина може являти собою будь-яку з трансгенних рослин, передбачених у рамках даного винаходу. Трансгенна рослина може включати будь-який вектор, касету експресії або виділену нуклеїнову кислоту або їх фрагмент.

Змішування одержаного способами генної інженерії рослинного матеріалу з перетворюючим у пульпу складом можна проводити за допомогою будь-якого комбінування одержаного

способами генної інженерії рослинного матеріалу з перетворюючим у пульпу складом. Змішування можна проводити шляхом струшування.

Перетворюючий у пульпу склад може являти собою речовину, яка руйнує лігнін, що зв'язує одне з одним лігноцелюлозні волокна в лігноцелюлозному рослинному матеріалі. Речовина може руйнувати лігнін без суттєвої деградації лігноцелюлозних волокон. Попередня обробка може приводити до часткового вивільнення ферментів, експресованих у модифікованих способами генетичної інженерії рослинах, і часткової деградації лігніну в лігноцелюлозному рослинному матеріалі.

Спосіб може включати активацію білка CWD до, у ході або після попередньої обробки. Спосіб може включати активацію білка CWD до, у ході або після забезпечення умов гідролізу. Активований білок CWD може являти собою перший білок, другий білок або третій білок, або будь-який додатковий фермент, що здійснює процесинг лігноцелюлози. Білок CWD можна модифікувати так, щоб він включав інтеїн. Інтеїн може бути злитий з ферментом CWD на кінці ферменту або у ферменті. Інтеїн може бути індукований до сплайсингу шляхом надання умов індукції. Умови індукції можуть являти собою конкретну температуру суміші. Умови індукції можуть являти собою температуру, забезпечену до, у ході або після однієї зі стадій попередньої обробки або забезпечення гідролізу. Модифіковані інтеїном ферменти й умови індукції сплайсингу інтеїнів, які можна використовувати як умови активації, описані в заявці США 10/886393, поданій 7 липня 2004 року, і PCT/US10/55746, поданій 5 листопада 2010 року, і PCT/US10/55669, поданій 5 листопада 2010 року, і PCT/US10/55751, поданій 5 листопада 2010 року, що включені в даний опис як посилання у повному обсязі.

Компонент може являти собою будь-яку частину, для якої бажаний процесинг. Компонент може являти собою лігноцелюлозний матеріал. Компонент може являти собою субстрат для будь-якого білка CWD, наведеного в даному описі. Компонент може являти собою субстрат для ксиланази, ендоглюканази, екзоглюканази, ферулоїлестерази. Компонент може являти собою частину, що включає субстрат для будь-якого білка CWD, наведеного в даному описі. Компонент може являти собою частину, що включає субстрат для ксиланази, ендоглюканази, екзоглюканази, ферулоїлестерази.

Спосіб також може включати додавання іншого рослинного матеріалу до, у ході або після змішування, попередньої обробки або надання умов гідролізу. Інший рослинний матеріал може являти собою будь-яку рослинну біомасу, целюлозний або лігноцелюлозний матеріал, відмінний від одержаного способами генної інженерії рослинного матеріалу. Інший рослинний матеріал може бути з установки по переробці біомаси. Інший рослинний матеріал може включати залишки лісівництва і сільськогосподарські залишки. Залишки лісівництва і сільськогосподарські залишки можуть являти собою, але не обмежуватися ними, кукурудзяну солому, макуху, пшеничну солому, відходи деревини, деревні обрізки, макулатуру і муніципальні тверді відходи (MSW). Інший рослинний матеріал може являти собою будь-яку енергетичну сільськогосподарську культуру. Енергетична сільськогосподарська культура може являти собою, але не обмежуватися ними, просо, сорго, цукровий буряк, цукрову тростину, міскантус і тополю.

Полінуклеотидна послідовність, кодуюча перший білок, другий білок, третій білок або будь-який додатковий фермент, може бути функціонально зв'язана з регуляторною послідовністю. У цьому контексті, функціонально зв'язаний означає, що регуляторний елемент передає свою функцію полінуклеотидній послідовності. У випадку регуляторного елемента, що являє собою промотор, промотор здатний контролювати експресію з полінуклеотидної послідовності, коли вони функціонально зв'язані. У випадку регуляторного елемента, що є термінатором, термінатор здатний здійснювати термінацію транскрипції з полінуклеотидної послідовності. Необмежувальні приклади регуляторних елементів представлені нижче.

Щонайменше один з першого білка, другого білка або третього білка може являти собою, але не обмежуватися ними, фермент, вибраний з XynA: бета-1,4-ксиланаза 229B з Dictyoglomus thermophilum (номер доступу Uniprot P77853); XynB: ендо-1,4-бета-ксиланаза з Thermomyces lanuginosus (номер доступу Uniprot O43097); EGA: ендо-бета-1,4-ендоглюканаза з Nasutitermes takasagoensis (номер доступу Uniprot O77044); EGB: ендо-бета-1,4-ендоглюканаза з Acidothermus cellulolyticus (номер доступу Uniprot P54583); AccA: ферулоїлестераза A з Apergillus niger (номер доступу Uniprot O42807); AccB: ферулоїлестераза B з Aspergillus niger (номер доступу Uniprot Q8WZI8); AccA/B: ферулоїлестераза A і ферулоїлестераза B з Aspergillus niger, EGC: ендо-бета-1,4-ендоглюканаза з Rhodothermus marinus (номер доступу Uniprot O33897); P40942: бета-1,4-ксиланаза з Clostridium stercorarium F9 (номер доступу Uniprot P40942); P40943: бета-1,4-ксиланаза з Geobacillus stearothermophilus T-6 (Bacillus stearothermophilus; номер доступу Uniprot P40943); O30700: бета-1,4-ксиланаза з Bacillus sp.

NG-27 (номер доступу Uniprot O30700); CBHA: целобіогідролаза A з Clostridium thermocellum (номер доступу Uniprot O68438); CBHB: целобіогідролаза B (SYT BD22308); або XynE: ксиланаза (EU591743).

Перший білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438] і SEQ ID NO: 11 [P54583]. Перший білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю SEQ ID NO: 12 [BD22308].

Другий білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з другою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438] і SEQ ID NO: 11 [P54583]. Другий білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю SEQ ID NO: 12 [BD22308].

Третій білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 97, 98, 99 або 100 % ідентичністю з третьою референсною послідовністю, вибраною з групи, що складається з SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EUS91743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8[P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438] і SEQ ID NO: 11 [P54583]. Третій білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю SEQ ID NO: 12 [BD22308].

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності може додатково включати першу націлюючу полінуклеотидну послідовність, кодуючу відповідний націлюючий пептид. Для одержаного способами генної інженерії рослинного матеріалу, позбавленого третьої полінуклеотидної послідовності, перша націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для одержаного способами генної інженерії рослинного матеріалу, позбавленого другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, перша націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Кожен відповідний націлюючий пептид може бути незалежно вибраний для кожної з першої, другої або третьої полінуклеотидної послідовності. Націлюючий пептид може бути злитим з першим білком, другим білком або третім білком. Кожен відповідний націлюючий пептид може бути незалежно вибраний з, але не обмежуючись ними, націлюючого на амілопласт сигналу, націлюючого на клітинну стінку пептиду, націлюючого на мітохондрії пептиду, сигналу локалізації в цитозолі, націлюючого на хлоропласти сигналу, націлюючого в ядро пептиду і націлюючого на вакуоль пептиду.

Перший націлюючий полінуклеотид може бути розташований вище першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності. Націлюючий пептид може мати щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, або 100 % ідентичність з однією з SEQ ID NO: 13 [BAASS], послідовності алейрону ячменю SEQ ID NO: 14 [HVAlePS], SEQ ID NO: 15 [PR1a], SEQ ID NO: 16 [послідовність гамма-зеїну xGZein27ss-02] або SEQ ID NO: 17 [Glu B4SP].

Перша націлююча полінуклеотидна послідовність у комбінації з однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом можуть кодувати амінокислотну послідовність із щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 18 [BAASS:P77853], SEQ ID NO: 19 [BAASS:O33897], SEQ ID NO: 20 [HVAlePS:NtEGm] і SEQ ID NO: 21 [BAAS:P77853:S158-30-108-35].

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності, крім того, може включати другу націлюючу полінуклеотидну послідовність, кодуючу карбоксинацілюючий пептид. Для одержаного способами генної інженерії рослинного матеріалу, позбавленого третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для одержаного способами генної інженерії рослинного матеріалу, позбавленого другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Карбоксинацілюючий пептид може бути вибраний з, але не обмежуючись ними, послідовності, що має щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичність з однією з SEQ ID NO: 22 [SEKDEL], укороченої SEQ ID NO: 23 [KDEL] або послідовності, що визначає вакуолярне сортування ячменю SEQ ID NO: 24 [HvVSD-01]. Карбоксинацілюючий пептид може бути злитий із щонайменше одним з першого білка, другого білка або третього білка.

Щонайменше один з першого білка, другого білка або третього білка може бути наданий без націлюючого пептиду для накопичення в цитоплазмі.

Перша націлююча полінуклеотидна послідовність і друга націлююча полінуклеотидна послідовність у комбінації з однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом можуть кодувати амінокислотну послідовність з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 25 [BAASS:AnfaeB:SEKDEL], SEQ ID NO: 26 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 27 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 28 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 29 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 30 [BAASS:O43097:SEKDEL] і SEQ ID NO: 31 [xGZein27ss-02:BD22308:HVVSD-01].

Щонайменше одна з першої, другої або третьої полінуклеотидної послідовності може кодувати "варіант" білка CWD. Амінокислотна послідовність варіанта білка CWD може відрізнятися делеціями, вставками, замінами амінокислотних послідовностей або іншими модифікаціями білка CWD. Варіант білка CWD може зберігати біологічну активність білка CWD. Зберігати біологічну активність, як використовують у рамках винаходу, означає, що варіант має щонайменше 60 % активності білка CWD, з якого він походить. Активність ксиланази можна оцінювати в аналізі з використанням субстрату Xylazyme AX, як розкрито в даному описі в підрозділі прикладу 1 даного опису за назвою "Аналіз ферментів соломи". Активність ендоглюканази можна оцінювати з використанням субстрату Cellazyme, як розкрито в даному описі в підрозділі прикладу 1 даного опису за назвою "Аналіз ферментів соломи". Активність екзоглюканази можна оцінювати з використанням флуоресцентного 4-метилумбеліферил-b-D-лактопіранозиду (4-MU), як описано в Harrison M.D. et al. 2011, "Accumulation of recombinant cellobiohydrolase and endogluconase in the leaves of mature transgenic sugar cane", Plant Biotechnology Journal, 9:884-896, включеній в даний опис як посилання в повному обсязі. Активність ферулоїлестерази можна оцінювати з використанням аналізу з використанням міченого pNP ферулату як субстрату (як описано в Hegde S. et al. 2009, "Single-step synthesis of 4-nitrophenyl ferulate for spectrophotometric assay of feruloyl esterases", Analytical Biochemistry, 387(1):128-129). Представлені вище випробування активності ксиланази, ендоглюканази, екзоглюканази або ферулоїлестерази можна використовувати для визначення того, чи є послідовність з менше ніж 100 % ідентичністю з деградуючою CWD білковою послідовністю, розкритою в даному описі, варіантом деградуючого CWD білка. Амінокислотну послідовність варіантів білка CWD, розкритих у даному описі, можна модифікувати відносно білка CWD, виходячи з подібності в гідрофобності, гідрофільності, розчинності, полярності амінокислотних залишків. Варіанти білка CWD, розкриті в даному описі, можуть відрізнятися після посттрансляційних модифікацій. Посттрансляційна модифікація, що забезпечує відмінність, може являти собою, але не обмежуватися ними, глікозилування, ацетилювання або фосфорилування. Варіант може бути одержаний за допомогою будь-яких засобів. Варіант може бути розроблений за допомогою сайт-направленного мутагенезу або ненаправленого мутагенезу. Для створення мутантів білка CWD, розкритого в даному описі, можна використовувати ПЛР зі зниженою точністю, і будь-який з аналізів, описаних вище, можна використовувати для оцінки того, чи є мутант варіантом.

Варіанти здійснення включають щонайменше один з першого білка, другого білка або третього білка, або їх варіантів, злитий з варіантами щонайменше одного з націлюючого пептиду або карбоксинацілюючого пептиду. Варіанти націлюючого пептиду або

карбоксинацілюючого пептиду будуть націлювати білок, з яким вони злиті, у ту ж область, що і референсна послідовність націлюючого пептиду або карбоксинацілюючого пептиду.

Варіанти інтеїну можуть бути надані в першому білку, другому білку або третьому білку. Варіант інтеїну може видалятися шляхом сплайсингу з білка, з яким він злитий.

Визначення процентної ідентичності двох амінокислотних послідовностей або двох послідовностей нуклеїнових кислот може включати вирівнювання і порівняння амінокислотних залишків або нуклеотидів у відповідних положеннях у двох послідовностях. Якщо всі положення в двох послідовностях зайняті ідентичними амінокислотними залишками або нуклеотидами, тоді послідовності називають ідентичними на 100 %. Процентну ідентичність можна вимірювати за допомогою алгоритму Smith Waterman (Smith T.F., Waterman M.S. 1981, "Identification of Common Molecular Subsequences", J. Mol. Biol. 147:195-197, що включена в даний опис як посилання в повному обсязі).

В одному варіанті здійснення полінуклеотидна послідовність, кодуюча білок, що має менше ніж 100 % ідентичність з цитованою амінокислотною референсною послідовністю, може кодувати варіант білка, що має референсну амінокислотну послідовність. В одному варіанті здійснення білок, що має менше ніж 100 % ідентичність з цитованою референсною амінокислотною послідовністю, може являти собою варіант білка, що має референсну амінокислотну послідовність. В одному варіанті здійснення полінуклеотидна послідовність, кодуюча білок, що має менше ніж 100 % ідентичність з білком, кодованим цитованою референсною послідовністю нуклеїнової кислоти, може кодувати варіант білка, кодованого референсною послідовністю.

Посилаючись на ФІГ. 1, проілюстрований спосіб одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. На ФІГ. 1 представлена схема технологічного процесу для об'єднаної попередньої обробки і гідролізу одержаного способами генної інженерії рослинного матеріалу. Одержаний способами генної інженерії рослинний матеріал або одержаний способами генної інженерії рослинний матеріал, змішаний з іншим рослинним матеріалом, можна додавати через подавальний пристрій 10 у реактор 20 для хімічної попередньої обробки і ферментативного розрідження; тобто способу конверсії твердої лігноцелюлозної біомаси в розріджений стан, придатний для подальшої обробки і гідролізу. У реакторі 20 одержаний способами генної інженерії рослинний матеріал або одержаний способами генної інженерії рослинний матеріал, змішаний з іншим рослинним матеріалом, можна змішувати з перетворюючим у пульпу складом. Перетворюючий у пульпу склад може включати щонайменше одну частину, що має іон, вибраний із групи, що складається з: сульфіту, бісульфіту, сульфату, карбонату, гідроксиду й оксиду. Щонайменше одна частина може додатково включати, але не обмежуватися ними, протиіон, вибраний із групи, що складається з: амонію, натрію, магнію і кальцію. Щонайменше одна частина може являти собою сіль. Сіль може являти собою, але не обмежуватися ними, сульфіт (SO_3^{2-}), бісульфіт (HSO_3^-), оксид (O^{2-}) і гідроксид (OH^-). Сіль може включати протиіон. Протиіон може являти собою, але не обмежуватися ними, натрій (Na^+), кальцій (Ca^{2+}), водень, калій (K^+), магній (Mg^{2+}) і амоній (NH_4^+). Перетворюючий у пульпу склад може включати щонайменше одне з оксиду кальцію (CaO), вапна або гідроксиду кальцію ($Ca(OH)_2$), гашеного вапна.

В одному варіанті здійснення перетворюючий у пульпу склад може включати щонайменше один з бісульфіту амонію і карбонату амонію. Бісульфіт амонію може бути в концентрації від 0,02 до 0,35М, і карбонат амонію може бути в концентрації від 0,025 до 0,25М. Перетворюючий у пульпу склад можна змішувати з одержаним способами генної інженерії рослинним матеріалом або одержаним способами генної інженерії рослинним матеріалом, змішаним з іншим рослинним матеріалом, при оптимальному співвідношенні рідини і твердої речовини в суміші. Суміш може мати співвідношення рідини і твердої речовини, вибране з величини, що менше або дорівнює одному з 10, 9, 8, 7, 6, 5, 4, 3, 2 або 1, або будь-якої іншої величини в діапазоні між будь-якими двома з зазначених вище (включаючи граничні значення). Наприклад, співвідношення рідини і твердої речовини може являти собою величину, меншу ніж будь-яке ціле або нецілe число, вибране з від 3 до 7. Співвідношення рідини і твердої речовини може дорівнювати 10, 9, 8, 7, 6, 5, 4, 3, 2 або 1, або будь-якій величині в діапазоні між будь-якими двома з зазначених вище величин (включаючи граничні значення). Наприклад, співвідношення рідини і твердої речовини може являти собою будь-яке ціле або нецілe число в діапазоні від 3 до 7.

Попередня обробка може включати інкубацію суміші протягом будь-якого періоду часу. Попередня обробка може включати інкубацію суміші протягом аж до 16 годин. Інкубацію можна проводити протягом більш тривалих або більш коротких періодів часу. Попередня обробка

може включати інкубацію суміші протягом періоду часу, що менше або дорівнює одному з 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 або 1 години.

Попередня обробка може включати забезпечення температури суміші від 40 °С до 95 °С. Температура суміші від 40 °С до 95 °С може дозволяти руйнування або видалення частин лігніну в лігноцелюлозному матеріалі в суміші без інактивації гідролітичних ферментів. Попередня обробка може включати забезпечення температури суміші 55 °С, 65 °С, 75 °С, 95 °С, менше 55 °С, менше 65 °С, менше 75 °С, менше 95 °С, менше 100 °С, від 40 °С до 55 °С, від 40 °С до 65 °С, від 40 °С до 75 °С, від 40 °С до 95 °С, від 40 °С до менше ніж 100 °С, від 55 °С до 65 °С, від 55 °С до 75 °С, від 55 °С до 95 °С, від 55 °С до менше ніж 100 °С, від 65 °С до 75 °С, від 65 °С до 95 °С, від 65 °С до менше ніж 100 °С, від 75 °С до 95 °С, від 75 °С до менше ніж 100 °С або від 95 °С до менше ніж 100 °С.

Попередня обробка може включати забезпечення pH суміші в діапазоні від 5,0 до 10. Попередня обробка може включати забезпечення pH суміші в діапазоні від 6,5 до 8,5. Забезпечуване значення pH суміші може складати 5,0, 5,5, 6,0, 7,0, 7,5, 8,0, 9,0, 9,5 або 10, або pH у діапазоні між будь-якими двома з зазначених вище величин pH (включаючи граничні значення). pH суміші в ході попередньої обробки може залежати від типу використовуваної хімічної речовини і/або типу використовуваного рослинного матеріалу. Забезпечення pH суміші може включати додавання модифікуючої pH хімічної речовини. Модифікуюча pH хімічна речовина може являти собою кислоту або луг.

Наприкінці стадії попередньої обробки суміш може включати рослинний матеріал, що частково деградував, і рідку фазу, називану частковим фільтратом, що може включати хімічні речовини з перетворюючого в пульпу складу і низьку концентрацію ферменту або ферментів CWD, що вивільняються з одержаного способами генної інженерії рослинного матеріалу. Спосіб може включати відділення часткового фільтрату від твердих речовин частково деградованого рослинного матеріалу в сепараторі 30. Розділення можна здійснювати різними шляхами. Розділення можна досягати шляхом осадження, фільтрації або центрифугування часткового фільтрату. Спосіб може включати щонайменше одне зі збирання або рециркуляції щонайменше частини часткового фільтрату в множинних раундах попередньої обробки. Після видалення часткового фільтрату концентрація твердих речовин у суміші може збільшуватися і вона може являти собою будь-яке ціле число або нецілe число в діапазоні від 2 % до 15 % (мас./об.) або будь-який діапазон у межах двох цілих величин з діапазону від 2 % до 15 % (мас./об.).

Спосіб може включати промивання попередньо обробленого одержаного способами генної інженерії рослинного матеріалу будь-якою придатною рідиною. Рідина може являти собою деіонізовану воду. Рідину можна видаляти центрифугуванням.

Крім того, спосіб включає очищення за допомогою механічного подрібнювання, яке проводять в очисному пристрої 40, відомим способом, таким як, але не обмежується ними, дефібрилювання, помел або дроблення.

Спосіб може включати перенесення очищеної попередньо обробленої біомаси в ємність 50 для оцукрювання. Гідроліз ферментом CWD, вивільненим з одержаного способами генної інженерії рослинного матеріалу, може відбуватися в ємності 5 для оцукрювання.

Забезпечення умов гідролізу може включати доведення суміші до рівня твердих речовин 2-25 %, до рівня, що дорівнює будь-якому цілому або нецілому числу в межах від 2 % до 25 % твердих речовин (включаючи граничні значення), або до рівня, що дорівнює будь-якому цілому або нецілому числу в діапазоні між двома цілими числами в межах від 2 % до 25 % твердих речовин. Забезпечення умов гідролізу може включати інкубацію суміші протягом періоду часу аж до 144 годин, періоду часу, вибраного з будь-якого цілого або нецілого числа аж до 144 годин, або періоду часу в межах діапазону між будь-якими двома цілими числами від нуля й аж до 144 годин. Забезпечення умов гідролізу може включати забезпечення температури суміші 100 °С або менше, 65 °С або менше, 50 °С або менше, від 48 °С до 50 °С, від 48 °С до 65 °С, від 48 °С до менше ніж 100 °С або від 48 °С до 100 °С. Забезпечення умов гідролізу може включати забезпечення pH у діапазоні від 4,8 до 5,0, pH 4,8, pH 4,9 або pH 5,0. Можна вибирати щонайменше одне з температури, pH або часу обробки, виходячи зі специфічної активності ферменту CWD в одержаному способами генної інженерії рослинному матеріалі.

Якщо одержаний способами генної інженерії рослинний матеріал включає множину ферментів CWD, умови, оптимальні для щонайменше одного з експресії, попередньої обробки або гідролізу кожним з множини ферментів CWD, можуть бути забезпечені послідовно. Умови гідролізу можуть включати забезпечення pH, оптимального для активності одного ферменту, а потім іншого pH, оптимального для активності іншого ферменту. Умови гідролізу можуть включати коректування температури в різні періоди часу для оптимальної активності кожного

ферменту. Наприклад, для ксиланази може вимагатися інші температура або pH, ніж для ендоглюканази.

Спосіб може включати додавання одного або декількох екзогенних ферментів до щонайменше одного з одержаного способами генної інженерії рослинного матеріалу, іншого рослинного матеріалу або суміші. Екзогенні ферменти можна додавати до, у процесі або після попередньої обробки. Екзогенні ферменти можна додавати до, у процесі або після забезпечення умов гідролізу. Екзогенні ферменти можна додавати в ємність 50 для оцукрювання. Один або декілька екзогенних ферментів можуть бути надані в коктейлі ферментів. Коктейль ферментів може включати один або декілька ферментів CWD. CWD, передбачений в одному варіанті здійснення в рамках даного винаходу, може являти собою, але не обмежуватися ними, фермент, що здійснює деградацію лігніну, фермент, що здійснює деградацію целюлози, або фермент, що здійснює деградацію геміцелюлози. Фермент CWD, передбачений в одному варіанті здійснення в рамках даного винаходу, може являти собою, але не обмежуватися ними, фермент, вибраний із глікозидаз, ксиланаз, целюлаз, ендоглюканаз, екзоглюканаз, целобіогідролаз, β-ксилозидаз, ферулоїлестераз і амілаз. Коктейль ферментів може включати целюлазу, виділену з Trichoderma reesii. Коктейль ферментів можна придбавати у постачальника. Коктейль ферментів може являти собою, але не обмежуватися ними, Accellerase™ 1000, Accellerase® 1500 і Accellerase® XY, доступні від Genencor International (Rochester, NY). Коктейль ферментів може являти собою Cellic. Коктейль ферментів може включати різні класи ферментів CWD. Можуть бути забезпечені оптимальні умови для різних класів ферментів CWD. Наприклад, температуру, pH і час обробки для гідролізу можна коректувати в процесі здійснення способу для забезпечення оптимальних умов для різних ферментів у коктейлі. Умови гідролізу можуть включати знижене навантаження зовнішніми ферментами, включеними в коктейль ферментів. Знижене навантаження може включати склади, що мають менше або позбавлені білка або білків CWD, експресованих в одержаному способами генної інженерії рослинному матеріалі. Наприклад, якщо трансгенна рослина експресує ксиланазу і ендоглюканазу, ці ферменти можна усувати з коктейлю ферментів, складеного для гідролізу одержаного способами генної інженерії рослинного матеріалу, що має трансгенну рослину.

Ефективність гідролізу можна оцінювати шляхом вимірювання розчинності рослинного матеріалу. Способи вимірювання розчинності рослинного матеріалу відомі в даній галузі і можуть включати визначення концентрацій моносахаридів і дисахаридів, наприклад за допомогою високоефективної рідинної хроматографії (ВЕРХ). Як розкрито в розділі "Приклади" даного опису, ВЕРХ можна проводити з використанням бінарного насоса Shimadzu LC-20 AD із програмним забезпеченням LC solutions (Shimadzu, Kyoto, Японія) і концентрацію цукру можна визначати з використанням колонки для визначення цукрів Aminex HPX-87P (Bio-Rad Laboratories). Доступні інші способи вимірювання розчинності рослинного матеріалу, наприклад, шляхом визначення зменшення маси, видалення лігніну або деацетилювання в попередньо обробленому рослинному матеріалі.

Крім того, спосіб може включати контактування суміші і/або продуктів гідролізу з ферментуючим організмом з одержанням біохімічного продукту. Після ферментативного гідролізу розчинні цукри можна виділяти і використовувати для продукування біохімічного продукту. Альтернативно в способі можна проводити одночасне оцукрювання і ферментацію розчинних цукрів у біохімічний продукт. Біохімічний продукт може являти собою, але не обмежуватися ними, бутан, бутандіол, бутадієнт, бутанол, ізобутанол, пропан, пропандіол, пропіл, пропанол, ізопропанол, метан, метанол, етанол, фенол, гліцерин, етилен, толуол, етил, бензол, стирол, ксилол, етиленгліколь, оксид етилену, мурашину кислоту, діоксид вуглецю, формальдегід, ацетальдегід, ацетон, вітамін, етан, пентан, гексан, гептан, октан, бензол, оцтову кислоту, сорбіт, арабініт, бурштинову кислоту, фумарову кислоту, яблучну кислоту, фурандикарбонову кислоту, аспарагінову кислоту, глукарову кислоту, глутамінову кислоту, ітаконову кислоту, левулінову кислоту, гідроксибутиролактон, гліцерин, сорбіт, ксиліт, арабініт, глюконову кислоту, молочну кислоту, малонову кислоту, пропіонову кислоту, лимонну кислоту, аконітинову кислоту, ксилонову кислоту, фурфураль, левоглюкозан, аланін, пролін, лізин, серин або треонін (див. Werpy T. and Petersen G., Top Value Added Chemicals From Biomass, Volume 1, Results of Screening for Potential Candidates from Sugars and Synthesis Gas, August 2004, Report, PNNL & NREL, що включена в даний опис як посилання в повному обсязі). Спосіб може включати одночасне оцукрювання і ферментацію розчинних цукрів з одержанням етанолу. Одночасне оцукрювання і ферментація для одержання етанолу можуть включати надання Saccharomyces cerevisiae D5A до, у процесі або після попередньої обробки або забезпечення умов гідролізу.

Конверсію цукрів у бажані біохімічні продукти можна проводити за допомогою будь-якого придатного ферментуючого організму. Ферментуючий організм можна вибирати, виходячи з бажаного біохімічного продукту. Ферментуючим організмом можуть бути дріжджі. Дріжджі можуть являти собою, але не обмежуватися ними, одні з Saccharomyces, Kluyveromyces, Pichia, Yarrowia, Spathaspora або Scheffersomyces ssp. Ферментуючим організмом може бути бактерія. Бактерія може являти собою, але не обмежуватися ними, Zymomonas, Escherichia, Bacillus, Lactobacillus або Clostridium ssp. Ферментуючий організм може являти собою організм дикого типу або модифікований способами генетичної інженерії рекомбінантний організм.

Варіант здійснення включає одержану способами генної інженерії рослину, що включає першу полінуклеотидну послідовність, кодуючу перший білок. Перший білок може являти собою білок CWD. Перший білок може являти собою модифікований інтеїном білок CWD. Перший білок може являти собою будь-який білок, описаний відносно способу продукування розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Перший білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з першою референсною послідовністю, вибраною із групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438] і SEQ ID NO: 11 [P54583]. Перший білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю SEQ ID NO: 12 [BD22308].

Одержана способами генної інженерії рослина, крім того, може включати другу полінуклеотидну послідовність, кодуючу другий білок. Другий білок може являти собою фермент CWD. Другий білок може являти собою модифікований інтеїном білок CWD. Другий білок може являти собою будь-який білок, описаний відносно способу одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Другий білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з другою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU 591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438], SEQ ID NO: 11 [P54583] і SEQ ID NO: 12 [BD22308]. SEQ ID NO, вибрана як друга референсна послідовність, може відрізнятися від SEQ ID NO, вибраної як перша референсна послідовність.

Одержана способами генної інженерії рослина, крім того, може включати третю полінуклеотидну послідовність, кодуючу третій білок. Третій білок може являти собою фермент CWD. Третій білок може являти собою модифікований інтеїном білок CWD. Третій білок може являти собою будь-який білок, описаний відносно способу одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Третій білок може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з третьою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438], SEQ ID NO: 11 [P54583] і SEQ ID NO: 12 [BD22308]. SEQ ID NO, вибрана як третя референсна послідовність, може відрізнятися від SEQ ID NO, вибраної як перша референсна послідовність. SEQ ID NO, вибрана як третя референсна послідовність, може відрізнятися від SEQ ID NO, вибраної як друга референсна послідовність.

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності в одержаній способами генної інженерії рослині може додатково включати першу націлюючу полінуклеотидну послідовність, кодуючу відповідний націлюючий пептид. Для одержаної способами генної інженерії рослини, позбавленої третьої полінуклеотидної послідовності, перша націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для одержаної способами генної інженерії рослини, позбавленої другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, перша націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Відповідний націлюючий пептид може бути незалежно вибраний з, але не обмежуючись ними, націлюючого на амілопласт сигналу,

націлюючого на клітинну стінку пептиду, націлюючого на мітохондрії пептиду, сигналу локалізації в цитозолі, націлюючого на хлоропласти сигналу, націлюючого в ядро пептиду і націлюючого на вакуоль пептиду.

Кожен відповідний націлюючий пептид може бути злитий з відповідним першим білком, другим білком або третім білком. Націлюючий пептид може мати щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичність з однією з SEQ ID NO: 13 [BAASS], SEQ ID NO: 14 [HvAleSP], SEQ ID NO: 1 [PR1a] 5, SEQ ID NO: 16 [xGZein27ss-02] або SEQ ID NO: 17 [GluB4SP].

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності в одержаній способами генної інженерії рослині може додатково включати другу націлюючу полінуклеотидну послідовність, кодуючу карбоксинацілюючий пептид. Для одержаної способами генної інженерії рослини, позбавленої третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для одержаного способами генної інженерії рослини, позбавленої другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Карбоксинацілюючий пептид може бути вибраний з, але не обмежуючись ними, послідовності, що має щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичність з однією з SEQ ID NO: 22 [SEKDEL], укороченої SEQ ID NO: 23 [KDEL] або SEQ ID NO: 24 [послідовність, що визначає вакуолярне сортування ячменю HvVSD-01]. Карбоксинацілюючий пептид може бути злитий щонайменше з одним з першого білка, другого білка або третього білка.

Одержана способами генної інженерії рослина може включати щонайменше одну полінуклеотидну послідовність, кодуючу амінокислотну послідовність, що включає, по суті складається з або складається з послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з послідовністю, вибраною з групи, що складається з: SEQ ID NO: 18 [BAASS:P77853], SEQ ID NO: 19 [BAASS:O33897], SEQ ID NO: 20 [HVAlePS:NtEGm], SEQ ID NO: 21 [BAASS:P77853:S158-30-108-35], SEQ ID NO: 25 [BAASS:AnfaeB:SEKDEL], SEQ ID NO: 26 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 27 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 28 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 29 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 30 [BAASS:O43097:SEKDEL] і SEQ ID NO: 31 [xGZein27ss-02:BD22308:HvVSD-01].

Одержана способами генної інженерії рослина може включати щонайменше одну амінокислотну послідовність, що включає, по суті складається або складається з послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з послідовністю, вибраною з групи, що складається з: SEQ ID NO: 18 [BAASS:P77853], SEQ ID NO: 19 [BAASS:O33897], SEQ ID NO: 20 [HVAlePS:NtEGm], SEQ ID NO: 21 [BAASS:P77853:S158-30-108-35], SEQ ID NO: 25 [BAASS:AnfaeB:SEKDEL], SEQ ID NO: 26 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 27 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 28 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 29 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 30 [BAASS:O43097:SEKDEL] і SEQ ID NO: 31 [xGZein27ss-02:BD22308:HvVSD-01].

Одержана способами генної інженерії рослина може являти собою трансгенну рослину, нащадка трансгенної рослини, низхідного родича трансгенної рослини або будь-яку частину зазначених вище. Одержана способами генної інженерії рослина може включати білок CWD, що не зустрічається в природі в рослині, або ген, кодуючий його. Білок CWD може являти собою модифікований інтеїном білок CWD. Трансгенна рослина може являти собою рослину будь-якого типу. Тип трансгенної рослини може являти собою кукурудзу, цукрову тростину, цукровий буряк, сорго, просо, міскантус, евкаліпт, вербу або тополю. Трансгенну рослину можна створювати відомими способами для експресії ферменту CWD або білка CWD у будь-якій формі. Рослину можна створювати за допомогою опосередковуваної Agrobacterium трансформації з використанням вектора, що включає полінуклеотидні послідовності, кодуючі фермент. Трансгенну рослину можна створювати іншими способами трансформації рослин, наприклад бомбардуванням частинками або прямим захопленням ДНК. Трансгенна рослина може включати будь-яку виділену нуклеїнову кислоту, амінокислотну послідовність, касету експресії або вектор, розкриті в даному описі.

В одному варіанті здійснення передбачена касета експресії, що включає щонайменше одну з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності, які кодують, відповідно, перший білок, другий білок і третій білок. Будь-який один або декілька з першого білка, другого білка або третього білка можуть

являти собою білок CWD. Будь-який один або декілька з першого білка, другого білка або третього білка можуть являти собою модифікований інтеїном білок CWD. Будь-який один або декілька з першого білка, другого білка або третього білка можуть являти собою один з білків, описаних відносно способу продукування розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу або одержаних способами генної інженерії рослин. Будь-який один або декілька з першого білка, другого білка або третього білка можуть являти собою ксиланазу, ендоглюканазу, екзоглюканазу, ферулоїлестеразу, модифіковану інтеїном ксиланазу, модифіковану інтеїном ендоглюканазу, модифіковану інтеїном екзоглюканазу або модифіковану інтеїном ферулоїлестеразу. Білок, вибраний як другий білок, може відрізнятися від білка, вибраного як перший білок. Білок, вибраний як третій білок, може відрізнятися від білка, вибраного як перший білок. Білок, вибраний як третій білок, може відрізнятися від білка, вибраного як перший білок. Білок, вибраний як третій білок, може відрізнятися від білка, вибраного як другий білок.

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності, кодуючої білок CWD у касеті експресії, може бути модифікована вставкою послідовності нуклеїнової кислоти, кодуючої інтеїн. Модифікований інтеїном полінуклеотид може кодувати модифікований інтеїном білок з модифікованою функцією. Модифікована функція може являти собою інактивацію білка CWD, поки інтеїн залишається злитим з білком CWD або в білку CWD. Можна індукувати сплайсинг інтеїну в модифікованому інтеїном білку у форму немодифікованого інтеїном білка. Умови індукції сплайсингу можуть являти собою, але не обмежуватися ними, забезпечення визначеної температури. Забезпечувана температура може являти собою температуру, яку забезпечують в умовах щонайменше одного з попередньої обробки або гідролізу, описаних відносно способу продукування розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу. Умови індукції можуть являти собою будь-які інші умови індукції, що відповідають вибраному інтеїну. Модифікований інтеїном білок може являти собою iXynA: тобто модифіковану інтеїном XynA. Модифікований інтеїном білок може являти собою модифіковану інтеїном P77853. Модифікований інтеїном білок може являти собою P77853:T134-100-101 або P77853:S158-30-108-35.

Один або декілька з першого білка, другого білка або третього білка в касеті експресії може включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1 [WT P77853], SEQ ID NO: 2 [AnfaeA], SEQ ID NO: 3 [AnfaeB], SEQ ID NO: 4 [NtEGm], SEQ ID NO: 5 [EU 591743], SEQ ID NO: 6 [O43097], SEQ ID NO: 7 [P77853:T134-100-101], SEQ ID NO: 8 [P77853:S158-30-108-35], SEQ ID NO: 9 [O33897], SEQ ID NO: 10 [O68438] і SEQ ID NO: 11 [P54583]. Один або декілька з першого білка, другого білка або третього білка можуть включати, по суті складатися або складатися з амінокислотної послідовності з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 98, 99 або 100 % ідентичністю з референсною послідовністю SEQ ID NO: 12 [BD22308].

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності в касеті експресії може додатково включати першу націлюючу полінуклеотидну послідовність, кодуючу відповідний націлюючий пептид. Для експресуючої конструкції, позбавленої третьої полінуклеотидної послідовність, перша націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для експресуючої конструкції, позбавленої другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, перша націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Кожен відповідний націлюючий пептид може бути незалежно вибраний для кожної з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності. Націлюючий пептид може бути злитий з першим білком, другим білком або третім білком. Кожен відповідний націлюючий пептид може бути незалежно вибраний з, але не обмежуючись ними, націлюючого на амілопласт сигналу, націлюючого на клітинну стінку пептиду, націлюючого на мітохондрії пептиду, сигналу локалізації в цитозолі, націлюючого на хлоропласти сигналу, націлюючого в ядро пептиду і націлюючого на вакуоль пептиду. Перший націлюючий полінуклеотид може бути розташований вище першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності. Націлюючий пептид може мати щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 97, 98, 99 або 100 % ідентичність одній з BAASS (SEQ ID NO: 13), послідовності алейрону ячменю HVAlePS (SEQ ID NO: 14), PR1a (SEQ ID NO: 15), послідовності гамма-зеїну xGZein27ss-02 (SEQ ID NO: 16) або GluB4SP (SEQ ID NO:

17). Перша націлююча полінуклеотидна послідовність у комбінації з однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом можуть кодувати амінокислотну послідовність з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 18 [BAASS:P77853], SEQ ID NO: 19 [BAASS:O33897], SEQ ID NO: 20 [HVAlePS:NtEGm] і SEQ ID NO: 21 [BAAS:P77853:S158-30-108-35].

Щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності в касеті експресії може, крім того, включати другу націлюючу полінуклеотидну послідовність, кодуючу карбоксинацілюючий пептид. Для експресуючої конструкції, позбавленої третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в щонайменше одну з першої полінуклеотидної послідовності або другої полінуклеотидної послідовності. Для експресуючої конструкції, позбавленої другої полінуклеотидної послідовності і третьої полінуклеотидної послідовності, друга націлююча полінуклеотидна послідовність може бути включена в першу полінуклеотидну послідовність. Карбоксинацілюючий пептид може бути вибраний з, але не обмежуючись ними, послідовності, що має щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичність з однією з SEKDEL (SEQ ID NO: 22), укороченої KDEL (SEQ ID NO: 23) або послідовності, що визначає вакуолярне сортування ячменю HvVSD-01 (SEQ ID NO: 24). Карбоксинацілюючий пептид може бути злитий із щонайменше одним з першого білка, другого білка або третього білка.

Касета експресії може бути організована так, щоб щонайменше один з першого білка, другого білка або третього білка був наданий без націлюючого пептиду для накопичення в цитоплазмі.

У касеті експресії перша націлююча полінуклеотидна послідовність і друга націлююча полінуклеотидна послідовність у комбінації з однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом можуть кодувати амінокислотну послідовність з щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 25 [BAASS: AnfaeB: SEKDEL], SEQ ID NO: 26 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 27 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 28 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 29 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 30 [BAASS:O43097:SEKDEL] і SEQ ID NO: 31 [xGZein27ss-02:BD22308:HvVSD-01].

Варіанти здійснення включають касету експресії, кодуючу щонайменше один з першого білка, другого білка або третього білка або їх варіанти, злиті з варіантами щонайменше одного з націлюючого пептиду або карбоксинацілюючого пептиду.

В одному варіанті здійснення полінуклеотидна послідовність, яка кодує білок у касеті експресії і має менше ніж 100 % ідентичність з цитованою референсною амінокислотною послідовністю, може кодувати варіант білка, що має референсну амінокислотну послідовність. В одному варіанті здійснення білок, що має менше ніж 100 % ідентичність з цитованою референсною амінокислотною послідовністю, може являти собою варіант білка, що має референсну амінокислотну послідовність. В одному варіанті здійснення полінуклеотидна послідовність, кодуюча білок, що має менше ніж 100 % ідентичність з білком, кодованим цитованою референсною послідовністю нуклеїнової кислоти, може кодувати варіант білка, кодований референсною послідовністю.

У касеті експресії щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності може бути здатна гібридизуватися з референсною послідовністю, кодуючою білок CWD або модифікований інтеїном білок CWD в одних з умов низької, помірної або високої жорсткості. Щонайменше один з першого, другого або третього полінуклеотидів може бути здатний гібридизуватися в одних з умов низької, помірної або високої жорсткості з нуклеїновою кислотою, що складається з референсної послідовності, вибраної з групи, що складається з: SEQ ID NO: 32 [WT P77853], SEQ ID NO: 33 [AnfaeA], SEQ ID NO: 34 [AnfaeB], SEQ ID NO: 35 [NtEGm], SEQ ID NO: 36 [EU591743], SEQ ID NO: 37 [O43097], SEQ ID NO: 38 [P77853:T134-100-101], SEQ ID NO: 39 [P77853:S158-30-108-35], SEQ ID NO: 40 [O33897], SEQ ID NO: 41 [O68438], SEQ ID NO: 42 [P54583] і SEQ ID NO: 43 [BD22308]. Касета експресії може включати полінуклеотидну послідовність, здатну гібридизуватися в одних з умов низької, помірної або високої жорсткості з нуклеїновою кислотою, що складається з референсної послідовності, вибраної з групи послідовностей, що складається з: SEQ ID NO: 52 [BAASS:P77853], SEQ ID NO: 53 [BAASS:O33897], SEQ ID NO: 54 [HVAlePS:NtEGm], SEQ ID NO: 55 [BAASS:P77853:S158-30-108-

35], SEQ ID NO: 56 [BAASS:AnfaeB:SEKDEL], SEQ ID NO: 57 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 58 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 59 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 60 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 61 [BAASS:O43097:SEKDEL] і SEQ ID NO: 62 [xGZein27ss-02:BD22308:HvVSD-01].

Касета експресії в рамках даного винаходу може включати регуляторний елемент.

В одному варіанті здійснення передбачений вектор, що включає будь-яку виділену нуклеїнову кислоту, полінуклеотидну послідовність або касету експресії, розкриті в даному описі. Варіанти здійснення в рамках даного винаходу включають плазміду, хромосому, мітохондріальну ДНК, пластидну ДНК, вірус або фрагмент нуклеїнової кислоти, що має щонайменше одну касету експресії, розкриту в даному описі, включену у них. Один з варіантів здійснення стосується вектора для експресії білків CWD у рослині. Один варіант здійснення стосується вектора для трансформації рослин. Вектор для трансформації рослин може являти собою, але не обмежуватися ними, вектор Т-ДНК, бінарний вектор або вектор-коінтеграт. Вектор для трансформації може включати будь-яку виділену нуклеїнову кислоту, полінуклеотидну послідовність або касету експресії, розкриті в даному описі.

Один з варіантів здійснення включає вектор експресії, що включає полінуклеотидну послідовність, здатну гібридизуватися в одних умовах з умов низької, помірної або високої жорсткості з нуклеїновою кислотою, яка складається з послідовності, що включає, по суті складається з або складається з послідовності, вибраної з групи, що складається з: SEQ ID NO: 63 [pAG 2015], SEQ ID NO: 64 [pAG2048], SEQ ID NO: 65 [pAG2049], SEQ ID NO: 66 [pAG2063], SEQ ID NO: 67 [pAG2069], SEQ ID NO: 68 [pAG2091], SEQ ID NO: 69 [pAG2092], SEQ ID NO: 70 [pAG2096], SEQ ID NO: 71 [pAG2201], SEQ ID NO: 72 [pAG2229], SEQ ID NO: 73 [pAG2233], SEQ ID NO: 74 [pAG2234], SEQ ID NO: 75 [pAG2242], SEQ ID NO: 76 [pAG2252], SEQ ID NO: 77 [pAG2253], SEQ ID NO: 78 [pAG2309], SEQ ID NO: 79 [pAG2310], SEQ ID NO: 80 [pAG2339], SEQ ID NO: 81 [pAG2342], SEQ ID NO: 82 [pAG2345] і SEQ ID NO: 83 [pAG2349].

Один з варіантів здійснення включає вектор експресії, який включає полінуклеотидну послідовність, що має щонайменше 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 або 100 % ідентичність з послідовністю, вибраною з групи, що складається з: SEQ ID NO: 63 [pAG2015], SEQ ID NO: 64 [pAG2048], SEQ ID NO: 65 [pAG2049], SEQ ID NO: 66 [pAG2063], SEQ ID NO: 67 [pAG2069], SEQ ID NO: 68 [pAG2091], SEQ ID NO: 69 [pAG2092], SEQ ID NO: 70 [pAG2096], SEQ ID NO: 71 [pAG2201], SEQ ID NO: 72 [pAG2229], SEQ ID NO: 73 [pAG2233], SEQ ID NO: 74 [pAG2234], SEQ ID NO: 75 [pAG 2242], SEQ ID NO: 76 [pAG2252], SEQ ID NO: 77 [pAG2253], SEQ ID NO: 78 [pAG2309], SEQ ID NO: 79 [pAG2310], SEQ ID NO: 80 [pAG2339], SEQ ID NO: 81 [pAG2342], SEQ ID NO: 82 [pAG2345] і SEQ ID NO: 83 [pAG2349].

Способи гібридизації й умови жорсткості відомі в даній галузі й описані в наступних книгах: Molecular Cloning, T. Maniatis, E.F. Fritsch and J. Sambrook, Cold Spring Harbor Laboratory, 1982, і Current Protocols in Molecular Biology, F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Siedman, J.A. Smith, K. Struhl, Volume 1, John Wiley & Sons, 2000, що включені в даний опис як посилання в повному обсязі.

Умови помірної жорсткості можуть бути наступними: фільтри, навантажені зразками ДНК, попередньо обробляють протягом 2-4 годин при 68 °С у розчині, що містить 6× цитратний сольовий буфер (SSC; Amresco, Inc., Solon, OH), 0,5 % додецилсульфат натрію (SDS; Amresco, Inc., Solon, OH), 5× розчин Денхардта (Amresco, Inc., Solon, OH) і денатуровану сперму лосося (Invitrogen Life Technologies, Inc. Carlsbad, CA). Гібридизацію проводять у тому ж розчині з наступними модифікаціями: 0,01М EDTA (Amresco, Inc., Solon, OH), 100 мкг/мл ДНК сперми лосося і 5-20 × 10^6 cpm ^{32}P-мічених або флуоресцентно мічених зондів. Фільтри інкубують у суміші для гібридизації протягом 16-20 годин, а потім промивають протягом 15 хвилин у розчині, що містить 2×SSC і 0,1 % SDS. Промивальний розчин заміняють для другого промивання розчином, що містить 0,1×SSC і 0,5 % SDS, і інкубують протягом додаткових 2 годин при температурі від 20 °С до 29 °С нижче Tm (температура плавлення в°С). Tm $Tm=81,5+16,61Log_{10}([Na^+]/(1,0+0,7[Na^+]))+0,41(\%[G+C])-(500/n)-P-F$. $[Na^+]$ = молярна концентрація іонів натрію. %[G+C] = відсоток основ G+C у послідовності ДНК. N = довжина послідовності ДНК у основах. P = поправка температури на % не відповідних пар основ (~1 °С на 1 % невідповідність). F = поправка на концентрацію формаміду (=0,63 °С на 1 % формамід). Фільтри експонують для проявлення в пристрої для візуалізації або за допомогою радіоавтографії. Умови низької жорсткості належать до умов гібридизації при низьких температурах, наприклад від 37 °С до 60 °С, і другого промивання з більш високою $[Na^+]$ (аж до 0,825М) і при температурі від 40 °С до 48 °С нижче Tm. Висока жорсткість належить до умов гібридизації при високих температурах, наприклад більше 68 °С, і другого промивання з $[Na^+]$ = від 0,0165 до 0,0330М при температурі від 5 °С до 10 °С нижче Tm.

Один варіант здійснення стосується виділеної послідовності нуклеїнової кислоти, яка має послідовність, що гібридизується в одних з умов низької, помірної або високої жорсткості з нуклеїновою кислотою, що складається з послідовності, вибраної з SEQ ID NO: 32 [WT P77853], SEQ ID NO: 33 [AnfaeA], SEQ ID NO: 34 [AnfaeB], SEQ ID NO: 35 [NtEGm], SEQ ID NO: 36 [EU591743], SEQ ID NO: 37 [O43097], SEQ ID NO: 38 [P77853:T134-100-101] і SEQ ID NO: 39 [P77853:S158-30-108-35], SEQ ID NO: 40 [O33897], SEQ ID NO: 41 [O68438], SEQ ID NO: 42 [P54583] і SEQ ID NO: 43 [BD22308], SEQ ID NO: 52 [BAASS:P77853], SEQ ID NO: 53 [BAASS:O33897], SEQ ID NO: 54 [HVAlePS:NtEGm], SEQ ID NO: 55 [BAASS:P77853:S158-30-108-35], SEQ ID NO: 56 [BAASS: AnfaeB: SEKDEL], SEQ ID NO: 57 [BAASS:AnfaeA:SEKDEL], SEQ ID NO: 58 [PR1a:NtEGm:SEKDEL], SEQ ID NO: 59 [BAASS:P77853:T134-100-101:SEKDEL], SEQ ID NO: 60 [HvAleSP:NtEGm:SEKDEL], SEQ ID NO: 61 [BAASS:O43097:SEKDEL] і SEQ ID NO: 62 [xGZein27ss-02:BD22308:HvVSD-01], SEQ ID NO: 63 [pAG2015], SEQ ID NO: 64 [pAG2048], SEQ ID NO: 65 [pAG2049], SEQ ID NO: 66 [pAG2063], SEQ ID NO: 67 [pAG2069], SEQ ID NO: 68 [pAG2091], SEQ ID NO: 69 [pAG2092], SEQ ID NO: 70 [pAG2096], SEQ ID NO: 71 [pAG2201], SEQ ID NO: 72 [pAG2229], SEQ ID NO: 73 [pAG2233], SEQ ID NO: 74 [pAG2234], SEQ ID NO: 75 [pAG 2242], SEQ ID NO: 76 [pAG2252], SEQ ID NO: 77 [pAG2253], SEQ ID NO: 78 [pAG2309], SEQ ID NO: 79 [pAG2310], SEQ ID NO: 80 [pAG2339], SEQ ID NO: 81 [pAG2342], SEQ ID NO: 82 [pAG2345] і SEQ ID NO: 83 [pAG2349], або послідовності, комплементарної їй.

Один варіант здійснення стосується фрагмента будь-якої з описаних вище виділених нуклеїнових кислот. Фрагмент може являти собою зонд або праймер для гібридизації. Зонд або праймер може мати будь-яку довжину. Зонд або праймер може мати довжину 6, 10, 15, 20, 25, 30, 35, 40, 45 або 50 нуклеотидів, або він може мати довжину в діапазоні між будь-якими з двох зазначених вище довжин (включаючи граничні значення). Фрагмент може мати довжину, меншу ніж повна довжина, і/або він може включати заміни або делеції в порівнянні з цитованою референсною послідовністю. Фрагмент може являти собою варіант цитованої референсної послідовності. Пептид, кодований фрагментом, може мати довжину, меншу ніж повна довжина, і/або він може включати заміни або делеції в порівнянні з амінокислотною послідовністю, кодованою цитованою референсною послідовністю. Пептид з довжиною, меншою ніж повна довжина, може являти собою варіант амінокислотної послідовності, кодований цитованою референсною послідовністю.

Касету експресії можна одержувати рекомбінантним шляхом відомими способами. Касета експресії може включати серію певних елементів нуклеїнової кислоти, що дозволяють транскрипцію конкретної нуклеїнової кислоти в клітині рослини або тканині рослини. Касета експресії може включати полінуклеотидну послідовність, кодуючу білок. Білок може являти собою фермент CWD або модифікований інтеїном фермент CWD. Фермент CWD може бути вибраний зі списку ферментів CWD, що складається з: ксиланаз, ендоглюканаз, екзоглюканаз, ксилозидаз, глюкозидаз і ферулоїлестераз.

Полінуклеотидна послідовність у касеті експресії, виділеній нуклеїновій кислоті, векторі або будь-якій іншій конструкції ДНК, розкритій в даному описі або використовуваній в способі, наданому в даному описі, може бути функціонально зв'язана з одним або декількома регуляторними елементами. Включений регуляторний елемент може являти собою промотор. Промотор може являти собою конститутивний промотор, що забезпечує транскрипцію полінуклеотидних послідовностей всюди в рослині в більшості клітин, тканин і органів і протягом багатьох, але не обов'язково всіх, стадій розвитку. Промотор може являти собою індуцибельний промотор, що ініціює транскрипцію полінуклеотидних послідовностей, тільки коли на нього впливає конкретний хімічний стимул або стимул навколишнього середовища. Промотор може бути специфічним для конкретної стадії розвитку, органа або тканини. Тканиноспецифічний промотор може бути здатний ініціювати транскрипцію в конкретній тканині рослини. Рослинна тканина, на яку може бути націлений тканиноспецифічний промотор, може являти собою, але не обмежуватися ними, стебло, листи, трихоми, пиляки або насіння. Конститутивний промотор, розкритий у даному описі, може являти собою промотор убіквітину 3 рису (OsUbi3P) або промотор актину 1 рису. Можна використовувати інші відомі конститутивні промотори і вони включають, але не обмежуються ними, промотор вірусу мозаїки цвітної капусти (CAMV) 35S, промотор вірусу жовтої кучерявості листів цеструму (CMP) або коротку версію CMP (CMPS), промотор малої субодиниці Rubisco (рибулозо-1,5-біфосфаткарбоксилази/оксигенази) і промотор убіквітину кукурудзи. Тканиноспецифічний промотор може включати специфічний для насіння промотор. Специфічний для насіння промотор може являти собою, але не обмежуватися ними, промотор GluB4 рису або промотор зеїну кукурудзи. Іншим регуляторним елементом, що може бути наданий, є послідовність термінатора, яка термінує транскрипцію. Послідовність термінатора може бути включена на 3'-кінці елемента транскрипції касети

експресії. Термінатор може походити з різних генів рослин. Термінатор може являти собою послідовність термінатора з генів нопалінсинтази або октопінсинтази Agrobacterium tumefaciens.

Вектори, що включають касету експресії, розкриту в даному описі, також можуть включати додаткові генетичні елементи як ділянки множинного клонування для полегшення молекулярного клонування і селективні маркери для полегшення селекції. Селективний маркер, що може бути включений у вектор, може являти собою ген фосфоманозоізомерази (PMI) з Escherichia coli, що надає трансформованій клітині здатність утилізувати манозу для росту. Селективні маркери, що можуть бути включені у вектор, включають, але не обмежуються ними, ген неоміцинфосфотрансферази (npt), що надає стійкість до канаміцину, ген гігроміцинфосфотрансферази (hpt), що надає стійкість до гігроміцину, і ген енолпірувілшикімат-3-фосфатсинтази, що надає стійкість до гліфосату.

В одному варіанті здійснення вектор можна конструювати так, щоб він включав полінуклеотидні послідовності, кодуючі множину ферментів CWD. Вектор, представлений у даному описі, може додатково включати полінуклеотидну послідовність, призначену для пригнічення гена або генів у рослині.

Вектор експресії можна вводити в придатні клітини-хазяїни, тканини, органи і/або організми. Придатними хазяїнами можуть бути двосім'ядольні (дводольні) або односім'ядольні (однодольні) рослини.

Додаткові варіанти здійснення в рамках даного винаходу можуть бути одержані шляхом доповнення одного з варіантів здійснення одним або декількома елементами з будь-якого одного або декількох інших варіантів здійснення, розкритих у даному описі, і/або шляхом заміни одного або декількох елементів з одного варіанта здійснення одним або декількома елементами з одного або декількох інших варіантів здійснення, розкритих у даному описі. Додаткові варіанти здійснення в рамках даного винаходу можуть бути описані, посилаючись на будь-який один із прикладених пунктів формули винаходу після п. 1, і шляхом прочитання вибраного пункту формули винаходу так, щоб він залежав від будь-якого одного або декількох попередніх пунктів формули винаходу.

ПРИКЛАДИ

Наступні необмежувальні приклади надані для ілюстрації конкретних варіантів здійснення. Варіанти здійснення можуть бути доповнені однією або декількома деталями з одного або декількох прикладів, наведених нижче, і/або один або декілька елементів з варіанта здійснення можуть бути замінені одним або декількома деталями з одного або декількох прикладів, наведених нижче.

Приклад 1. Матеріали і способи

Вектори

Конструкція вектора, розкритого в даному описі, основана на проміжній плазміді pSB11, доступній від Japan Tobacco і описаній в міжнародних заявках № PCT/US10/55746, поданій 5 листопада 2010 року, PCT/US10/55669, поданій 5 листопада 2010 року, і PCT/US10/55751, поданій 5 листопада 2010, що включені в даний опис як посилання в повному обсязі. У короткому викладі, плазміду pSB11, використовувану для клонування, кон'югують з акцепторним вектором pSB1, обеззброєною Ti-плазмідою, шляхом гомологічної рекомбінації з використанням ділянок cos і ori, присутніх як у pSB11, так і в pSB1. Вбудований вектор містить гени вірулентності, такі як virB, virC і virG, необхідні для перенесення Т-ДНК, і його можна використовувати для трансформації рослин. Базовий вектор для трансформації включає касету експресії, що містить ген man A, кодуючий PMI, під контролем промотору CMPS, пізніше замінюваного промотором OsUbi3P. Цей базовий вектор використовували для одержання векторів, наведених нижче, які використовували для трансформації рослин і експресії ферментів, деградуючих клітинну стінку, у рослинах:

1) pAG2015(SEQ ID NO: 63): OsUbi3P:P77853;

2) pAG2048 (SEQ ID NO: 64): OsUbi3P:HvAleSP:NtEGm між промотором Ubi3 рису, злитим з вакуолею;

3) pAG2049 (SEQ ID NO: 65): OsUbi3P:HvAleSP:NtEGm:SEKDEL;

4) pAG2063 (SEQ ID NO: 66): OsUbi3P:BAASS:O43097:SEKDEL;

5) pAG2069 (SEQ ID NO: 67): OsUbi3P:O68438;

6) pAG2091 (SEQ ID NO: 68): OsUbi3P:BAASS:AnfaeA:SEKDEL+OsUbi3P:BAASS:P77853;

7) pAG2092 (SEQ ID NO: 69): OsUbi3P:BAASS:AnfaeB:SEKDEL+OsUbi3P:BAASS:P77853;

8) pAG2096 (SEQ ID NO: 70): OsUbi3P:BAASS:AnfaeA:SEKDEL+OsUbi3P:BAASS:AnfaeB:SEKDEL+OsUbi3P:BAASS:P77853;

9) pAG2201 (SEQ ID NO: 71): OsUbi3P:ZmUBQm:P77853;

10) pAG2229 (SEQ ID NO: 72): OsUbi3P:BAASS:P77853:T134-100-101:SEKDEL (модифікована інтеїном ксиланаза);
11) pAG2233 (SEQ ID NO: 73): OsUbi3P:P77853:S158-30-108-35;
12) pAG2234(SEQ ID NO: 74): OsUbi3P:BAASS:P77853:S158-30-108-35;
13) pAG2242 (SEQ ID NO: 75): OsUbi3P:PR1aSP:NtEGm:SEKDEL+OsUbi3P:ZmUBQm:P77853;
14) pAG2252 (SEQ ID NO: 76): OsUbi3P:O33897 (ендоглюканаза);
15) pAG2253 (SEQ ID NO: 77): OsUbi3P:BAASS:O33897;
16) pAG2309 (SEQ ID NO: 78): OsUbi3P:HvAleSP:NtEGm+OsUbi3P:P77853;
17) pAG2310 (SEQ ID NO: 79): OsUbi3P:EU591743 (ксиланаза);
18) pAG2339 (SEQ ID NO: 80): OsUbi3P:O68438+OsUbi3P:BAASS:O33897+OsUbi3P:EU591743;
19) pAG2342 (SEQ ID NO: 81): OsUbi3P:HvAleSP:NtEGm:SEKDEL+OsUbi3P:P77853;
20) pAG2345 (SEQ ID NO: 82): OsUbi3P:O68438+OsUbi3P:HvAleSP:NtEGm:SEKDEL+OsUbi3P:BAASS:O43097:SEKDEL;
21) pAG2349 (SEQ ID NO: 83): ZmUbilP:ZmKozak:xGZein27ss-02:BD22308:HvVSD-01+OsUbi3P:HvAleSP:NtEGm:SEKDEL+OsUbi3P:BAASS:O43097:SEKDEL;
22) pAG2042 (SEQ ID NO: 84): P54583 (ендоглюканаза EGB).

Одержання трансгенних рослин кукурудзи

Способи трансформації кукурудзи і проса описані в міжнародних заявках № PCT/US10/55746, поданій 5 листопада 2010 року, PCT/US10/55669, поданій 5 листопада 2010 року, PCT/US10/55751, поданій 5 листопада 2010 року, і Gray et al. 2011, Plant Biotech. J. 9:1100, усі з який включені в даний опис як посилання в повному обсязі. У короткому викладі, ембріогенний калюс з кукурудзи дикого типу AxB інокулювали клітинами Agrobacterium LBA4404, що містять відповідну плазміду для трансформації. Опосередковувану Agrobacterium трансформацію незрілих ембріонів кукурудзи проводили, як описано раніше (Negrotto D. et al. 2000, Plant Cell. Rep. 19:798; Ishida Y. et al. 1996, Nat. Biotech. 14:745). Касети експресії для генів ферментів клонували в ділянки KpnI-EcoRI вектора pAG2004 (SEQ ID NO: 85) з одержанням проміжного вектора, здатного рекомбінувати з вектором pSB1 у схрещуванні з трьома батьківськими організмами в Agrobacterium tumefaciens штаму LBA4404 з використанням методик, описаних раніше (Ishida Y. et al. 1996, Nat. Biotech. 14:745; Hiei Y. et al. 1994, Plant J. 6:271; Hiei Y. and Komari T. 2006, Plant Cell Tissue Organ Cult. 85:27; Komari T. et al. 1996, Plant J. 10:165). Вихідні рослини кукурудзи (Zea mays культиварів HiII, A188 або B73) вирощували в теплиці при 16 годинах денного світла при 28 °C. Незрілі зиготичні ембріони виділяли з зерен і інокулювали розчином Agrobacterium, які містили гени, що представляють інтерес. Після інокуляції незрілі ембріони вирощували в процесі культивування тканин протягом 10-12 тижнів. З добре розвинених паростків з листами і коренями брали зразки для аналізу способом ПЛР для ідентифікації трансгенних рослин, які містять гени, що представляють інтерес. Позитивні по ПЛР і укорінені рослини промивали водою, щоб змити агарове середовище, і пересаджували в ґрунт і вирощували в теплиці для одержання насіння і соломи.

Конкретні трансгенні рослини позначені в даному описі за допомогою позначення ферменту (наприклад, "P77853", "P40942", "O30700", "NtEGm" і т. д.) або трансгенного контролю (наприклад, "TGC" і т. д.), з наступним числом у тисячах, що позначає плазміду, використану для одержання трансгенної рослини (наприклад, "2014", "2015", "2229", "2092" і т. д.). Іноді додані додаткові символи, але позначення i) ферменту або контролю і ii) плазміди зрозуміло в контексті. Згадувані плазміди позначають як pAGXXXX. Наприклад, позначення "2229", "2252", "2253", "2092", "2096" або "2042" у назві трансгенної рослини означає, що трансгенна рослина була одержана шляхом трансформації за допомогою "pAG2229", "pAG2252", "pAG2253", "pAG2092", "pAG2096" або "pAG2042", відповідно. Може бути наведене посилання на включені послідовності, позначені назвами плазмід, для визначення послідовностей, використаних для одержання конкретної трансгенної рослини.

Для одержання трансгенних рослин проса, насіння з Panicum virgatum, cv. Alamo використовували для ініціації ліній ембріогенного калюсу, надалі використаних для трансформації з використанням Agrobacterium LBA4404, що містить плазміду pSB1. Присутність гена, що представляє інтерес, підтверджували способом ПЛР з використанням специфічних для гена праймерів.

Наступні трансгенні рослини, експресуючі фермент CWD або ферменти CWD, і контрольні рослини використовували для об'єднаної попередньої обробки і гідролізу:

1) рослина кукурудзи дикого типу, використана як негативні контролі (AxB; BxA);

2) рослини кукурудзи, трансформовані порожнім вектором, використані як негативні контролі (TGC.4000.12; TGC.4000.11; TGC.2004.8.02; TGC.2004.8.04; TGC.2243.01);

3) трансгенна рослина кукурудзи XynA.2015.05, одержана шляхом трансформації pAG2015 і експресуюча ксиланазу XynA (P77853);

4) трансгенна рослина кукурудзи другого покоління XynA.2015.5T1, одержана шляхом трансформації pAG2015 і експресуюча ксиланазу XynA (P77853);

5) трансгенна рослина кукурудзи XynB.2063.17, одержана шляхом трансформації pAG2063 і експресуюча ксиланазу XynB (O43097);

6) трансгенні рослини кукурудзи EGA.2049.02 і EGA.2049.10, одержані шляхом трансформації pAG2049 і експресуючі ендоглюканазу EGA (NtEG);

7) трансгенна рослина кукурудзи EGB.2042.03, одержана шляхом трансформації pAG2042 і експресуюча ендоглюканазу EGB (P54583);

8) трансгенна рослина кукурудзи EGC.2253.4b, одержана шляхом трансформації pAG2253 і експресуюча ендоглюканазу EGC (O33897);

9) трансгенна рослина кукурудзи EGA/XynA.2242.09, одержана шляхом трансформації pAG2242 і експресуюча ендоглюканазу EGA (NtEG) і ксиланазу XynA (P77853);

10) трансгенна рослина кукурудзи другого покоління від рослини 9, вище, називана EGA/XynA.2242.09.16T1 і експресуюча ендоглюканазу EGA (NtEG) і ксиланазу XynA (P77853);

11) трансгенна рослина кукурудзи XynA/AccB.2092.103, одержана шляхом трансформації pAG2092 і експресуюча ксиланазу XynA (P77853) і ферулоїлестеразу B з Aspergillus niger;

12) трансгенні рослини кукурудзи XynA/AccA/B.2096.01 і XynA/AccA/B.2096.05, одержані шляхом трансформації pAG2096 і експресуючі ксиланазу XynA (P77853), ферулоїлестеразу A з Aspergillus niger і ферулоїлестеразу B з Aspergillus niger;

13) трансгенна рослина кукурудзи CBHA.2069.3.17, одержана шляхом трансформації pAG2069 і експресуюча екзоглюканазу CBH (O68438);

14) трансгенні рослини проса XynA.pv2015.3c і XynA.pv2015.4c, одержані шляхом трансформації pAG2015 і експресуючі ксиланазу XynA (P77853);

15) трансгенна рослина кукурудзи iXynA.2229.110, одержана шляхом трансформації pAG2229 і експресуюча модифіковану інтеїном ксиланазу XynA (P77853);

16) трансгенні рослини кукурудзи XynA/EGA.2309.54 і XynA/EGA.2309.107, одержані шляхом трансформації pAG2309 і експресуючі XynA (P77853), ендоглюканазу EGA (NtEGm);

17) трансгенна рослина кукурудзи XynA/EGA.2342.105, одержана шляхом трансформації pAG2342 і експресуюча XynA (P77853) і EGA (NtEGm);

18) трансгенні рослини кукурудзи XynE/EGC/CBHA.2339.03, XynE/EGC/CBHA.2339.04 і XynE/EGC/CBHA.2339.05, одержані шляхом трансформації pAG2339 і експресуючі XynE (EU591743), ендоглюканазу EGC (O33897) і CBHA (O68438);

19) трансгенна рослина кукурудзи XynB/EGA/CBHA.2345.116, одержана шляхом трансформації pAG2345 і експресуюча XynB (O43097), ендоглюканазу EGA (NtEGm) і CBHA (O68438);

20) трансгенні рослини кукурудзи XynB/EGA/CBHB.2349.55 і XynB/EGA/CBHB.2349.56, одержані шляхом трансформації pAG2349 і експресуючі XynB (O43097), ендоглюканазу EGA (NtEG), CBHB (BD22308) і ZmUbilP:ZmKozak:xGZein27ss-02:BD22308:HvVSD-01:NosT.

Солома з рослин

Солому зібраної в теплиці кукурудзи сушили в циркуляторі повітря при 37 °C протягом 1-2 тижнів. Після висушування солому нарізали вручну на фрагменти 1,0-1,5 дюйма (2,5-3,75 см), а потім подрібнювали з використанням млина UDY (Model 014, UDY Corporation, Fort Collins, CO) з 0,5-мм ситом.

Одержання екстрактів білків рослин

Окремі подрібнені зерна або 20 мг подрібненої соломи ресуспендували в буфері для екстракції білка, що включає 100 мМ фосфат натрію (pH 6,5), етилендіамінтетраоцтову кислоту (EDTA; 1 мМ), Triton X-100 (0,1 %, об./об.) і фенілметансульфонілфторид (PMSF; 0,1 мМ). Ресуспендовані зразки тканини ретельно перемішували, а потім нерозчинний матеріал осаджували центрифугуванням. Рідину супернатанту, що містить розчинний білок, переносили в нову пробірку.

Хімічні реагенти і ферменти

Стандарти цукрів (глюкоза, ксилоза, арабіноза, галактоза, маноза і целобіоза) придбавали від Acros Organics (Morris Plains, NJ). Всі інші хімічні реагенти, використані в цьому іспиті, придбавали від Sigma-Aldrich (St. Louis, MO). Усі з ендоглюканази (C8546), β-глікозидази (49291) і ендоксиланази (X2753) для одержання власного коктейлю придбавали від Sigma (St. Louis, MO). Целобіогідролазу (CBHI) (EC 3.2.1.91) і β-ксилозидазу (EC 3.2.1.37) придбавали від Megazyme (Wicklow, Ірландія). Accellerase® 1500 і Accellerase® XY були люб'язними

подарунками Genencor International (Rochester, NY). Дріжджі Saccharomyces cereuisiae штаму D5A придбавали в American Type Culture Collection (ATCC) (Manassas, VA).

Аналізи ферментів соломи

Білок екстрагували з 15 мг соломи в 500 мкл буфера для екстракції (100 мМ натрій-фосфатний буфер, pH 6,5; NaOAc, pH 4,5; або Tris, pH 8,0, EDTA (10 мМ), і Triton X-100 (0,1 %) після інкубації протягом 30 хвилин при кімнатній температурі. Солому осаджували центрифугуванням. Супернатант збирали і переносили в нову пробірку Eppendorf. Для аналізів ферментів 50 мкл екстракту білків ресуспендували в буфері. Як правило, буфер включав Xylazyme у 100 мМ фосфаті Na, pH 6,5, або Cellazyme у 100 мМ NaOAc, pH 4,5. Таблетки Xylazyme AX або Cellazyme використовували відповідним чином для кожної пробірки в аналізі ферментів. Реакційні суміші інкубували при температурі аналізу (звичайно приблизно 50-60 °C, залежно від досліджуваного ферменту) доти, поки в рідині супернатанту не ставав помітним синій колір. Кількість синього барвника кількісно визначали шляхом вимірювання поглинання реакційної суміші при 590 нм. Контролі для цієї реакції включали одержані за допомогою мікроорганізмів ферменти й екстракти з рослин дикого типу. Субстрати для гідролізу також можна визначати з використанням кон'югованого з AZCL субстрату (Megazyme) замість таблеток Xylazyme AX і Cellazyme. Використання кон'югованого з AZCL субстрату дозволяє оптимізацію як досліджуваного об'єму соломи, так і концентрації субстрату.

Виявлення активності ксиланази

Розчинні білки аналізували з використанням Xylazyme AX (Megazyme, Bray, Co. Wicklow, Ірландія) як субстрату у реакційних сумішах об'ємом 0,5 мл при 50 °C у буфері HEPES (100 мМ, pH 8,0) для BSX або у фосфаті натрію (100 мМ, pH 6,5) для XynB. Для зупинення реакцій Xylazyme AX, до реакційних сумішей додавали 1 мл 2 % (мас./об.) Tris-основи. Нерозчинний матеріал з реакції Xylazyme AX осаджували центрифугуванням, і поглинання реакційної суміші в 100 мкл супернатанту вимірювали в трьох екземплярах спектрофотометрично при 590 нм. Для кількісного визначення рівнів накопичення BSX або XynB будували калібрувальні криві шляхом інкубації відомих кількостей очищених, одержаних з мікроорганізмів BSX або XynB, розведених у буфері для аналізу, з таблетками Xylazyme AX, одночасно з аналізом з Xylazyme AX, у якому використовувався матеріал трансгенних рослин.

У таблиці 1 нижче продемонстрована активність ферментів, виявлена в трансгенних рослинах. Як зазначено, активність ферменту була виявлена для декількох ксиланаз, ендоглюканаз, целобіогідролаз і ферулоїлестераз. Для кожної трансгенної події виявлену активність ферментів також підтверджували за допомогою вестерн-блотингу. "N/A" стосується ще не проведеного аналізу.

Таблиця 1

Трансгенна рослина	Ксиланаза	Ендоглюканаза	CBH	Acc A або B
AxB	-	-	-	-
BxA	-	-	-	-
TGC.2243.01	-	-	-	-
TGC.4000.12	-	-	-	-
TGC.4000.11	-	-	-	-
TGC.2004.8.02	-	-	-	-
TGC.2004.8.04	-	-	-	-
TGC.2243.01	-	-	-	-
XynA.2015.05	+	-	-	-
XynA.2015.5T1	+	-	-	-
XynB.2063.17	+	-	-	-
EGA.2049.02	-	+	-	-
EGA.2049.10	-	+	-	-
EGB.2042.03	-	+	-	-
EGC.2253.4b	-	+	-	-
EGA/XynA.2242.09	+	+	-	-
EGA/XynA.2242.09.16T1	+	+	-	-
XynA/AccB.2092.103	+	-	-	+
XynA/AccA/B.2096.01	+	-	-	+
XynA/AccA/B.2096.05	+	-	-	+
XynA.pv2015.3c	+	-	-	-
XynA.pv2015.4c	+	-	-	-
iXynA.2229.110	+	-	-	-
XynA/EGA.2309.54	+	+	-	-
XynA/EGA.2309.107	+	+	-	-
XynA/EGA.2342.105	+	+	-	-
XynE/EGC/CBHA.2339.03	+	+	N/A	-
XynE/EGC/CBHA.2339.04	+	+	N/A	-
XynE/EGC/CBHA.2339.05	+	+	N/A	-
XynB/EGA/CBHA.2345.116	+	+	N/A	-
XynB/EGA/CBHB.2349.55	+	+	+	-
XynB/EGA/CBHB.2349.56	+	+	+	-

Аналіз вуглеводного складу біомаси

Перед аналізом вуглеводного складу 3,0 г висушеної повітрям подрібненої соломи в двох екземплярах кип'ятили зі зворотним холодильником з 90 % (об./об.) етанолом з використанням скляної системи екстракції Soxhlet (Fisher Scientific, Pittsburg, PA) для видалення екстрагованих етанолом матеріалів згідно зі стандартами NREL (NREL/TP-510-42619). Екстракти, що містять етанол, упарювали у вакуумі з використанням роторного випарника, обладнаного водяною банею, установленою на 40 °C (Heidolph LR4000 G5B, IL USA). Склад екстракту визначали по масі твердих речовин у колбі після сушіння в сушильній шафі при 50 °C протягом 48 годин.

Солому, що не містить екстракту, піддавали двостадійному кислотному гідролізу (NREL/TP-510-42618), спочатку гідролізуючи при 30 °C за допомогою 1,5 мл 72 % (об./об.) H_2SO_4 на 0,16-0,18 г (повітряно-суха маса) протягом 60 хв., а потім при 121 °C протягом 1 години з додаванням 42,0 мл води. Після кислотного гідролізу додавали гідроксид натрію і гідроксид кальцію для доведення pH до значення від 4,0 до 9,0 і всі зразки фільтрували через 0,2-мкм PVDF-фільтри (Fisher Scientific, Pittsburg, PA) для аналізу за допомогою високоефективної рідинної хроматографії (ВЕРХ).

Об'єднаний спосіб з помірною попередньою обробкою й оцукрюванням

Щоб оцінити ефект експресованих рослиною ферментів CWD на гідроліз соломи, був розроблений об'єднаний спосіб, що включає м'яку попередню обробку з наступним ферментативним гідролізом без промивання біомаси між стадіями/детоксикації. В об'єднаному процесі усунуті які-небудь стадії промивання/розділення/детоксикації і він дозволяє інтегровану попередню обробку й одночасний процес оцукрювання і ферментації (SSF).

Помірна попередня обробка. Була розроблена ефективна м'яка попередня обробка, яка може досягти деяких ефектів попередньої обробки біомаси, але не інактивує гідролітичні ферменти в рослині. Хімічним реагентом для попередньої обробки була суміш 0,02-0,18М бісульфіту амонію і 0,025-0,20М карбонату амонію з pH від 5,0 до 9,0, переважно близько 8,10. Для оцінки гідролізу соломи рослини, 20,0 мг подрібненої кукурудзяної соломи додавали в 2-мл мікроцентрифужні пробірки з хімічним розчином для попередньої обробки зі співвідношенням рідини і твердої речовини (L/S), що складає 10 або менше (переважно 3-6). Суміш для попередньої обробки інкубували в пристрої для струшування при 350 об./хв. і температурі 40 °C-95 °C протягом 0-16 годин. Для подрібненої і неподрібненої соломи, після механічного очищення або дефібрилювання йшла попередня обробка хімічними реагентами. Далі попередньо оброблений матеріал піддавали ферментативному гідролізу без промивання між стадіями.

Ферментативний гідроліз. Попередньо оброблену солому піддавали ферментативному гідролізу в полібуфері Бріттона-Робінсона (40 мМ фосфат, 40 мМ ацетат, 40 мМ борат) з азидом натрію. Ферментативний гідроліз проводили при вмісті твердих речовин 2 % (мас./об.), pH 4,9, 50 °C у пристрої для струшування New Brunswick (New Brunswick Scientific, New Jersey, США) при 250 об./хв. протягом різних періодів часу (0-144 годин). Як коктейль #1 додавали 0,5 мкМ ендоглюканазу, 0,1 мкМ целобіогідролазу (CBHI), 0,05 мкМ β-глікозидазу і 0,5 мкМ ендоксиланазу в розрахунку на 10,0 мг соломи з реакційним об'ємом 1 мл. Коктейль 5# являв собою коктейль #1 з додаванням 0,1 мкМ β-ксилозидази. У сукупності, паралельно проводили три типи ферментативного гідролізу: без коктейлю ферментів (NCt), повний коктейль ферментів (FCt) і коктейль ферментів, позбавлений експресованого в рослині ферменту (Ct-PE), наприклад коктейль ферментів, позбавлений ендоксиланази (Ct-Xyn) або ендоглюканази (Ct-EG), або обох з них (Ct-EG-Xyn), залежно від ферменту, експресованого в рослинах. Accellerase® 1500 додавали в кількості 0,2 мл/г сухої маси і Accellerase® XY додавали в кількості 0,1 мл/г сухої маси. Виходи глюкози і ксилози (% від теоретичної величини) виражали як відсоток загального вмісту глюкози і ксилози в кожному субстраті. Планки похибки на прикладених кресленнях являють собою стандартне відхилення від середнього значення в повтореннях аналізу.

Одночасне оцукрювання і ферментація (SSF)

Інокулят одержували шляхом вирощування штаму дріжджів Saccharomyces cerevisiae D5A до OD_{600} 0,5 у YPD (10 г/л дріжджового екстракту, 20 г/л пептону і 20 г/л декстрози) при 30 °C і 250 об./хв. Клітини збирали центрифугуванням (3000 g протягом 5 хв.) і ресуспендували в 1X YP (10 г/л дріжджового екстракту і 20 г/л пептону).

Експерименти SSF проводили в двох екземплярах у 250-мл скляних колбах Ерленмейєра з робочим об'ємом 50 мл, що містили 3,0-4,0 г (маса сухої речовини) попередньо обробленої біомаси, буфер Бріттона-Робінсона, 10 × YP (100 г/л дріжджового екстракту і 200 г/л пептону), інокулят і гідролітичні ферменти. Колби закривали гумовою пробкою з повітряною пробкою. Експерименти починали шляхом додавання дріжджового інокуляту і ферментів (Accellerase® 1500 у кількості 10 FPU/г маси сухої речовини і Accellerase® XY у кількості 0,1 мл/г маси сухої речовини), і інкубували при 35 °C і 120 об./хв. протягом 7 діб. Узяття зразків проводили через 0, 24, 48, 72, 144 і 168 годин і аналізували відносно етанолу і цукрів.

Аналіз ферментованих цукрів і етанолу

Зразки гідролізату нагрівали при 90 °C протягом 20 хв., а потім центрифугували при 10000 g, після чого супернатанти очищали, пропускаючи через 0,20-мкм PVDF-фільтри (каталожний номер: 09-910-13, Fisher Scientific, Pittsburg, PA). Концентрації моносахаридів і дисахаридів визначали високоефективною рідинною хроматографією (ВЕРХ) з використанням бінарного насоса Shimadzu LC-20 AD із програмним забезпеченням LC solutions (Shimadzu, Kyoto, Японія). Концентрації цукрів визначали з використанням колонки для цукрів Aminex HPX-87P (Bio-Rad Laboratories, Hercules, CA), що діє при 0,6 мл/хв. і 80 °C з дегазованою водою як рухомою фазою. Концентрацію етанолу у ферментаційному бульйоні аналізували з використанням кислотної колонки Aminex HPX-87H (Bio-Rad Laboratories, Hercules, CA), що діє при 0,6 мл/хв., 60 °C з 0,004М сірчаною кислотою як рухомою фазою. Площі піків для всіх зразків, проаналізованих за допомогою детектора RI (RID 10AD), інтегрували і величини порівнювали зі стандартними кривими для кількісного визначення.

Приклад 2. Аналіз вуглеводного складу соломи з рослин

Солому з трансгенних рослин охарактеризовували з точки зору її складу структурних вуглеводів і вмісту цукрів для дослідження яких-небудь значних змін, викликаних генетичною модифікацією. У таблиці 2 представлені результати аналізу структурних вуглеводів з узятих випадковим чином зразків випадків трансгенної і нетрансгенної кукурудзи і проса. Вміст глюкану і ксилану з набору трансгенних рослин, як експресуючих один або декілька ферментів CWD, так і позбавлених трансгену, що кодує фермент CWD (трансгенний контроль TGC), є подібним з контрольними рослинами дикого типу (AxB).

Таблиця 2

Вміст глюкану і ксилану в трансгенних рослинах (експресуючі CWD або TGC) проти нетрансгенних рослин дикого типу (AxB)

Солома з рослин	# випадків (n)	Глюкан (г/100 г соломи)	Ксилан (г/100 г соломи)
Контролі у вигляді кукурудзи дикого типу	4	31,51±0,33	17,12±0,66
Контролі у вигляді трансгенної кукурудзи	8	30,16±1,02	15,99±1,14
Трансгенна кукурудза з ферментами	18	31,40±1,52	16,59±1,34

Т-критерій Стьюдента для даних, представлених у таблиці 2, показує відсутність значимих відмінностей у кількості глюкану між випадками трансгенної кукурудзи, експресуючої ферменти CWD, і кукурудзою дикого типу AxB, або між випадками трансгенної кукурудзи, експресуючої ферменти CWD, і трансгенними контрольними випадками, що не експресують фермент CWD, зі значенням P 0,90 і 0,14, відповідно. Відповідні значення P, виходячи з t-критерію, для вмісту ксилану, складають 0,57 і 0,36, відповідно.

Таким чином, експресія ферментів у рослинах забезпечує можливість одержання не тільки недорогих ферментів, але також сировини біомаси з гідролітичними ознаками для недорогого продукування ферментованих цукрів.

Приклад 3. Ефект експресованих рослиною ферментів CWD на гідроліз біомаси

Методологія оцінки гідролізу біомаси рослин

Однією з цілей експресії ферментів CDW у рослинах є усунення або зниження важкості умов хімічної попередньої обробки для переробки лігноцелюлозної біомаси.

Для оцінки ефектів експресованих рослиною ферментів CWD на гідроліз біомаси був розроблений об'єднаний процес з помірною хімічною попередньою обробкою (pH 5,0-9,0, 55 °C протягом 16 годин) з наступним ферментативним гідролізом (pH 4,9, 50 °C протягом 72 годин) без промивання між стадіями, і він був вибраний як стандартна методика для початкового скринінгу соломи з рослин. У цьому способі для оцінки використовували власні коктейлі ферментів (коктейль #1 і коктейль #5). Власний коктейль являє собою комбінацію індивідуальних ферментних компонентів, що дозволяє виключення кожного з компонентів, залежно від типу ферменту(ів), експресованого в рослині. Для кожної соломи з трансгенної рослини і соломи з рослини дикого типу або трансгенної контрольної рослини, паралельно проводили наступну обробку для ферментативного гідролізу: без коктейлю ферментів (NCt), повний коктейль (FCt) і коктейль, позбавлений ферменту, експресованого в рослині (Ct-PE; наприклад, коктейль, позбавлений ксиланази (Ct-Xyn), коктейль, позбавлений ендоглюканази (Ct-EG), або коктейль, позбавлений як ксиланази, так і ендоглюканази (Ct-Xyn-EG)).

Для визначення того, який фермент або ферменти підтримують високу ефективність гідролізу, два критерії використовували для оцінки характеристик обробки трансгенних випадків, експресуючих ферменти CWD, при початковому скринингу:

1) загальний вихід цукрів після гідролізу повним коктейлем (FCt) (висота (1) на ФІГ. 2A, 2B);

2) відмінність у виході цукрів між способами гідролізу, що залучають повний коктейль (FCt) і коктейль ферментів без ферменту, що експресується в рослині (Ct-PE) (висота (2) на ФІГ. 2A, 2B).

Загальна одержана кількість цукрів (висота (1)) після обробки є критерієм, який необхідно враховувати, оскільки це прямо впливає на вихід кінцевих продуктів, продуктивність і операційні витрати. За допомогою експресії в рослинах ферментів CWD було продемонстровано, що експресуючі ферменти трансгенні рослини досягали кращого загального гідролізу, ніж контрольна рослина, у тих же умовах обробки, що продемонстровано за допомогою загальних

виходів глюкози і ксилози на ФІГ. 2. Другий критерій - відмінність у виході цукрів між FCt і Ct-PE (висота (2)), відображує ефект експресованих рослиною ферментів на гідроліз. Було виявлено, що, при використанні цих трансгенних рослин як сировини біомаси, зовнішні ферменти в коктейлі ферментів можуть бути частково або повністю замінені ферментом CWD або ферментами CWD, експресованими в трансгенних рослинах, при досягненні подібного або рівного гідролізу, на що вказувала менша зміна або відсутність відмінностей у виході цукрів між гідролізом FCt і Ct-PE (ФІГ. 2B).

Оцінка гідролізу соломи з рослин

З використанням двох критеріїв селекції, зазначених вище, проводили ферментативний гідроліз зразків соломи за допомогою власного коктейлю для скринінгу ефективності різних трансгенних рослин кукурудзи, експресуючих ферменти CWD. Виходячи з результатів цього скринінгу, були ідентифіковані випадки трансгенних рослин з найкращою ефективністю, і вони включали експресуючі ксиланазу трансгенні рослини XynA.2015.05 і XynB.2063.17; експресуючі ендоглюканазу трансгенні рослини EGA.2049.10 і EGB.2042.03 і трансгенні рослини, експресуючі множину ферментів - XynA/AccA/B.2096.01, XynA/AccB.2092.103, EGA/XynA.2242.09, XynB/EGA/CBHA.2345.116, XynB/EGA/CBHB.2349.55 XynB/EGA/CBHB.2349.56 і XynB/EGA/CBHB.2349.229.

На ФІГ. 2 проілюстрований вихід глюкози (ФІГ. 2A) і ксилози (ФІГ. 2B) після гідролізу матеріалу з трансгенної рослини, експресуючої ксиланазу A (XynA; XynA.2015/05), і трансгенної контрольної рослини TGC.4000.12, яка не експресує фермент CWD (TGC.4000.12). Дані по виходу цукрів із трансгенного випадку XynA.2015.05 і трансгенного контролю після ферментативного гідролізу за допомогою FCt (власний коктейль #5), NCt і Ct-Xyn (власний коктейль #5, позбавлений ксиланази A) оцінювали з використанням описаних вище критеріїв селекції. Було показано, що величина загального виходу глюкози після обробки FCt (критерій 1) є більш високою, ніж різниця величин виходу глюкози між обробками FCt і Ct-Xyn (критерій 2) як для XynA.2015.05, так і для TGC.4000.12. Величина загального виходу глюкози і ксилози для всіх обробок була більш високою для трансгенного випадку XynA.2015.05, ніж для контрольної рослини TGC.4000.12. Цікаво, що відмінність у виходах глюкози і ксилози після обробки повним коктейлем ферментів і коктейлем ферментів без експресованої рослиною ксиланази A була дуже невеликою. Виходячи з цих результатів, ксиланаза A, експресована в рослині, була практично настільки ж ефективна відносно гідролізу соломи рослин, як і ксиланаза A, надана в повному коктейлі. Виходячи з критеріїв вибору 1 і 2, трансгенний випадок XynA.2015.05, показаний на ФІГ. 2, був ідентифікований як добре здійснюючий гідроліз.

Рослини, експресуючі ксиланазу

Відомо, що ксилан є переважною геміцелюлозою у твердій деревині, сільськогосподарських залишках, біомасі і багаторічних травах. Ксилан являє собою гетерополімерний біополімер, що складається з повторюваного кістяка ксилози з β-1,4-зв'язками, із приєднаними до нього групами гілок, і він може бути зшитим з лігніном ароматичними складними ефірами (Dodd D. and Cann I.O. 2009, Glob Change Biol. Bioenergy, 1:2). Руйнування й усунення ксилану сприяє гідролізу целюлози до ферментованих цукрів. У типовому коктейлі гідролітичних ферментів ксиланази є основним класом ферментів CWD, необхідним для гідролізу полімерів геміцелюлози, оскільки вони відіграють ключову роль у забезпеченні більшої доступності целюлози для ферментативного гідролізу. Посилаючись на ФІГ. 3B, ФІГ. 5B-5D і ФІГ. 7B, випадки трансгенних рослин, експресуючі XynA або XynB (XynB.2063.17, XynA/Acc/A/B.2096.01 і XynA.2015.05T1), продемонстрували на 29,80-172,1 % більш високу конверсію при гідролізі Ct-Xyn, ніж контрольні рослини, що вказує на поліпшений ефект ксиланази, експресованої у рослинах, на гідроліз ксилану біомас. Аналогічно, ці трансгенні рослини також демонструють на 50,1-93,5 % більш високу конверсію глюкану при гідролізі Ct-Xyn, ніж контрольні рослини.

На ФІГ. 3A, 3B проілюстровані виходи глюкози (ФІГ. 3A) і ксилози (ФІГ. 3B) з попередньо обробленої тканини трансгенної рослини, експресуючої ксиланазу B (XynB.2063.17), і контролю дикого типу (AxB) після гідролізу власним коктейлем #1 (FCt), коктейлем #1, позбавленим ксиланази B (Ct-Xyn), і без коктейлю (NCt). Результати обробки Ct-Xyn продемонстрували на 63,2 % більш високий вихід глюкози і на 109,4 % більш високий вихід ксилози з трансгенного випадку XynB.2063.17, ніж з контрольної рослини AxB. Збільшений гідроліз ксилану у випадку XynB.2063.17 також був очевидний з невеликої відмінності у величинах виходу ксилози між обробками FCt і Ct-Xyn (критерій 2).

Ці результати показують несподівано високу ефективність ксиланази B, експресованої у рослинах, відносно гідролізу соломи в порівнянні з ферментом, наданим у повному коктейлі.

Рослини, експресуючі целюлозу

Відомо, що лігноцелюлозна біомаса складається з матриксу з множиною біополімерів, що переплітаються (целюлоза, геміцелюлози, лігнін і екстрактивні речовини), які вимагають декілька різних класів ферментів у великих кількостях, щоб ефективно вивільняти ферментовані цукри. Серед них ключовим ферментом є целюлаза. Три типи целюлаз: ендоглюканаза, екзоглюканаза і β-глюкозидаза, діють разом, здійснюючи гідроліз целюлози в глюкозу. У процесі гідролізу ендоглюканаза руйнує зшивки між ланцюгами целюлози, у той час як екзоглюканаза гідролізує окремі ланцюги глюкану і β-глюкозидаза руйнує продукти екзоглюканази до мономерів глюкози (Sticklen M.B. 2008, Nature Reviews Genetics, 9:433).

На ФІГ. 4А, 4В показані результати гідролізу для трансгенних рослин, експресуючих ендоглюканази. На цих фігурах проілюстрований вихід глюкози з трансгенних випадків EGA.2049.02 і EGA.2049.10, що експресують ендоглюканазу А (EGA), і трансгенної контрольної рослини TGC.4000.12, яка позбавлена цього ферменту (ФІГ. 4А). На цих фігурах також проілюстрований вихід глюкози з трансгенного випадку EGB.2042.03, що експресує ендоглюканазу В (EGB), і трансгенної контрольної рослини TGC.2004.8.02 (ФІГ. 4В). Гідролітичні обробки проводили з повним коктейлем ферментів #1 (FCt), коктейлем ферментів, позбавленим ендоглюканази (Ct-EG), і без ферментів (NCt). Для експресуючих EGA випадків кукурудзи, як EGA.2049.02, так і EGA.2049.10, досягали на 48,9-126,9 % більш високої конверсії глюкану в порівнянні з трансгенною контрольною рослиною (TGC.4000.12) (ФІГ. 4А). Відмінність у виходах глюкози між гідролізом Ct-EG і FCt є незначною для EGA.2049.10 і приблизно на 29,1 % нижче для TGC.4000.12. Подібні спостереження, виходячи з критерію 2, були зроблені для трансгенного випадку EGA.2049.02. Посилаючись на ФІГ. 4В, експресуюча EGB трансгенна рослина EGB.2042.0 демонструє на 63,6 % більш високу конверсію глюкану після гідролізу Ct-EG, ніж трансгенний контроль TGC.2004.8.02. Несподівано, вихід глюкози після гідролізу EGB.2042.03 був тільки на 12,2 % нижче, ніж після гідролізу FCt, у порівнянні з на 23,0 % більш низькою величиною після відповідної обробки TGC.2004.8.02. Ці дані показують приблизно на 50,0 % кращий гідроліз для експресуючої EGB рослини, ніж для контрольної рослини.

Рослини, експресуючі множину ферментів

Для розробки ефективної і недорогої системи продукування ферментів для швидкої і менш дорогої деполімеризації біомаси, декілька ферментів, використовуваних у коктейлі гідролітичних ферментів, експресували в кукурудзі.

На ФІГ. 5А-5D і ФІГ. 6А, 6В показані результати гідролізу трансгенних рослин XynA/AccA/B.2096.05, XynA/AccA/B.2096.01, EGA/XynA.2242.09, XynB/EGA/CBHB.2349.56, XynB/EGA/CBHB.2349.55 і XynB/EGA/CBHA.2345.116, що експресують множину ферментів.

На ФІГ. 5А, 5В проілюстровані дані ферментативного гідролізу попередньо оброблених трансгенних рослин кукурудзи XynA/AccA/B.2096.01, XynA/AccA/B.2096.05, експресуючих ксиланазу А (XynA) і допоміжні ферменти (Acc), і трансгенної контрольної рослини TGC.2004.8.02 після обробки повним коктейлем #1 (FCt), коктейлем #1 без ксиланази (Ct-Xyn) і без коктейлю (NCt). Вихід глюкози (ФІГ. 5А) після гідролізу Ct-Xyn трансгенних випадків XynA/AccA/B.2096.01, XynA/AccA/B.2096.05, відповідно, на 80,4 % і 93,5 % перевищує вихід глюкози з контрольної рослини TGC.2004.8.02.

Посилаючись на ФІГ. 5С, несподівано більш високий вихід глюкози після гідролізу EGA/XynA.2242.09 за допомогою Ct-Xyn-EG може бути пояснений синергічним гідролітичним ефектом. Аналогічно, ефективність конверсії ксилану, виходячи з виходу ксилози (ФІГ. 5В), після гідролізу CT-Xyn трансгенних тканин з XynA/AccA/B.2096.01, XynA/AccA/B.2096.05, відповідно, на 143,4 % і 172,1 % перевищує ефективність конверсії ксилану в контрольній рослині TGC.2004.8.02. Виявлена висока ефективність конверсії ксилану для цих трансгенних випадків також може бути приписана синергічному ефекту декількох ферментів.

На ФІГ. 5 проілюстрований вихід глюкози (ФІГ. 5С) і ксилози (ФІГ. 5D) для випадку трансгенної кукурудзи EGA/XynA.2242.09, одночасно експресуючої ендоглюканазу А (EGA) і ксиланазу А (XynA), після ферментативних обробок повним коктейлем #1 (FCt), коктейлем #1, позбавленим ксиланази (Ct-Xyn), коктейлем #1, позбавленим ендоглюканази (Ct-EG), і коктейлем #1, позбавленим ксиланази і ендоглюканази (Ct-Xyn-EG). Експресія XynA у рослині приводить до збільшених виходів глюкози і ксилози для EGA/XynA.2242.09 після гідролізу. Наприклад, при обробці Ct-Xyn трансгенні випадки продемонстрували на 50,1 % більш високу ефективність конверсії глюкану (ФІГ. 5С) і на 29,8 % більш високу ефективність конверсії ксилану (ФІГ. 5D) відносно контрольної рослини. Експресія EGA у рослинах приводить до збільшеної ефективності гідролізу глюкану, про яку свідчить відмінність у виходах глюкози між обробками FCt, Ct-EG і Ct-Xyn-EG.

На ФІГ.6А, 6В проілюстрований вихід глюкози (ФІГ. 6А) і ксилози (ФІГ. 6В) з 1) трансгенних випадків, одночасно експресуючих три ферменти CWD (XynB/EGA/CBHB.2349.56 і

XynB/EGA/CBHB.2349.55, які одночасно експресують ксиланазу B, ендоглюканазу A (EGA), целобіогідролазу B, і XynB/EGA/CBHA.2345.116, що одночасно експресує ксиланазу B, ендоглюканазу A (EGA) і целобіогідролазу A), і 2) контрольної рослини дикого типу AxB. Показані результати обробки повним коктейлем (FCt) і без ферментів (NCt). Попередня обробка включала 0,17M бісульфіт амонію і карбонат амонію (BSC), співвідношення рідини і твердої речовини, що дорівнює 10, при 55 °C протягом 17 годин. Ферментативний гідроліз соломи проводили при 50 °C, pH 5,0, протягом трьох діб з використанням 0,2/0,1 мл Accellerase® 1500/XY на грам соломи. Посилаючись на ФІГ. 6, результати показують, що виходи глюкози і ксилози з трансгенних рослин, експресуючих три ферменти CWD, були значно більш високими, ніж з контрольної рослини дикого типу. Несподівано, випадок з найбільшою ефективністю XynB/EGA/CBHB.2345.56 показав на 43,6 % більш високий вихід глюкози і на 117,6 % більш високий вихід ксилози, ніж негативний контроль (AxB), після дуже помірної попередньої хімічної обробки.

Рослини другого покоління, експресуючі ферменти CWD

Рослина першого покоління (T0) XynA.2015.05 була ідентифікована як хороший кандидат для гідролізу. Для подальшої оцінки цього випадку і відповідної конструкції ферменту (XynA), насіння цього випадку висівали з одержанням потомства другого покоління T1. Оцінка результатів гідролізу для рослин XynA.2015.05 T0 і T1 представлена на ФІГ. 7B.

Два критерії, використані для оцінки рослин T0, також використовували для оцінки ефективності гідролізу для рослин, одержаних у T1, і для демонстрації того, що ферменти можуть бути ефективними в цьому виді. На ФІГ. 7B показані виходи глюкози (ФІГ. 7A) з рослин проса, одержаних з використанням pAG2015, і виходи ксилози (ФІГ. 7B) із трансгенного випадку T0 XynA.2015.05T0, експресуючого ксиланазу A, трансгенного випадку T1 XynA.2015.05T1, експресуючого ксиланазу A, і трансгенної контрольної рослини, позбавленої ферменту TGC.4000.11. Гідроліз проводили повним коктейлем (FCt), коктейлем, позбавленим ксиланази (Ct-Xyn), і без обробки коктейлем ферментів (NCt). Для обробки Ct-Xyn трансгенний випадок першого покоління XynA.2015.05T1 продемонстрував на 55,3 % більш високий гідроліз глюкану і на 101,6 % більш високий гідроліз ксилану, виходячи з виходів глюкози і ксилози, аналогічно тому, як і випадок першого покоління XynA.2015.05T0, і більш високий вихід цукрів для контрольної рослини TGC.4000.12.

Ці дані показують, що ферменти, експресовані в рослинах, є успадковуваними і зберігають активність у наступних поколіннях трансгенних рослин.

Різні види рослин

На доповнення до трансгенних випадків кукурудзи, одержували рослини проса, експресуючі ксиланазу A, шляхом трансформації вектором pAG2015. Коли ксиланазу A з XynA.2015.05, що добре здійснює гідроліз, експресували в просі шляхом трансформації тією же конструкцією (pAG2015), одержане трансгенне просо XynA.pv2015.3c також продемонструвало кращу конверсію глюкану і ксилану відносно контрольного проса Alamo. На ФІГ. 8 проілюстровані виходи глюкози (ФІГ. 8A) і ксилози (ФІГ. 8B) після гідролізу попередньо оброблених випадків трансгенного проса XynA.pv2015.3c, XynA.pv2015.4c і контрольної рослини проса дикого типу (Alamo) при обробці повним коктейлем (FCt), коктейлем ферментів, позбавленим ксиланази (Ct-Xyn), і без ферментів (NCT). Обидва трансгенних випадки XynA.pv2015.3c і XynA.pv2015.4c показали кращий гідроліз, ніж контрольне просо (Alamo). Випадок з найбільшою ефективністю XynA.pv2015.3c продемонстрував приблизно на 30,0 % більш високу ефективність конверсії глюкану і на 50,0 % більш високу ефективність конверсії ксилану в порівнянні з контрольною рослиною. Ці дані показують, що одна і таж ознака гідролізу може бути збережена серед видів, експресуючих ферменти в рослинах.

Рослини, експресуючі модифіковані інтеїном ферменти

Щоб уникнути шкідливих ефектів накопичення в рослинах ферментів CWD на ріст рослин, були розроблені і експресовані в рослинах модифіковані інтеїном ферменти для досягнення бажаної ефективності гідролізу без виникнення фенотипічних аномалій у рослинах. На ФІГ. 9 проілюстрований вихід глюкози (ФІГ. 9A) і ксилози (ФІГ. 9B) після гідролізу попередньо обробленої рослини iXynA.2229.110, експресуючої модифіковану інтеїном XynA, і контрольної рослини дикого типу AxB при обробці FCt, Ct-Xyn і NCt. Температура попередньої обробки вище 50 °C індукувала сплайсинг інтеїну в iXynA.2229.110. Гідроліз за допомогою Ct-Xyn демонструє на 66,0 % більш високу ефективність конверсії глюкану і на 57,3 % більш високу ефективність конверсії ксилану для iXynA.2229.110, ніж для контрольної рослини AxB. Усі трансгенні рослини iXynA.2229.110 були нормальними.

Дані аналізу складу вуглеводів показали відсутність значимих відмінностей у кількостях глюкану і ксилану між трансгенними рослинами, експресуючими гідролітичний фермент або

ферменти, і контрольними рослинами. Результати гідролізу продемонстрували, що трансгенні рослини, які експресують ферменти CWD, досягали аж до на 141 % більш високого виходу глюкози і на 172 % більш високого виходу ксилози в порівнянні з контрольними рослинами після ферментативного гідролізу в експериментальних умовах.

Приклад 4. Гідроліз з перебігом часу

Щоб краще оцінити ефект експресованих у рослинах ферментів на гідроліз, проводили оцінку гідролізу з перебігом часу для трансгенних рослин-кандидатів. На ФІГ. 10 порівнюється гідроліз з перебігом часу повним коктейлем (FCt) трансгенної рослини XynB.2063.17, XynA.2015.05T1 і контрольної рослини TGC.2004.8.02. Кінетика гідролізу узгоджується з типовим профілем: після швидкого початкового гідролізу іде повільна фаза зростання і кінцеве плато. Гідроліз уповільнюється через 24 години і вирівнюється через 48 годин. Трансгенні рослини, експресуючі ксиланази в рослинах, демонструють стабільно кращий гідроліз, ніж контрольна рослина, з перебігом часу, як очевидно з на 30,0-40,0 % більш високих виходів глюкози для гідролізу протягом 3 діб. Експресію ксиланази в рослинах можна вважати попередньою ферментативною обробкою для підвищення гідролізу як геміцелюлози, так і целюлози, у біомасі.

На ФІГ. 11 проілюстрований ферментативний гідроліз з перебігом часу трансгенної рослини EGA.2049.10, експресуючої ендоглюканазу A (EGA), у порівнянні з трансгенною контрольною рослиною TGC.4000.11, з використанням FCt і Ct-EG. Ефект експресованої у рослинах ендоглюканази A на гідроліз був продемонстрований по відмінності у виході глюкози з цих рослин протягом гідролізу Ct-EG з перебігом часу. З перебігом часу, експресуючі ендоглюканазу рослини EGA.2049.10 продемонстрували на 48,9 % і 63,6 % стабільно більш високу конверсію глюкану після гідролізу Ct-EG, ніж контрольні рослини TGC.4000.11. Несподівано, був досягнутий більш ефективний і швидкий гідроліз трансгенних рослин, експресуючих ендоглюканазу.

На ФІГ. 12 проілюстрований ферментативний гідроліз з перебігом часу трансгенного випадку EGA/XynA.2242.09, експресуючого EGA і XynA, і трансгенної контрольної рослини TGC.4000.11. Трансгенний випадок EGA/XynA.2242.09 демонструє стабільно більш високі виходи глюкози (ФІГ. 12A) і виходи ксилози (ФІГ. 12B) при використанні FCt, Ct-EG і Ct-Xyn, у порівнянні з контрольною рослиною TGC.4000.11. Ці дані демонструють, що ці високі виходи цукрів є наслідком одночасної експресії ендоглюканази і ксиланази в рослинах.

На ФІГ. 13 проілюстрований ферментативний гідроліз з перебігом часу трансгенної рослини XynA/AccB.2092.103, експресуючої XynA і допоміжний фермент (FAE B Aspergillus niger), і трансгенної контрольної рослини TGC.4000.11 з використанням обробки FCt і Ct-Xyn. Гідроліз Ct-Xyn попередньо обробленої соломи з трансгенної рослини XynA/AccB.2092.103 досягав більше ніж на 30 % більш високої ефективності конверсії глюкану (ФІГ. 13A) і більше ніж на 24 % більш високої ефективності конверсії ксилану (ФІГ. 13B), ніж у випадку трансгенної контрольної рослини, з перебігом часу.

Посилаючись на ФІГ. 12 і ФІГ. 13, кращий гідроліз трансгенних рослин, експресуючих множину ферментів, також спостерігали з точки зору виходів глюкози і ксилози протягом гідролізу XynA/AccB.2092.103 і EGA/XynA.2242.09. Кращий гідроліз може бути пояснений синергічним ефектом дії декількох ферментів.

Посилаючись на ФІГ. 10, 11, 12A і 13A, результати показують, що, на доповнення до більш високих виходів гідролізу, досягнутих шляхом експресії ферментів CWD у рослинах, профілі кінетики трансгенних випадків, експресуючих ферменти CWD, протягом гідролізу демонструють більш високий початковий нахил зміни виходу глюкози в порівнянні з контрольними рослинами, що вказує на більш швидкий початковий гідроліз.

Посилаючись на ФІГ. 10, на додаток до кращого гідролізу, трансгенні рослини, експресуючі гідролітичні ферменти, також демонструють більш швидкий початковий гідроліз, ніж контрольна рослина (ФІГ. 10). Після експресії в рослинах гідролітичні ферменти накопичувалися в клітинах рослин. У ході обробки вони можуть почати функціонувати безпосередньо in situ без потреби в транспортуванні на тривалі відстані і розподілі. Таким чином, очікується, що ефективність цих ферментів є високою внаслідок низького опору масовому перенесенню, і очікується, що вона знизитися при неселективному зв'язуванні ферментів у рослинах з лігніном і іншими молекулами, що не є мішенями. Надекспресія ферментів, деградуючих рослинну біомасу, у рослинах, очевидно, не приводить до зменшення кількості целюлози, а замість цього знижує інтеркаляцію ксилоглюкану з наступною необоротною модифікацією стінки. Усі ці фактори можуть приводити до більш високого гідролізу рослин, експресуючих ферменти.

Приклад 5. Підвищення гідролізу шляхом збільшення температури попередньої обробки

За допомогою хімічних реагентів для попередньої обробки відносно висока температура попередньої обробки, як правило, забезпечує кращі ефекти попередньої обробки на гідроліз. Для дослідження ефекту температури попередньої обробки, ідентифіковані найкращі здійснюючі гідроліз рослини піддавали попередній обробці при 65 °C і 75 °C. Вихід глюкози при гідролізі цих рослин представлений на ФІГ. 14.

Було виявлено, що термічна стабільність деяких ферментів CWD, включаючи ендоглюканазу A (O77044), може бути збільшена при експресії в рослинах. Термічну стабільність високотермостабільних ферментів, таких як ендоглюканаза C сімейства 12 (O33897), можна далі підвищувати після експресії в рослинах, забезпечуючи можливість використання підвищених температур попередньої обробки в ході обробки для досягнення збільшеного гідролізу. На ФІГ. 14 проілюстрований ефект температур попередньої обробки на вихід глюкози трансгенних випадків з найкращою ефективністю XynB.2063 (експресує ксиланазу B), XynA.2015.05T1 (експресує ксиланазу A), EGA.2049.10 (експресує ендоглюканазу A), XynA/AccB.2092.103 (експресує ксиланазу A і допоміжний фермент B), XynA/EGA.2242.09 (експресує ксиланазу A і ендоглюканазу A) і iXynA.2229.110 (експресує модифіковану інтеїном ксиланазу A) проти контрольної рослини дикого типу AxB і трансгенної контрольної рослини, позбавленої ферментів TGC.4000.11. Гідроліз рослин проводили з використанням FCt при температурі 65 °C або 75 °C. Було показано, що збільшення температури попередньої обробки від 65 °C до 75 °C підвищувало вихід гідролізу глюкози для контрольних рослин AxB і TGC.4000.11, але не для трансгенних рослин, експресуючих ферменти. Цей факт може бути пояснений насиченням експресованих у рослинах ферментів на доступному субстраті біомаси. При температурі попередньої обробки 65 °C, усі трансгенні рослини, експресуючі ферменти, продемонстрували на 11,0-33,4 % більш високий гідроліз у порівнянні з контрольною рослиною дикого типу AxB і трансгенним контролем TGC.4000.11, досягаючи 80-84 % від теоретичного виходу глюкози.

На ФІГ. 15 проілюстрований вихід глюкози з трансгенної рослини проса EGC.2253.4b, експресуючої високотермостабільну ендоглюканазу A (EGC), після попередньої обробки при температурах 65 °C, 75 °C і 95 °C і ферментативного гідролізу. Як показано, вихід глюкози з трансгенного випадку EGC.2253.4b був стабільно більш високим, ніж з контрольної рослини Alamo, досягаючи відповідних рівнів конверсії при попередній обробці при 95 °C, що складають 89,4 % і 71,4 %.

На ФІГ. 16A, 16B проілюстрований ефект температури і часу попередньої обробки на вихід глюкози (ФІГ. 16A) і ксилози (ФІГ. 16B) із трансгенного випадку XynA/EGA.2342.05 і трансгенної контрольної рослини, позбавленої ферментів, TGC.2243.01 після ферментативного гідролізу. Попередню обробку проводили з використанням 0,175M бісульфіту амонію і карбонату амонію при pH 8,1 і температурах 55 °C, 60 °C і 75 °C протягом 2, 4, 6 і 16 годин при кожній температурі і 3000 об./хв. Ферментативний гідроліз проводили з використанням 0,2 мл Accellerase® 1500/г попередньо оброблені соломи і 0,1 мл Accellerase® XY/г соломи, з вмістом твердих речовин 2 %, pH ~5,0, 1× полібуфера Бріттона-Робінсона (BR; кінцеве значення pH 4,9), 0,02 % азиду натрію, при 50 °C протягом 72 годин, при 250 об./хв. Як показано, для всіх температур і періодів часу попередньої обробки, трансгенний випадок XynA/EGA.2342.105, експресуючий ксиланазу A і ендоглюканазу A, мав стабільно кращу ефективність, ніж контрольна рослина TGC.2243.01. Ці дані також показують, що збільшення часу попередньої обробки збільшує вихід як глюкози, так і ксилози, з рослин практично лінійно і що попередня обробка протягом 16 годин значно підвищує гідроліз трансгенних і контрольних рослин.

На ФІГ. 17A, 17B проілюстрований ефект температури попередньої обробки на вихід глюкози (ФІГ. 17A) і ксилози (ФІГ. 17B) із трансгенних випадків XynA/EGA.2242.09.01 і XynA/EGA.2242.09.07 (експресуючих ксиланазу A і ендоглюканазу A) і контрольних рослин дикого типу AxB і трансгенних контролів TGC.4000.11. Рослини піддавали ферментативному гідролізу з використанням повного коктейлю ферментів Accellerase при температурах від 55 °C до 65 °C і 75 °C. Кількості ферментів, що додаються, включали 0,2 мл Accellerase® 1500 на грам попередньо обробленої соломи і 0,1 мл Accellerase® XY на грам попередньо обробленої соломи. Трансгенні рослини, експресуючі ферменти CWD, досягали аж до 83,5-89,1 % виходу глюкози і 50,0-64,3 % виходу ксилози в порівнянні з рослинами, що досягли тільки 63,0-76,6 % виходу глюкози і 35,7-45,3 % виходу ксилози.

Значний гідроліз може бути досягнутий при збільшенні температури попередньої обробки до 65 °C. Для деяких типів соломи кукурудзи, збільшення температури до 75 °C підвищувало гідроліз контрольних рослин, але не трансгенних рослин, експресуючих ферменти. Посилаючись на ФІГ. 15, для трансгенного проса EGC.2253.4b, експресуючого високотермостабільну ендоглюканазу C (EGC), несподівано було виявлено, що збільшений

гідроліз був значно більш високим при збільшенні температури попередньої обробки, особливо при 95 °С, ніж для контрольної рослини Alamo.

У цілому, попередня обробка соломи з трансгенних рослин з найбільшою ефективністю при підвищених температурах (65 °С і 75 °С) забезпечувала вихід гідролізу глюкози більш 80 %, що відповідає на 25 % більш високому гідролізу в порівнянні з контрольними рослинами, що не експресують гетерологічні гідролітичні ферменти. Таким чином, експресія високотермостабільних гідролітичних ферментів у рослинах забезпечить більше можливості для досягнення гідролізу заданих компонентів.

Приклад 6. Зниження навантаження ферментом і здатність до ферментації

Оскільки трансгенні рослини, експресуючі ферменти CWD, продемонстрували більш високий вихід гідролізу і більш швидку кінетику в ході гідролізу в порівнянні з контрольними рослинами в подібних умовах обробки, досліджували зниження навантаження екзогенних ферментів. На ФІГ. 18 проілюстровані виходи глюкози з трансгенних рослин XynB.2063.17, експресуючої ксиланазу B, XynA.2015.5T1, експресуючої ксиланазу A, і контрольної рослини TGC.2004.8.4 після гідролітичних обробок з використанням власного коктейлю #1 з різним навантаженням, таким як повний коктейль (FCt), 75 % коктейль (0,75 FCt), 50 % коктейль (0,50 FCt), 25 % коктейль (0,25 FCt), 10 % коктейль (0,10 FCt) і без ферментів (0 FCt). Дані демонструють здійсненність зниження навантаження екзогенних ферментів з використанням трансгенного матеріалу, експресуючого ферменти CWD, без зниження виходу цукрів. Наприклад, трансгенний випадок XynA.2015.15T1 досяг більше ніж 60 % конверсії глюкану з використанням навантаження 0,75 FCt, що є подібною з приблизно 65 % конверсією глюкану, досягнутою з використанням навантаження FCt. Навпаки, трансгенний контроль TGC.2004.8.4 досяг конверсії глюкану приблизно 50,0 % при використанні навантаження FCt і конверсії глюкану приблизно 40 % при використанні навантаження 0,75 FCt. Ці дані показують, що гідроліз рослин, експресуючих ферменти CWD, був більш ефективним, ніж гідроліз контрольних рослин, і при більш низьких навантаженнях зовнішніх ферментів.

На ФІГ. 19 проілюстровані виходи глюкози з трансгенних рослин XynE/EGC/CBHA.2339.03, XynE/EGC/CBHA.2339.04 і XynE/EGC/CBHA.2339.05 (експресуючих ксиланазу E, ендоглюканазу C і целобіогідролазу A) і контрольної рослини дикого типу BxA після ферментативного гідролізу з обробкою повним коктейлем (FCt), 20 % повним коктейлем (0,2 FCt) і без ферментів (NCt). Дані показують, що 50-70 % вихід глюкози може бути досягнутий із трансгенних рослин XynE/EGC/CBHA.2339.03, XynE/EGC/CBHA.2339.04 і XynE/EGC/CBHA.2339.05 з використанням тільки 20 % навантаження повного коктейлю, що на 35-77 % перевищує вихід глюкози з контрольних рослин.

Несподівано, вихід глюкози з трансгенного випадку XynE/EGC/CBHA.2339.03 після гідролізу з тільки 20 % навантаженням повного коктейлю, усе ще був на 15 % вище виходу глюкози з контрольної рослини BxA після гідролітичної обробки з використанням повного коктейлю.

На ФІГ. 20A, 20B ілюструється ефект зниження навантаження зовнішніх ферментів на вихід глюкози з трансгенних випадків XynA/EGA.2309.54 і XynA/EGA.2309.107 (експресуючих ксиланазу A і ендоглюканазу A) і контрольних рослин BxA після гідролізу повним коктейлем (FCt), 20 % повним коктейлем (0,2FCt) і без ферментів (NCt). Були досягнуті 60-70 % виходи глюкози з трансгенних рослин, експресуючих ферменти, з використанням 20 % навантаження повних коктейлів у порівнянні з приблизно 80 % виходами глюкози, досягнутими за допомогою повного коктейлю. Навпаки, рослини негативного контролю забезпечували тільки 53 % глюкози при обробці 0,2 FCt і 70 % глюкози при обробці FCt. Посилаючись на ФІГ. 20B, були досягнуті приблизно 38 % виходи ксилози з трансгенних рослин, експресуючих ферменти CWD, при обробці 0,2 FCt, у порівнянні з виходами ксилози приблизно 47 % при обробці FCt і значно більш низькими виходами ксилози для рослин негативного контролю при обробці як FCt, так і 0,2 FCt. При 20 % FCt для XynA/EGA.2309.54 і XynA/EGA.2309.107 може бути досягнутий на 16-29,3 % більш високий вихід глюкози, а також може бути досягнутий на 27,6-31 % більш високий вихід ксилози, у порівнянні з негативним контролем.

На ФІГ. 21A, 21B проілюстрований ефект зниження навантаження зовнішніми ферментами на виходи глюкози (ФІГ. 21A) і ксилози (ФІГ. 21B) із трансгенних рослин EGA/XynA.2242.09.16 (експресуючих ендоглюканазу CBHA.2069.1.3) і трансгенного контролю TGC.4000.11 після гідролізу без використання ферментів (0), з використанням 20 % повного коктейлю (0,2 FCt), 40 % повного коктейлю (0,4 FCt), 60 % повного коктейлю (0,6 FCt), 80 % повного коктейлю (0,8 FCt) і повного коктейлю (Accellerase® 1500/XY). Попередню обробку проводили з використанням 0,25M бісульфіту амонію і карбонату амонію (pH 8,56) при співвідношенні рідини і твердої речовини, що дорівнює 7, при 75 °С протягом 20 годин. Ферментативний гідроліз проводили з використанням вмісту твердих речовин приблизно 2 %, pH 5,0, при 50 °С протягом

трьох діб. Несподівано, дані показують, що гідроліз 60 % FCt може забезпечувати 80 % від теоретичного виходу глюкози і приблизно 60 % від теоретичного виходу ксилози для обох трансгенних рослин, експресуючих ферменти, у той же час тільки 60 % і 49 %, відповідно, для негативного контролю TGC.4000.11.

Ці результати демонструють потенціал до зниження навантаження зовнішніх ферментів і одночасного досягнення ефективного гідролізу соломи з рослин з використанням трансгенних рослин, експресуючих ферменти CWD.

Для оцінки здатності до ферментації гідролізатів, одержаних із трансгенних рослин, експресуючих ферменти CWD, проводили експеримент по одночасному оцукрюванню і ферментації (SSF) з використанням Saccharomyces cerevisiae D5A. На ФІГ. 22 проілюстроване продукування етанолу в ході SSF із трансгенних рослин EGA.2049.10 (експресуючої ендоглюканазу A) і EGA/XynA.2242.09 (експресуючої ендоглюканазу A і ксиланазу A) і контрольних рослин AxB і TGC.4000.11 при вмісті твердих речовин у біомасі 6 %. Ці дані показують, що досягали приблизно на 77,8 % більш високої конверсії целюлози з використанням трансгенних рослин, експресуючих ферменти CWD, у порівнянні з контрольними рослинами. Крім того, у випадку біомаси з помірною попередньою обробкою, SSF експресуючих ферменти рослин EGA.2049.10 і EGA/XynA.2242.09 забезпечувала продукування етанолу в концентрації 8,0 г/л або вихід етанолу 65 %, у порівнянні з 4,5 г/л або виходом етанолу 42 % для контрольних рослин, що відповідає 55 % підвищенню продукції етанолу.

Збільшений гідроліз має потенціал до того, щоб перетворюватися в зниження екзогенного навантаження ферментами при збереженні подібного гідролізу, як показано на ФІГ. 18-21. Експресія ферментів CWD у рослинах продемонструвала можливість продукувати недорогий цукор з культур або біомаси, експресуючих ферменти CWD, для продукування біопалив, біохімічних речовин і біоматеріалів. Перевага швидкого початкового гідролізу також забезпечує можливість досягнення подібного або кращого гідролізу при меншому операційному часі, що є перевагою для процесу одночасного оцукрювання і ферментації (ФІГ. 22), і можливість зниження потреби в продуктивності обладнання й операційних витрат.

Продукування ферментів, деградуючих клітинну стінку, у рослинах, являє собою засіб для зниження витрат, асоційованих із продукуванням ферментованих цукрів з біомаси за допомогою прямого гідролізу трансгенних рослин.

Приклад 7. Термохімічний ефект: солюбілізація біомаси після помірної попередньої обробки

На ФІГ. 23 проілюстроване розчинення біомаси після помірної попередньої обробки, виходячи зі зниження маси трансгенної рослини, експресуючої екзоглюканазу CBHA (CBHA.2069.3.17), і контрольної рослини дикого типу (AxB) після попередньої обробки деіонізованою водою (DI), 0,06M або 0,25M розведеною сірчаною кислотою (DSA) або 0,25M бісульфітом амонію і 0,23M карбонатом амонію (BSC; pH 8,1). Попередню обробку проводили при температурі 75 °C, співвідношенні рідини і твердих речовин, що дорівнює 8, протягом 16 годин. На цій фігурі представлені дані зниження маси для зразків висушеної в сушильній шафі кукурудзяної соломи, одержаної з трансгенної рослини і контрольної рослини дикого типу. Вимірювання проводили після промивання і центрифугування попередньо обробленої біомаси. У порівнянні з попередньою обробкою DI, попередня обробка 0,25M BSC (pH 8,1) приводить до на 17,5 % більшого зниження маси біомаси для трансгенної рослини (CBHA.2069.3.17) і на 12,2 % більшого зниження маси для нетрансгенної контрольної рослини (AxB). Трансгенна рослина, експресуюча екзоглюканазу (CBHA.2069.3.17), демонструє більше зниження маси біомаси після всіх попередніх обробок у порівнянні з нетрансгенною контрольною рослиною.

Приклад 8. Термохімічний ефект: зниження маси біомаси і видалення лігніну після помірної попередньої обробки

Зниження маси біомаси і видалення лігніну визначали для зразків висушеної в сушильній шафі соломи, одержаних із трансгенної контрольної рослини TGC.2209, після попередньої обробки деіонізованою водою (DI) і 0,17M бісульфітом амонію з 0,165M карбонатом амонію (BSC) при pH 8,1, зі співвідношенням рідини і твердих речовин, що дорівнює 8, при температурі 75 °C, протягом 16 годин. Вимірювання проводили після промивання і центрифугування попередньо обробленої біомаси. У таблиці 3 показано, що попередня обробка 0,17M BSC (pH 8,1) приводить до більшого зниження маси біомаси і до більшого видалення лігніну в порівнянні з попередньою обробкою DI-водою.

Таблиця 3

Зниження маси біомаси після попередньої обробки і нерозчинного в кислоті лігніну в попередньо обробленій біомасі

Хімічні речовини для попередньої обробки	Зниження маси біомаси (% від маси соломи)	Нерозчинний у кислоті лігнін (% від попередньо обробленої біомаси)
DI-вода	34,6	10,0
0,17M BSC	39,9	14,6

Приклад 9. Термохімічний ефект: деацетилювання при помірній попередній обробці

На ФІГ. 24A, 24B проілюстрований ефект температури і часу на деацетилювання рослинної біомаси, оцінений для висушеної в сушильній шафі кукурудзяної соломи, що походить з трансгенної рослини, експресуючої целобіогідролазу A (CBHA.2063.3.17), і нетрансгенної контрольної рослини (AxB). Попередня обробка включала деіонізовану воду (DI), 0,06M розведену сірчану кислоту (DSA), 25M DSA або 0,25M бісульфіт амонію з 0,23M карбонатом амонію (BSC) і її проводили при pH 8,1, співвідношенні рідини і твердої речовини, що дорівнює 8, при температурах 75 °C або 95 °C протягом 16 годин або 85 °C протягом 7 годин. Концентрацію оцтової кислоти (HAc) визначали за допомогою аналізу BEPX зразків фільтрату, одержаних з попередньо обробленої біомаси, з використанням кислотної колонки HPX-87H Column (Bio-Rad Laboratories, Hercules, CA), що діє при 0,6 мл/хв., 60 °C з 0,004M сірчаною кислотою як рухомою фазою. Попередня обробка 0,25M BSC (pH 8,1) приводила до суттєвого деацетилювання в порівнянні з попередньою обробкою DI-водою і 0,06M DSA.

Приклад 10. Термохімічний ефект: невелика/відсутність деградації цукрів (фурфураль і HMF) при помірній попередній обробці

На ФІГ. 25A, 25B ілюструється вихід продуктів деградації цукрів гідроксиметилфурфуралю (HMF) і фурфуралю в зразках висушеної в сушильній шафі соломі з трансгенної рослини, експресуючої екзоглюканазу CBHA (CBHA.2069.3.17), і нетрансгенної контрольної рослини (AxB) після попередньої обробки деіонізованою водою (DI), 0,06M розведеною сірчаною кислотою (DSA), 0,25 DSA або 0,25M бісульфітом амонію і 0,23M карбонатом амонію (BSC; pH 8,1). Попередню обробку проводили при температурах 75 °C і 95 °C протягом 16 годин і 85 °C протягом 7 годин. Концентрації HMF і фурфуралю у фільтраті попередньо обробленої біомаси вимірювали за допомогою аналізу BEPX із використанням кислотної колонки HPX-87H Column (Bio-Rad Laboratories, Hercules, CA), що діє при 0,6 мл/хв., 60 °C з 0,004M сірчаною кислотою як рухомою фазою.

Дані показують, що від незначної кількості до відсутності HMF або фурфуралю було виявлено в зразках із трансгенної рослини, експресуючої екзоглюканазу CBHA (CBHA.2069.3.17), або нетрансгенної контрольної рослини (AxB) після попередньої обробки 0,25M BSC (pH 8,1) у порівнянні з попередньою обробкою деіонізованою водою (DI), 0,06M розведеною сірчаною кислотою (DSA) або 0,25M DSA, яка привела до деградації цукрів і виявлення HMF і фурфуралю в зразках.

Приклад 11. Аутогідроліз внаслідок ферментів у рослинах

Приклади аутогідролізу рослин, експресуючих ферменти, деградуючі клітинну стінку, представлені на ФІГ. 26. На цій фігурі проілюстровані результати для попередньо обробленої рослини кукурудзи, експресуючої XynB (XynB.2063.15) окремо, і рослини, одночасно експресуючої три ферменти; XynA і допоміжні ферменти A і B (XynA/AccA/B.2096.01). Результати порівнювали з попередньо обробленою нетрансгенною контрольною рослиною (AxB). Попередню обробку проводили з використанням 0,17M бісульфіту амонію і 0,165M карбонату амонію (pH 8,1), співвідношення рідини і твердої речовини (L/S), що дорівнює 10, і при 55 °C протягом 16 годин. Ферментативний гідроліз здійснювали за допомогою ферментів, продукованих у рослинах, при вмісті твердих речовин 2 %, без коктейлю ферментів (NCt), що додається ззовні, і при температурі 50 °C, pH 5,0 протягом 3 діб. Наступний кислотний гідроліз проводили при pH менше 1,0 і температурі 121 °C протягом 60 хвилин. Обидві трансгенні рослини, експресуючі ксиланазу, виявляють у 3-5 разів більш високий вихід ксилози внаслідок аутогідролізу в порівнянні з контрольною рослиною AxB.

На ФІГ. 27 проілюстровані результати аутогідролізу для рослини кукурудзи, одночасно експресуючої ксиланазу B, ендоглюканазу і CHBA (XynB/EGA/CBHA.2345.116), і рослини, експресуючої ксиланазу B, ендоглюканазу і CBHB (XynB/EGA/CBHB.2349.55), у порівнянні з

нетрансгенною контрольною рослиною (AxB). Рослини попередньо обробляли 0,17М бісульфітом амонію і 0,165М карбонатом амонію BSC (pH 8,1) з L/S, що дорівнює 10, при 55 °С протягом 16 годин. Між попередньою обробкою і гідролізом ферментами, експресованими в рослині, не застосовували проміжного промивання. Гідроліз здійснювали в наступних умовах: вміст твердих речовин 2 %, без коктейлю (NCt), 50 °С і pH 5,0 протягом 3 діб. Трансгенні рослини, експресуючі ксиланазу й інші целюлази, демонструють значно більш високий вихід ксилози внаслідок аутогідролізу в порівнянні з контрольною рослиною AxB.

Приклад 12. Ефект механічного дефібрилювання на обробку неподрібненої соломи

Дані про ефект механічного дефібрилювання на виходи глюкози і ксилози після обробки неподрібненої соломи представлені на ФІГ. 28А, 28В. Дані були одержані з експериментів з попередньо обробленою трансгенною рослиною, експресуючою ендоглюканазу і ксиланазу А (EGA/XynA.2242.09T1), проти трансгенної контрольної рослини TGC.4000. Попередню обробку проводили з використанням 0,25М бісульфіту амонію з 0,23М карбонатом амонію (pH 8,56, L/S, що дорівнює 8, протягом 16 годин, з наступним механічним дефібрилюванням (вміст твердих речовин 6 %). Ферментативний гідроліз здійснювали за допомогою ферментів, продукованих у рослинах, при вмісті твердих речовин 4 %, з використанням Accelerase ® 1500/XY (0,2/0,1 мл на г соломи) при температурі 50 °С протягом аж до 3 діб і при pH 5,0. Трансгенна рослина кукурудзи, одночасно експресуюча ендоглюканазу і ксиланазу, виявляє стабільно більш високі виходи глюкози і ксилози в порівнянні з контрольною рослиною TGC.4000 з перебігом часу, досягаючи виходів глюкози і ксилози 83 % і 63 %, відповідно, при гідролізі протягом 3 діб.

Посилання, цитовані протягом цієї заявки, включені для всіх цілей, представлених у даному описі й у самих посиланнях, як якби кожне посилання було повністю представлене. Для демонстрації, конкретні з цих посилань цитовані в конкретних місцях даного опису. Цитування посилання в конкретному місці вказує на порядок (порядки), у якому включені ідеї посилання. Однак цитування посилання в конкретному місці не обмежує порядок, у якому усі з ідей цитованого посилання включені для будь-яких цілей.

Таким чином, зрозуміло, що даний винахід не обмежується конкретними описаними варіантами здійснення, а охоплює всі модифікації, що знаходяться в межах суті й обсягу винаходу, як визначено в прикладеній формулі винаходу, у представленому вище описі і/або показано на прикладених кресленнях.

СПИСОК ПОСЛІДОВНОСТЕЙ

<110> Agrivida, Inc.

<120> ОБ'ЄДНАНІ ПОПЕРЕДНЯ ОБРОБКА І ГІДРОЛІЗ РОСЛИННОЇ БІОМАСИ, ЕКСПРЕСУЮЧОЇ ФЕРМЕНТИ, ДЕГРАДУЮЧІ КЛІТИННУ СТІНКУ

<130> AGR-PT015.1WO

<150> 61/449,769
<151> 2011-03-07

<160> 85

<170> PatentIn version 3.5

<210> 1
<211> 337
<212> Білок
<213> Dictyoglomus thermophilum, амінокислота, P77853

<400> 1

Met Gln Thr Ser Ile Thr Leu Thr Ser Asn Ala Ser Gly Thr Phe Asp
1 5 10 15

Gly Tyr Tyr Tyr Glu Leu Trp Lys Asp Thr Gly Asn Thr Thr Met Thr
20 25 30

Val Tyr Thr Gln Gly Arg Phe Ser Cys Gln Trp Ser Asn Ile Asn Asn
35 40 45

Ala Leu Phe Arg Thr Gly Lys Lys Tyr Asn Gln Asn Trp Gln Ser Leu
50 55 60

Gly Thr Ile Arg Ile Thr Tyr Ser Ala Thr Tyr Asn Pro Asn Gly Asn
65 70 75 80

Ser Tyr Leu Cys Ile Tyr Gly Trp Ser Thr Asn Pro Leu Val Glu Phe
85 90 95

Tyr Ile Val Glu Ser Trp Gly Asn Trp Arg Pro Pro Gly Ala Thr Ser
100 105 110

Leu Gly Gln Val Thr Ile Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr
115 120 125

Thr Arg Val Asn Gln Pro Ser Ile Val Gly Thr Ala Thr Phe Asp Gln
130 135 140

Tyr Trp Ser Val Arg Thr Ser Lys Arg Thr Ser Gly Thr Val Thr Val
145 150 155 160

Thr Asp His Phe Arg Ala Trp Ala Asn Arg Gly Leu Asn Leu Gly Thr

165 170 175

Ile Asp Gln Ile Thr Leu Cys Val Glu Gly Tyr Gln Ser Ser Gly Ser
180 185 190

Ala Asn Ile Thr Gln Asn Thr Phe Ser Gln Gly Ser Ser Ser Gly Ser
195 200 205

Ser Gly Gly Ser Ser Gly Ser Thr Thr Thr Thr Arg Ile Glu Cys Glu
210 215 220

Asn Met Ser Leu Ser Gly Pro Tyr Val Ser Arg Ile Thr Asn Pro Phe
225 230 235 240

Asn Gly Ile Ala Leu Tyr Ala Asn Gly Asp Thr Ala Arg Ala Thr Val
245 250 255

Asn Phe Pro Ala Ser Arg Asn Tyr Asn Phe Arg Leu Arg Gly Cys Gly
260 265 270

Asn Asn Asn Asn Leu Ala Arg Val Asp Leu Arg Ile Asp Gly Arg Thr
275 280 285

Val Gly Thr Phe Tyr Tyr Gln Gly Thr Tyr Pro Trp Glu Ala Pro Ile
290 295 300

Asp Asn Val Tyr Val Ser Ala Gly Ser His Thr Val Glu Ile Thr Val
305 310 315 320

Thr Ala Asp Asn Gly Thr Trp Asp Val Tyr Ala Asp Tyr Leu Val Ile
325 330 335

Gln

<210> 2
<211> 261
<212> Білок
<213> Aspergillus niger, амінокислотна Anfae A

<400> 2

Met Ala Ser Thr Gln Gly Ile Ser Glu Asp Leu Tyr Asn Arg Leu Val
1 5 10 15

Glu Met Ala Thr Ile Ser Gln Ala Ala Tyr Ala Asp Leu Cys Asn Ile
20 25 30

Pro Ser Thr Ile Ile Lys Gly Glu Lys Ile Tyr Asn Ala Gln Thr Asp
35 40 45

```
Ile Asn Gly Trp Ile Leu Arg Asp Asp Thr Ser Lys Glu Ile Ile Thr
    50                  55                  60

Val Phe Arg Gly Thr Gly Ser Asp Thr Asn Leu Gln Leu Asp Thr Asn
65                  70                  75                  80

Tyr Thr Leu Thr Pro Phe Asp Thr Leu Pro Gln Cys Asn Asp Cys Glu
                85                  90                  95

Val His Gly Gly Tyr Tyr Ile Gly Trp Ile Ser Val Gln Asp Gln Val
            100                 105                 110

Glu Ser Leu Val Lys Gln Gln Ala Ser Gln Tyr Pro Asp Tyr Ala Leu
        115                 120                 125

Thr Val Thr Gly His Ser Leu Gly Ala Ser Met Ala Ala Leu Thr Ala
    130                 135                 140

Ala Gln Leu Ser Ala Thr Tyr Asp Asn Val Arg Leu Tyr Thr Phe Gly
145                 150                 155                 160

Glu Pro Arg Ser Gly Asn Gln Ala Phe Ala Ser Tyr Met Asn Asp Ala
                165                 170                 175

Phe Gln Val Ser Ser Pro Glu Thr Thr Gln Tyr Phe Arg Val Thr His
            180                 185                 190

Ser Asn Asp Gly Ile Pro Asn Leu Pro Pro Ala Asp Glu Gly Tyr Ala
        195                 200                 205

His Gly Gly Val Glu Tyr Trp Ser Val Asp Pro Tyr Ser Ala Gln Asn
    210                 215                 220

Thr Phe Val Cys Thr Gly Asp Glu Val Gln Cys Cys Glu Ala Gln Gly
225                 230                 235                 240

Gly Gln Gly Val Asn Asp Ala His Thr Thr Tyr Phe Gly Met Thr Ser
                245                 250                 255

Gly Ala Cys Thr Trp
            260
```

<210> 3
<211> 505
<212> Білок
<213> Aspergillus niger, амінокислотна Anfae B

<400> 3

```
Met Ala Thr Asp Pro Phe Gln Ser Arg Cys Asn Glu Phe Gln Asn Lys
1               5                  10                  15

Ile Asp Ile Ala Asn Val Thr Val Arg Ser Val Ala Tyr Val Ala Ala
            20                  25                  30

Gly Gln Asn Ile Ser Gln Ala Glu Val Ala Ser Val Cys Lys Ala Ser
        35                  40                  45

Val Gln Ala Ser Val Asp Leu Cys Arg Val Thr Met Asn Ile Ser Thr
    50                  55                  60

Ser Asp Arg Ser His Leu Trp Ala Glu Ala Trp Leu Pro Arg Asn Tyr
65                  70                  75                  80

Thr Gly Arg Phe Val Ser Thr Gly Asn Gly Gly Leu Ala Gly Cys Val
                85                  90                  95

Gln Glu Thr Asp Leu Asn Phe Ala Ala Asn Phe Gly Phe Ala Thr Val
            100                 105                 110

Gly Thr Asn Gly Gly His Asp Gly Asp Thr Ala Lys Tyr Phe Leu Asn
        115                 120                 125

Asn Ser Glu Val Leu Ala Asp Phe Ala Tyr Arg Ser Val His Glu Gly
    130                 135                 140

Thr Val Val Gly Lys Gln Leu Thr Gln Leu Phe Tyr Asp Glu Gly Tyr
145                 150                 155                 160

Asn Tyr Ser Tyr Tyr Leu Gly Cys Ser Thr Gly Gly Arg Gln Gly Tyr
                165                 170                 175

Gln Gln Val Gln Arg Phe Pro Asp Asp Tyr Asp Gly Val Ile Ala Gly
            180                 185                 190

Ser Ala Ala Met Asn Phe Ile Asn Leu Ile Ser Trp Gly Ala Phe Leu
        195                 200                 205

Trp Lys Ala Thr Gly Leu Ala Asp Asp Pro Asp Phe Ile Ser Ala Asn
    210                 215                 220

Leu Trp Ser Val Ile His Gln Glu Ile Val Arg Gln Cys Asp Leu Val
225                 230                 235                 240

Asp Gly Ala Leu Asp Gly Ile Ile Glu Asp Pro Asp Phe Cys Ala Pro
                245                 250                 255
```

```
Val Ile Glu Arg Leu Ile Cys Asp Gly Thr Thr Asn Gly Thr Ser Cys
            260                 265                 270

Ile Thr Gly Ala Gln Ala Ala Lys Val Asn Arg Ala Leu Ser Asp Phe
        275                 280                 285

Tyr Gly Pro Asp Gly Thr Val Tyr Tyr Pro Arg Leu Asn Tyr Gly Gly
    290                 295                 300

Glu Ala Asp Ser Ala Ser Leu Tyr Phe Thr Gly Ser Met Tyr Ser Arg
305                 310                 315                 320

Thr Glu Glu Trp Tyr Lys Tyr Val Val Tyr Asn Asp Thr Asn Trp Asn
                325                 330                 335

Ser Ser Gln Trp Thr Leu Glu Ser Ala Lys Leu Ala Leu Glu Gln Asn
            340                 345                 350

Pro Phe Asn Ile Gln Ala Phe Asp Pro Asn Ile Thr Ala Phe Arg Asp
        355                 360                 365

Arg Gly Gly Lys Leu Leu Ser Tyr His Gly Thr Gln Asp Pro Ile Ile
    370                 375                 380

Ser Ser Thr Asp Ser Lys Leu Tyr Tyr Arg Arg Val Ala Asn Ala Leu
385                 390                 395                 400

Asn Ala Ala Pro Ser Glu Leu Asp Glu Phe Tyr Arg Phe Phe Gln Ile
                405                 410                 415

Ser Gly Met Gly His Cys Gly Asp Gly Thr Gly Ala Ser Tyr Ile Gly
            420                 425                 430

Gln Gly Tyr Gly Thr Tyr Thr Ser Lys Ala Pro Gln Val Asn Leu Leu
        435                 440                 445

Arg Thr Met Val Asp Trp Val Glu Asn Gly Lys Ala Pro Glu Tyr Met
    450                 455                 460

Pro Gly Asn Lys Leu Asn Ala Asn Gly Ser Ile Glu Tyr Met Arg Lys
465                 470                 475                 480

His Cys Arg Tyr Pro Lys His Asn Ile His Thr Gly Pro Gly Asn Tyr
                485                 490                 495

Thr Asp Pro Asn Ser Trp Thr Cys Val
            500                 505
```

<210> 4
<211> 433
<212> БІЛОК
<213> Nasutitermes takasagoensis, амінокислотна послідовність NtEGM

<400> 4

Met Ala Tyr Asp Tyr Lys Gln Val Leu Arg Asp Ser Leu Leu Phe Tyr
1 5 10 15

Glu Ala Gln Arg Ser Gly Arg Leu Pro Ala Asp Gln Lys Val Thr Trp
20 25 30

Arg Lys Asp Ser Ala Leu Asn Asp Gln Gly Asp Gln Gly Gln Asp Leu
35 40 45

Thr Gly Gly Tyr Phe Asp Ala Gly Asp Phe Val Lys Phe Gly Phe Pro
50 55 60

Met Ala Tyr Thr Ala Thr Val Leu Ala Trp Gly Leu Ile Asp Phe Glu
65 70 75 80

Ala Gly Tyr Ser Ser Ala Gly Ala Leu Asp Asp Gly Arg Lys Ala Val
85 90 95

Lys Trp Ala Thr Asp Tyr Phe Ile Lys Ala His Thr Ser Gln Asn Glu
100 105 110

Phe Tyr Gly Gln Val Gly Gln Gly Asp Ala Asp His Ala Phe Trp Gly
115 120 125

Arg Pro Glu Asp Met Thr Met Ala Arg Pro Ala Tyr Lys Ile Asp Thr
130 135 140

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

Ala Ala Ser Ile Val Phe Arg Asn Val Asp Gly Thr Tyr Ser Asn Asn
165 170 175

Leu Leu Thr His Ala Arg Gln Leu Phe Asp Phe Ala Asn Asn Tyr Arg
180 185 190

Gly Lys Tyr Ser Asp Ser Ile Thr Asp Ala Arg Asn Phe Tyr Ala Ser
195 200 205

Ala Asp Tyr Arg Asp Glu Leu Val Trp Ala Ala Ala Trp Leu Tyr Arg
210 215 220

```
Ala Thr Asn Asp Asn Thr Tyr Leu Asn Thr Ala Glu Ser Leu Tyr Asp
225                 230                 235                 240

Glu Phe Gly Leu Gln Asn Trp Gly Gly Gly Leu Asn Trp Asp Ser Lys
                245                 250                 255

Val Ser Gly Val Gln Val Leu Leu Ala Lys Leu Thr Asn Lys Gln Ala
            260                 265                 270

Tyr Lys Asp Thr Val Gln Ser Tyr Val Asn Tyr Leu Ile Asn Asn Gln
        275                 280                 285

Gln Lys Thr Pro Lys Gly Leu Leu Tyr Ile Asp Met Trp Gly Thr Leu
    290                 295                 300

Arg His Ala Ala Asn Ala Ala Phe Ile Met Leu Glu Ala Ala Glu Leu
305                 310                 315                 320

Gly Leu Ser Ala Ser Ser Tyr Arg Gln Phe Ala Gln Thr Gln Ile Asp
                325                 330                 335

Tyr Ala Leu Gly Asp Gly Gly Arg Ser Phe Val Cys Gly Phe Gly Ser
            340                 345                 350

Asn Pro Pro Thr Arg Pro His His Arg Ser Ser Ser Cys Pro Pro Ala
        355                 360                 365

Pro Ala Thr Cys Asp Trp Asn Thr Phe Asn Ser Pro Asp Pro Asn Tyr
    370                 375                 380

His Val Leu Ser Gly Ala Leu Val Gly Gly Pro Asp Gln Asn Asp Asn
385                 390                 395                 400

Tyr Val Asp Asp Arg Ser Asp Tyr Val His Asn Glu Val Ala Thr Asp
                405                 410                 415

Tyr Asn Ala Gly Phe Gln Ser Ala Leu Ala Ala Leu Val Ala Leu Gly
            420                 425                 430

Tyr
```

<210> 5
<211> 194
<212> Білок
<213> Невідомо

<220>
<223> Кінцева GH11 із невідомої бактерії, амінокислотна.

EU591743

<400> 5

Met Ala Gln Thr Cys Leu Thr Ser Pro Gln Thr Gly Phe His Asn Gly
1 5 10 15

Phe Phe Tyr Ser Phe Trp Lys Asp Ser Pro Gly Thr Val Asn Phe Cys
20 25 30

Leu Leu Glu Gly Gly Arg Tyr Thr Ser Asn Trp Ser Gly Ile Asn Asn
35 40 45

Trp Val Gly Gly Lys Gly Trp Gln Thr Gly Ser Arg Arg Asn Ile Thr
50 55 60

Tyr Ser Gly Ser Phe Asn Thr Pro Gly Asn Gly Tyr Leu Ala Leu Tyr
65 70 75 80

Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr Tyr Val Val Asp Ser Trp
85 90 95

Gly Ser Trp Arg Pro Pro Gly Ser Asp Gly Thr Phe Leu Gly Thr Val
100 105 110

Asn Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Ala Gln Arg Val Asn
115 120 125

Ala Pro Ser Ile Ile Gly Asn Ala Thr Phe Tyr Gln Tyr Trp Ser Val
130 135 140

Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr Thr Gly Asn His Phe
145 150 155 160

Asp Ala Trp Ala Ser Val Gly Leu Asn Leu Gly Thr His Asn Tyr Gln
165 170 175

Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser Ser Asp Ile Thr
180 185 190

Val Ser

<210> 6
<211> 207
<212> PRT
<213> Thermomyces lanuginosus, AMINOKWASOWA, O43097

<400> 6

Met Phe Pro Ala Gly Asn Ala Thr Glu Leu Glu Lys Arg Gln Thr Thr

```
1               5                   10                  15

Pro Asn Ser Glu Gly Trp His Asp Gly Tyr Tyr Tyr Ser Trp Trp Ser
            20                  25                  30

Asp Gly Gly Ala Gln Ala Thr Tyr Thr Asn Leu Glu Gly Gly Thr Tyr
        35                  40                  45

Glu Ile Ser Trp Gly Asp Gly Gly Asn Leu Val Gly Gly Lys Gly Trp
    50                  55                  60

Asn Pro Gly Leu Asn Ala Arg Ala Ile His Phe Glu Gly Val Tyr Gln
65                  70                  75                  80

Pro Asn Gly Asn Ser Tyr Leu Ala Val Tyr Gly Trp Thr Arg Asn Pro
                85                  90                  95

Leu Val Glu Tyr Tyr Ile Val Glu Asn Phe Gly Thr Tyr Asp Pro Ser
            100                 105                 110

Ser Gly Ala Thr Asp Leu Gly Thr Val Glu Cys Asp Gly Ser Ile Tyr
        115                 120                 125

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

Gln Thr Phe Asp Gln Tyr Trp Ser Val Arg Gln Asp Lys Arg Thr Ser
145                 150                 155                 160

Gly Thr Val Gln Thr Gly Cys His Phe Asp Ala Trp Ala Arg Ala Gly
                165                 170                 175

Leu Asn Val Asn Gly Asp His Tyr Tyr Gln Ile Val Ala Thr Glu Gly
            180                 185                 190

Tyr Phe Ser Ser Gly Tyr Ala Arg Ile Thr Val Ala Asp Val Gly
        195                 200                 205
```

<210> 7
<211> 761
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, модифікована
інтеіном мікобактерій, амінокислотна, P77853:T134-100-101

<400> 7

```
Met Gln Thr Ser Ile Thr Leu Thr Ser Asn Ala Ser Gly Thr Phe Asp
1               5                   10                  15
```

```
Gly Tyr Tyr Tyr Glu Leu Trp Lys Asp Thr Gly Asn Thr Thr Met Thr
            20                  25                  30

Val Tyr Thr Gln Gly Arg Phe Ser Cys Gln Trp Ser Asn Ile Asn Asn
        35                  40                  45

Ala Leu Phe Arg Thr Gly Lys Lys Tyr Asn Gln Asn Trp Gln Ser Leu
    50                  55                  60

Gly Thr Ile Arg Ile Thr Tyr Ser Ala Thr Tyr Asn Pro Asn Gly Asn
65                  70                  75                  80

Ser Tyr Leu Cys Ile Tyr Gly Trp Ser Thr Asn Pro Leu Val Glu Phe
                85                  90                  95

Tyr Ile Val Glu Ser Trp Gly Asn Trp Arg Pro Pro Gly Ala Cys Leu
            100                 105                 110

Ala Glu Gly Ser Leu Val Leu Asp Ala Ala Thr Gly Gln Arg Val Pro
        115                 120                 125

Ile Glu Lys Val Arg Pro Gly Met Glu Val Phe Ser Leu Gly Pro Asp
    130                 135                 140

Tyr Arg Leu Tyr Arg Val Pro Val Leu Glu Val Leu Glu Ser Gly Val
145                 150                 155                 160

Arg Glu Val Val Arg Leu Arg Thr Arg Ser Gly Arg Thr Leu Val Leu
                165                 170                 175

Thr Pro Asp His Pro Leu Leu Thr Pro Glu Gly Trp Lys Pro Leu Cys
            180                 185                 190

Asp Leu Pro Leu Gly Thr Pro Ile Ala Val Pro Ala Glu Leu Pro Val
        195                 200                 205

Ala Gly His Leu Ala Pro Pro Glu Glu Arg Val Thr Leu Leu Ala Leu
    210                 215                 220

Leu Leu Gly Asp Gly Asn Thr Lys Leu Ser Gly Arg Arg Gly Thr Arg
225                 230                 235                 240

Pro Asn Ala Phe Phe Tyr Ser Lys Asn Pro Glu Leu Leu Ala Ala Tyr
                245                 250                 255

Arg Arg Cys Ala Glu Ala Leu Gly Ala Lys Val Lys Ala Tyr Val His
            260                 265                 270
```

```
Pro Thr Thr Gly Val Val Thr Leu Ala Thr Leu Ala Pro Arg Pro Gly
        275                 280                 285

Ala Gln Asp Pro Val Lys Arg Leu Val Val Glu Ala Gly Met Val Ala
    290                 295                 300

Lys Ala Glu Glu Lys Arg Val Pro Glu Glu Val Phe Arg Tyr Arg Arg
305                 310                 315                 320

Glu Ala Leu Ala Leu Phe Leu Gly Arg Leu Phe Ser Thr Asp Gly Ser
                325                 330                 335

Val Glu Lys Lys Arg Ile Ser Tyr Ser Ser Ala Ser Leu Gly Leu Ala
            340                 345                 350

Gln Asp Val Ala His Leu Leu Leu Arg Leu Gly Ile Thr Ser Gln Leu
        355                 360                 365

Arg Ser Arg Gly Pro Arg Ala His Glu Val Leu Ile Ser Gly Arg Glu
    370                 375                 380

Asp Ile Leu Arg Phe Ala Glu Leu Ile Gly Pro Tyr Leu Leu Gly Ala
385                 390                 395                 400

Lys Arg Glu Arg Leu Ala Ala Leu Glu Ala Glu Ala Arg Arg Arg Leu
                405                 410                 415

Pro Gly Gln Gly Trp His Leu Arg Leu Val Leu Pro Ala Val Ala Tyr
            420                 425                 430

Arg Val Gly Glu Ala Glu Arg Arg Ser Gly Phe Ser Trp Ser Glu Ala
        435                 440                 445

Gly Arg Arg Val Ala Val Ala Gly Ser Cys Leu Ser Ser Gly Leu Asn
    450                 455                 460

Leu Lys Leu Pro Arg Arg Tyr Leu Ser Arg His Arg Leu Ser Leu Leu
465                 470                 475                 480

Gly Glu Ala Phe Ala Asp Pro Gly Leu Glu Ala Leu Ala Glu Gly Gln
                485                 490                 495

Val Leu Trp Asp Pro Ile Val Ala Val Glu Pro Ala Gly Lys Ala Arg
            500                 505                 510

Thr Phe Asp Leu Arg Val Pro Pro Phe Ala Asn Phe Val Ser Glu Asp
        515                 520                 525
```

```
Leu Val Val His Asn Thr Val Pro Leu Gly Gln Val Thr Ile Asp Gly
    530                 535                 540

Gly Thr Tyr Asp Ile Tyr Arg Thr Thr Arg Val Asn Gln Pro Ser Ile
545                 550                 555                 560

Val Gly Thr Ala Thr Phe Asp Gln Tyr Trp Ser Val Arg Thr Ser Lys
                565                 570                 575

Arg Thr Ser Gly Thr Val Thr Val Thr Asp His Phe Arg Ala Trp Ala
            580                 585                 590

Asn Arg Gly Leu Asn Leu Gly Thr Ile Asp Gln Ile Thr Leu Cys Val
        595                 600                 605

Glu Gly Tyr Gln Ser Ser Gly Ser Ala Asn Ile Thr Gln Asn Thr Phe
    610                 615                 620

Ser Gln Gly Ser Ser Ser Gly Ser Ser Gly Gly Ser Ser Gly Ser Thr
625                 630                 635                 640

Thr Thr Thr Arg Ile Glu Cys Glu Asn Met Ser Leu Ser Gly Pro Tyr
                645                 650                 655

Val Ser Arg Ile Thr Asn Pro Phe Asn Gly Ile Ala Leu Tyr Ala Asn
            660                 665                 670

Gly Asp Thr Ala Arg Ala Thr Val Asn Phe Pro Ala Ser Arg Asn Tyr
        675                 680                 685

Asn Phe Arg Leu Arg Gly Cys Gly Asn Asn Asn Asn Leu Ala Arg Val
    690                 695                 700

Asp Leu Arg Ile Asp Gly Arg Thr Val Gly Thr Phe Tyr Tyr Gln Gly
705                 710                 715                 720

Thr Tyr Pro Trp Glu Ala Pro Ile Asp Asn Val Tyr Val Ser Ala Gly
                725                 730                 735

Ser His Thr Val Glu Ile Thr Val Thr Ala Asp Asn Gly Thr Trp Asp
            740                 745                 750

Val Tyr Ala Asp Tyr Leu Val Ile Gln
        755                 760
```

<210> 8
<211> 760

<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, модифікована
інтеїном комамаєа, амінокислотна P77853:S158-30-108-35

<400> 8

Met Gln Thr Ser Ile Thr Leu Thr Ser Asn Ala Ser Gly Thr Phe Asp
1 5 10 15

Gly Tyr Asn Tyr Glu Leu Trp Lys Asp Thr Gly Asn Thr Thr Met Thr
20 25 30

Val Tyr Thr Gln Gly Arg Phe Ser Cys Gln Trp Ser Asn Ile Asn Asn
35 40 45

Ala Leu Phe Arg Thr Gly Lys Lys Tyr Asn Gln Asn Trp Gln Ser Leu
50 55 60

Gly Thr Ile Arg Ile Thr Tyr Ser Ala Thr Tyr Asn Pro Asn Gly Asn
65 70 75 80

Ser Tyr Met Cys Ile Tyr Gly Trp Ser Thr Asn Pro Leu Val Glu Phe
85 90 95

Tyr Ile Val Glu Ser Trp Gly Asn Trp Arg Pro Pro Gly Ala Thr Ser
100 105 110

Leu Gly Gln Val Thr Ile Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr
115 120 125

Thr Arg Val Asn Gln Pro Cys Leu Ala Glu Gly Ser Leu Val Leu Asp
130 135 140

Ala Ala Thr Gly Gln Arg Val Pro Ile Glu Lys Val Arg Pro Gly Met
145 150 155 160

Glu Val Phe Ser Leu Gly Pro Asp Tyr Arg Leu Tyr Arg Val Pro Val
165 170 175

Leu Glu Val Leu Glu Ser Gly Val Gly Glu Val Val Arg Leu Arg Thr
180 185 190

Arg Ser Gly Arg Thr Leu Val Leu Thr Pro Glu His Pro Leu Leu Thr
195 200 205

Pro Glu Gly Trp Lys Pro Leu Cys Asp Leu Pro Leu Gly Thr Pro Ile
210 215 220

```
Ala Val Pro Ala Glu Leu Pro Val Ala Gly His Leu Ala Pro Pro Glu
225                 230                 235                 240

Glu Arg Val Met Leu Leu Ala Leu Leu Leu Gly Asp Gly Asn Thr Lys
                245                 250                 255

Leu Ser Gly Arg Arg Gly Thr Arg Pro Asn Ala Phe Phe Tyr Ser Lys
            260                 265                 270

Asp Ser Glu Leu Leu Ala Ala Tyr Arg Arg Cys Ala Glu Ala Leu Gly
        275                 280                 285

Ala Lys Val Lys Ala Tyr Val His Pro Thr Thr Gly Val Val Thr Leu
    290                 295                 300

Ala Thr Leu Ala Pro Arg Pro Gly Ala Gln Asp Pro Val Lys Arg Leu
305                 310                 315                 320

Val Val Glu Ala Gly Met Val Ala Lys Val Glu Glu Lys Arg Val Pro
                325                 330                 335

Glu Glu Val Phe Arg Tyr Arg Arg Glu Ala Leu Ala Leu Phe Leu Gly
            340                 345                 350

Arg Leu Phe Ser Thr Asp Gly Ser Val Glu Lys Lys Arg Ile Ser Tyr
        355                 360                 365

Ser Ser Ala Ser Leu Gly Leu Ala Gln Asp Val Ala His Leu Leu Leu
    370                 375                 380

Arg Leu Gly Ile Thr Ser Gln Leu Arg Ser Arg Gly Pro Arg Ala His
385                 390                 395                 400

Lys Val Leu Ile Ser Gly Arg Glu Asp Ile Leu Arg Phe Ala Glu Leu
                405                 410                 415

Ile Gly Pro Tyr Leu Leu Gly Ala Lys Arg Glu Arg Leu Ala Ala Leu
            420                 425                 430

Glu Ala Glu Ala Arg Arg Arg Leu Pro Gly Gln Gly Trp His Leu Arg
        435                 440                 445

Pro Val Leu Pro Ala Val Ala Tyr Arg Val Ser Glu Ala Lys Arg Arg
    450                 455                 460

Ser Gly Phe Ser Trp Ser Glu Ala Gly Arg Arg Val Ala Val Ala Gly
465                 470                 475                 480
```

```
Ser Cys Leu Ser Ser Gly Leu Asn Leu Lys Leu Pro Arg Arg Tyr Leu
                485                 490                 495

Ser Arg His Arg Leu Ser Leu Leu Gly Glu Ala Phe Ala Asp Pro Gly
            500                 505                 510

Leu Glu Ala Leu Ala Glu Gly Gln Val Leu Trp Asp Pro Ile Val Ala
        515                 520                 525

Val Glu Pro Ala Gly Lys Ala Arg Thr Phe Asp Leu Arg Val Pro Pro
    530                 535                 540

Phe Ala Asn Phe Val Ser Glu Asp Leu Val Val His Asn Ser Ile Val
545                 550                 555                 560

Gly Thr Ala Thr Phe Asp Gln Tyr Trp Ser Val Arg Thr Ser Lys Arg
                565                 570                 575

Thr Ser Gly Thr Val Thr Val Thr Asp His Phe Arg Ala Trp Ala Asn
            580                 585                 590

Arg Gly Leu Asn Leu Gly Thr Ile Asp Gln Ile Thr Leu Cys Val Glu
        595                 600                 605

Gly Tyr Gln Ser Ser Gly Ser Ala Asn Ile Thr Gln Asn Thr Phe Ser
    610                 615                 620

Gln Gly Ser Ser Ser Gly Ser Ser Gly Gly Ser Ser Gly Ser Thr Thr
625                 630                 635                 640

Thr Thr Arg Ile Glu Cys Glu Asn Met Ser Leu Ser Gly Pro Tyr Val
                645                 650                 655

Ser Arg Ile Thr Asn Pro Phe Asn Gly Ile Ala Leu Tyr Ala Asn Gly
            660                 665                 670

Asp Thr Ala Arg Ala Thr Val Asn Phe Pro Ala Ser Arg Asn Tyr Asn
        675                 680                 685

Phe Arg Leu Arg Gly Cys Gly Asn Ser Asn Asn Leu Ala Arg Val Asp
    690                 695                 700

Leu Arg Ile Asp Gly Arg Thr Val Gly Thr Phe Tyr Tyr Gln Gly Thr
705                 710                 715                 720

Tyr Pro Trp Glu Ala Pro Ile Asp Asn Val Tyr Val Ser Ala Gly Ser
                725                 730                 735
```

His Thr Val Glu Ile Thr Val Thr Ala Asp Asn Gly Thr Arg Asp Val
740 745 750

Tyr Ala Asp Tyr Leu Val Ile Gln
755 760

<210> 9
<211> 244
<212> БЕЛОК
<213> Rhodothermus marinus, аминокислотная, 033897

<400> 9

Met Cys Asp Trp Leu Phe Pro Asp Gly Asp Asn Gly Lys Glu Pro Glu
1 5 10 15

Pro Glu Pro Glu Pro Thr Val Glu Leu Cys Gly Arg Trp Asp Ala Arg
20 25 30

Asp Val Ala Gly Gly Arg Tyr Arg Val Ile Asn Asn Val Trp Gly Ala
35 40 45

Glu Thr Ala Gln Cys Ile Glu Val Gly Leu Glu Thr Gly Asn Phe Thr
50 55 60

Ile Thr Arg Ala Asp His Asp Asn Gly Asn Asn Val Ala Ala Tyr Pro
65 70 75 80

Ala Ile Tyr Phe Gly Cys His Trp Ala Pro Ala Arg Ala Ile Arg Asp
85 90 95

Cys Ala Ala Arg Ala Gly Ala Val Arg Arg Ala His Glu Leu Asp Val
100 105 110

Thr Pro Ile Thr Thr Gly Arg Trp Asn Ala Ala Tyr Asp Ile Trp Phe
115 120 125

Ser Pro Val Thr Asn Ser Gly Asn Gly Tyr Ser Gly Gly Ala Glu Leu
130 135 140

Met Ile Trp Leu Asn Trp Asn Gly Gly Val Met Pro Gly Gly Ser Arg
145 150 155 160

Val Ala Thr Val Glu Leu Ala Gly Ala Thr Trp Glu Val Trp Tyr Ala
165 170 175

Asp Trp Asp Trp Asn Tyr Ile Ala Tyr Arg Arg Thr Thr Pro Thr Thr
180 185 190

Ser Val Ser Glu Leu Asp Leu Lys Ala Phe Ile Asp Asp Ala Val Ala
195 200 205

Arg Gly Tyr Ile Arg Pro Glu Trp Tyr Leu His Ala Val Glu Thr Gly
210 215 220

Phe Glu Leu Trp Glu Gly Gly Ala Gly Leu Arg Thr Ala Asp Phe Ser
225 230 235 240

Val Thr Val Gln

<210> 10
<211> 869
<212> Білок
<213> Clostridium thermocellum, амінокислотна, Q68438

<400> 10

Met Leu Glu Asp Lys Ser Pro Lys Leu Pro Asp Tyr Lys Asn Asp Leu
1 5 10 15

Leu Tyr Glu Arg Thr Phe Asp Glu Gly Leu Cys Phe Pro Trp His Thr
20 25 30

Cys Glu Asp Ser Gly Gly Lys Cys Asp Phe Ala Val Val Asp Val Pro
35 40 45

Gly Glu Pro Gly Asn Lys Ala Phe Arg Leu Thr Val Ile Asp Lys Gly
50 55 60

Gln Asn Lys Trp Ser Val Gln Met Arg His Arg Gly Ile Thr Leu Glu
65 70 75 80

Gln Gly His Thr Tyr Thr Val Arg Phe Thr Ile Trp Ser Asp Lys Ser
85 90 95

Cys Arg Val Tyr Ala Lys Ile Gly Gln Met Gly Glu Pro Tyr Thr Glu
100 105 110

Tyr Trp Asn Asn Asn Trp Asn Pro Phe Asn Leu Thr Pro Gly Gln Lys
115 120 125

Leu Thr Val Glu Gln Asn Phe Thr Met Asn Tyr Pro Thr Asp Asp Thr
130 135 140

Cys Glu Phe Thr Phe His Leu Gly Gly Glu Leu Ala Ala Gly Thr Pro
145 150 155 160

Tyr Tyr Val Tyr Leu Asp Asp Val Ser Leu Tyr Asp Pro Arg Phe Val

165 170 175

Lys Pro Val Glu Tyr Val Leu Pro Gln Pro Asp Val Arg Val Asn Gln
180 185 190

Val Gly Tyr Leu Pro Phe Ala Lys Lys Tyr Ala Thr Val Val Ser Ser
195 200 205

Ser Thr Ser Pro Leu Lys Trp Gln Leu Leu Asn Ser Ala Asn Gln Val
210 215 220

Val Leu Glu Gly Asn Thr Ile Pro Lys Gly Leu Asp Lys Asp Ser Gln
225 230 235 240

Asp Tyr Val His Trp Ile Asp Phe Ser Asn Phe Lys Thr Glu Gly Lys
245 250 255

Gly Tyr Tyr Phe Lys Leu Pro Thr Val Asn Ser Asp Thr Asn Tyr Ser
260 265 270

His Pro Phe Asp Ile Ser Ala Asp Ile Tyr Ser Lys Met Lys Phe Asp
275 280 285

Ala Leu Ala Phe Phe Tyr His Lys Arg Ser Gly Ile Pro Ile Glu Met
290 295 300

Pro Tyr Ala Gly Gly Glu Gln Trp Thr Arg Pro Ala Gly His Ile Gly
305 310 315 320

Val Ala Pro Asn Lys Gly Asp Thr Asn Val Pro Thr Trp Pro Gln Asp
325 330 335

Asp Glu Tyr Ala Gly Arg Pro Gln Lys Tyr Tyr Thr Lys Asp Val Thr
340 345 350

Gly Gly Trp Tyr Asp Ala Gly Asp His Gly Lys Tyr Val Val Asn Gly
355 360 365

Gly Ile Ala Val Trp Thr Leu Met Asn Met Tyr Glu Arg Ala Lys Ile
370 375 380

Arg Gly Ile Ala Asn Gln Gly Ala Tyr Lys Asp Gly Gly Met Asn Ile
385 390 395 400

Pro Glu Arg Asn Asn Gly Tyr Pro Asp Ile Leu Asp Glu Ala Arg Trp
405 410 415

Glu Ile Glu Phe Phe Lys Lys Met Gln Val Thr Glu Lys Glu Asp Pro

420 425 430

Ser Ile Ala Gly Met Val His His Lys Ile His Asp Phe Arg Trp Thr
435 440 445

Ala Leu Gly Met Leu Pro His Glu Asp Pro Gln Pro Arg Tyr Leu Arg
450 455 460

Pro Val Ser Thr Ala Ala Thr Leu Asn Phe Ala Ala Thr Leu Ala Gln
465 470 475 480

Ser Ala Arg Leu Trp Lys Asp Tyr Asp Pro Thr Phe Ala Ala Asp Cys
485 490 495

Leu Glu Lys Ala Glu Ile Ala Trp Gln Ala Ala Leu Lys His Pro Asp
500 505 510

Ile Tyr Ala Glu Tyr Thr Pro Gly Ser Gly Gly Pro Gly Gly Gly Pro
515 520 525

Tyr Asn Asp Asp Tyr Val Gly Asp Glu Phe Tyr Trp Ala Ala Cys Glu
530 535 540

Leu Tyr Val Thr Thr Gly Lys Asp Glu Tyr Lys Asn Tyr Leu Met Asn
545 550 555 560

Ser Pro His Tyr Leu Glu Met Pro Ala Lys Met Gly Glu Asn Gly Gly
565 570 575

Ala Asn Gly Glu Asp Asn Gly Leu Trp Gly Cys Phe Thr Trp Gly Thr
580 585 590

Thr Gln Gly Leu Gly Thr Ile Thr Leu Ala Leu Val Glu Asn Gly Leu
595 600 605

Pro Ser Ala Asp Ile Gln Lys Ala Arg Asn Asn Ile Ala Lys Ala Ala
610 615 620

Asp Lys Trp Leu Glu Asn Ile Glu Glu Gln Gly Tyr Arg Leu Pro Ile
625 630 635 640

Lys Gln Ala Glu Asp Glu Arg Gly Gly Tyr Pro Trp Gly Ser Asn Ser
645 650 655

Phe Ile Leu Asn Gln Met Ile Val Met Gly Tyr Ala Tyr Asp Phe Thr
660 665 670

Gly Asn Ser Lys Tyr Leu Asp Gly Met Gln Asp Gly Met Ser Tyr Leu

675 680 685

Leu Gly Arg Asn Gly Leu Asp Gln Ser Tyr Val Thr Gly Tyr Gly Glu
690 695 700

Arg Pro Leu Gln Asn Pro His Asp Arg Phe Trp Thr Pro Gln Thr Ser
705 710 715 720

Lys Lys Phe Pro Ala Pro Pro Pro Gly Ile Ile Ala Gly Gly Pro Asn
725 730 735

Ser Arg Phe Glu Asp Pro Thr Ile Thr Ala Ala Val Lys Lys Asp Thr
740 745 750

Pro Pro Gln Lys Cys Tyr Ile Asp His Thr Asp Ser Trp Ser Thr Asn
755 760 765

Glu Ile Thr Ile Asn Trp Asn Ala Pro Phe Ala Trp Val Thr Ala Tyr
770 775 780

Leu Asp Glu Ile Asp Leu Ile Thr Pro Pro Gly Gly Val Asp Pro Glu
785 790 795 800

Glu Pro Glu Val Ile Tyr Gly Asp Cys Asn Gly Asp Gly Lys Val Asn
805 810 815

Ser Thr Asp Ala Val Ala Leu Lys Arg Tyr Ile Leu Arg Ser Gly Ile
820 825 830

Ser Ile Asn Thr Asp Asn Ala Asp Val Asn Ala Asp Gly Arg Val Asn
835 840 845

Ser Thr Asp Leu Ala Ile Leu Lys Arg Tyr Ile Leu Lys Glu Ile Asp
850 855 860

Val Leu Pro His Lys
865

<210> 11
<211> 522
<212> [illegible]
<213> Acidothermus cellulolyticus, [illegible], P54583

<400> 11

Met Ala Gly Gly Gly Tyr Trp His Thr Ser Gly Arg Glu Ile Leu Asp
1 5 10 15

Ala Asn Asn Val Pro Val Arg Ile Ala Gly Ile Asn Trp Phe Gly Phe
20 25 30

```
Glu Thr Cys Asn Tyr Val Val His Gly Leu Trp Ser Arg Asp Tyr Arg
        35                  40                  45

Ser Met Leu Asp Gln Ile Lys Ser Leu Gly Tyr Asn Thr Ile Arg Leu
    50                  55                  60

Pro Tyr Ser Asp Asp Ile Leu Lys Pro Gly Thr Met Pro Asn Ser Ile
65                  70                  75                  80

Asn Phe Tyr Gln Met Asn Gln Asp Leu Gln Gly Leu Thr Ser Leu Gln
                85                  90                  95

Val Met Asp Lys Ile Val Ala Tyr Ala Gly Gln Ile Gly Leu Arg Ile
            100                 105                 110

Ile Leu Asp Arg His Arg Pro Asp Cys Ser Gly Gln Ser Ala Leu Trp
        115                 120                 125

Tyr Thr Ser Ser Val Ser Glu Ala Thr Trp Ile Ser Asp Leu Gln Ala
    130                 135                 140

Leu Ala Gln Arg Tyr Lys Gly Asn Pro Thr Val Val Gly Phe Asp Leu
145                 150                 155                 160

His Asn Glu Pro His Asp Pro Ala Cys Trp Gly Cys Gly Asp Pro Ser
                165                 170                 175

Ile Asp Trp Arg Leu Ala Ala Glu Arg Ala Gly Asn Ala Val Leu Ser
            180                 185                 190

Val Asn Pro Asn Leu Leu Ile Phe Val Glu Gly Val Gln Ser Tyr Asn
        195                 200                 205

Gly Asp Ser Tyr Trp Trp Gly Gly Asn Leu Gln Gly Ala Gly Gln Tyr
    210                 215                 220

Pro Val Val Leu Asn Val Pro Asn Arg Leu Val Tyr Ser Ala His Asp
225                 230                 235                 240

Tyr Ala Thr Ser Val Tyr Pro Gln Thr Trp Phe Ser Asp Pro Thr Phe
                245                 250                 255

Pro Asn Asn Met Pro Gly Ile Trp Asn Lys Asn Trp Gly Tyr Leu Phe
            260                 265                 270

Asn Gln Asn Ile Ala Pro Val Trp Leu Gly Glu Phe Gly Thr Thr Leu
        275                 280                 285
```

```
Gln Ser Thr Thr Asp Gln Thr Trp Leu Lys Thr Leu Val Gln Tyr Leu
    290                 295                 300

Arg Pro Thr Ala Gln Tyr Gly Ala Asp Ser Phe Gln Trp Thr Phe Trp
305                 310                 315                 320

Ser Trp Asn Pro Asp Ser Gly Asp Thr Gly Gly Ile Leu Lys Asp Asp
                325                 330                 335

Trp Gln Thr Val Asp Thr Val Lys Asp Gly Tyr Leu Ala Pro Ile Lys
            340                 345                 350

Ser Ser Ile Phe Asp Pro Val Gly Ala Ser Ala Ser Pro Ser Ser Gln
        355                 360                 365

Pro Ser Pro Ser Val Ser Pro Ser Pro Ser Pro Ser Pro Ser Ala Ser
    370                 375                 380

Arg Thr Pro Thr Pro Thr Pro Thr Pro Thr Ala Ser Pro Thr Pro Thr
385                 390                 395                 400

Leu Thr Pro Thr Ala Thr Pro Thr Pro Thr Ala Ser Pro Thr Pro Ser
                405                 410                 415

Pro Thr Ala Ala Ser Gly Ala Arg Cys Thr Ala Ser Tyr Gln Val Asn
            420                 425                 430

Ser Asp Trp Gly Asn Gly Phe Thr Val Thr Val Ala Val Thr Asn Ser
        435                 440                 445

Gly Ser Val Ala Thr Lys Thr Trp Thr Val Ser Trp Thr Phe Gly Gly
    450                 455                 460

Asn Gln Thr Ile Thr Asn Ser Trp Asn Ala Ala Val Thr Gln Asn Gly
465                 470                 475                 480

Gln Ser Val Thr Ala Arg Asn Met Ser Tyr Asn Asn Val Ile Gln Pro
                485                 490                 495

Gly Gln Asn Thr Thr Phe Gly Phe Gln Ala Ser Tyr Thr Gly Ser Asn
            500                 505                 510

Ala Ala Pro Thr Val Ala Cys Ala Ala Ser
        515                 520
```

<210> 12
<211> 513

<212> Білок
<213> Penicillium funiculosum, амінокислотна. RD22308

<400> 12

Met Gln Gln Ile Gly Thr Tyr Thr Ala Glu Thr His Pro Ser Leu Ser
1 5 10 15

Trp Ser Thr Cys Lys Ser Gly Gly Ser Cys Thr Thr Asn Ser Gly Ala
20 25 30

Ile Thr Leu Asp Ala Asn Trp Arg Trp Val His Gly Val Asn Thr Ser
35 40 45

Thr Asn Cys Tyr Thr Gly Asn Thr Trp Asn Thr Ala Ile Cys Asp Thr
50 55 60

Asp Ala Ser Cys Ala Gln Asp Cys Ala Leu Asp Gly Ala Asp Tyr Ser
65 70 75 80

Gly Thr Tyr Gly Ile Thr Thr Ser Gly Asn Ser Leu Arg Leu Asn Phe
85 90 95

Val Thr Gly Ser Asn Val Gly Ser Arg Thr Tyr Leu Met Ala Asp Asn
100 105 110

Thr His Tyr Gln Ile Phe Asp Leu Leu Asn Gln Glu Phe Thr Phe Thr
115 120 125

Val Asp Val Ser His Leu Pro Cys Gly Leu Asn Gly Ala Leu Tyr Phe
130 135 140

Val Thr Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Asn Asn Lys
145 150 155 160

Ala Gly Ala Gln Tyr Gly Val Gly Tyr Cys Asp Ser Gln Cys Pro Arg
165 170 175

Asp Leu Lys Phe Ile Ala Gly Gln Ala Asn Val Glu Gly Trp Thr Pro
180 185 190

Ser Ser Asn Asn Ala Asn Thr Gly Leu Gly Asn His Gly Ala Cys Cys
195 200 205

Ala Glu Leu Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala Leu Thr
210 215 220

Pro His Pro Cys Asp Thr Pro Gly Leu Ser Val Cys Thr Thr Asp Ala
225 230 235 240

```
Cys Gly Gly Thr Tyr Ser Ser Asp Arg Tyr Ala Gly Thr Cys Asp Pro
                245                 250                 255

Asp Gly Cys Asp Phe Asn Pro Tyr Arg Leu Gly Val Thr Asp Phe Tyr
            260                 265                 270

Gly Ser Gly Lys Thr Val Asp Thr Thr Lys Pro Ile Thr Val Val Thr
        275                 280                 285

Gln Phe Val Thr Asp Asp Gly Thr Ser Thr Gly Thr Leu Ser Glu Ile
    290                 295                 300

Arg Arg Tyr Tyr Val Gln Asn Gly Val Val Ile Pro Gln Pro Ser Ser
305                 310                 315                 320

Lys Ile Ser Gly Val Ser Gly Asn Val Ile Asn Ser Asp Phe Cys Asp
                325                 330                 335

Ala Glu Ile Ser Thr Phe Gly Glu Thr Ala Ser Phe Ser Lys His Gly
            340                 345                 350

Gly Leu Ala Lys Met Gly Ala Gly Met Glu Ala Gly Met Val Leu Val
        355                 360                 365

Met Ser Leu Trp Asp Asp Tyr Ser Val Asn Met Leu Trp Leu Asp Ser
    370                 375                 380

Thr Tyr Pro Thr Asn Ala Thr Gly Thr Pro Gly Ala Ala Arg Gly Ser
385                 390                 395                 400

Cys Pro Thr Thr Ser Gly Asp Pro Lys Thr Val Glu Ser Gln Ser Gly
                405                 410                 415

Ser Ser Tyr Val Thr Phe Ser Asp Ile Arg Val Gly Pro Phe Asn Ser
            420                 425                 430

Thr Phe Ser Gly Gly Ser Ser Thr Gly Gly Ser Ser Thr Thr Thr Ala
        435                 440                 445

Ser Gly Thr Thr Thr Thr Lys Ala Ser Ser Thr Ser Thr Ser Ser Thr
    450                 455                 460

Ser Thr Gly Thr Gly Val Ala Ala His Trp Gly Gln Cys Gly Gly Gln
465                 470                 475                 480

Gly Trp Thr Gly Pro Thr Thr Cys Ala Ser Gly Thr Thr Cys Thr Val
                485                 490                 495
```

Val Asn Pro Tyr Tyr Ser Gln Cys Leu Asp Glu Leu Lys Ala Glu Ala
500 505 510

Lys

<210> 13
<211> 26
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, найближчий пептид, амінокислотна, BAASS

<400> 13

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Val
20 25

<210> 14
<211> 43
<212> Білок
<213> Hordeum vulgare, амінокислотна, HvAleSP

<400> 14

Met Ala His Ala Arg Val Leu Leu Leu Ala Leu Ala Val Leu Ala Thr
1 5 10 15

Ala Ala Val Ala Val Ala Ser Ser Ser Ser Phe Ala Asp Ser Asn Pro
20 25 30

Ile Arg Pro Val Thr Asp Arg Ala Ala Ser Thr
35 40

<210> 15
<211> 30
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, найближчий пептид, амінокислотна, PR1a

<400> 15

Met Gly Phe Val Leu Phe Ser Gln Leu Pro Ser Phe Leu Leu Val Ser
1 5 10 15

Thr Leu Leu Leu Phe Leu Val Ile Ser His Ser Cys Arg Ala
20 25 30

<210> 16
<211> 19
<212> Білок
<213> Zea mays, Найближчий пептид, амінокислотна, G2ein27ss-02

<400> 16

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
1 5 10 15

Ala Thr Ser

<210> 17
<211> 24
<212> Білок
<213> Oryza sativa, найближчий пептид, амінокислотна, GluB4SP

<400> 17

Met Ala Thr Ile Ala Phe Ser Arg Leu Ser Ile Tyr Phe Cys Val Leu
1 5 10 15

Leu Leu Cys His Gly Ser Met Ala
20

<210> 18
<211> 361
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, химерна, злита з найближчим пептидом, амінокислотна, BAASS:P77853

<400> 18

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Gln Thr Ser Ile Thr Leu Thr
20 25 30

Ser Asn Ala Ser Gly Thr Phe Asp Gly Tyr Tyr Tyr Glu Leu Trp Lys
35 40 45

Asp Thr Gly Asn Thr Thr Met Thr Val Tyr Thr Gln Gly Arg Phe Ser
50 55 60

Cys Gln Trp Ser Asn Ile Asn Asn Ala Leu Phe Arg Thr Gly Lys Lys
65 70 75 80

Tyr Asn Gln Asn Trp Gln Ser Leu Gly Thr Ile Arg Ile Thr Tyr Ser

85 90 95

Ala Thr Tyr Asn Pro Asn Gly Asn Ser Tyr Leu Cys Ile Tyr Gly Trp
100 105 110

Ser Thr Asn Pro Leu Val Glu Phe Tyr Ile Val Glu Ser Trp Gly Asn
115 120 125

Trp Arg Pro Pro Gly Ala Thr Ser Leu Gly Gln Val Thr Ile Asp Gly
130 135 140

Gly Thr Tyr Asp Ile Tyr Arg Thr Thr Arg Val Asn Gln Pro Ser Ile
145 150 155 160

Val Gly Thr Ala Thr Phe Asp Gln Tyr Trp Ser Val Arg Thr Ser Lys
165 170 175

Arg Thr Ser Gly Thr Val Thr Val Thr Asp His Phe Arg Ala Trp Ala
180 185 190

Asn Arg Gly Leu Asn Leu Gly Thr Ile Asp Gln Ile Thr Leu Cys Val
195 200 205

Glu Gly Tyr Gln Ser Ser Gly Ser Ala Asn Ile Thr Gln Asn Thr Phe
210 215 220

Ser Gln Gly Ser Ser Ser Gly Ser Ser Gly Gly Ser Ser Gly Ser Thr
225 230 235 240

Thr Thr Thr Arg Ile Glu Cys Glu Asn Met Ser Leu Ser Gly Pro Tyr
245 250 255

Val Ser Arg Ile Thr Asn Pro Phe Asn Gly Ile Ala Leu Tyr Ala Asn
260 265 270

Gly Asp Thr Ala Arg Ala Thr Val Asn Phe Pro Ala Ser Arg Asn Tyr
275 280 285

Asn Phe Arg Leu Arg Gly Cys Gly Asn Asn Asn Asn Leu Ala Arg Val
290 295 300

Asp Leu Arg Ile Asp Gly Arg Thr Val Gly Thr Phe Tyr Tyr Gln Gly
305 310 315 320

Thr Tyr Pro Trp Glu Ala Pro Ile Asp Asn Val Tyr Val Ser Ala Gly
325 330 335

Ser His Thr Val Glu Ile Thr Val Thr Ala Asp Asn Gly Thr Trp Asp

340 345 350

Val Tyr Ala Asp Tyr Leu Val Ile Gln
355 360

<210> 19
<211> 270
<212> БІЛОК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, ендоглюканаза, злита з целюлозним пептидом, амінокислотна, BAASS:033897

<400> 19

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Gln Val Cys Asp Trp Leu Phe
20 25 30

Pro Asp Gly Asp Asn Gly Lys Glu Pro Glu Pro Glu Pro Glu Pro Thr
35 40 45

Val Glu Leu Cys Gly Arg Trp Asp Ala Arg Asp Val Ala Gly Gly Arg
50 55 60

Tyr Arg Val Ile Asn Asn Val Trp Gly Ala Glu Thr Ala Gln Cys Ile
65 70 75 80

Glu Val Gly Leu Glu Thr Gly Asn Phe Thr Ile Thr Arg Ala Asp His
85 90 95

Asp Asn Gly Asn Asn Val Ala Ala Tyr Pro Ala Ile Tyr Phe Gly Cys
100 105 110

His Trp Ala Pro Ala Arg Ala Ile Arg Asp Cys Ala Ala Arg Ala Gly
115 120 125

Ala Val Arg Arg Ala His Glu Leu Asp Val Thr Pro Ile Thr Thr Gly
130 135 140

Arg Trp Asn Ala Ala Tyr Asp Ile Trp Phe Ser Pro Val Thr Asn Ser
145 150 155 160

Gly Asn Gly Tyr Ser Gly Gly Ala Glu Leu Met Ile Trp Leu Asn Trp
165 170 175

Asn Gly Gly Val Met Pro Gly Gly Ser Arg Val Ala Thr Val Glu Leu
180 185 190

Ala Gly Ala Thr Trp Glu Val Trp Tyr Ala Asp Trp Asp Trp Asn Tyr
195 200 205

Ile Ala Tyr Arg Arg Thr Thr Pro Thr Thr Ser Val Ser Glu Leu Asp
210 215 220

Leu Lys Ala Phe Ile Asp Asp Ala Val Ala Arg Gly Tyr Ile Arg Pro
225 230 235 240

Glu Trp Tyr Leu His Ala Val Glu Thr Gly Phe Glu Leu Trp Glu Gly
245 250 255

Gly Ala Gly Leu Arg Thr Ala Asp Phe Ser Val Thr Val Gln
260 265 270

<210> 20
<211> 475
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, ендоглюканаза, злита з найближчим пептидом, амінокислотна, HvAleSP:NtEGM

<400> 20

Met Ala His Ala Arg Val Leu Leu Leu Ala Leu Ala Val Leu Ala Thr
1 5 10 15

Ala Ala Val Ala Val Ala Ser Ser Ser Ser Phe Ala Asp Ser Asn Pro
20 25 30

Ile Arg Pro Val Thr Asp Arg Ala Ala Ser Thr Ala Tyr Asp Tyr Lys
35 40 45

Gln Val Leu Arg Asp Ser Leu Leu Phe Tyr Glu Ala Gln Arg Ser Gly
50 55 60

Arg Leu Pro Ala Asp Gln Lys Val Thr Trp Arg Lys Asp Ser Ala Leu
65 70 75 80

Asn Asp Gln Gly Asp Gln Gly Gln Asp Leu Thr Gly Gly Tyr Phe Asp
85 90 95

Ala Gly Asp Phe Val Lys Phe Gly Phe Pro Met Ala Tyr Thr Ala Thr
100 105 110

Val Leu Ala Trp Gly Leu Ile Asp Phe Glu Ala Gly Tyr Ser Ser Ala
115 120 125

```
Gly Ala Leu Asp Asp Gly Arg Lys Ala Val Lys Trp Ala Thr Asp Tyr
    130                 135                 140

Phe Ile Lys Ala His Thr Ser Gln Asn Glu Phe Tyr Gly Gln Val Gly
145                 150                 155                 160

Gln Gly Asp Ala Asp His Ala Phe Trp Gly Arg Pro Glu Asp Met Thr
                165                 170                 175

Met Ala Arg Pro Ala Tyr Lys Ile Asp Thr Ser Arg Pro Gly Ser Asp
            180                 185                 190

Leu Ala Gly Glu Thr Ala Ala Ala Leu Ala Ala Ala Ser Ile Val Phe
        195                 200                 205

Arg Asn Val Asp Gly Thr Tyr Ser Asn Asn Leu Leu Thr His Ala Arg
    210                 215                 220

Gln Leu Phe Asp Phe Ala Asn Asn Tyr Arg Gly Lys Tyr Ser Asp Ser
225                 230                 235                 240

Ile Thr Asp Ala Arg Asn Phe Tyr Ala Ser Ala Asp Tyr Arg Asp Glu
                245                 250                 255

Leu Val Trp Ala Ala Ala Trp Leu Tyr Arg Ala Thr Asn Asp Asn Thr
            260                 265                 270

Tyr Leu Asn Thr Ala Glu Ser Leu Tyr Asp Glu Phe Gly Leu Gln Asn
        275                 280                 285

Trp Gly Gly Gly Leu Asn Trp Asp Ser Lys Val Ser Gly Val Gln Val
    290                 295                 300

Leu Leu Ala Lys Leu Thr Asn Lys Gln Ala Tyr Lys Asp Thr Val Gln
305                 310                 315                 320

Ser Tyr Val Asn Tyr Leu Ile Asn Asn Gln Gln Lys Thr Pro Lys Gly
                325                 330                 335

Leu Leu Tyr Ile Asp Met Trp Gly Thr Leu Arg His Ala Ala Asn Ala
            340                 345                 350

Ala Phe Ile Met Leu Glu Ala Ala Glu Leu Gly Leu Ser Ala Ser Ser
        355                 360                 365

Tyr Arg Gln Phe Ala Gln Thr Gln Ile Asp Tyr Ala Leu Gly Asp Gly
    370                 375                 380
```

```
Gly Arg Ser Phe Val Cys Gly Phe Gly Ser Asn Pro Pro Thr Arg Pro
385                 390                 395                 400

His His Arg Ser Ser Ser Cys Pro Pro Ala Pro Ala Thr Cys Asp Trp
                405                 410                 415

Asn Thr Phe Asn Ser Pro Asp Pro Asn Tyr His Val Leu Ser Gly Ala
            420                 425                 430

Leu Val Gly Gly Pro Asp Gln Asn Asp Asn Tyr Val Asp Asp Arg Ser
        435                 440                 445

Asp Tyr Val His Asn Glu Val Ala Thr Asp Tyr Asn Ala Gly Phe Gln
    450                 455                 460

Ser Ala Leu Ala Ala Leu Val Ala Leu Gly Tyr
465                 470                 475


<210>  21
<211>  784
<212>  Білок
<213>  Штучна послідовність

<220>
<223>  Синтетична послідовність, модифікована інтеїном
       ксиланеза, злита з нанілежним пептидом, амінокислотна,
       BAASS:P77853:S158-30-108-35

<400>  31

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1               5                   10                  15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Gln Thr Ser Ile Thr Leu Thr
            20                  25                  30

Ser Asn Ala Ser Gly Thr Phe Asp Gly Tyr Asn Tyr Glu Leu Trp Lys
        35                  40                  45

Asp Thr Gly Asn Thr Thr Met Thr Val Tyr Thr Gln Gly Arg Phe Ser
    50                  55                  60

Cys Gln Trp Ser Asn Ile Asn Asn Ala Leu Phe Arg Thr Gly Lys Lys
65                  70                  75                  80

Tyr Asn Gln Asn Trp Gln Ser Leu Gly Thr Ile Arg Ile Thr Tyr Ser
                85                  90                  95

Ala Thr Tyr Asn Pro Asn Gly Asn Ser Tyr Met Cys Ile Tyr Gly Trp
            100                 105                 110
```

```
Ser Thr Asn Pro Leu Val Glu Phe Tyr Ile Val Glu Ser Trp Gly Asn
        115                 120                 125

Trp Arg Pro Pro Gly Ala Thr Ser Leu Gly Gln Val Thr Ile Asp Gly
    130                 135                 140

Gly Thr Tyr Asp Ile Tyr Arg Thr Thr Arg Val Asn Gln Pro Cys Leu
145                 150                 155                 160

Ala Glu Gly Ser Leu Val Leu Asp Ala Ala Thr Gly Gln Arg Val Pro
                165                 170                 175

Ile Glu Lys Val Arg Pro Gly Met Glu Val Phe Ser Leu Gly Pro Asp
            180                 185                 190

Tyr Arg Leu Tyr Arg Val Pro Val Leu Glu Val Leu Glu Ser Gly Val
        195                 200                 205

Gly Glu Val Val Arg Leu Arg Thr Arg Ser Gly Arg Thr Leu Val Leu
    210                 215                 220

Thr Pro Glu His Pro Leu Leu Thr Pro Glu Gly Trp Lys Pro Leu Cys
225                 230                 235                 240

Asp Leu Pro Leu Gly Thr Pro Ile Ala Val Pro Ala Glu Leu Pro Val
                245                 250                 255

Ala Gly His Leu Ala Pro Pro Glu Glu Arg Val Met Leu Leu Ala Leu
            260                 265                 270

Leu Leu Gly Asp Gly Asn Thr Lys Leu Ser Gly Arg Arg Gly Thr Arg
        275                 280                 285

Pro Asn Ala Phe Phe Tyr Ser Lys Asp Ser Glu Leu Leu Ala Ala Tyr
    290                 295                 300

Arg Arg Cys Ala Glu Ala Leu Gly Ala Lys Val Lys Ala Tyr Val His
305                 310                 315                 320

Pro Thr Thr Gly Val Val Thr Leu Ala Thr Leu Ala Pro Arg Pro Gly
                325                 330                 335

Ala Gln Asp Pro Val Lys Arg Leu Val Val Glu Ala Gly Met Val Ala
            340                 345                 350

Lys Val Glu Glu Lys Arg Val Pro Glu Glu Val Phe Arg Tyr Arg Arg
        355                 360                 365
```

```
Glu Ala Leu Ala Leu Phe Leu Gly Arg Leu Phe Ser Thr Asp Gly Ser
    370                 375                 380

Val Glu Lys Lys Arg Ile Ser Tyr Ser Ser Ala Ser Leu Gly Leu Ala
385                 390                 395                 400

Gln Asp Val Ala His Leu Leu Leu Arg Leu Gly Ile Thr Ser Gln Leu
                405                 410                 415

Arg Ser Arg Gly Pro Arg Ala His Lys Val Leu Ile Ser Gly Arg Glu
            420                 425                 430

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

Lys Arg Glu Arg Leu Ala Ala Leu Glu Ala Glu Ala Arg Arg Arg Leu
    450                 455                 460

Pro Gly Gln Gly Trp His Leu Arg Pro Val Leu Pro Ala Val Ala Tyr
465                 470                 475                 480

Arg Val Ser Glu Ala Lys Arg Arg Ser Gly Phe Ser Trp Ser Glu Ala
                485                 490                 495

Gly Arg Arg Val Ala Val Ala Gly Ser Cys Leu Ser Ser Gly Leu Asn
            500                 505                 510

Leu Lys Leu Pro Arg Arg Tyr Leu Ser Arg His Arg Leu Ser Leu Leu
        515                 520                 525

Gly Glu Ala Phe Ala Asp Pro Gly Leu Glu Ala Leu Ala Glu Gly Gln
    530                 535                 540

Val Leu Trp Asp Pro Ile Val Ala Val Glu Pro Ala Gly Lys Ala Arg
545                 550                 555                 560

Thr Phe Asp Leu Arg Val Pro Pro Phe Ala Asn Phe Val Ser Glu Asp
                565                 570                 575

Leu Val Val His Asn Ser Ile Val Gly Thr Ala Thr Phe Asp Gln Tyr
            580                 585                 590

Trp Ser Val Arg Thr Ser Lys Arg Thr Ser Gly Thr Val Thr Val Thr
        595                 600                 605

Asp His Phe Arg Ala Trp Ala Asn Arg Gly Leu Asn Leu Gly Thr Ile
    610                 615                 620
```

```
Asp Gln Ile Thr Leu Cys Val Glu Gly Tyr Gln Ser Ser Gly Ser Ala
625                 630                 635                 640

Asn Ile Thr Gln Asn Thr Phe Ser Gln Gly Ser Ser Ser Gly Ser Ser
                645                 650                 655

Gly Gly Ser Ser Gly Ser Thr Thr Thr Thr Arg Ile Glu Cys Glu Asn
            660                 665                 670

Met Ser Leu Ser Gly Pro Tyr Val Ser Arg Ile Thr Asn Pro Phe Asn
        675                 680                 685

Gly Ile Ala Leu Tyr Ala Asn Gly Asp Thr Ala Arg Ala Thr Val Asn
    690                 695                 700

Phe Pro Ala Ser Arg Asn Tyr Asn Phe Arg Leu Arg Gly Cys Gly Asn
705                 710                 715                 720

Ser Asn Asn Leu Ala Arg Val Asp Leu Arg Ile Asp Gly Arg Thr Val
                725                 730                 735

Gly Thr Phe Tyr Tyr Gln Gly Thr Tyr Pro Trp Glu Ala Pro Ile Asp
            740                 745                 750

Asn Val Tyr Val Ser Ala Gly Ser His Thr Val Glu Ile Thr Val Thr
        755                 760                 765

Ala Asp Asn Gly Thr Arg Asp Val Tyr Ala Asp Tyr Leu Val Ile Gln
    770                 775                 780
```

<210> 22
<211> 6
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, націлюючий пептид, амінокислотна, SEKDEL

<400> 22

```
Ser Glu Lys Asp Glu Leu
1               5
```

<210> 23
<211> 4
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, націлюючий пептид, амінокислотна, KDEL

<400> 23

Lys Asp Glu Leu
1

<210> 24
<211> 8
<212> Білок
<213> Hordeum vulgare, амінокислотна, HVVSD

<400> 24

Asp Glu Leu Lys Ala Glu Ala Lys
1 5

<210> 25
<211> 536
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, ферулілестераза, злита з найбільшими пептидами, амінокислотна, SAASS:AnfaeB:SEKDEL

<400> 25

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Val Ala Thr Asp Pro Phe Gln
20 25 30

Ser Arg Cys Asn Glu Phe Gln Asn Lys Ile Asp Ile Ala Asn Val Thr
35 40 45

Val Arg Ser Val Ala Tyr Val Ala Ala Gly Gln Asn Ile Ser Gln Ala
50 55 60

Glu Val Ala Ser Val Cys Lys Ala Ser Val Gln Ala Ser Val Asp Leu
65 70 75 80

Cys Arg Val Thr Met Asn Ile Ser Thr Ser Asp Arg Ser His Leu Trp
85 90 95

Ala Glu Ala Trp Leu Pro Arg Asn Tyr Thr Gly Arg Phe Val Ser Thr
100 105 110

Gly Asn Gly Gly Leu Ala Gly Cys Val Gln Glu Thr Asp Leu Asn Phe
115 120 125

Ala Ala Asn Phe Gly Phe Ala Thr Val Gly Thr Asn Gly Gly His Asp
130 135 140

```
Gly Asp Thr Ala Lys Tyr Phe Leu Asn Asn Ser Glu Val Leu Ala Asp
145                 150                 155                 160

Phe Ala Tyr Arg Ser Val His Glu Gly Thr Val Val Gly Lys Gln Leu
                165                 170                 175

Thr Gln Leu Phe Tyr Asp Glu Gly Tyr Asn Tyr Ser Tyr Tyr Leu Gly
            180                 185                 190

Cys Ser Thr Gly Gly Arg Gln Gly Tyr Gln Gln Val Gln Arg Phe Pro
        195                 200                 205

Asp Asp Tyr Asp Gly Val Ile Ala Gly Ser Ala Ala Met Asn Phe Ile
    210                 215                 220

Asn Leu Ile Ser Trp Gly Ala Phe Leu Trp Lys Ala Thr Gly Leu Ala
225                 230                 235                 240

Asp Asp Pro Asp Phe Ile Ser Ala Asn Leu Trp Ser Val Ile His Gln
                245                 250                 255

Glu Ile Val Arg Gln Cys Asp Leu Val Asp Gly Ala Leu Asp Gly Ile
            260                 265                 270

Ile Glu Asp Pro Asp Phe Cys Ala Pro Val Ile Glu Arg Leu Ile Cys
        275                 280                 285

Asp Gly Thr Thr Asn Gly Thr Ser Cys Ile Thr Gly Ala Gln Ala Ala
    290                 295                 300

Lys Val Asn Arg Ala Leu Ser Asp Phe Tyr Gly Pro Asp Gly Thr Val
305                 310                 315                 320

Tyr Tyr Pro Arg Leu Asn Tyr Gly Gly Glu Ala Asp Ser Ala Ser Leu
                325                 330                 335

Tyr Phe Thr Gly Ser Met Tyr Ser Arg Thr Glu Glu Trp Tyr Lys Tyr
            340                 345                 350

Val Val Tyr Asn Asp Thr Asn Trp Asn Ser Ser Gln Trp Thr Leu Glu
        355                 360                 365

Ser Ala Lys Leu Ala Leu Glu Gln Asn Pro Phe Asn Ile Gln Ala Phe
    370                 375                 380

Asp Pro Asn Ile Thr Ala Phe Arg Asp Arg Gly Gly Lys Leu Leu Ser
385                 390                 395                 400
```

```
Tyr His Gly Thr Gln Asp Pro Ile Ile Ser Ser Thr Asp Ser Lys Leu
                405                 410                 415

Tyr Tyr Arg Arg Val Ala Asn Ala Leu Asn Ala Ala Pro Ser Glu Leu
            420                 425                 430

Asp Glu Phe Tyr Arg Phe Phe Gln Ile Ser Gly Met Gly His Cys Gly
        435                 440                 445

Asp Gly Thr Gly Ala Ser Tyr Ile Gly Gln Gly Tyr Gly Thr Tyr Thr
    450                 455                 460

Ser Lys Ala Pro Gln Val Asn Leu Leu Arg Thr Met Val Asp Trp Val
465                 470                 475                 480

Glu Asn Gly Lys Ala Pro Glu Tyr Met Pro Gly Asn Lys Leu Asn Ala
                485                 490                 495

Asn Gly Ser Ile Glu Tyr Met Arg Lys His Cys Arg Tyr Pro Lys His
            500                 505                 510

Asn Ile His Thr Gly Pro Gly Asn Tyr Thr Asp Pro Asn Ser Trp Thr
        515                 520                 525

Cys Val Ser Glu Lys Asp Glu Leu
    530                 535
```

<210> 26
<211> 292
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, феруловілестераза, злита з
N-кінцевими пептидами, амінокислотна, BAASS:AnfaeA:SEKDEL

<400> 26

```
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1               5                   10                  15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Val Ala Ser Thr Gln Gly Ile
            20                  25                  30

Ser Glu Asp Leu Tyr Asn Arg Leu Val Glu Met Ala Thr Ile Ser Gln
        35                  40                  45

Ala Ala Tyr Ala Asp Leu Cys Asn Ile Pro Ser Thr Ile Ile Lys Gly
    50                  55                  60
```

```
Glu Lys Ile Tyr Asn Ala Gln Thr Asp Ile Asn Gly Trp Ile Leu Arg
65                  70                  75                  80

Asp Asp Thr Ser Lys Glu Ile Ile Thr Val Phe Arg Gly Thr Gly Ser
                85                  90                  95

Asp Thr Asn Leu Gln Leu Asp Thr Asn Tyr Thr Leu Thr Pro Phe Asp
            100                 105                 110

Thr Leu Pro Gln Cys Asn Asp Cys Glu Val His Gly Gly Tyr Tyr Ile
        115                 120                 125

Gly Trp Ile Ser Val Gln Asp Gln Val Glu Ser Leu Val Lys Gln Gln
    130                 135                 140

Ala Ser Gln Tyr Pro Asp Tyr Ala Leu Thr Val Thr Gly His Ser Leu
145                 150                 155                 160

Gly Ala Ser Met Ala Ala Leu Thr Ala Ala Gln Leu Ser Ala Thr Tyr
                165                 170                 175

Asp Asn Val Arg Leu Tyr Thr Phe Gly Glu Pro Arg Ser Gly Asn Gln
            180                 185                 190

Ala Phe Ala Ser Tyr Met Asn Asp Ala Phe Gln Val Ser Ser Pro Glu
        195                 200                 205

Thr Thr Gln Tyr Phe Arg Val Thr His Ser Asn Asp Gly Ile Pro Asn
    210                 215                 220

Leu Pro Pro Ala Asp Glu Gly Tyr Ala His Gly Gly Val Glu Tyr Trp
225                 230                 235                 240

Ser Val Asp Pro Tyr Ser Ala Gln Asn Thr Phe Val Cys Thr Gly Asp
                245                 250                 255

Glu Val Gln Cys Cys Glu Ala Gln Gly Gly Gln Gly Val Asn Asp Ala
            260                 265                 270

His Thr Thr Tyr Phe Gly Met Thr Ser Gly Ala Cys Thr Trp Ser Glu
        275                 280                 285

Lys Asp Glu Leu
    290
```

<210> 27
<211> 468
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, ендоглюканаза, злита з
націлюючими пептидами, амінокислотна, PRla:NtEGM:SEKDEL

<400> 27

Met Gly Phe Val Leu Phe Ser Gln Leu Pro Ser Phe Leu Leu Val Ser
1 5 10 15

Thr Leu Leu Leu Phe Leu Val Ile Ser His Ser Cys Arg Ala Ala Tyr
20 25 30

Asp Tyr Lys Gln Val Leu Arg Asp Ser Leu Leu Phe Tyr Glu Ala Gln
35 40 45

Arg Ser Gly Arg Leu Pro Ala Asp Gln Lys Val Thr Trp Arg Lys Asp
50 55 60

Ser Ala Leu Asn Asp Gln Gly Asp Gln Gly Gln Asp Leu Thr Gly Gly
65 70 75 80

Tyr Phe Asp Ala Gly Asp Phe Val Lys Phe Gly Phe Pro Met Ala Tyr
85 90 95

Thr Ala Thr Val Leu Ala Trp Gly Leu Ile Asp Phe Glu Ala Gly Tyr
100 105 110

Ser Ser Ala Gly Ala Leu Asp Asp Gly Arg Lys Ala Val Lys Trp Ala
115 120 125

Thr Asp Tyr Phe Ile Lys Ala His Thr Ser Gln Asn Glu Phe Tyr Gly
130 135 140

Gln Val Gly Gln Gly Asp Ala Asp His Ala Phe Trp Gly Arg Pro Glu
145 150 155 160

Asp Met Thr Met Ala Arg Pro Ala Tyr Lys Ile Asp Thr Ser Arg Pro
165 170 175

Gly Ser Asp Leu Ala Gly Glu Thr Ala Ala Ala Leu Ala Ala Ala Ser
180 185 190

Ile Val Phe Arg Asn Val Asp Gly Thr Tyr Ser Asn Asn Leu Leu Thr
195 200 205

His Ala Arg Gln Leu Phe Asp Phe Ala Asn Asn Tyr Arg Gly Lys Tyr
210 215 220

Ser Asp Ser Ile Thr Asp Ala Arg Asn Phe Tyr Ala Ser Ala Asp Tyr

225 230 235 240

Arg Asp Glu Leu Val Trp Ala Ala Ala Trp Leu Tyr Arg Ala Thr Asn
245 250 255

Asp Asn Thr Tyr Leu Asn Thr Ala Glu Ser Leu Tyr Asp Glu Phe Gly
260 265 270

Leu Gln Asn Trp Gly Gly Gly Leu Asn Trp Asp Ser Lys Val Ser Gly
275 280 285

Val Gln Val Leu Leu Ala Lys Leu Thr Asn Lys Gln Ala Tyr Lys Asp
290 295 300

Thr Val Gln Ser Tyr Val Asn Tyr Leu Ile Asn Asn Gln Gln Lys Thr
305 310 315 320

Pro Lys Gly Leu Leu Tyr Ile Asp Met Trp Gly Thr Leu Arg His Ala
325 330 335

Ala Asn Ala Ala Phe Ile Met Leu Glu Ala Ala Glu Leu Gly Leu Ser
340 345 350

Ala Ser Ser Tyr Arg Gln Phe Ala Gln Thr Gln Ile Asp Tyr Ala Leu
355 360 365

Gly Asp Gly Gly Arg Ser Phe Val Cys Gly Phe Gly Ser Asn Pro Pro
370 375 380

Thr Arg Pro His His Arg Ser Ser Ser Cys Pro Pro Ala Pro Ala Thr
385 390 395 400

Cys Asp Trp Asn Thr Phe Asn Ser Pro Asp Pro Asn Tyr His Val Leu
405 410 415

Ser Gly Ala Leu Val Gly Gly Pro Asp Gln Asn Asp Asn Tyr Val Asp
420 425 430

Asp Arg Ser Asp Tyr Val His Asn Glu Val Ala Thr Asp Tyr Asn Ala
435 440 445

Gly Phe Gln Ser Ala Leu Ala Ala Leu Val Ala Leu Gly Tyr Ser Glu
450 455 460

Lys Asp Glu Leu
465

<210> 28

<211> 791
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, ксиланаза, злита з
націлюючими пептидами, амінокислотна,
BAASS:P77853:T134-100-101:SEKDEL

<400> 28

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Glu Thr Ser Ile Thr Leu Thr
20 25 30

Ser Asn Ala Ser Gly Thr Phe Asp Gly Tyr Tyr Tyr Glu Leu Trp Lys
35 40 45

Asp Thr Gly Asn Thr Thr Met Thr Val Tyr Thr Gln Gly Arg Phe Ser
50 55 60

Cys Gln Trp Ser Asn Ile Asn Asn Ala Leu Phe Arg Thr Gly Lys Lys
65 70 75 80

Tyr Asn Gln Asn Trp Gln Ser Leu Gly Thr Ile Arg Ile Thr Tyr Ser
85 90 95

Ala Thr Tyr Asn Pro Asn Gly Asn Ser Tyr Leu Cys Ile Tyr Gly Trp
100 105 110

Ser Thr Asn Pro Leu Val Glu Phe Tyr Ile Val Glu Ser Trp Gly Asn
115 120 125

Trp Arg Pro Pro Gly Ala Cys Leu Ala Glu Gly Ser Leu Val Leu Asp
130 135 140

Ala Ala Thr Gly Gln Arg Val Pro Ile Glu Lys Val Arg Pro Gly Met
145 150 155 160

Glu Val Phe Ser Leu Gly Pro Asp Tyr Arg Leu Tyr Arg Val Pro Val
165 170 175

Leu Glu Val Leu Glu Ser Gly Val Arg Glu Val Val Arg Leu Arg Thr
180 185 190

Arg Ser Gly Arg Thr Leu Val Leu Thr Pro Asp His Pro Leu Leu Thr
195 200 205

Pro Glu Gly Trp Lys Pro Leu Cys Asp Leu Pro Leu Gly Thr Pro Ile

210 215 220

Ala Val Pro Ala Glu Leu Pro Val Ala Gly His Leu Ala Pro Pro Glu
225 230 235 240

Glu Arg Val Thr Leu Leu Ala Leu Leu Leu Gly Asp Gly Asn Thr Lys
245 250 255

Leu Ser Gly Arg Arg Gly Thr Arg Pro Asn Ala Phe Phe Tyr Ser Lys
260 265 270

Asn Pro Glu Leu Leu Ala Ala Tyr Arg Arg Cys Ala Glu Ala Leu Gly
275 280 285

Ala Lys Val Lys Ala Tyr Val His Pro Thr Thr Gly Val Val Thr Leu
290 295 300

Ala Thr Leu Ala Pro Arg Pro Gly Ala Gln Asp Pro Val Lys Arg Leu
305 310 315 320

Val Val Glu Ala Gly Met Val Ala Lys Ala Glu Glu Lys Arg Val Pro
325 330 335

Glu Glu Val Phe Arg Tyr Arg Arg Glu Ala Leu Ala Leu Phe Leu Gly
340 345 350

Arg Leu Phe Ser Thr Asp Gly Ser Val Glu Lys Lys Arg Ile Ser Tyr
355 360 365

Ser Ser Ala Ser Leu Gly Leu Ala Gln Asp Val Ala His Leu Leu Leu
370 375 380

Arg Leu Gly Ile Thr Ser Gln Leu Arg Ser Arg Gly Pro Arg Ala His
385 390 395 400

Glu Val Leu Ile Ser Gly Arg Glu Asp Ile Leu Arg Phe Ala Glu Leu
405 410 415

Ile Gly Pro Tyr Leu Leu Gly Ala Lys Arg Glu Arg Leu Ala Ala Leu
420 425 430

Glu Ala Glu Ala Arg Arg Arg Leu Pro Gly Gln Gly Trp His Leu Arg
435 440 445

Leu Val Leu Pro Ala Val Ala Tyr Arg Val Gly Glu Ala Glu Arg Arg
450 455 460

Ser Gly Phe Ser Trp Ser Glu Ala Gly Arg Arg Val Ala Val Ala Gly

465 470 475 480

Ser Cys Leu Ser Ser Gly Leu Asn Leu Lys Leu Pro Arg Arg Tyr Leu
485 490 495

Ser Arg His Arg Leu Ser Leu Leu Gly Glu Ala Phe Ala Asp Pro Gly
500 505 510

Leu Glu Ala Leu Ala Glu Gly Gln Val Leu Trp Asp Pro Ile Val Ala
515 520 525

Val Glu Pro Ala Gly Lys Ala Arg Thr Phe Asp Leu Arg Val Pro Pro
530 535 540

Phe Ala Asn Phe Val Ser Glu Asp Leu Val Val His Asn Thr Val Pro
545 550 555 560

Leu Gly Gln Val Thr Ile Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr
565 570 575

Thr Arg Val Asn Gln Pro Ser Ile Val Gly Thr Ala Thr Phe Asp Gln
580 585 590

Tyr Trp Ser Val Arg Thr Ser Lys Arg Thr Ser Gly Thr Val Thr Val
595 600 605

Thr Asp His Phe Arg Ala Trp Ala Asn Arg Gly Leu Asn Leu Gly Thr
610 615 620

Ile Asp Gln Ile Thr Leu Cys Val Glu Gly Tyr Gln Ser Ser Gly Ser
625 630 635 640

Ala Asn Ile Thr Gln Asn Thr Phe Ser Gln Gly Ser Ser Ser Gly Ser
645 650 655

Ser Gly Gly Ser Ser Gly Ser Thr Thr Thr Thr Arg Ile Glu Cys Glu
660 665 670

Asn Met Ser Leu Ser Gly Pro Tyr Val Ser Arg Ile Thr Asn Pro Phe
675 680 685

Asn Gly Ile Ala Leu Tyr Ala Asn Gly Asp Thr Ala Arg Ala Thr Val
690 695 700

Asn Phe Pro Ala Ser Arg Asn Tyr Asn Phe Arg Leu Arg Gly Cys Gly
705 710 715 720

Asn Asn Asn Asn Leu Ala Arg Val Asp Leu Arg Ile Asp Gly Arg Thr

725 730 735

Val Gly Thr Phe Tyr Tyr Gln Gly Thr Tyr Pro Trp Glu Ala Pro Ile
740 745 750

Asp Asn Val Tyr Val Ser Ala Gly Ser His Thr Val Glu Ile Thr Val
755 760 765

Thr Ala Asp Asn Gly Thr Trp Asp Val Tyr Ala Asp Tyr Leu Val Ile
770 775 780

Gln Ser Glu Lys Asp Glu Leu
785 790

<210> 29
<211> 481
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, химерна, злита з
наліпковими пептидами, амінокислотна, HvAleSP:NtEGM:SEKDEL

<400> 29

Met Ala His Ala Arg Val Leu Leu Leu Ala Leu Ala Val Leu Ala Thr
1 5 10 15

Ala Ala Val Ala Val Ala Ser Ser Ser Ser Phe Ala Asp Ser Asn Pro
20 25 30

Ile Arg Pro Val Thr Asp Arg Ala Ala Ser Thr Ala Tyr Asp Tyr Lys
35 40 45

Gln Val Leu Arg Asp Ser Leu Leu Phe Tyr Glu Ala Gln Arg Ser Gly
50 55 60

Arg Leu Pro Ala Asp Gln Lys Val Thr Trp Arg Lys Asp Ser Ala Leu
65 70 75 80

Asn Asp Gln Gly Asp Gln Gly Gln Asp Leu Thr Gly Gly Tyr Phe Asp
85 90 95

Ala Gly Asp Phe Val Lys Phe Gly Phe Pro Met Ala Tyr Thr Ala Thr
100 105 110

Val Leu Ala Trp Gly Leu Ile Asp Phe Glu Ala Gly Tyr Ser Ser Ala
115 120 125

Gly Ala Leu Asp Asp Gly Arg Lys Ala Val Lys Trp Ala Thr Asp Tyr
130 135 140

```
Phe Ile Lys Ala His Thr Ser Gln Asn Glu Phe Tyr Gly Gln Val Gly
145                 150                 155                 160

Gln Gly Asp Ala Asp His Ala Phe Trp Gly Arg Pro Glu Asp Met Thr
                165                 170                 175

Met Ala Arg Pro Ala Tyr Lys Ile Asp Thr Ser Arg Pro Gly Ser Asp
            180                 185                 190

Leu Ala Gly Glu Thr Ala Ala Ala Leu Ala Ala Ala Ser Ile Val Phe
        195                 200                 205

Arg Asn Val Asp Gly Thr Tyr Ser Asn Asn Leu Leu Thr His Ala Arg
    210                 215                 220

Gln Leu Phe Asp Phe Ala Asn Asn Tyr Arg Gly Lys Tyr Ser Asp Ser
225                 230                 235                 240

Ile Thr Asp Ala Arg Asn Phe Tyr Ala Ser Ala Asp Tyr Arg Asp Glu
                245                 250                 255

Leu Val Trp Ala Ala Ala Trp Leu Tyr Arg Ala Thr Asn Asp Asn Thr
            260                 265                 270

Tyr Leu Asn Thr Ala Glu Ser Leu Tyr Asp Glu Phe Gly Leu Gln Asn
        275                 280                 285

Trp Gly Gly Gly Leu Asn Trp Asp Ser Lys Val Ser Gly Val Gln Val
    290                 295                 300

Leu Leu Ala Lys Leu Thr Asn Lys Gln Ala Tyr Lys Asp Thr Val Gln
305                 310                 315                 320

Ser Tyr Val Asn Tyr Leu Ile Asn Asn Gln Gln Lys Thr Pro Lys Gly
                325                 330                 335

Leu Leu Tyr Ile Asp Met Trp Gly Thr Leu Arg His Ala Ala Asn Ala
            340                 345                 350

Ala Phe Ile Met Leu Glu Ala Ala Glu Leu Gly Leu Ser Ala Ser Ser
        355                 360                 365

Tyr Arg Gln Phe Ala Gln Thr Gln Ile Asp Tyr Ala Leu Gly Asp Gly
    370                 375                 380

Gly Arg Ser Phe Val Cys Gly Phe Gly Ser Asn Pro Pro Thr Arg Pro
385                 390                 395                 400
```

```
His His Arg Ser Ser Ser Cys Pro Pro Ala Pro Ala Thr Cys Asp Trp
                405                 410                 415

Asn Thr Phe Asn Ser Pro Asp Pro Asn Tyr His Val Leu Ser Gly Ala
            420                 425                 430

Leu Val Gly Gly Pro Asp Gln Asn Asp Asn Tyr Val Asp Asp Arg Ser
        435                 440                 445

Asp Tyr Val His Asn Glu Val Ala Thr Asp Tyr Asn Ala Gly Phe Gln
    450                 455                 460

Ser Ala Leu Ala Ala Leu Val Ala Leu Gly Tyr Ser Glu Lys Asp Glu
465                 470                 475                 480

Leu
```

<210> 30
<211> 238
<212> Білок
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, комплекс, злиття з націлюючими пептидами, амінокислотна, BAASS:043097:SEKDEL

<400> 30

```
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1               5                   10                  15

Leu Ser Ala Ser Leu Ala Ser Gly Gln Val Phe Pro Ala Gly Asn Ala
            20                  25                  30

Thr Glu Leu Glu Lys Arg Gln Thr Thr Pro Asn Ser Glu Gly Trp His
        35                  40                  45

Asp Gly Tyr Tyr Tyr Ser Trp Trp Ser Asp Gly Gly Ala Gln Ala Thr
    50                  55                  60

Tyr Thr Asn Leu Glu Gly Gly Thr Tyr Glu Ile Ser Trp Gly Asp Gly
65                  70                  75                  80

Gly Asn Leu Val Gly Gly Lys Gly Trp Asn Pro Gly Leu Asn Ala Arg
                85                  90                  95

Ala Ile His Phe Glu Gly Val Tyr Gln Pro Asn Gly Asn Ser Tyr Leu
            100                 105                 110
```

```
Ala Val Tyr Gly Trp Thr Arg Asn Pro Leu Val Glu Tyr Tyr Ile Val
        115                 120                 125

Glu Asn Phe Gly Thr Tyr Asp Pro Ser Ser Gly Ala Thr Asp Leu Gly
    130                 135                 140

Thr Val Glu Cys Asp Gly Ser Ile Tyr Arg Leu Gly Lys Thr Thr Arg
145                 150                 155                 160

Val Asn Ala Pro Ser Ile Asp Gly Thr Gln Thr Phe Asp Gln Tyr Trp
                165                 170                 175

Ser Val Arg Gln Asp Lys Arg Thr Ser Gly Thr Val Gln Thr Gly Cys
            180                 185                 190

His Phe Asp Ala Trp Ala Arg Ala Gly Leu Asn Val Asn Gly Asp His
        195                 200                 205

Tyr Tyr Gln Ile Val Ala Thr Glu Gly Tyr Phe Ser Ser Gly Tyr Ala
    210                 215                 220

Arg Ile Thr Val Ala Asp Val Gly Ser Glu Lys Asp Glu Leu
225                 230                 235
```

<210> 31
<211> 531
<212> Вірус
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, целобіогідролаза, злита з
наступними пептидами, амінокислотна,
G2ein27ss-02:BD22308:HvVSD-01

<400> 31

```
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
1               5                   10                  15

Ala Thr Ser Gln Gln Ile Gly Thr Tyr Thr Ala Glu Thr His Pro Ser
            20                  25                  30

Leu Ser Trp Ser Thr Cys Lys Ser Gly Gly Ser Cys Thr Thr Asn Ser
        35                  40                  45

Gly Ala Ile Thr Leu Asp Ala Asn Trp Arg Trp Val His Gly Val Asn
    50                  55                  60

Thr Ser Thr Asn Cys Tyr Thr Gly Asn Thr Trp Asn Thr Ala Ile Cys
65                  70                  75                  80
```

```
Asp Thr Asp Ala Ser Cys Ala Gln Asp Cys Ala Leu Asp Gly Ala Asp
                85                  90                  95

Tyr Ser Gly Thr Tyr Gly Ile Thr Thr Ser Gly Asn Ser Leu Arg Leu
            100                 105                 110

Asn Phe Val Thr Gly Ser Asn Val Gly Ser Arg Thr Tyr Leu Met Ala
        115                 120                 125

Asp Asn Thr His Tyr Gln Ile Phe Asp Leu Leu Asn Gln Glu Phe Thr
    130                 135                 140

Phe Thr Val Asp Val Ser His Leu Pro Cys Gly Leu Asn Gly Ala Leu
145                 150                 155                 160

Tyr Phe Val Thr Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Asn
                165                 170                 175

Asn Lys Ala Gly Ala Gln Tyr Gly Val Gly Tyr Cys Asp Ser Gln Cys
            180                 185                 190

Pro Arg Asp Leu Lys Phe Ile Ala Gly Gln Ala Asn Val Glu Gly Trp
        195                 200                 205

Thr Pro Ser Ser Asn Asn Ala Asn Thr Gly Leu Gly Asn His Gly Ala
    210                 215                 220

Cys Cys Ala Glu Leu Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
225                 230                 235                 240

Leu Thr Pro His Pro Cys Asp Thr Pro Gly Leu Ser Val Cys Thr Thr
                245                 250                 255

Asp Ala Cys Gly Gly Thr Tyr Ser Ser Asp Arg Tyr Ala Gly Thr Cys
            260                 265                 270

Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Leu Gly Val Thr Asp
        275                 280                 285

Phe Tyr Gly Ser Gly Lys Thr Val Asp Thr Thr Lys Pro Ile Thr Val
    290                 295                 300

Val Thr Gln Phe Val Thr Asp Asp Gly Thr Ser Thr Gly Thr Leu Ser
305                 310                 315                 320

Glu Ile Arg Arg Tyr Tyr Val Gln Asn Gly Val Val Ile Pro Gln Pro
                325                 330                 335
```

Ser Ser Lys Ile Ser Gly Val Ser Gly Asn Val Ile Asn Ser Asp Phe
340 345 350

Cys Asp Ala Glu Ile Ser Thr Phe Gly Glu Thr Ala Ser Phe Ser Lys
355 360 365

His Gly Gly Leu Ala Lys Met Gly Ala Gly Met Glu Ala Gly Met Val
370 375 380

Leu Val Met Ser Leu Trp Asp Asp Tyr Ser Val Asn Met Leu Trp Leu
385 390 395 400

Asp Ser Thr Tyr Pro Thr Asn Ala Thr Gly Thr Pro Gly Ala Ala Arg
405 410 415

Gly Ser Cys Pro Thr Thr Ser Gly Asp Pro Lys Thr Val Glu Ser Gln
420 425 430

Ser Gly Ser Ser Tyr Val Thr Phe Ser Asp Ile Arg Val Gly Pro Phe
435 440 445

Asn Ser Thr Phe Ser Gly Gly Ser Ser Thr Gly Gly Ser Ser Thr Thr
450 455 460

Thr Ala Ser Gly Thr Thr Thr Thr Lys Ala Ser Ser Thr Ser Thr Ser
465 470 475 480

Ser Thr Ser Thr Gly Thr Gly Val Ala Ala His Trp Gly Gln Cys Gly
485 490 495

Gly Gln Gly Trp Thr Gly Pro Thr Thr Cys Ala Ser Gly Thr Thr Cys
500 505 510

Thr Val Val Asn Pro Tyr Tyr Ser Gln Cys Leu Asp Glu Leu Lys Ala
515 520 525

Glu Ala Lys
530

<210> 32
<211> 1014
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ксиланазу, оптимізована для експресії в рослинах, нуклеїнова кислота, P77853

<400> 32

atgcaaacaa gcattactct gacatccaac gcatcggta cgtttgacgg ttactattac 60

gaactctgga aggatactgg caatacaaca atgacggtct acactcaagg tcgcttttcc 120

tgccagtggt cgaacatcaa taacgcgttg tttaggaccg ggaagaaata caaccagaat 180

tggcagtctc ttggcacaat ccggatcacg tactctgcga cttacaaccc aaacgggaac 240

tcctacttgt gtatctatgg ctggtctacc aacccattgg tcgagttcta catcgttgag 300

tcctgggggа actggagacc gcctggtgcc acgtccctgg gccaagtgac aatcgatggc 360

gggacctacg acatctatag gacgacacgc gtcaaccagc cttccattgt ggggacagcc 420

acgttcgatc agtactggag cgtgcgcacc tctaagcgga cttcaggaac agtgaccgtg 480

accgatcact tccgcgcctg ggcgaaccgg ggcctgaacc tcggcacaat agaccaaatt 540

acattgtgcg tggagggtta ccaaagctct ggatcagcca acatcaccca gaacaccttc 600

tctcagggct cttcttccgg cagttcgggt ggctcatccg gctccacaac gactactcgc 660

atcgagtgtg agaacatgtc cttgtccgga ccctacgtta gcaggatcac caatcccttt 720

aatggtattg cgctgtacgc caacggagac acagccgcg ctaccgttaa cttccccgca 780

agtcgcaact acaatttccg cctgcggggt tgcggcaaca acaataatct tgccccgtgtg 840

gacctgagga tcgacggacg gaccgtcggg accttttatt accagggcac ataccctgg 900

gaggccccaa ttgacaatgt ttatgtcagt gcggggagtc atacagtcga aatcactgtt 960

actgcggata acggcacatg ggacgtgtat gccgactacc tggtgataca gtga 1014

<210> 33
<211> 786
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча феруполестеразу, оптимізована для експресії в рослинах, нуклеїнова кислота, Anfae A

<400> 33
atggcatcca cgcagggtat ctctgaggac ctgtataatc gcctcgtgga aatggccaca 60

atttcacaag cggcttacgc agatctttgt aatatcccct cgacaattat caagggagag 120

aaaatctata acgcccagac tgacatcaac ggctggatac tgcgggatga cacgagcaag 180

gaaattatca cagtctttag agggaccggt tccgatacaa atttgcagtt ggacacgaat 240

tacacactga cccccttcga tactctccct caatgcaacg actgtgaggt tcacggtggg 300

tactatattg gctggatctc tgttcaagac caagtcgagt cacttgttaa gcagcaagcg 360

tcgcagtacc cggactacgc attgacggtg acagggcaca gcctgggtgc ctcgatggca 420

gcgctcaccg ctgcccagct ttctgcaacc tacgataatg tcaggctgta cactttcgga 480

gaaccacgct caggcaacca agcgtttgct tcgtatatga acgacgcttt ccaggttagc 540

tccccgaga cgacacaata ctttcgggtg acccattcta acgacgggat tcctaacctc 600

ccgccagccg acgaaggtta cgcacacggg ggtgtcgagt actggtcagt ggaccctac 660

agcgcgcaga atactttcgt ttgcacgggc gatgaggtcc agtgctgtga agctcaaggg 720

ggtcagggag tgaatgatgc acacacaacc tatttcggaa tgacttccgg ggcttgcacg 780

tggtga 786

<210> 34
<211> 1518
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча феруліл естеразу, оптимізована для експресії в рослинах, нуклеїнова кислота, AnFae B

<400> 34
atggctactg atcccttcca gtctcgctgt aacgagtttc aaaacaaaat tgatatcgca 60

aacgtcacgg ttaggtcggt cgcctacgtc gcggctggtc agaatatctc tcaagcagaa 120

gtcgcttcgg tttgcaaggc aagcgtccag gcctccgtgg acctctgtag agtgacaatg 180

aatattagca cctcggatag gagccacctt tgggccgagg cttggctgcc tcgcaattat 240

actggacggt tcgtctccac tgggaacggg ggcctggctg gatgcgtgca ggaaacagac 300

ttgaatttcg ctgccaactt tggcttcgca accgttggca ctaatggagg gcacgatgga 360

gacacggcga aatactttct gaataactct gaggtcctcg ctgatttcgc ctatcgcagc 420

gttcacgagg gcactgtggt cgggaagcaa cttacccagt tgttctacga cgaaggttac 480

aattatttcg actatctggg ctgcagcacc ggaggcaggc aaggctatca gcaagtgcag 540

cgctttccgg atgactacga tggtgttatt gccggaagtg cagcgatgaa ctttatcaat 600

ttgatttctt gggcgcttt tattttggaag gccacgggac tcgcagacga cccagatttt 660

atctcagcga acctctggag cgttatccac caggagattg tgcggcagtg tgacttggtt 720

gatggtgctc ttgacggaat tattgaagat cctgacttct gcgcccctgt tatcgagaga 780

ctgatttgtg atgggactac caacggcacc tcctgtatca ctggggcaca agcggctaaa 840

gtgaataggg ccctctcgga cttttatggt ccggatggga cagtctacta tccacgcttg 900

aactacggag gggaggcaga ctctgcgagc ctgtacttca caggctccat gtattcacgg 960

accgaagagt ggtacaagta cgtcgtgtac aacgatacta actggaatag ttctcaatgg 1020

acgctcgaaa gcgctaagtt ggccttggag cagaaccct ttaatattca ggcattcgac 1080

cctaatatca cagcgttcag ggaccgcggt ggaaaactct tgagctacca cggcacccaa 1140

gatccgatta tttcatcgac tgacagcaag ttgtattaca gaagagttgc taacgccctt 1200

aacgctgccc catccgaact cgacgagttt tatcgcttct ttcagatctc tgggatgggt 1260

cactgcggcg acggaaccgg ggcttcatac attggccagg gatacggcac gtatacctcg 1320

aaggccccac aggtcaactt gcttcggact atggtggatt gggttgagaa tggaaaagct 1380

ccggaatata tgcctgggaa taagttgaac gccaacggtt cgatcgagta tatgagaaag 1440

cactgtcgtt acccaaaaca caatattcac acgggcccag gtaactacac ggaccctaac 1500

tcctggacct gcgtctga 1518

<210> 35
<211> 1302
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ензоплоканазу, оптимізована для експресії в рослинах, нуклеїнова кислота, NtEGM

<400> 35
atggcttacg actacaagca ggtgttgcgg gactcgctac tattctatga ggccagaga 60

tccggccggc tcccagccga ccagaaggtc acgtggagga aggatagcgc gctgaatgac 120

cagggtgacc agggacaaga cttgaccggc ggcctacttg acgctgggga cttcgtcaag 180

ttcgggttcc ccatggctta taccgcaacc gtgctggcat ggggcctcat agattttgag 240

gccggctaca gcagtgccgg ggccttggat gatggacgga aggctgtcaa atgggccacc 300

gactatttca taaaggccca cacaagtcaa aatgagttct atggtcaggt cggccagggt 360

gacgccgatc acgctttctg gggaagacca gaggatatga cgatggcgcg cccggcgtac 420

aagatagaca cctcaaggcc tggctctgat ctggcaggcg agacagcggc tgctcttgcc 480

gctgcttcaa tcgtgttccg gaacgtcgat ggcacttact caaataacct gttaacacac 540

gctcgccagc tattcgactt cgcgaacaac taccggggaa agtatagtga ctctattact 600

gacgcaagaa atttctacgc aagcgcagac tacagagacg agttggtttg ggctgctgcg 660

tggttataca gagcgaccaa cgacaacacc tacctcaaca ctgctgagtc actgtacgat 720

gagtttgggc tacagaactg gggggggggc ctgaactggg atagcaaggt gtctggcgtg 780

caggtgttgt tggccaagct taccaataag caggcctaca aggacacggt gcagtcttac 840

gtcaattacc taattaataa ccagcagaag actcccaagg gcctcctcta catcgacatg 900

tggggcaccc ttcgccacgc tgccaacgcc gcattcatca tgctcgaagc cgccgagctg 960

ggcttgtccg cctcctctta tagacagttc gcgcaaacgc aaatcgacta cgccctgggc 1020

gatggtggcc gctcctttgt gtgcgggttc gggagtaatc ctcctacgag accgcaccac 1080

agatcctcgt cgtgcccgcc agctcccgct acttgcgact ggaatacatt caactcacct 1140

gacccaaact accacgtcct ctctggggcc ctagtgggcg gacctgatca gaatgacaac 1200

tacgtcgatg accgttcaga ctatgttcac aacgaagtcg ccactgatta caacgccggt 1260

ttccagtccg cgttagctgc tttggtggcc cttggttact ga 1302

<210> 36
<211> 585
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ксиланазу, оптимізована для експресії в рослинах, нуклеїнова кислота EU591743

<400> 36
atggcccaaa catgtctgac cagcccacaa acaggattcc acaacggctt cttttactcg 60

ttttggaaag attctcctgg tactgttaac ttttgccttc ttgaaggggg aagatatacc 120

tccaattggt ctgggataaa caattgggtc ggtggcaagg gctggcagac cggtagtagg 180

cgcaacatca cgtattccgg gagcttcaat acaccagggg atggatacct ggctctctac 240

ggttggacca ccaacccgtt ggttgagtac tatgttgtcg atagctgggg ctcctggcgc 300

cctcccggaa gtgatgggac atttcttggc acagtgaact cagacggcgg cacgtatgac 360

atctacaggg cgcaaagagt caacgcaccg agcatcattg gcaatgccac cttctaccag 420

tattggtccg tgcggcagtc caagcgtgtc ggtgggacaa tcacgactgg caaccacctc 480

gacgcgtggg ccagcgtggg cctcaacctc gggactcata actaccagat aatggctact 540

gagggttacc agtcgtcagg atcttcagac attacggtgt catga 585

<210> 37
<211> 624
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ксиланазу, оптимізована для експресії в рослинах, нуклеїнова кислота, O43097

<400> 37
atgttcccag ctggaaacgc aacggaattg gagaaaagac aaaccacccc taactctgag 60

ggctggcatg acggatacta ctactcttgg tggagcgatg gtggtgcaca ggccacctat 120

acaaacctcg aaggcggcac ttatgagatt tcatggggtg acggtggcaa ccttgtcggc 180

ggaaaggggt ggaaccccgg acttaacgcc agggcaatcc acttcgaagg ggtgtaccag 240

cccaatggca actcatacct ggccgtctac gggtggacgc gcaatccgct ggttgagtac 300

tatatcgtgg agaatttcgg aacttatgac cctagctccg gtgccacgga cctcgggaca 360

gtcgagtgtg acggaagcat ctacaggctg ggtaaaacta cccgcgttaa tgctccatcg 420

atcgacggca cgcaaacatt tgatcaatac tggtccgtgc ggcaggataa gaggacaagc 480

ggcacagttc agacgggttg ccactttgat gcctgggcaa gagcggggct caatgtgaat 540

ggggaccact actatcagat tgtggcgacc gagggctatt tctccagtgg ctatgcgcgt 600

ataaccgtcg ctgatgttgg atga 624

<210> 38
<211> 2286
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодувча модифікована інтеїном кукурудзи, оптимізована для експресії в рослинах, нуклеїнова кислота, P77853:T134-100-101

<400> 38
atgcaaacaa gcattactct gacatccaac gcatccggta cgtttgacgg ttactattac 60

gaactctgga aggatactgg caatacaaca atgacggtct acactcaagg tcgcttttcc 120

tgccagtggt cgaacatcaa taacgcgttg tttaggaccg ggaagaaata caaccagaat 180

tggcagtctc ttggcacaat ccggatcacg tactctgcga cttacaaccc aaacgggaac 240

tcctacttgt gtatctatgg ctggtctacc aacccattgg tcgagttcta catcgttgag 300

tcctgggggg actggagacc gcctggtgcc tgcctggccg agggctcgct cgtcttggac 360

gcggctaccg ggcagagggt ccctatcgaa aaggtgcgtc cggggatgga agttttctcc 420

ttgggacctg attacagact gtatcgggtg cccgttttgg aggtccttga gagcggggtt 480

agggaagttg tgcgcctcag aactcggtca gggagaacgc tggtgttgac accagatcac 540

ccgcttttga ccccccgaagg ttggaaacct ctttgtgacc tcccgcttgg aactccaatt 600

gcagtccccg cagaactgcc tgtggcgggc cacttggccc cacctgaaga acgtgttacg 660

ctcctggctc ttctgttggg ggatgggaac acaaagctgt cgggtcggag aggtacacgt 720

cctaatgcct tcttctacag caaaaacccc gaattgctcg cggcttatcg ccggtgtgca 780

gaagccttgg gtgcaaaggt gaaagcatac gtcccacccga ctacgggggt ggttacactc 840

gcaaccctcg ctccacgtcc tggagctcaa gatcctgtca aacgcctcgt tgtcgaggcg 900

ggaatggttg ctaaagccga agagaagagg gtcccggagg aggtgtttcg ttaccggcgt 960

gaggcgttgg ccctttcctt gggccgtttg ttctcgacag acggctctgt tgaaaagaag 1020

aggatctctt attcaagtgc cagtttggga ctggcccagg atgtcgcaca tctcttgctg 1080

cgccttggaa ttacatctca actccgttcg agagggccac gggctcacga ggtttctata 1140

tcgggccgcg aggatatttt gcggtttgct gaacttatcg gaccctacct cttgggggcc 1200

aagagggaga gacttgcagc gctggaagct gaggcccgca ggcgtttgcc tggacaggga 1260

tggcacttgc ggcttgttct tcctgccgtg gcgtacagag tgggcgaggc ggaaaggcgc 1320

tcgggatttt cgtggagtga agccggtcgg cgcgtcgcag ttgcgggacc gtgtttgtca 1380

tctggactca acctcaaatt gcccagacgc taccttttcc ggcaccggtt gtcgctgctc 1440

ggtgaggctt ttgccgaccc tgggctggaa gcgctcgcgg aaggccaagt gctctgggac 1500

cctattgttg ctgtcgaacc ggccggtaag gcgagaacat tcgacttgcg cgttccaccc 1560

tttgcaaact tcgtgagcga ggacctggtg gtgcataaca ccgtcccctt gggccaagtg 1620

acaatcgatg gcgggacctа cgacatctat aggacgacac gcgtcaacca gccttccatt 1680

gtggggcaag ccacgttcga tcagtactgg agcgtgcgca cctctaagcg gacttcagga 1740

acagtgaccg tgaccgatca cttccgcgcc tgggcgaacc ggggcctgaa cctcggcaca 1800

atagaccaaa ttacattgtg cgtggagggt taccaaagct ctggatcagc caacatcacc 1860

cagaacacct tcttcaggg ctcttcttcc ggcagttcgg gtggctcatc cggctccaca 1920

acgactactc gcatcgagtg tgagaacatg tccttgtccg gaccctacgt tagcaggatc 1980

accaatcccc ttaatggtat tgcgctgtac gccaacggag acacagcccg cgctaccgtt 2040

aacttccccg caagtcgcaa ctacaatttc cgcctgcggg gttgcggcaa caacaataat 2100

cttgcccgtg tggacctgag gatcgacgga cggaccgtcg ggaccttta ttaccagggc 2160

acataccct gggaggcccc aattgacaat gtttatgtca gtgcggggag tcatacagtc 2220

gaaatcactg ttactgcgga taacggcaca tgggacgtgt atgccgacta cctggtgata 2280

cagtga 2286

<210> 39
<211> 2283
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, копуюча модифікована
інсектицидний білок, оптимізована для експресії в рослинах,
нуклеїнова кислота, P77853:S158-30-108-35

<400> 39
atgcaaacaa gcattactct gacatccaac gcatccggta cgtttgacgg ttacaattac 60

gaactctgga aggatactgg caatacaaca atgacggtct acactcaagg tcgcttttcc 120

tgccagtggt cgaacatcaa taacgcgttg tttaggaccg ggaagaaata caaccagaat 180

tggcagtctc tcggcacaat ccggatcacg tactccgcga cttacaaccc aaacgggaac 240

tcctacatgt gtatctatgg ctggtctacc aacccattgg tcgagttcta catcgttgag 300

tcctgggggа actggagacc gcctggtgcc acgtccctgg gccaagtgac aatcgatggc 360

gggacctacg acatctatag gacgacacgc gtcaatcagc cttgcctggc cgagggctcg 420

ctcgtcttgg acgcggctac cgggcagagg gtcctatcg aaaaggtgcg tcccgggatg 480

gaagttttct ccttgggacc tgattacaga ctgtatcggg tgccccgtttt ggaggtcctt 540

gagagcgggg ttggggaagt tgtgcgcctc cgaactcggt cagggagaac gctggtgttg 600

acaccagaac accgctttt gaccccgaa ggttggaaac ctctttgtga cctcccgctt 660

ggaactccaa ttgcagtccc cgcagaactg cctgtggcgg gccacttggc cccacctgaa 720

gaacgtgtta tgctcctggc tcttctgttg ggggatggga acacaaagct gtcgggtcgg 780

agaggtacac gtcctaatgc cttcttctac agcaaagact ccgaattgct cgcggcttat 840

cgccggtgtg cagaagcctt gggtgcaaag gtgaaagcat acgtccaccc gactacgggg 900

gtggttacac tcgcaaccct cgctccacgt cctggagctc aagatcctgt caaacgcctc 960

gttgtcgagg cgggaatggt tgctaaagtc gaagagaaga gggtcccgga ggaggtgttt 1020

cgttaccggc gtgaggcgtt ggcccttttc ttgggccgtt tgttctcgac agacggctct 1080

gttgaaaaga agaggatctc ttattcaagt gccagtttgg gactggccca ggatgtcgca 1140

catctcttgc tgcgccttgg aattacatct caactccgtt cgagagggcc acgggctcac 1200

aaggttctta tatcgggccg cgaggatatt ttgcggtttg ctgaacttat cggaccctac 1260

ctcttggggg ccaagaggga gagacttgca gcgctggaag ctgaggcccg caggcgtttg 1320

cctggacagg gatggcactt gcggcctgtt cttcctgccg tggcgtacag agtgagcgag 1380

gctaaaaggc gctcgggatt ttcgtggagt gaagccggtc ggcgcgtcgc agttgcggga 1440

tcgtgtttgt catctggact caacctcaaa ttgcccagac gctacctttc tcggcaccgg 1500

ttgtcgctgc tcggtgaggc ttttgccgac cctgggctgg aagcgctcgc ggaaggccaa 1560

gtgctctggg accctattgt tgctgtcgaa ccggccggta aggcgagaac attcgacttg 1620

cgcgtcccac cctttgcaaa cttcgtgagc gaggacctgg tggtgcataa ctccattgtg 1680

gggacagcca cgttcgatca gtactggagc gtgcgcacct ctaagcggac ttcaggaaca 1740

gtgaccgtga ccgatcactt ccgcgcctgg gcgaaccggg gcctgaacct cggcacaata 1800

gaccaaatta cattgtgcgt ggagggttac caaagctctg gatcagccaa catcacccag 1860

aacaccttct ctcagggctc ttcttccggc agttcgggtg gctcatccgg ctccacaacg 1920

actactcgca tcgagtgtga gaacatgtcc ttgtccggac cctacgttag caggatcacc 1980

aatcccttta atggtattgc gctgtacgcc aacggagaca cagcccgcgc taccgttaac 2040

ttccccgcaa gtcgcaacta caatttccgc ctgcggggtt gcggcaacag caataatctt 2100

gcccgtgtgg acctgaggat cgacggacgg accgtcggga cctttatta ccagggcaca 2160

tacccctggg aggccccaat tgacaatgtt tatgtcagtg cggggagtca tacagtcgaa 2220

atcactgtta ctgcggataa cggcacacgg gacgtgtatg ccgactacct ggtgatacag 2280

tga 2283

<210> 40
<211> 735
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ендоглюканазу.

оптимізована для експресії в рослинах, нуклеїнова кислота, 033897

<400> 40
atgtgcgatt ggctgttccc tgatggagat aatggaaagg agcccgaacc agaaccagaa 60

ccaacggtag agctctgcgg aagatgggat gcaagagacg tagcgggagg aagatataga 120

gttattaaca atgtgtgggg ggccgaaaca gcacagtgta tagaagtagg tcttgaaaca 180

ggtaatttta cgataacaag agccgatcat gacaatggta ataatgttgc agcatatcca 240

gcaatttact tcggatgtca ttgggcacca gcgagagcaa taagggattg tgccgcgcgt 300

gcgggagcgg ttaggagagc acacgaattg gatgttacac caattaccac gggaagatgg 360

aacgcagctt acgatatatg gtttagtccg gtaacaaact ctggtaacgg gtactcggga 420

ggcgccgaac ttatgatatg gctgaattgg aatggtggcg taatgccagg aggatcacgg 480

gtagcaactg tcgaattggc tggagcgaca tgggaagtgt ggtatgcaga ttgggattgg 540

aattacatcg catatagaag aacgactccg acaacctcag tgagtgaact tgacttgaaa 600

gcctttattg atgatgcagt agcgagagga tacataaggc cagaatggta tctgcatgca 660

gtggaaacgg gatttgaatt gtgggaaggg ggggctgggt tgaggacagc agattttagc 720

gtaactgtac agtga 735

<210> 41
<211> 2610
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча целобіогідролазу, оптимізована для експресії в рослинах, нуклеїнова кислота, 068438

<400> 41
atgctggagg acaagtctcc caaactgcct gattataaga acgaccttct gtacgaacgc 60

acattcgacg aggggctctg cttccgtgg cacacgtgcg aagattcagg agggaaatgc 120

gattttgccg tggtcgacgt tccaggcgag cctgggaaca aggcgttcag gctcactgtt 180

atcgataagg gtcagaacaa gtggtcggtc caaatgagac acggggtat cacgttggag 240

caggggcaca tatacaccgt tcggtttact atctggagcg acaagagctg ccgcgtgtat 300

gccaaaatcg gccaaatggg tgaaccctac acggagtact ggaacaataa ctggaatccg 360

ttcaacctca ctccggggca gaaattgacg gtggaacaga actttactat gaattatccc 420

acggacgaca cgtgtgagtt taccttccac ttgggagggg aactggcagc cgggacccct 480

tactacgtgt acctcgacga cgtttctctt tacgatcccc gctttgtcaa gccagtggaa 540

tacgtcctgc ctcaacggga tgtcagggtt aatcaagttg gatacctccc ttttgctaag 600

aaatatgcta ctgtcgtgtc atcgagcacg tcccattga agtggcaact tctgaatagt 660

gcaaaccaag ttgtcttgga gggcaatacа atcccaagg gactggacaa agattcacaa 720

gactacgttc attggatcga tttctcgaac tttaagaccg aaggcaaggg gtactatttc 780

aagttgccca ctgtgaactc cgatactaac tactcccacc cgtttgatat ttctgcagat 840

atctattcaa agatgaagtt cgacgcgctc gctttctttt accataaaag gtcgggaata 900

ccaatcgaga tgccctacgc cgggggagag cagtggacaa ggcccgcagg gcacattggt 960

gtcgcgccga acaagggcga cacgaatgtg ccaacttggc cccaggatga cgaatatgct 1020

ggacgccccc agaaatacta tacgaaagac gtgaccggcg ggtggtacga tgccggtgac 1080

cacggcaagt acgtcgtgaa cgggggtatc gcagtttgga cccttatgaa tatgtacgag 1140

agagcaaaga ttagaggaac cgctaaccag ggtgcctaca aagatggagg aatgaatatc 1200

ccggaaagga ataacggcta tcctgatatt ctggacgagg ccagatggga gatcgaattt 1260

tttaagaaga tgcaagtcac tgagaaagaa gatccgtcga ttgcaggtat ggtgcaccac 1320

aagatccacg atttcaggtg gacggcgctc ggaatgttgc ctcacgagga cccccagcca 1380

cgctaccttc ggcccgtcag cacagcggca accctgaatt tcgcagcgac cctcgctcag 1440

tctgccagat tgtggaagga ttacgacccg acttttgcag cggactgcct tgagaaagct 1500

gaaattgcct ggcaagcagc actcaaacac ccggacatct acgctgagta cacgccagga 1560

agcggtgggc cgggtggagg tccttataat gacgattatg tcggggacga gttctactgg 1620

gccgcttgtg aactctatgt gacaaccggt aaggatgagt acaagaatta cttgatgaat 1680

agtccgcact atctggaaat gccagcgaag atgggcgaga acggaggggc taacggcgag 1740

gacaacggtc tctggggctg ctttacttgg ggaacgacac aggggttggg tacaattacc 1800

cttgccctcg ttgaaaacgg cctcccttcg gcggatattc aaaaggcccg caacaatatc 1860

gctaaagccg cagataagtg gcttgagaat attgaagaac aaggttacgg cctgcctatc 1920

aaacaagcgg aggatgaacg gggcggatac cgtggggtа gtaattcttt cattctcaac 1980

cagatgatcg tcatgggcta cgcttacgac ttcacgggaa acagcaagta tcttgacggg 2040

atgcaggacg gcatgtccta cctgctcggt agaaacggac ttgatcaatc gtacgttact 2100

gggtacgggg agaggccact tcagaacccc cacgaccgct tttggacccc tcaaacttcg 2160

aagaaattcc cggccccacc ccctggtatt atcgcaggcg ggccgaatag ccggtttgaa 2220

gacccaacga tcactgcagc ggttaagaag gatacacccc cgcagaagtg ctatattgac 2280

cacaccgatt cctggtctac taacgagatc acgattaatt ggaacgcccc cttcgcgtgg 2340

gtcacagcgt atctggacga aattgacttg attaccccac ccggcggagt ggaccctgaa 2400

gagccggaag ttatctacgg tgattgtaac ggcgacggaa aggttaattc gaccgatgct 2460

gtggccctta aaaggtatat cctccgcagc ggtatctcga tcaacacgga caacgcggac 2520

gttaatgcag atggtcgcgt gaatagcact gacctcgcta ttttgaagcg ctatattttg 2580

aaggagatcg atgttcttcc tcacaagtga 2610

<210> 42
<211> 1569
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ендоглюканазу, оптимізована для експресії в рослинах, нуклеїнова кислота, P54583

<400> 42
atggctggag gaggatactg gcacacttcc ggcagggaga tcctcgacgc aaataacgtt 60

ccagtcagaa tcgccgggat taattggttt ggcttcgaaa cgtgtaacta cgtggttcac 120

ggcctgtggt ctcgggatta cagatcaatg ctcgaccaga tcaaatcctt ggggtataat 180

acaattaggc tgccctacag cgatgacatt cttaagcctg gaaccatgcc gaactcgatt 240

aatttctacc aaatgaacca ggatctgcag ggattgacct ctctgcaggt tatggacaag 300

atcgtggcgt acgccggcca aatcgggctc agaattattt tggatcggca caggccagac 360

tgctcaggtc agtcggccct gtggtacaca agctccgtgt cagaggcaac atggatttca 420

gatcttcaag ccctcgcaca acgctataaa ggcaacccca cggttgtggg attcgacctt 480

cacaacgaac ctcacgatcc ggcctgttgg ggctgcgggg acccttcgat cgactggaga 540

ctggcagcgg agagggctgg taacgccgtt ctcagcgtca atcccaactt gctgatcttt 600

gtggagggag ttcagtccta caacggcgat tcttactggt ggggcggaaa tctccaaggc 660

gcagggcagt atcctgtcgt gcttaacgtt ccgaatcgcc tggtctactc agcacacgac 720

tacgcgacta gcgtgtaccc acagacgtgg ttctccgatc ccacatttcc taacaatatg 780

ccgggaatct ggaacaagaa ttggggttac ttgtttaacc aaaacattgc tccagcttgg 840

ttgggtgaat ttggcaccac tcttcagtcg acgacagacc aaacctggct gaaaacccta 900

gtccagtatt tgcggccaac tgctcagtac ggagcagatt cttttcaatg gacgttctgg 960

tcttggaatc ctgactccgg ggatacaggc ggtatcctga aagacgattg gcagaccgtg 1020

gacactgtta aggacgggta cttggcgccg attaaaagct cgatctttga ccccagtcggc 1080

gctagcgctt ccccatcttc acaaccttcg ccgagcgtca gccccagccc aagcccaagc 1140

ccgtctgcca gcagaacccc cactcccaca cctaccccca cggcctcacc aactccgacg 1200

ctcactccta cggcgacgcc aacaccaact gcttcaccca ctcctagccc caccgcagcg 1260

agcggggcta ggtgcaccgc ttcttaccag gtcaactctg actggggtaa tggcttcacc 1320

gtgactgtgg cggtcactaa ctcaggaagc gtcgcgacga aaacctggac tgtgtcctgg 1380

acgttcgggg gcaaccaaac aatcaccaac agctggaacg ctgcagttac gcagaatggg 1440

caaagcgtca cggcgcgcaa tatgagctac aacaacgtga ttcaaccagg ccagaatacc 1500

acattcggtt ttcaagcaag ctataccggg tcaaacgctg cccaactgt cgcttgtgct 1560

gcctcatga 1569

<210> 43
<211> 1542
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча целобіогідролазу, оптимізована для експресії в рослинах, нуклеїнова кислота, BD22308

<400> 43
atgcagcaga tcggcacctа caccgccgag acccacccaa gcctgagctg gtccacctgc 60

aagagcggcg gttcctgcac gaccaacagc ggcgccatca ccctgatgc gaactggcgc 120

tgggtgcacg gcgtgaacac cagcaccaac tgctacacgg gtaacacgtg gaacaccgcc 180

atctgcgaca cggacgcttc ctgcgcccag gactgcgcgc ttgatggcgc cgactactcc 240

ggcacctacg gcatcaccac ctcgggcaac agcctgcgcc tgaacttcgt gaccggcagc 300

aatgtgggca gccgcaccta cctgatggcc gacaacaccc actaccagat cttcgacctg 360

ctgaaccagg agttcacctt caccgtcgac gtgtcccacc tgcctgcgg cctgaacggc 420

gccctctact tcgtgacgat ggacgccgac ggcggcgtgt ccaagtaccc gaacaacaag 480

gctggcgccc agtacggtgt gggctactgc gacagccagt gccccgaggga cctgaagttc 540

atcgccggcc aggccaacgt ggagggctgg acccccgagca gcaacaacgc caacaccggc 600

ctgggcaacc acggcgcctg ctgcgccgag ctggacatct gggaggccaa cagcatcagc 660

gaggccctga ccccacaccc atgcgacacc ccaggcctgt ctgtgtgcac caccgacgcc 720

tgcggcggca cctactccag cgaccgctac gcggcaccct gcgacccaga cggctgcgac 780

ttcaacccgt accgcctggg cgtgaccgac ttctacggca gcggcaagac cgtggacacc 840

accaagccga tcaccgtggt gacccagttc gtgaccgacg acggcaccag caccggcacc 900

ctgagcgaga tccgccgcta ctacgtccag aacggcgtgg tgatcccgca gccgagcagc 960

aagatcagcg gcgtgtccgg caacgtgatc aacagcgact tctgcgacgc cgagatcagc 1020

accttcggcg agaccgccag cttcagcaag cacggcggcc tggccaagat gggcgctggc 1080

atggaagccg gcatggtgct ggtgatgagc ctgtgggacg actactccgt gaacatgctg 1140

tggctggaca gcacctaccc gaccaacgcc accgggacgc caggcgctgc caggggcagc 1200

tgcccaacca cctcgggcga ccccaagacc gtcgagagcc agagcggcag cagctacgtg 1260

accttcagcg acatccgcgt gggcccgttc aactccacgt tcagcggtgg ctctagcacg 1320

ggcggctcct ccaccaccac cgccagcggc accaccacca ccaaggcctc cagcacgtct 1380

actagctcca cctctaccgg caccggcgtt gctgcccatt ggggccagtg cggtggccag 1440

ggctggacgg gtccaacgac ttgcgcctcc ggcaccacct gcaccgtggt caatccgtac 1500

tactcccagt gcctggacga gctgaaggcc gaggccaagt ga 1542

<210> 44
<211> 78
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча націлюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, BAASS

<400> 44
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaagtc 78

<210> 45
<211> 129
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча націлюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, HvAleSP

<400> 45
atggcccacg ccgcgctcct cctcctgggg ctcgccgtcc tggccaccgc cgccgtcgcc 60

gtcgcctcct cctcctcctt cgccgactcc aaccccgatc gcccggtgac cgaccgcgcc 120

gcctccacc 129

<210> 46
<211> 90
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча націлюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, PR1a

<400> 46
atgggcttcg tgctcttctc ccagctgcct tccttccttc ttgtctccac ccagctcttg 60

ttcctcgtga tctcccactc ctgccgcgcc 90

<210> 47
<211> 87
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча націлюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, GZein27ss-02

<400> 47
atgcgcgtgc tgctcgtggc crtggccctg ctggctcttg ctgccagcgc cacctct 57

<210> 48
<211> 72
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ацилюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, GluB4SP

<400> 48
atggccacca tgctttttct ccgcttgtcc atctacttct gcgtgcttct cctgtgccac 60

ggctccatgg cc 72

<210> 49
<211> 18
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ацилюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, SEKDEL

<400> 49
agcgagaagg acgagctg 18

<210> 50
<211> 12
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ацилюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, KDEL

<400> 50
aaggacgagc tg 12

<210> 51
<211> 24
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ацилюючий пептид, оптимізована для експресії в рослинах, нуклеїнова кислота, HyVSD

<400> 51
gacgagctga aggccgaggc caag 24

<210> 52

<211> 1086
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча кислазу, злиту
з націлюючим пептидом, оптимізована для експресії
в рослинах, нуклеїнова кислота, BAASS:P77853

<400> 52
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaacaaac aagcattact ctgacatcca acgcatccgg tacgtttgac 120

ggttactatt acgaactctg gaaggatact ggcaatacaa caatgacggt ctacactcaa 180

ggtcgctttt cctgccagtg gtcgaacatc aataacgcgt tgtttaggac cgggaagaaa 240

tacaaccaga attggcagtc tcttggcaca atccggatca cgtactctgc gacttacaac 300

ccaaacggga actcctactt gtgtatctat ggctggtcta ccaacccatt ggtcgagttc 360

tacatcgttg agtcctgggg gaactggaga ccgcctggtg ccacgtccct gggccaagtg 420

acaatcgatg gcgggaccta cgacatctat aggacgacac gcgtcaacca gccttccatt 480

gtggggacag ccacgttcga tcagtactgg agcgtgcgca cctctaagcg gacttcagga 540

acagtgaccg tgaccgatca cttccgcgcc tgggcgaacc ggggcctgaa cctcggcaca 600

atagaccaaa ttacattgtg cgtggagggt taccaaagct ctggatcagc caacatcacc 660

cagaacacct tctctcaggg ctcttcttcc ggcagttcgg gtggctcatc cggctccaca 720

acgactactc gcatcgagtg tgagaacatg tccttgtccg gaccctacgt tagcaggatc 780

accaatccct ttaatggtat tgcgctgtac gccaacggag acacagcccg cgctaccgtt 840

aacttccccg caagtcgcaa ctacaatttc cgcctgcggg gttgcggcaa caacaataat 900

cttgcccgtg tggacctgag gatcgacgga cggaccgtcg ggacctttta ttaccagggc 960

acataccccc gggaggcccc aattgacaat gtttatgtca gtgcggggag tcatacagtc 1020

gaaatcactg ttactgcgga taacggcaca tgggacgtgt atgccgacta cctggtgata 1080

cagtga 1086

<210> 53
<211> 813
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ендоглюканазу,
злиту з націлюючим пептидом, оптимізована для експресії
в рослинах, нуклеїнова кислота, BAASS:033897

<400> 53
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaacaggt gtgcgattgg ctgttcccctg atggagataa tggaaaggag 120

cccgaaccag aaccagaacc aacggtagag ctctgcggaa gatgggatgc aagagacgta 180

gcgggaggaa gatatagagt tattaacaat gtgcgggggg ccgaaacagc acagtgtata 240

gaagtaggtc ttgaaacagg taattttacg ataacaagag ccgatcatga caatggtaat 300

aatgttgcag catatccagc aatttacttc ggatgtcatt gggcaccagc gagagcaata 360

agggattgtg ccgcgcgtgc gggagcggtt aggagagcac acgaattgga tgttacacca 420

attaccacgg gaagatggaa cgcagcttac gatatatggt ttagtccggt aacaaactct 480

ggtaacgggt actcgggagg cgccgaactt atgatatggc tgaattggaa tggtggcgta 540

atgccaggag gatcacgggt agcaactgtc gaattggctg gagcgacatg ggaagtgtgg 600

tatgcagatt gggattggaa ttacatcgca tatagaagaa cgactccgac aacctcagtg 660

agtgaacttg acttgaaagc ctttattgat gatgcagtag cgagaggata cataaggcca 720

gaatggtatc tgcatgcagt ggaaacggga tttgaattgt gggaaggggg ggctgggttg 780

aggacagcag attttagcgt aactgtacag tga 813

<210> 54
<211> 1428
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, похідна ендоглюканазу, злиту з найлюким пептидом, оптимізована для експресії в рослинах, нуклеїнова кислота, HvAleSP:NtEGM

<400> 54
atggcccacg cccgcgtcct cctcctggcg ctcgccgtcc tggccaccgc cgccgtcgcc 60

gtcgcctcct cctcctcctt cgccgactcc aacccgatcc gcccggtgac cgaccgcgcc 120

gcctccaccg cttacgacta caagcaggtg ttgcgggact cgctactatt ctatgaggcc 180

cagagatccg gccggctccc agccgaccag aaggtcacgt ggaggaagga tagcgcgctg 240

aatgaccagg gtgaccaggg acaagacttg accggcggct actttgacgc tggggacttc 300

gtcaagttcg ggttccccat ggcttatacc gcaaccgtgc tggcatgggg cctcatagac 360

tttgaggccg gctacagcag tgccggggcc ttggatgatg gacggaaggc tgtcaaatgg 420

gccaccgact atttcataaa ggcccacaca agtcaaaatg agttctatgg tcaggtcggc 480

cagggtgacg ccgatcacgc tttctgggga agaccagagg atatgacgat ggcgcgcccg 540

gcgtacaaga tagacacctc aaggcctggc tctgatctgg caggcgagac agcggctgct 600

cttgccgctg cttcaatcgt gttccggaac gtcgatggca cttactcaaa taacctgtta 660

acacacgctc gccagctatt cgacttcgcg aacaactacc ggggaaagta tagtgactct 720

attactgacg caagaaattt ctacgcaagc gcagactaca gagacgagtt ggtttgggct 780

gctgcgtggt tatacagagc gaccaacgac aacacctacc tcaacactgc tgagtcactg 840

tacgatgagt ttgggctaca gaactggggg gggggcctga actgggatag caaggtgtct 900

ggcgtgcagg tgttgttggc caagcttacc aataagcagg cctacaagga cacggtgcag 960

tcttacgtca attacctaat taataaccag cagaagactc ccaagggcct cctctacatc 1020

gacatgtggg gcacccttcg ccacgctgcc aacgccgcat tcatcatgct cgaagccgcc 1080

gagctgggct tgtccgcctc ctcttataga cagttcgcgc aaacgcaaat cgactacgcc 1140

ctgggcgatg gtggccgctc ctttgtgtgc gggttcggga gcaatcctcc tacgagaccg 1200

caccacagat cctcgtcgtg cccgccagct cccgctactt gcgactggaa tacattcaac 1260

tcacctgacc caaactacca cgtcctctct ggggcctcag tgggcggacc tgatcagaat 1320

gacaactacg tcgatgaccg ttcagactat gttcacaacg aagtcgccac tgattacaac 1380

gcgggtttcc agtccgcgtt agctgctttg gtggccttg gttactga 1428

<210> 55
<211> 2355
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча
модифіковану інтеїном коканазу, злиту з
націлюючим пептидом, оптимізована для експресії
в рослинах, нуклеїнова кислота, BAASS:P77853:S158-30-108-35

<400> 55
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaacaaac aagcattact ctgacatcca acgcatccgg tacgtttgac 120

ggttacaatt acgaactctg gaaggatact ggcaatacaa caatgacggt ctacactcaa 180

ggtcgctttt cctgccagtg gtcgaacatc aataacgcgt tgtttaggac cgggaagaaa 240

tacaaccaga attggcagtc tctcggcaca atccggatca cgtactctgc gacttacaac 300

ccaaacggga actcctacat gtgtatctat ggctggtcta ccaacccatt ggtcgagttc 360

tacatcgttg agtcctgggg gaactggaga ccgcctggtg ccacgtccct gggccaagtg 420

acaatcgatg cgcgggacct cgacatctat aggacgacac gcgtcaacca gccttgcctg 480

gccgagggct cgctcgtctt ggacgcggct accgggcaga gggtccctac cgaaaaggtg 540

cgtccgggga tggaagtttt ctccttggga cctgattaca gactgtatcg ggtgccccgtt 600

ttggaggtcc ttgagagcgg ggttggggaa gttgtgcgcc tccgaactcg gtcagggaga 660

acgctggtgt tgacaccaga acaccccgctt ttgaccccccg aaggttggaa accctctttgt 720

gacctcccgc ttggaactcc aattgcagtc cccgcagaac tgcctgtggc gggccacttg 780

gccccacctg aagaacgtgt tatgctcctg gctcttctgt tgggggatgg gaacacaaag 840

ctgtcgggtc ggagaggtac acgtcctaat gccttcttct acagcaaaga ctccgaattg 900

ctcgcggctt atcgccggtg tgcagaagcc ttgggtgcaa aggtgaaagc atacgtccac 960

```
ccgactacgg gggtggttac actcgcaacc ctcgctccac gtcctggagc tcaagatcct    1020
gtcaaacgcc tcgttgtcga ggcgggaatg gttgctaaag tcgaagagaa gagggtcccg    1080
gaggaggtgt ttcgttaccg gcgtgaggcg ttgggccttt tcttgggccg tttgttctcg    1140
acagacggct ctgttgaaaa gaagaggatc tcttattcaa gtgccagttt gggactggcc    1200
caggatgtcg cacatctctt gctgcgcctt ggaattacat ctcaactccg ttcgagaggg    1260
ccacgggctc acaaggttct tatatcgggc cgcgaggata ttttgcggtt tgctgaactt    1320
atcggaccct acctcttggg ggccaagagg gagagacttg cagcgctgga agctgaggcc    1380
cgcaggcgtt tgcctggaca gggatggcac ttgcggcctg ttcttcctgc cgtggcgtac    1440
agagtgagcg aggctaaaag gcgctcggga ttttcgtgga gtgaagccgg tcggcgcgtc    1500
gcagttgcgg gatcgtgttt gtcatctgga ctcaacctca aattgcccag acgctacctt    1560
tctcggcacc ggttgtcgct gctcggtgag gcttttgccg accctgggct ggaagcgctc    1620
gcggaaggcc aagtgctctg ggaccctatt gttgctgtcg aaccggccgg taaggcgaga    1680
acattcgact tgcgcgtccc accctttgca aacttcgtga gcgaggacct ggtggtgcat    1740
aactccattg tggggacagc cacgttcgat cagtactgga gcgtgcgcac ctctaagcgg    1800
acttcaggaa cagtgaccgt gaccgatcac ttccgcgcct gggcgaaccg gggcctgaac    1860
ctcggcacaa tagaccaaat tacattgtgc gtggagggtt accaaagctc tggatcagcc    1920
aacatcaccc agaacacctt ctctcagggc tcttcttccg gcagttcggg tggctcatcc    1980
ggctccacaa cgactactcg catcgagtgt gagaacatgt ccttgtccgg accctacgtt    2040
agcaggatca ccaatccctt taatggtatt gcgctgtacg ccaacggaga cacagccgcc    2100
gctaccgtta acttcccgcc aagtcgcaac tacaatttcc gcctgcgggg ttgcggcaac    2160
agcaataatc ttgcccgtgt ggacctgagg atcgacggac ggaccgtcgg gaccttttat    2220
taccagggca ataccccctg ggaggccccc attgacaatg tttatgtcag tgcggggagt    2280
catacagtcg aaatcactgt tactgcggat aacggcacac gggacgtgta tgccgactac    2340
ctggtgatac agtga                                                    2355

<210>  56
<211>  1611
<212>  ДНК
<213>  Штучна послідовність

<220>
<223>  Синтетична послідовність, кодуюча
       фередоксинестеразу, злиту з націльовими
       пептидами, оптимізована для експресії в рослинах,
       нуклеїнова кислота, EAASS:AnfaeB:SEKDEL

<400>  56
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc      60
```

ttggcctccg ggcaagtcgc tactgatccc ttccagtctc gctgtaacga gtttcaaaac 120

aaaattgata tcgcaaacgt cacggttagg tcggtcgcct acgtcgcggc tggtcagaat 180

atctctcaag cagaagtcgc ttcggtttgc aaggcaagcg tccaggcctc cgtggacctc 240

tgtagagtga caatgaatat tagcacctcg gataggagcc acctttgggc cgaggcttgg 300

ctgcctcgca attatactgg acggttcgtc tccactggga acgggggcct ggctggatgc 360

gtgcaggaaa cagacttgaa tttcgctgcc aactttggct tcgcaacgct tggcactaat 420

ggagggcacg atggagcacc ggcgaaatac tttctgaata actctgaggt cctcgctgat 480

ttcgcctatc gcagcgttca cgagggcact gtggtcggga agcaacttac ccagttgttc 540

tacgacgaag gttacaatta ttcgtactat ctgggctgca gcaccggagg caggcaaggc 600

tatcagcaag tgcagcgctt tccggatgac tacgatggtg ttattgccgg aagtgcagcg 660

atgaacttta tcaatttgat ttcttggggc gcttttcttt ggaaggccac gggactcgca 720

gacgacccag attttatctc agcgaacctc tggagcgtta tccaccagga gattgtgcgg 780

cagtgtgact tggttgatgg tgctcttgac ggaattattg aagatcctga cttctgcgcc 840

cctgttatcg agagactgat ttgtgatggg actaccaacg gcacctcctg tatcactggg 900

gcacaagcgg ctaaagtgaa tagggcctc tcggactttt atggtccgga tgggacagtc 960

tactatccac gcttgaacta cggaggggag gcagactctg cgagcctgta cttcacaggc 1020

tccatgtatt cacggaccga agagtggtac aagtacgtcg tgtacaacga tactaactgg 1080

aatagttctc aatggacgct cgaaagcgct aagttggcct tggagcagaa ccccttaat 1140

attcaggcat tcgaccctaa tatcacagcg ttcagggacc gcggtggaaa actcttgagc 1200

taccacggca cccaagatcc gattatttca tcgactgaca gcaagttgta ttacagaaga 1260

gttgctaacg cccttaacgc tgcccatcc gaactcgacg agttttatcg cttctttcag 1320

atctctggga tgggtcactg cggcgacgga accggggctt catacattgg ccagggatac 1380

ggcacgtata cctcgaaggc cccacaggtc aacttgcttc ggactatggt ggattgggtt 1440

gagaatggaa aagctccgga atatatgcct gggaataagt tgaacgccaa cggttcgatc 1500

gagtatatga gaaagcactg tcgttaccca aaacacaata ttcacacggg cccaggtaac 1560

tacacggacc ctaactcctg gacctgcgtc agcgagaagg acgagctgtg a 1611

<210> 87
<211> 879
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча
феруліоїлестеразу, злиту з нашаруючими
пептидами, оптимізована для експресії в рослинах,
нуклеїнова кислота, BAASS:AnfaeA:SEKDEL

<400> 57
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaagtcgc atccacgcag ggtatctctg aggacctgta taatcgcctc 120

gtggaaatgg ccacaatttc acaagcggct tacgcagatc tttgtaatat ccctccgaca 180

attatcaagg gagagaaaat ctataacgcc cagactgaca tcaacggctg gatactgcgg 240

gatgacacga gcaaggaaat tatcacagtc tttagaggga ccggttccga tacaaatttg 300

cagttggaca cgaattacac actgaccccc ttcgatactc tccctcaatg caacgactgt 360

gaggttcacg gtgggtacta tattggctgg atcttctgtt caagaccaagt cgagtcactt 420

gttaagcagc aagcgtcgca gtacccggac tacgcattga cggtgacagg gcacagcctg 480

ggtgcctcga tggcagcgct caccgctgcc cagctttctg caacctacga taatgtcagg 540

ctgtacactt tcggagaacc acgctcaggc aaccaagcgt ttgcttcgta tatgaacgac 600

gctttccagg ttagctcccc cgagacgaca caatactttc gggtgaccca ttctaacgac 660

gggattccta acctcccgcc agccgacgaa ggttacgcac acggggggtgt cgagtactgg 720

tcagtggacc cctacagcgc gcagaatact ttcgtttgca cgggcgatga ggtccagtgc 780

tgtgaagctc aaggggggtca gggagtgaat gatgcacaca caacctattt cggaatgact 840

tccggggctt gcacgtggag cgagaaggac gagctgtga 879

<210> 58
<211> 1407
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ендоглюканазу, злиту з націлюючими пептидами, оптимізована для експресії в рослинах, Нуклеїнова кислота, PR1a:NtEGM:SEKDEL

<400> 58
atgggcttcg tgctcttctc ccagctgcct tccttccttc ttgttctccac cctgctcttg 60

ttcctcgtga tctcccactc ctgccgcgcc gcttacgact acaagcaggt gttgcgggac 120

tcgctactat tctatgaggc ccagagatcc ggccggctcc cagccgacca gaaggtcacg 180

tggaggaagg atagcgcgct gaatgaccag ggtgaccagg gacaagactt gaccggcggc 240

tactttgacg ctggggactt cgtcaagttc gggttccccca tggcttatac cgcaaccgtg 300

ctggcatggg gcctcataga ttttgaggcc ggctacagca gtgccggggc cttggatgat 360

ggacggaagg ctgtcaaatg ggccaccgac tatttcataa aggcccacac aagtcaaaat 420

gagttctatg gtcaggtcgg ccagggtgac gccgatcacg ctttctgggg aagaccagag 480

gatatgacga tggcgcgccc ggcgtacaag atagacacct caaggcctgg ctctgatctg 540

gcaggcgaga cagcggctgc tcttgccgct gcttcaatcg tgttccggaa cgtcgatggc 600

acttactcaa ataacctgtt aacacacgct cgccagctat tcgacttcgc gaacaactac 660

cggggaaagt atagtgactc tattactgac gcaagaaatt tctacgcaag cgcagactac 720

agagacgagt tggtttgggc tgctgcgtgg ttatacagag cgaccaacga caacacctac 780

ctcaacactg ctgagtcact gtacgatgag tttgggctac agaactgggg gggggcctg 840

aactgggata gcaaggtgtc tggcgtgcag gtgttgttgg ccaagcttac caataagcag 900

gcctacaagg acacggtgca gtcttacgtc aattacctaa ttaataacca gcagaagact 960

cccaagggcc tcctctacat cgacatgtgg ggcaccctc gccacgctgc caacgccgca 1020

ttcatcatgc tcgaagccgc cgagctgggc ttgtccgcct cctcttatag acagttcgcg 1080

caaacgcaaa tcgactacgc cctgggcgat ggtggccgct cctttgtgtg cgggttcggg 1140

agtaatcctc ctacgagacc gcaccacaga tcctcgtcgt gcccgccagc tcccgctact 1200

tgcgactgga atacattcaa ctcacctgac ccaaactacc acgtcctctc tggggcccta 1260

gtgggcggac ctgatcagaa tgacaactac gtcgatgacc gttcagacta tgttcacaac 1320

gaagtcgcca ctgattacaa cgcgggtttc cagtccgcgt tagctgcttt ggtggccctt 1380

ggttacagcg agaaggacga gctgtga 1407

<210> 59
<211> 2376
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча
модифіковане інтеїном ксиланазу, злиту з
націлюючими пептидами, оптимізовано для експресії
в рослинах, нуклеїнова кислота,
BAASS:P77853:T134-100-101:SEKDEL

<400> 59
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaacaaac aagcattact ctgacatcca acgcatccgg tacgtttgac 120

ggttactatt acgaactctg gaaggatact ggcaatacaa caatgacggt ctacactcaa 180

ggtcgctttt cctgccagtg gtcgaacatc aataacgcgt tgtttaggac cgggaagaaa 240

tacaaccaga attggcagtc tcttggcaca atccggatca cgtactctgc gacttacaac 300

ccaaacggga actcctactt gtgtatctat ggctggtcta ccaaccccatt ggtcgagttc 360

tacatcgttg agtcctgggg gaactggaga ccgcctggtg cctgcctggc cgagggctcg 420

ctcgtcttgg acgcggctac cgggcagagg gtccctatcg aaaaggtgcg tccgggatg 480

gaaagttttct ccttgggacc tgattacaga ctgtatcggg tgcccgtttt ggaggtcctt 540

gagagcgggg ttagggaagt tgtgcgcctc agaactcggt cagggagaac gctggtgttg 600

acaccagatc accgctttt gaccccgaa ggttggaaac ctctttgtga cctccgcct 660

ggaactccaa ttgcagtccc cgcagaactg cctgtggcgg gccacttggc cccacctgaa 720

gaacgtgtta cgctcctggc tcttctgttg ggggatggga acacaaagct gtcgggtcgg 780

agaggtacac gtcctaatgc cttcttctac agcaaaaacc ccgaattgct cgcggcttat 840

cgccggtgtg cagaagcctt gggtgcaaag gtgaaagcat acgtccaccc gactacgggg 900

gtggttacac tcgcaaccct cgctccacgt cctggagctc aagatcctgt caaacgcctc 960

gttgtcgagg cgggaatggt tgctaaagcc gaagagaaga gggtccccgg ggaggtgttt 1020

cgttaccggc gtgaggcgtt ggccctttte ttgggccgtt tgttctcgac agacggctct 1080

gttgaaaaga agaggatctc ttattcaagt gccagtttgg gactggccca ggatgtcgca 1140

catctcttgc tgcgccttgg aattacatct caactccgtt cgagagggcc acgggctcac 1200

gaggttctta tatcgggccg cgaggatatt ttgcggtttg ctgaacttat cggaccctac 1260

ctcttggggg ccaagaggga gagacttgca gcgctggaag ctgaggcccg caggcgtttg 1320

cctggacagg gatggcactt gcggcttgtt cttcctgccg tggcgtacag agtgggcgag 1380

gcggaaaggc gctcgggatt ttcgtggagt gaagccggtc ggcgcgtcgc agttgcggga 1440

tcgtgtttgt catctggact caacctcaaa ttgcccagac gctacctttc tcggcaccgg 1500

ttgtcgctgc tcggtgaggc ttttgccgac cctgggctgg aagcgctcgc ggaaggccaa 1560

gtgctctggg accctattgt tgctgtcgaa ccggccggta aggcgagaac attcgacttg 1620

cgcgttccac cctttgcaaa cttcgtgagc gaggacctgg tggtgcataa caccgtcccc 1680

ctgggccaag tgacaatcga tggcgggacc tacgacatct ataggacgac acgcgtcaac 1740

cagccttcca ttgtggggac agccacgttc gatcagtact ggagcgtgcg cacctctaag 1800

cggacttcag gaacagtgac cgtgaccgat cacttccgcg cctgggcgaa ccggggcctg 1860

aacctcggca caatagacca aattacattg tgcgtggagg gttaccaaag ctctggatca 1920

gccaacatca cccagaacac cttctctcag ggctcttctt ccggcagttc gggtggctca 1980

tccggctcca caacgactac tcgcatcgag tgtgagaaca tgtccttgtc cggaccctac 2040

gttagcagga tcaccaatcc ctttaatggt attgcgctgt acgccaacgg agacacagcc 2100

cgcgctaccg ttaacttccc cgcaagtcgc aactacaatt tccgcctgcg gggttgcggc 2160

aacaacaata atcttgcccg tgtggacctg aggatcgacg gacggaccgt cgggaccttt 2220

tattaccagg gcacataccc ctgggaggcc ccaattgaca atgtttatgt cagtgcgggg 2280

agtcatacag tcgaaatcac tgttactgcg gataacggca catgggacgt gtatgccgac 2340

taactggtga tacagagcga gaaggacgag ctgtga 2376

<210> 60
<211> 1446
<212> ДНК
<213> Штучна послідовність

<220>

<223> Синтетична послідовність, кодуюча ендоглюканазу, злиту з нецільовими пептидами, оптимізована для експресії в рослинах, нуклеїнова кислота, HvAleSP:NtEGM:SEKDEL

<400> 60
atggcccacg cccgcgtcct cctcctggcg ctcgccgtcc tggccaccgc cgccgtcgcc 60

gtcgcctcct cctcctcctt cgccgactcc aaccccgatcc gcccggtgac cgaccgcgcc 120

gcctccaccg cttacgacta caagcaggtg ttgcgggact cgctactatt ctatgaggcc 180

cagagatccg gcgggctccc agccgaccag aaggtcacgt ggaggaagga tagcgcgctg 240

aatgaccagg gtgaccaggg acaagacttg accggcggct actttgacgc tggggacttc 300

gtcaagttcg ggttccccat ggcttatacc gcaaccgtgc tggcatgggg cctcatagat 360

tttgaggccg gctacagcag tgccggggcc ttggatgatg gacggaaggc tgtcaaatgg 420

gccaccgact atttcataaa ggcccacaca agtcaaaatg agttctatgg tcaggtcggc 480

cagggtgacg ccgatcacgc tttctgggga agaccagagg atatgacgat ggcgcgcccg 540

gcgtacaaga tagacacctc aaggcctggc tctgatctgg caggcgagac agcggctgct 600

cttgccgctg cttcaatcgt gttccggaac gtcgatggca cttactcaaa taacctgtta 660

acacacgctc gccagctatt cgacttcgcg aacaactacc ggggaaagta tagtgactct 720

attactgacg caagaaattt ctacgcaagc gcagactaca gagacgagtt ggtttgggct 780

gctgcgtggt tatacagagc gaccaacgac aacacctacc tcaacactgc tgagtcactg 840

tacgatgagt ttgggctaca gaactggggg gggggcctga actgggatag caaggtgtct 900

ggcgtgcagg tgttgttggc caagcttacc aataagcagg cctacaagga cacggtgcag 960

tcttacgtca attacctaat taataaccag cagaagactc ccaagggcct cctctacatc 1020

gacatgtggg gcacccttcg ccacgctgcc aacgccgcat tcatcatgct cgaagccgcc 1080

gagctgggct tgtccgcctc ctcttataga cagttcgcgc aaacgcaaat cgactacgcc 1140

ctgggcgatg gtggccgctc ctttgtgtgc gggttcggga gtaatcctcc tacgagaccg 1200

caccacagat cctcgtcgtg cccgccagct cccgctactt gcgactggaa tacattcaac 1260

tcacctgacc caaactacca cgtcctctct ggggccctag tgggcggacc tgatcagaat 1320

gacaactacg tcgatgaccg ttcagactat gttcacaacg aagtcgccac tgattacaac 1380

gcgggtttcc agtccgcgtt agctgctttg gtggcccttg gttacagcga gaaggacgag 1440

ctgtga 1446

<210> 61
<211> 717
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча ксиланазу, злиту з нецільовими пептидами, оптимізована для

експресії в рослинах, нуклеїнова кислота,
BAASS:C43097:SEXDEL

<400> 61
atggcgaaca aacatttgtc cctctccctc ttcctcgtcc tccttggcct gtcggccagc 60

ttggcctccg ggcaagtctt cccagctgga aacgcaacgg aattggagaa aagacaaacc 120

acccctaact ctgagggctg gcatgacgga tactactact cttggtggag cgatggtggt 180

gcacaggcca cctatacaaa cctcgaaggc ggcacttatg agatttcatg gggtgacggt 240

ggcaaccttg tcggcggaaa ggggtggaac cccggactta acgccagggc aatccacttc 300

gaaggggtgt accagcccaa tggcaactca tacctggccg tctacgggtg gacgcgcaat 360

ccgctggttg agtactatat cgtggagaat ttcggaactt atgaccctag ctccggtgcc 420

acggacctcg ggacagtcga gtgtgacgga agcatctaca ggctgggtaa aactacccgc 480

gttaatgctc catcgatcga cggcacgcaa acatttgatc aatactggtc cgtgcggcag 540

gataagagga caagcggcac agttcagacg ggttgccact ttgatgcctg ggcaagagcg 600

gggctcaatg tgaatgggga ccactactat cagattgtgg cgaccgaggg ctatttctcc 660

agtggctatg cgcgtataac cgtcgctgat gttggaagcg agaaggacga gctgtga 717

<210> 62
<211> 1596
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, кодуюча
целобіогідролазу, злиту з націлюючими
пептидами, оптимізована для експресії в рослинах,
нуклеїнова кислота, GZein27ss-02:BD22308:HvVSD-01

<400> 62
atgcgcgtgc tgctcgtggc cctggccctg ctggctcttg ctgccagcgc cacctctcag 60

cagatcggca cctacaccgc cgagacccac ccaagcctga gctggtccac ctgcaagagc 120

ggcggttcct gcacgaccaa cagcggcgcc atcaccctlg atgcgaactg gcgctgggtg 180

cacggcgtga acaccagcac caactgctac acgggtaaca cgtggaacac cgccatctgc 240

gacacggacg cttcctgcgc ccaggactgc gcgcttgatg gcgccgacta ctccggcacc 300

tacggcatca ccacctccgg caacagcctg cgcctgaact tcgtgaccgg cagcaatgtg 360

ggcagccgca cctacctgat ggccgacaac acccactacc agatcttcga cctgctgaac 420

caggagttca ccttcaccgt cgacgtgtcc cacctgccct gcggcctgaa cggcgccctc 480

tacttcgtga cgatggacgc cgacggcggc gtgtccaagt accccgaaca a caaggctggc 540

gcccagtacg gtgtgggcta ctgcgacagc cagtgcccga gggacctgaa gttcatcgcc 600

ggccaggcca acgtggaggg ctggaccccg agcagcaaca acgccaacac cggcctgggc 660

aaccacggcg cctgctgcgc cgagctggac atctgggagg ccaacagcat cagcgaggcc 720

ctgaccccac acccatgcga cacccaggc ctgtctgtgt gcaccacgga cgcctgcggc 780

ggcacctact ccagcgaccg ctacgccggc acctgcgacc cagacggctg cgacttcaac 840

ccgtaccgcc tgggcgtgac cgacttctac ggcagcggca agaccgtgga caccaccaag 900

ccgatcaccg tggtgaccca gttcgtgacc gacgacggca ccagcaccgg cacccctgagc 960

gagatccgcc gctactacgt ccagaacggc gtggtgatcc cgcagccgag cagcaagatc 1020

agcggcgtgt ccggcaacgt gatcaacagc gaccttctgcg acgccgagat cagcaccttc 1080

ggcgagaccg ccagcttcag caagcacggc ggcctggcca agatgggcgc tggcatggaa 1140

gccggcatgg tgctggtgat gagcctgtgg gacgactact ccgtgaacat gctgtggctg 1200

gacagcacct acccgaccaa cgccaccggg acgccaggcg ctgccagggg cagctgccca 1260

accacctcgg gcgaccccaa gaccgtcgag agccagagcg gcagcagcta cgtgaccttc 1320

agcgacatcc gcgtgggccc gttcaactcc acgttcagcg gtggctctag cacgggcggc 1380

tctccaccca ccaccgccag cggcaccacc accaccaagg cctccagcac gtctactagc 1440

tccacctcta ccggcaccgg cgttgctgcc cattggggcc agtgcggtgg ccagggctgg 1500

acgggtccaa cgacttgcgc ctccggcacc acctgcaccg tggtcaatcc gtactactcc 1560

cagtgcctgg acgagctgaa ggccgaggcc aagtga 1596

<210> 63
<211> 13321
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2015

<400> 63
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgccccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg ccccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatcttaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgat ttttttcacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaatttttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700
gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760
atttctgaat tttaccccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820
aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880
ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940
aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000
cagttgttga cccagccggc gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060
gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120
gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180
cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240
ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300
atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360
caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480
tgggtttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600
ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgatgtcc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560

tgtcgtaatt ggatgaacta ttatcgtagt tgctcttaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgcccgga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga tasgcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaacoggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgtttcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc cagggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480

tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
ccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggg gggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400

taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg ttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg ccctttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accaccccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gcccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320

gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttcccdagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 11220
tgtttcgata tattacccta atgatgtgtc agatggggat tgtttcgata tattacccta 11280
atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400
ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ccggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta ccgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640
gaactgtagt gctgatgttg tactgtgttt ttttaccccgt tttatttgct ttactcgtgc 11700
aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000
tcttgcctct tccagatcca gataatgcaa acaagcatta ctctgacatc caacgcatcc 12060
ggtacgtttg acggttacta ttacgaactc tggaaggata ctggcaatac aacaatgacg 12120
gtctacactc aaggtcgctt ttcctgccag tggtcgaaca tcaataacgc gttgtttagg 12180
accgggaaga aatacaacca gatttggcag tctcttggca caatccggat cacgtactct 12240

gcgacttaca acccaaacgg gaactcctac ttgtgtatct atggctggtc taccaaccca 12300

ttggtcgagt tctacatcgt tgagtcctgg gggaactgga gaccgcctgg tgccacgtcc 12360

ctgggccaag tgacaatcga tggcgggacc tacgacatct ataggacgac acgcgtcaac 12420

cagccttcca ttgtggggac agccacgttc gatcagtact ggagcgtgcg cacctctaag 12480

cggacttcag gaacagtgac cgtgaccgat cacttccgcg cctgggcgaa ccgggcctg 12540

aacctcggca caatagacca aattacattg tgcgtggagg gttaccaaag ctctggatca 12600

gccaacatca cccagaacac cttctctcag ggctcttctt ccggcagttc gggtggctca 12660

tccggctcca caacgactac tcgcatcgag tgtgagaaca tgtccttgtc cggaccctac 12720

gttagcagga ccaccaatcc ctttaatggt attgcgctgt acgccaacgg agacacagcc 12780

cgcgctaccg ttaacttccc cgcaagtcgc aactacaatt tccgcctgcg gggttgcggc 12840

aacaacaata atcttgcccg tgtggacctg aggatcgacg gacggaccgt cgggaccttt 12900

tattaccagg gcacataccc ctgggaggcc ccaattgaca atgtttatgt cagtgcgggg 12960

agtcatacag tcgaaatcac tgttactgcg gataacggca catgggacgt gtatgccgac 13020

tacctggtga tacagtgacc taggtccccg aatttccccg atcgttcaaa catttggcaa 13080

taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgactatcat ataatttctg 13140

ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg 13200

gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag 13260

cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaattg 13320

g 13321

<210> 64
<211> 13735
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2048

<400> 64
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgccccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg cgcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgactttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg cacccgatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt tttttccccc ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cactttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg actttccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaaa cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcacgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020

actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080

tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140

acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200

acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260

cgactagatg ttgaggccta acattttatt agaggagcagg ctagttgctt agatacatga 4320

tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaatagа tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240

tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160

gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gagggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg cccttttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaatc 9780
tgattatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accaccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080

attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gcccccccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcca ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttcccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggccgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggagggcg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220
tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280
atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400
ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ctgcctagtt 11640
gaactgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700
aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatggcc cacgcccgcg tcctcctcct ggcgctcgcc 12060

gtcctggcca ccgccgccgt cgccgtcgcc tcctcctcct ccttcgccga ctccaacccg 12120

atccgcccgg tgaccgaccg cgccgcctcc accgcttacg actacaagca ggtgttgcgg 12180

gactcgctac tattctatga ggcccagaga tcccggccgc tcccagccga ccagaaggtc 12240

acgtggagga aggatagcgc gctgaatgac cagggtgacc agggacaaga cttgaccggc 12300

ggctactttg acgctgggga cttcgtcaag ttcgggttcc ccatggctta taccgcaacc 12360

gtgctggcat ggggcctcat agattttgag gccggctaca gcagtgccgg ggccttggat 12420

gatggacgga aggctgtcaa atgggccacc gactatttca taaaggcccc cacaagtcaa 12480

aatgagttct atggtcaggt cggccagggt gacgccgatc acgctttctg gggaagacca 12540

gaggatatga cgatggcgcg ccccggcgtac aagatagaca cctcaaggcc tggctctgat 12600

ctggcaggcg agacagcggc tgctcttgcc gctgcttcaa tcgtgttccg gaacgtcgat 12660

ggcacttact caaataacct gttaacacac gctcgccagc tattcgactt cgcgaacaac 12720

taccggggaa agtatagtga ctctattact gacgcaagaa atttctacgc aagcgcagac 12780

tacagagacg agttggtttg ggctgctgcg tggttataca gagcgaccaa cgacaacacc 12840

tacctcaaca ctgctgagtc actgtacgat gagtttgggc tacagaactg gggggggggc 12900

ctgaactggg atagcaaggt gtctggcgtg caggtgttgt tggccaagct taccaataag 12960

caggctaca aggacacggt gcagtcttac gtcaattacc taattaataa ccagcagaag 13020

actcccaagg gcctcctcta catcgacatg tggggcaccc ttcgccacgc tgccaacgcc 13080

gcattcatca tgctcgaagc cgccgagctg ggccttgtccg cctcctctta tagacagttc 13140

gcgcaaacgc aaatcgacta cgccctgggc gatggtggcc gctcctttgt gtgcgggttc 13200

gggagtaatc ctcctacgag accgcaccac agatcctcgt cgtgcccgcc agctcccgct 13260

acttgcgact ggaatacatt caactcacct gacccaaact accacgtcct cctgggggcc 13320

ctagtgggcg gacctgatca gaatgacaac tacgtcgatg accgttcaga ctatgttcac 13380

aacgaagtcg ccactgatta caacgcgggt ttccagtccg cgttagctgc tttggtggcc 13440

cttggttact gacctaggtc cccgaatttc cccgatcgtt caaacatttg gcaataaagt 13500

ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 13560

tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 13620

atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 13680

aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcggga attgg 13735

<210> 65
<211> 13753
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2049

<400> 65
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaacccc tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg ccccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

ccacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttaccogtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct gggcagcaaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgacttctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggcttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcctgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgct gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

acccggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaacgctg acaggctgtc gatcttgaga 4020
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440
tcctttgccc agatgtggaa aagaagttcg ttgccccatc gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640

cggcgatcac cgcttcccte atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cactttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccccttta ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaaccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560

gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620

attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680

agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740

atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800

cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040

gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100

gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160

gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220

aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280

aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340

caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400

taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460

gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580

aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640

cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700

tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760

gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940

ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000

cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060

agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120

gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180

tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240

acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300

gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360

caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420

atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480

agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttcacg cccttttaaa tattcgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agctcaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cgggttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gcctggtttt gtctaggccc aatatccccgt tcatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ctaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctcttcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 11220
tgtttcgata tattaccccta atgatgtgtc agatggggat tgtttcgata tattaccccta 11280
atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640
gaactgtagt gctgatgttg tactgtgttt ttttaccccgt tttatttgct ttactcgtgc 11700
aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttcct acatagatct 11760
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaatttgt 11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000
tcttgcctct tccagatcca gataatggcc cacgcccgcg tcctcctcct ggcgctcgcc 12060
gtcctggcca ccgccgccgt cgccgtcgcc tcctcctcct ccttcgccga ctccaacccg 12120
atccgcccgg tgaccgaccg cgccgcctcc accgcttacg actacaagca ggtgttgcgg 12180
gactcgctac tattctatga ggcccagaga tccggccggc tcccagccga ccagaaggtc 12240
acgtggagga aggatagcgc gctgaatgac cagggtgacc agggacaaga cttgaccggc 12300
ggctactttg acgctgggga cttcgtcaag ttcgggttcc ccatggctta taccgcaacc 12360
gtgctggcat ggggcctcat agattttgag gccggctaca gcagtgccgg ggccttggat 12420
gatggacgga aggctgtcaa atgggccacc gactatttca taaaggccca cacaagtcaa 12480
aatgagttct atggtcaggt cggccagggt gacgccgatc acgctttctg gggaagacca 12540
gaggatatga cgatggcgcg cccggcgtac aagatagaca cctcaaggcc tggctctgat 12600
ctggcaggcg agacagcggc tgctcttgcc gctgcttcaa tcgtgttccg gaacgtcgat 12660
ggcacttact caaataacct gttaacacac gctcgccagc tattcgactt cgcgaacaac 12720
taccggggaa agtatagtga ctctattact gacgcaagaa atttctacgc aagcgcagac 12780
tacagagacg agttggtttg ggctgctgcg tggttataca gagcgaccaa cgacaacacc 12840
tacctcaaca ctgctgagtc actgtacgat gagtttgggc tacagaactg gggggggggc 12900
ctgaactggg atagcaaggt gtctggcgtg caggtgttgt tggccaagct taccaataag 12960
caggcctaca aggacacggt gcagtcttac gtcaattacc taattaataa ccagcagaag 13020
actcccaagg gcctcctcta catcgacatg tggggcaccc ttcgccacgc tgccaacgcc 13080
gcattcatca tgctcgaagc cgccgagctg ggcttgtccg cctcctctta tagacagttc 13140
gcgcaaacgc aaatcgacta cgccctgggc gatggtggcc gctctttgt gtgcgggttc 13200
gggagtaatc ctctacgag accgcaccac agatcctcgt cgtgcccgcc agctcccgct 13260
acttgcgact ggaatacatt caactcaccc gacccaaact accacgtcct ctctgggcgc 13320

ctagtgggcg gacctgatca gaatgacaac tacgtcgatg accgttcaga ctatgttcac 13380

aacgaagtcg ccactgatta caacgcgggt ttccagtccg cgttagctgc tttggtggcc 13440

cctggttaca gcgagaagga cgagctgtga cctaggtccc cgaatttccc cgatcgttca 13500

aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 13560

atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 13620

tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 13680

aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 13740

gatcgggaat tgg 13753

<210> 66
<211> 13024
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2063

<400> 66
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg cacccgatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctcttccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat ccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ctatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaacgtg acaggctgtc gatcttgaga 4020

actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080

tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140

acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200

acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260

cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320

tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380

cgcagaagct cccatctttg ccgccataga cgccgcgccc ccgttttggg gtgtagaaca 4440

tccttttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500

cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560

tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620

tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680

gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740

cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800

tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860

gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920

attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980

tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040

cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100

gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160

catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220

gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280

cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340

gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400

cttctacagc gcggagaatc tcgctcgctc cagggcaagc cgaagtttcc aaaaggtcgt 5460

tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520

tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580

acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640

cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700

tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760

tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820

gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880

atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940

gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000

cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060

agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggctcca 6120

cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180

caatgaaatg gtctctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240

tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300

aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg ccagatttgg 6360

taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420

atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480

tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540

acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600

agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660

acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780

aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840

gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900

ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccccc gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaaccegac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg gggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccccc gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata ataccgcgcc acatagcaga actttaaaag 8820

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttacttttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg ccctttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agctcaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcgg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgccgtgc cctctgcgcg tgggagggag aggcgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gcccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttcccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggcgc tgccattggg ccgtacgcgt ggcggtgacc 10740

gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800

ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860

aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920

gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980

gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040

tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100

agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160

ttgttctgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 11220

tgtttcgata tattacccta atgatgtgtc agatggggat tgtttcgata tattacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaactgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatggcg aacaaacatt tgtccctctc cctcttcctc 12060

gtcctccttg gcctgtcggc cagcttggcc tccgggcaag tcttcccagc tggaaacgca 12120

acggaattgg agaaaagaca aaccacccct aactctgagg gctggcatga cggatactac 12180

tactcttggt ggagcgatgg tggtgcacag gccacctata caaacctcga aggcggcact 12240

tatgagattt catggggtga cggtggcaac cttgtcggcg gaaaggggtg gaacccggga 12300

cttaacgcca gggcaatcca cttcgaaggg gtgtaccagc ccaatggcaa ctcatacctg 12360

gccgtctacg ggtggacgcg caatccgctg gttgagtact atatcgtgga gaatttcgga 12420

acttatgacc ctagctccgg tgccacggac ctcgggacag tcgagtgtga cggaagcatc 12480

tacaggctgg gtaaaactac ccgcgttaat gctccatcga tcgacggcac gcaaacattt 12540

gatcaatact ggtccgtgcg gcaggataag aggacaagcg gcacagttca gacgggttgc 12600

cactttgatg cctgggcaag agcggggctc aatgtgaatg gggaccacta ctatcagatt 12660

gtggcgaccg agggctattt ctccagtggc tatgcgcgta taaccgtcgc tgatgttgga 12720

agcgagaagg acgagctgtg acctaggtcc ccgaatttcc ccgatcgttc aaacatttgg 12780

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 12840

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 12900

tgggtttta tgattagagt cccgcaatta tacatttaat acgcgatagа aaacaaaata 12960

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 13020

ttgg 13024

<210> 67
<211> 14917
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2069

<400> 67
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaacccа tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg cccegccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg cctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260
gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320
tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380
tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440
ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500
ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560
ttggaattgt gttcttagtt ttttttcccc ggatttggat tggaattgtg tggagctggg 1620
ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680
tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740
actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800
agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860
tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920
tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980
tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040
tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100
ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160
actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220
gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280
gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340
gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400
ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460
gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520
acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580
ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640
ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700
gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760
atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820
aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880
ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940
aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000
cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060
gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180
cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240
ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300
atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360
caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480
tgggtttta tgattagagt cccgcaatta tacatttaat acgcgatagg aaacaaaata 3540
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600
ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagcttgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaacg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagccc acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040

cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctcta caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960

ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaaccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg gggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccct gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata ataccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttttcacg ccctttttaaa tatccgatta tttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agctctaactc gagttactgg tacgtaccaa atccacggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgccccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gcccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttcccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800

ttaatcgcat agcgagtcga atcgaccgaa ggggagggag agcgaagctt tgcgttctct 10860

aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920

gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980

gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040

tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100

agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160

ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220

tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaactgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc tttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatgctg gaggacaagt ctcccaaact gcctgattat 12060

aagaacgacc ttctgtacga acgcacattc gacgaggggc tctgcttccc gtggcacacg 12120

tgcgaagatt caggagggaa atgcgatttt gccgtggtcg acgttccagg cgagcctggg 12180

aacaaggcgt tcaggctcac tgttatcgat aagggtcaga acaagtggtc ggtccaaatg 12240

agacaccggg gtatcacgtt ggagcagggg cacacataca ccgttcggtt tactatctgg 12300

agcgacaaga gctgccgcgt gtatgccaaa atcggccaaa tgggtgaacc ctacacggag 12360

tactggaaca ataactggaa tccgttcaac ctcactccgg ggcagaaatt gacggtggaa 12420

cagaacttta ctatgaatta tcccacggac gacacgtgtg agtttacctt ccacttggga 12480

ggggaactgg cagccgggac ccttactac gtgtacctcg acgacgtttc tctttacgat 12540

cccccgtttg tcaagccagt ggaatacgtc ctgcctcaac cggatgtcag ggttaatcaa 12600

gttggatacc tccctttgc taagaaatat gctactgtcg tgtcatcgag cacgtcccca 12660

ttgaagtggc aacttctgaa tagtgcaaac caagttgtct tggagggcaa tacaatcccc 12720

aagggactgg acaaagattc acaagactac gttcattgga tcgatttctc gaactttaag 12780
accgaaggca aggggtacta tttcaagttg cccactgtga actccgatac taactactcc 12840
cacccgtttg atatttctgc agatatctat tcaaagatga agttcgacgc gctcgctttc 12900
ttttaccata aaaggtcggg aataccaatc gagatgccct acgccggggg agagcagtgg 12960
acaaggcccg cagggcacat tggtgtcgcg ccgaacaagg gcgacacgaa tgtgccaact 13020
tggccccagg atgacgaata tgctggacgc ccccagaaat actatacgaa agacgtgacc 13080
ggcgggtggt acgatgccgg tgaccacggc aagtacgtcg tgaacggggg tatcgcagtt 13140
tggaccctta tgaatatgta cgagagagca aagattagag gaatcgctaa ccagggtgcc 13200
tacaaagatg gaggaatgaa tatcccggaa aggaataacg gctatcctga tattctggac 13260
gaggccagat gggagatcga attttttaag aagatgcaag tcactgagaa agaagatccg 13320
tcgattgcag gtatggtgca ccacaagatc cacgatttca ggtggacggc gctcggaatg 13380
ttgctcacg aggacccccc gccacgctac cttcggcccg tcagcacagc ggcaaccctg 13440
aatttcgcag cgaccctcgc tcagtctgcc agattgtgga aggattacga cccgactttt 13500
gcagcggact gccttgagaa agctgaaatt gcctggcaag cagcactcaa acacccggac 13560
atctacgctg agtacacgcc aggaagcggt gggccgggtg gaggtcctta taatgacgat 13620
tatgtcgggg acgagttcta ctgggccgct tgtgaactct atgtgacaac cggtaaggat 13680
gagtacaaga attacttgat gaatagtccg cactatctgg aaatgccagc gaagatgggc 13740
gagaacggag gggctaacgg cgaggacaac ggtctctggg gctgctttac ttggggaacg 13800
acacaggggt tgggtacaat tacccttgcc ctcgttgaaa acggcctccc ttcggcggat 13860
attcaaaagg ccgcaacaa tatcgctaaa gccgcagata agtggcttga gaatattgaa 13920
gaacaaggtt acccgcctgcc tatcaaacaa gcggaggatg aacggggcgg atacccgtgg 13980
ggtagtaatt ctttcattct caaccagatg atcgtcatgg gctacgctta cgacttcacg 14040
ggaaacagca agtatcttga cgggatgcag gacggcatgt cctacctgct cggtagaaac 14100
ggacttgatc aatcgtacgt tactgggtac ggggagaggc cacttcagaa ccccacgac 14160
cgcttttgga cccctcaaac ttcgaagaaa ttcccggccc cacccccgg tattatcgca 14220
ggcgggccga atagccggtt tgaagatcca acgatcactg cagcggttaa gaaggataca 14280
ccccgcaga agtgctatat tgaccacacc gattcctggt ctactaacga gatcacgatt 14340
aattggaacg cccccttcgc gtgggtcaca gcgtatctgg acgaaattga cttgattacc 14400
ccacccggcg gagtggaccc tgaagagccg gaagttatct acggtgattg taacggcgac 14460
ggaaaggtta attcgaccga tgctgtggcc cttaaaaggt atatcctccg cagcggtatc 14520
tcgatcaaca cggacaacgc ggacgttaat gcagatggtc gcgtgaatag cactgacctc 14580
gctattttga agcgctatat tttgaaggag atcgatgttc ttcctcacaa gtgacctagg 14640

tcccgaatt tcccgatcg ttcaaacatt tggcaataaa gtttcttaag attgaatcct 14700

gttgccggtc ttgcgatgat tatcatataa tttctgttga attacgttaa gcatgtaata 14760

attaacatgt aatgcatgac gttatttatg agatgggttt ttatgattag agtcccgcaa 14820

ttatacattt aatacgcgat agaaaacaaa atatagcgcg caaactagga taaattatcg 14880

cgcgcggtgt catctatgtt actagatcgg gaattgg 14917

<210> 68
<211> 16551
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2091

<400> 68
catggaatca aggtaccgtc gactctagta acggccgcca gtgtgctgga attaattcgg 60

cttgtcgacc acccaacccc atatcgacag aggatgtgaa gaacaggtaa atcacgcaga 120

agaacccatc tctgatagca gctatcgatt agaacaacga atccatattg ggtccgtggg 180

aaatacttac tgcacaggaa gggggcgatc tgacgaggcc ccgccacogg cctcgacccg 240

aggccgaggc cgacgaagcg ccggcgagta cggcgccgcg gcggcctctg cccgtgccct 300

ctgcgcgtgg gagggagagg ccgcggtggt gggggcgcgc gcgcgcgcgc gcgcagctgg 360

tgcggcggcg cgggggtcag ccgccgagcc ggcggcgacg gaggagcagg gcggcgtgga 420

cgcgaacctc cgatcggttg gtcagagtgc gcgagttggg cttagccaat taggtctcaa 480

caatctattg ggccgtaaaa ttcatgggcc ctggtttgtc taggcccaat atcccgttca 540

tttcagccca caaatatttc cccagaggat tattaaggcc cacacgcagc ttatagcaga 600

tcaagtacga tgtttcctga tcgttggatc ggaaacgtac ggtcttgatc aggcatgccg 660

acttcgtcaa agagaggcgg catgacctga cgcggagttg gttccgggca ccgtctggat 720

ggtccgtaccg ggaccggaca cgtgtcgcgc ctccaactac atggacacgt gtggtgctgc 780

cattgggccg tacgcgtggc ggtgaccgca ccggatgctg cctcgcaccg ccttgcccac 840

gctttatata gagaggtttt ctctccatta atcgcatagc gagtcgaatc gaccgaaggg 900

gaggggagc gaagctttgc gttctctaat cgcctcgtca aggtaactaa tcaatcacct 960

cgtcctaatc ctcgaatctc tcgtggtgcc cgtctaatct cgcgattttg atgctcgtgg 1020

tggaaagcgt aggaggatcc cgtgcgagtt agtctcaatc tctcagggtt tcgtgcgatt 1080

ttagggtgat ccacctctta atcgagttac ggtttcgtgc gattttaggg taatcctctt 1140

aatctctcat tgatttaggg tttcgtgaga atcgaggtag ggatctgtgt tatttatatc 1200

gatctaatag atggattggt tttgagattg ttctgtcaga tggggattgt ttcgatatat 1260

taccctaatg atgtgtcaga tggggattgt ttcgatatat taccctaatg atgtgtcaga 1320

tggggattgt ttcgatatat taccctaatg atggataata agagtagttc acagttatgt 1380
tttgatcctg ccacatagtt tgagttttgt gatcagattt agttttactt atttgtgctt 1440
agttcggatg ggattgttct gatattgttc caatagatga atagcttcgtt aggttaaaat 1500
ctttaggttg agttaggcga cacatagttt atttcctctg gatttggatt ggaattgtgt 1560
tcttagtttt tttccccttgg atttggattg gaattgtgtg gagctgggtt agagaattac 1620
atctgtatcg tgtacaccta cttgaactgt agagcttggg ttctaaggtc aatttaatct 1680
gtattgtatc tggctctttg cctagttgaa ctgtagtgct gatgttgtac tgtgtttttt 1740
tacccgtttt atttgcttta ctcgtgcaaa tcaaatctgt cagatgctag aactaggtgg 1800
ctttatttctg tgttcttaca tagatctgtt gtcctgtagt tacttatgtc agttttgtta 1860
ttatctgaag atatttttgg ttgttgcttg ttgatgtggt gtgagctgtg agcagcgctc 1920
ttatgattaa tgatgctgtc caattgtagt gtagtatgat gtgattgata tgttcatcta 1980
ttttgagctg acagtaccga tatcgtagga tctggtgcca acttattctc cagctgcttt 2040
ttttacctta tgttaattcc aatcctttct tgcctcttcc agatccagat aacggcgaac 2100
aaacatttgt cccttctcct cttcctcgtc ctccttggcc tgtcggccag cttggcctcc 2160
gggcaacaaa caagcattac tctgacatcc aacgcatccg gtacgtttga cggttactat 2220
tacgaactct ggaaggatac tggcaatata acaatgacgg tctacactca aggtcgcttt 2280
tcctgccagt ggtcgaacat caataacgcg ttgtttagga ccgggaagaa atacaaccag 2340
aattggcagt ctcttggcac aatccggatc acgtactctg cgacttacaa cccaaacggg 2400
aactcctact tgtgtatcta tggctggtct accaacccat tggtcgagtt ctacatcgtt 2460
gagtcctggg ggaactggag accgcctggt gccacgtccc tgggccaagt gacaatcgat 2520
ggcgggacct acgacatcta taggacgaca cgcgtcaacc agccttccat tgtggggaca 2580
gccacgttcg atcagtactg gagcgtgcgc acctctaagc ggacttcagg aacagtgacc 2640
gtgaccgatc acttccgcgc ctgggcgaac cggggcctga acctcggcac aatagaccaa 2700
attacattgt gcgtggaggg ttaccaaagc tctggatcag ccaacatcac ccagaacacc 2760
ttctctcagg gctcttcttc cggcagttcg ggtggctcat ccggctccac aacgactact 2820
cgcatcgagt gtgagaacat gtccttgtcc ggaccctacg ttagcaggat caccaatccc 2880
tttaatggta ttgcgctgta cgccaacgga gacacagccc gcgctaccgt taacttcccc 2940
gcaagtcgca actacaattt ccgcctgcgg ggttgcggca acaacaataa tcttgcccgt 3000
gtggacctga ggatcgacgg acggaccgtc gggaccttct attaccaggg cacataccgc 3060
tgggaggccc caattgacaa tgtttatgtc agtgcgggga gtcatacagt cgaaatcact 3120
gttactgcgg ataacggcac atgggacgtg tatgccgact acctggtgat acagtgacct 3180
aggtccccga atttccccga tcgttcaaac atttggcaat aaagtttctt aagattgaat 3240

cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta 3300
ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg 3360
caattatata tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 3420
tcgcgcgcgg tgtcatctat gttactagat cgggaattgg aattcatact aaagcttgca 3480
tgcctgcagg tcgactctag taacggcgcc cagtgtgctg gaattaattc ggcttgtcga 3540
ccacccaacc ccatatcgac agaggatgtg aagaacaggt aaatcacgca gaagaaccca 3600
tctctgatag cagctatcga ttagaacaac gaatccatat tgggtccgtg ggaaatactt 3660
actgcacagg aaggggcga tctgacgagg ccccgccacc ggcctcgacc cgaggccgag 3720
gccgacgaag cgccggcgag tacggcgccg cggcggcctc tgcccgtgcc ctctgcgcgt 3780
gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc gcgcgcagct ggtgcggcgg 3840
cgcgggggtc agccgccgag ccggcggcga cggaggagca gggcggcgtg gacgcgaact 3900
tccgatcggt tggtcagagt gcgcgagttg ggcttagcca attaggtctc aacaatctat 3960
tgggccgtaa aattcatggg ccctggtttg tctaggccca atatcccgtt catttcagcc 4020
cacaaatatt tccccagagg attattaagg cccacacgca gcttatagca gatcaagtac 4080
gatgtttcct gatcgttgga tcggaaacgt acggtcttga tcaggcatgc cgacttcgtc 4140
aaagagaggc ggcatgacct gacgcggagt tggttccggg caccgtctgg atggtcgtac 4200
cgggaccgga cacgtgtcgc gcctccaact acatggacac gtgtggtgct gccattgggc 4260
cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac cgccttgccc acgctttata 4320
tagagaggtt ttctctccat taatcgcata gcgagtcgaa tcgaccgaag gggaggggga 4380
gcgaagcttt gcgttctcta atcgcctcgt caaggtaact aatcaatcac ctcgtcctaa 4440
tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt tgatgctcgt ggtggaaagc 4500
gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg tttcgtgcga ttttagggtg 4560
atccacctct taatcgagtt acggtttcgt gcgattttag ggtaatcctc ttaatctctc 4620
attgatttag ggtttcgtga gaatcgaggt agggatctgt gttatttata tcgatctaat 4680
agatggattg gttttgagat tgttctgtca gatggggatt gtttcgatat attaccctaa 4740
tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatgtgtca gatggggatt 4800
gtttcgatat attaccctaa tgatggataa taagagtagt tcacagttat gttttgatcc 4860
tgccacatag tttgagtttt gtgatcagat ttagttttac ttatttgtgc ttagttcgga 4920
tgggattgtt ctgatattgt tccaatagat gaatagctcg ttaggttaaa atctttaggt 4980
tgagttaggc gacacatagt ttatttcctc tggatttgga ttggaattgt gttcttagtt 5040
tttttccact ggatttggat tggaattgtg tggagctggg ttagagaatt acatctgtat 5100
cgtgtacacc tacttgaact gtagagcttg ggttctaagg tcaatttaat ctgtattgta 5160

tctggctctt tgcctagttg aactgtagtg ctgatgttgt actgtgtttt tttacccgtt 5220

ttatttgctt tactcgtgca aatcaaatct gtcagatgct agaactaggt ggctttattc 5280

tgtgttcttt catagatctg ttgtcctgta gttacttatg tcagttttgt tattatctga 5340

agatattttt ggttgttgct tgttgatgtg gtgtgagctg tgagcagcgc tcttatgatt 5400

aatgatgctg tccaattgta gtgtagtatg atgtgattga tatgttcatc tattttgagc 5460

tgacagtacc gatatcgtag gatctggtgc caacttattc tccagctgct ttttttacc 5520

tatgttaatt ccaatccttt cttgcctctt ccagatccag ataatgcaga aactcattaa 5580

ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg actgaacttt atggtatgga 5640

aaatccgtcc agccagccga tggccgagct gtggatgggc gcacatccga aaagcagttc 5700

acgagtgcag aatgccgccg gagatatcgt ttcactgcgt gatgtgattg agagtgataa 5760

atcgactctg ctcggagagg ccgttgccaa acgctttggc gaactgcctt tcctgttcaa 5820

agtattatgc gcagcacagc cactctccat tcaggttcat ccaaacaaac acaattctga 5880

aatcggtttt gccaaagaaa atgccgcagg tatcccgatg gatgccgccg agcgtaacta 5940

taaagatcct aaccacaagc cggagctggt ttttgcgctg acgcctttcc ttgcgatgaa 6000

cgcgtttcgt gaatttttcg agattgtctc cctactccag ccggtcgcag gtgcacatcc 6060

ggcgattgct cacttttttac aacagcctga tgccgaacgt ttaagcgaac tgttcgccag 6120

cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg gcgattttaa aatcggccct 6180

cgatagccag cagggtgaac cgtggcaaac gattcgttta atttctgaat tttacccgga 6240

agacagcggt ctgttctccc cgctattgct gaatgtggtg aaattgaacc ctggcgaagc 6300

gatgttcctg ttcgctgaaa cacgcacgc ttacctgcaa ggcgtggcgc tggaagtgat 6360

ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct aaatacattg atattccgga 6420

actggttgcc aatgtgaaat tcgaagccaa accggctaac cagttgttga cccagccggt 6480

gaaacaaggt gcagaactgg acttcccgat tccagtggat gattttgcct tctcgctgca 6540

tgaccttagt gataaagaaa ccaccattag ccagcagagt gccgccattt tgttctgcgt 6600

cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta cagcttaaac cgggtgaatc 6660

agcgtttatt gccgccaacg aatcaccggt gactgtcaaa ggccacggcc gtttagcgcg 6720

tgtttacaac aagctgtaag agcttactga aaaaattaac atctcttgct aagctgggag 6780

ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg caataaagtt tcttaagatt 6840

gaatcctgtt gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca 6900

tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggtttta tgattagagt 6960

cccgcaatta tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa 7020

attatcgcgc gcggtgtcat ctatgttact agatcgggaa ttggcgagct cgaattaatt 7080

cagtacatta aaaacgtccg caatgtgtta ttaagttgtc taagcgtcaa tttgtttaca 7140
ccacaatata tcctgccacc agccagccaa cagctccccg accggcagct cggcacaaaa 7200
tcaccactcg atacaggcag cccatcagtc cgggacggcg tcagcgggag agccgttgta 7260
aggcggcaga ctttgctcat gttaccgatg ctattcggaa gaacggcaac taagctgccg 7320
ggtttgaaac acggatgatc tcgcggaggg tagcatgttg attgtaacga tgacagagcg 7380
ttgctgcctg tgatcaaata tcatctccct cgcagagatc cgaattatca gccttcttat 7440
tcatttctcg cttaaccgtg acaggctgtc gatcttgaga actatgccga cataatagga 7500
aatcgctgga taaagccgct gaggaagctg agtggcgcta tttctttaga agtgaacgtt 7560
gacgatcgtc gaccgtaccc cgatgaatta attcggacgt acgttctgaa cacagctgga 7620
cacttacttg ggcgattgtc atacatgaca tcaacaatgt acccgtttgt gtaaccgtct 7680
cttggaggtt cgtatgacac tagtggttcc cctcagcttg cgactagatg ttgaggccta 7740
acattttatt agagagcagg ctagttgctt agatacatga tcttcaggcc gttatctgtc 7800
agggcaagcg aaaattggcc atttatgacg accaatgccc cgcagaagct cccatctttg 7860
ccgccataga cgccgcgccc ccctttttggg gtgtagaaca tcctttgcc agatgtggaa 7920
aagaagttcg ttgtcccatt gttggcaatg acgtagtagc cggcgaaagt gcgagaccca 7980
tttgcgctat atataagcct acgatttccg ttgcgactat tgtcgtaatt ggatgaacta 8040
ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac tattgtcgta attgcttatg 8100
gagttgtcgt agttgcttgg agaaatgtcg tagttggatg gggagtagtc ataggaaaga 8160
cgagcttcat ccactaaaac aattggcagg tcagcaagtg cctgccccga tgccatcgca 8220
agtacgaggc ttagaaccac cttcaacaga tcgcgcatag tcttccccag ctctctaacg 8280
cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac gaattgttag acattatttg 8340
ccgactacct tggtgatctc gcctttcacg tagtgaacaa attcttccaa ctgatctgcg 8400
cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc tagcttcaag tatgacgggc 8460
tgatactggg ccggcaggcg ctccattgcc cagtcggcag cgacatcctt cggcgcgatt 8520
ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa gcactacatt tcgctcatcg 8580
ccagcccagt cgggcggcga gttccatagc gttaaggttt catttagcgc ctcaaatagа 8640
tcctgttcag gaaccggatc aaagagttcc tccgccgctg gacctaccaa ggcaacgcta 8700
tgttctcttg cttttgtcag caagatagcc agatcaatgt cgatcgtggc tggctcgaag 8760
atacctgcaa gaatgtcatt gcgctgccat tctccaaatt gcagttcgcg cttagctgga 8820
taacgccacg gaatgatgtc gtcgtgcaca acaatggtga cttctacagc gcggagaatc 8880
tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt tgatcaaagc tcgccgcgtt 8940
gtttcatcaa gccttacggt caccgtaacc agcaaatcaa tatcactgtg tggcttcagg 9000

ccgccatcca ctgcggagcc gtacaaatgt acggccagca acgtcggttc gagatggcgc 9060
tcgatgacgc caactacctc tgatagttga gtcgatactt cggcgatcac cgcttccctc 9120
atgatgttta actcctgaat taagccgcgc cgcgaagcgg tgtcggcttg aatgaattgt 9180
taggcgtcat cctgtgctcc cgagaaccag taccagtaca tcgctgtttc gttcgagact 9240
tgaggtctag ttttatacgt gaacaggtca atgccgccga gagtaaagcc acattttgcg 9300
tacaaattgc aggcaggtac attgttcgtt tgtgtctcta atcgtatgcc aaggagctgt 9360
ctgcttagtg cccacttttt cgcaaattcg atgagactgt gcgcgactcc tttgcctcgg 9420
tgcgtgtgcg acacaacaat gtgttcgata gaggctagat cgttccatgt tgagttgagt 9480
tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat agcaagcaga gtcttcatca 9540
gagtcatcat ccgagatgta atccttccgg taggggctca cacttctggt agatagttca 9600
aagccttggt cggataggtg cacatcgaac acttcacgaa caatgaaatg gttctcagca 9660
tccaatgttt ccgccacctg ctcagggatc accgaaatct tcatatgacg cctaacgcct 9720
ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa aaggcgtgac aggtttgcga 9780
atccgttgct gccacttgtt aaccctttg ccagatttgg taactataat ttatgttaga 9840
ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg atttcaggaa agtaaacatc 9900
accttccggc tcgatgtcta ttgtagatat atgtagtgta tctacttgat cgggggatct 9960
gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag ctcccggaga 10020
cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag 10080
cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc acgtagcgat agcggagtgt 10140
atactggctt aactatgcgg catcagagca gattgtactg agagtgcacc atatgcggtg 10200
tgaaataccg cacagatgcg taaggagaaa ataccgcatc aggcgctctt ccgcttcctc 10260
gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa 10320
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 10380
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 10440
ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 10500
aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc 10560
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc 10620
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 10680
tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа tccggtaact atcgtcttga 10740
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 10800
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta 10860
cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 10920

agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg 10980
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 11040
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 11100
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 11160
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 11220
agcgatctgt ctatttcgtt catccatagt tgcctgactc ccgtcgtgt agataactac 11280
gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc 11340
accggctcca gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg 11400
tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag 11460
tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctgcagggg gggggggggg 11520
ggggttccat tgttcattcc acggacaaaa acagagaaag gaaacgacag aggccaaaaa 11580
gctcgctttc agcacctgtc gtttcctttc ttttcagagg gtattttaaa taaaaacatt 11640
aagttatgac gaagaagaac ggaaacgcct taaaccggaa aattttcata aatagcgaaa 11700
acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg aaaggacccg taaagtgata 11760
atgattatca tctacatatc acaacgtgcg tggaggccat caaaccacgt caaataatca 11820
attatgacgc aggtatcgta ttaattgatc tgcatcaact taacgtaaaa acaacttcag 11880
acaatacaaa tcagcgacac tgaatacggg gcaacctcat gtcccccccc cccccccctt 11940
gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 12000
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 12060
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 12120
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 12180
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 12240
acacgggata taccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 12300
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 12360
actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 12420
aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 12480
ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 12540
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 12600
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 12660
aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg acgatcttgc tgcgttcgga 12720
tattttcgtg gagttcccgc cacagacccg gattgaaggc gagatccagc aactcgcgcc 12780
agatcatcct gtgacggaac tttggcgcgt gatgactggc caggacgtcg gccgaaagag 12840

cgacaagcag atcacgcttt tcgacagcgt cggatttgcg atcgaggatt tttcggcgct 12900
gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc agcccactcg accttctagc 12960
cgacccagac gagccaaggg atcttttttgg aatgctgctc cgtcgtcagg ctttccgacg 13020
tttgggtggt tgaacagaag tcattatcgc acggaatgcc aagcactccc gaggggaacc 13080
ctgtggttgg catgcacata caaatggacg aacggataaa ccttttcacg ccctttttaaa 13140
tatccgatta ttctaataaa cgctcttttc tcttaggttt acccgccaat atatcctgtc 13200
aaacactgat agtttaaact gaaggcggga aacgacaacc tgatcatgag cggagaatta 13260
agggagtcac gttatgaccc ccgccgatga cgcgggacaa gccgttttac gtttggaact 13320
gacagaaccg caacgttgaa ggagccactc agcttaatta agctctaactc gagcacccaa 13380
cccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catctctgat 13440
agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca 13500
ggaaggggc gatctgacga ggccccgcca ccggcctcga cccgaggccg aggccgacga 13560
agcgccggcg agtacggcgc cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga 13620
gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg 13680
tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 13740
gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 13800
aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 13860
tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 13920
ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg tcaaagagag 13980
gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 14040
gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 14100
tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 14160
ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 14220
ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 14280
tctctcgtgg tgcccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 14340
atcccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg tgatccacct 14400
cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 14460
agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta atagatggat 14520
tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 14580
cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 14640
atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 14700
agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 14760

ttctgatatt gctccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 14820

gcgacacata gtttatctcc tctggatttg gattggaatt gtgttcttag ttttttcccc 14880

ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt atcgtgtaca 14940

cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 15000

tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt ttttttaccg ttttatttgc 15060

tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttct 15120

tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 15180

ctggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 15240

tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga gctgacagta 15300

ccgatatcgt aggatctggt gccaacttat tctccagctg ctttttttta cctatgttaa 15360

ttccaatcct ttcttgcctc ttccagatcc agataatggc gaacaaacat ttgtcctct 15420

ccctcttcct cgtcctcctt ggcctgtcgg ccagcttggc ctccgggcaa gtcgcatcca 15480

cgcagggtat ctctgaggac ctgtataatc gcctcgtgga aatggccaca atttcacaag 15540

cggcttacgc agatctttgt aatatccctt cgacaattat caagggagag aaaatctata 15600

acgcccagac tgacatcaac ggctggatac tgcgggatga cacgagcaag gaaattatca 15660

cagtctttag agggaccggt tccgatacaa atttgcagtt ggacacgaat tacacactga 15720

cccccttcga tactctcctt caatgcaacg actgtgaggt tcacggtggg tactatattg 15780

gctggatctc tgttcaagac caagtcgagt cacttgttaa gcagcaagcg tcgcagtacc 15840

cggactacgc attgacggtg acagggcaca gcctgggtgc ctcgatggca gcgctcaccg 15900

ctgcccagct ttctgcaacc tacgataatg tcaggctgta cactttcgga gaaccacgct 15960

caggcaacca agcgtttgct tcgtatatga acgacgcttt ccaggttagc tcccccgaga 16020

cgacacaata ctttcgggtg acccattcta acgacgggat tcctaacctc ccgccagccg 16080

acgaaggtta cgcacacggg ggtgtcgagt actggtcagt ggaccccctac agcgcgcaga 16140

atactttcgt ttgcacgggc gatgaggtcc agtgctgtga agctcaaggg ggtcagggag 16200

tgaatgatgc acacacaacc tatttcggaa tgacttccgg ggcttgcacg tggagcgaga 16260

aggacgagct gtgatccccg aatttccccg atcgttcaaa catttggcaa taaagtttct 16320

taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg ttgaattacg 16380

ttaagcatgt aataattaac atgtaatgca tgacgttatt catgagatgg gttttatga 16440

ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact 16500

aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaattg c 16551

<210> 69
<211> 17296

<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2092

<400> 69
cgtcgacttct agtaacggcc gccagtgtgc tggaattaat tcggcttgtc gaccacccaa 60

ccccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catctctgat 120

agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca 180

ggaaqggggc gatctgacga ggccccgcca ccggcctcga ccgaggccg aggccgacga 240

agcgccggcg agtacggcgc cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga 300

gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg 360

tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 420

gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 480

aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 540

tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 600

ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg tcaaagagag 660

gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 720

gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 780

tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 840

ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 900

ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 960

tctctcgtgg tgccccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 1020

atccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg tgatccacct 1080

cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 1140

agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta atagatggat 1200

tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 1260

cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 1320

atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 1380

agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 1440

ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 1500

gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag ttttttcccc 1560

ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt atcgtgtaca 1620

cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 1680

tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt cttttacccg ctttatttgc 1740

tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttct 1800

tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 1860

ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 1920

tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga gctgacagta 1980

ccgatatcgt aggatctggt gccaacttat tctccagctg ctttttttta cctatgttaa 2040

ttccaatcct ttcttgcctc ttccagatcc agataatggc gaacaaacat ttgtcccttt 2100

ccctcttcct cgtcctcctt ggcctgtcgg ccagcttggc ctccgggcaa caaacaagca 2160

ttactctgac atccaacgca tccggtacgt ttgacggtta ctattacgaa ctctggaagg 2220

atactggcaa tacaacaatg acggtctaca ctcaaggtcg cttttcctgc cagtggtcga 2280

acatcaataa cgcgttgttt aggaccggga agaaatacaa ccagaattgg cagtctcttg 2340

gcacaatccg gatcacgtac tctgcgactt acaacccaaa cgggaactcc tacttgtgta 2400

tctatggctg gtctaccaac ccattggtcg agttctacat cgttgagtcc tgggggaact 2460

ggagaccgcc tggtgccacg tccctgggcc aagtgacaat cgatggcggg acctacgaca 2520

tctataggac gacacgcgtc aaccagcctt ccattgtggg gacagccacg ttcgatcagt 2580

actggagcgt gcgcacctct aagcggactt caggaacagt gaccgtgacc gatcacttcc 2640

gcgcctgggc gaaccggggc ctgaacctcg gcacaataga ccaaattaca ttgtgcgtgg 2700

agggttacca aagctctgga tcagccaaca tcacccagaa caccttctct cagggctctt 2760

cttccggcag ttcgggtggc tcatccggct ccacaacgac tactcgcatc gagtgtgaga 2820

acatgtcctt gtccggaccc tacgttagca ggatcaccaa tccctttaat ggtattgcgc 2880

tgtacgccaa cggagacaca gcccgcgcta ccgttaactt ccccgcaagt cgcaactaca 2940

atttccgcct gcggggttgc ggcaacaaca ataatcttgc ccgtgtggac ctgaggatcg 3000

acggacggac cgtcgggacc ttttattacc agggcacata ccctgggag gcccaattg 3060

acaatgttta tgtcagtgcg gggagtcata cagtcgaaat cactgttact gcggacaacg 3120

gcacatggga cgtgtatgcc gactacctgg tgatacagtg acctaggtcc ccgaatttcc 3180

ccgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg 3240

cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat 3300

gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 3360

acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat 3420

ctatgttact agatcgggaa ttggaattca tactaaagct tgcatgcctg caggtcgact 3480

ctagtaacgg ccgccagtgt gctggaatta attcggcttg tcgaccaccc aaccccatat 3540

cgacagagga tgtgaagaac aggtaaatca cgcagaagaa cccatctctg atagcagcta 3600

tcgattagaa caacgaatcc atattgggtc cgtgggaaat acttactgca caggaagggg 3660

gcgatctgac gaggccccgc cacggcctc gacccgaggc cgaggccgac gaagcgccgg 3720
cgagtacggc gccgcggcgg cctctgcccg tgccctctgc gcgtgggagg gagaggccgc 3780
ggtggtgggg gcgcgcgcgc gcgcgcgcgc agctggtgcg gcggcgcggg ggtcagccgc 3840
cgagccggcg gcgacggagg agcagggcgg cgtggacgcg aacttccgat cggttggtca 3900
gagtgcgcga gttgggctta gccaattagg tctcaacaat ctattgggcc gtaaaattca 3960
tgggccctgg tttgtctagg cccaatatcc cgttcatttc agcccacaaa tatttcccca 4020
gaggattatt aaggcccaca cgcagcttat agcagatcaa gtacgatgtt tcctgatcgt 4080
tggatcggaa acgtacggtc ttgatcaggc atgccgactt cgtcaaagag aggcggcatg 4140
acctgacgcg gagttggttc cgggcaccgt ctggatggtc gtaccgggac cggacacgtg 4200
tcgcgcctcc aactacatgg acacgtgtgg tgctgccatt cggccgtacg cgtggcggtg 4260
accgcaccgg atgctgcctc gcaccgcctt gcccacgctt tatatagaga ggttttctct 4320
ccattaatcg catagcgagt cgaatcgacc gaaggggagg gggagcgaag ctttgcgttc 4380
tctaatcgcc tcgtcaaggt aactaatcaa tcacctcgtc ctaatcctcg aatctctcgt 4440
ggtgcccgtc taatctcgcg attttgatgc tcgtggtgga aagcgtagga ggatcccgtg 4500
cgagttagtc tcaatctctc agcgtttcgt gcgatttttag ggtgatccac ctcttaatcg 4560
agttacggtt tcgtgcgatt ttagggtaat cctcttaatc tctcattgat ttagggtttc 4620
gtgagaatcg aggtagggat ctgtgttatt tatatcgatc taatagatgg attggttttg 4680
agattgttct gtcagatggg gattgtttcg atatattacc ctaatgatgt gtcagatggg 4740
gattgtttcg atatattacc ctaatgatgt gtcagatggg gattgtttcg atatattacc 4800
ctaatgatgg ataataagag tagttcacag ttatgttttg atcctgccac atagtttgag 4860
ttttgtgatc agatttagtt ttacttattt gtgcttagtt cggatgggat tgttctgata 4920
ttgttccaat agatgaatag ctcgttaggt taaaatcttt aggttgagtt aggcgacaca 4980
tagtttattt cctctggatt tggattggaa ttgtgttctt agtttttttc ccctggattt 5040
ggattggaat tgtgtggagc tgggttagag aattacatct gtatcgtgta cacttacttg 5100
aactgtagag cttgggttct aaggtcaatt taatctgtat tgtatctggc tctttgccta 5160
gttgaactgt agtgctgatg ttgtactgtg ttttttacc cgttttattt gctttactcg 5220
tgcaaatcaa atctgtcaga tgctagaact aggtggcttt attctgtgtt cttacataga 5280
tctgttgtcc tgtagttact tatgtcagtt ttgttattat ctgaagatat ttttggttgt 5340
tgcttgttga tgtggtgtga gctgtgagca gcgctcttat gattaatgat gctgtccaat 5400
tgtagtgtag tatgatgtga ttgatatgtt catctatttt gagctgacag tacccgatatc 5460
gtaggatctg gtgccaactt attctccagc tgcttttttt tacctatgtt aattccaatc 5520
ctttcttgcc tcttccagat ccagataatg cagaaactca ttaactcagt gcaaaactat 5580

gcctgggyca gcaaaacggc gttgactgaa ctttatggta tggaaaatcc gtccagccag 5640
ccgatggccg agctgtggat gggcgcacat ccgaaaagca gttcacgagt gcagaatgcc 5700
gccggagata tcgtttcact gcgtgatgtg attgagagtg ataaatcgac tctgctcgga 5760
gaggccgttg ccaaacgctt tggcgaactg cctttcctgt tcaaagtatt atgcgcagca 5820
cagccactct ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg ttttgccaaa 5880
gaaaatgccg caggtatccc gatggatgcc gccgagcgta actataaaga tcctaaccac 5940
aagccggagc tggttttgc gctgacgcct ttccttgcga tgaacgcgct tcgtgaattt 6000
tccgagattg tctcctact ccagccggtc gcaggtgcac atccggcgat tgctcacttt 6060
ttacaacagc ctgatgccga acgtttaagc gaactgttcg ccagcctgtt gaatatgcag 6120
ggtgaagaaa aatccccgcc gctggcgatt ttaaaatcgg ccctcgatag ccagcagggt 6180
gaaccgtggc aaacgattcg tttaatttct gaattttacc cggaagacag cggtctgttc 6240
tcccccgctat tgctgaatgt ggtgaaattg aaccctggcg aagcgatgtt cctgttcgct 6300
gaaacaccgc acgcttacct gcaaggcgtg gcgctggaag tgatggcaaa ctccgataac 6360
gtgctgcgtg cgggtctgac gcctaaatac attgatattc cggaactggt tgccaatgtg 6420
aaattcgaag ccaaaccggc taaccagttg ttgacccagc cggtgaaaca aggtgcagaa 6480
ctggacttcc cgattccagt ggatgatttt gccttctcgc tgcatgacct tagtgataaa 6540
gaaaccacca ttagccagca gagtgccgcc attttgttct gcgtcgaagg cgatgcaacg 6600
ttgtggaaag gttctcagca gttacagctt aaaccgggtg aatcagcgtt tattgccgcc 6660
aacgaatcac cggtgactgt caaaggccac ggccgtttag cgcgtgttta caacaagctg 6720
taagagctta ctgaaaaaat taacatctct tgctaagctg ggagctctag atccccgaat 6780
ttccccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt 6840
cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg 6900
taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt 6960
taatacgcga tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg 7020
tcatctatgt tactagatcg ggaattggcg agctcgaatt aattcagtac attaaaaacg 7080
tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc 7140
caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca ctcgatacag 7200
gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg cagactttgc 7260
tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg aaacacggat 7320
gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg cctgtgatca 7380
aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt ctcgcttaac 7440
cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc tggataaagc 7500

cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat cgtcgaccgt 7560
acccegatga attaattcgg acgtacgttc tgaacacagc tggatactta cttgggcgat 7620
tgtcatacat gacatcaaca atgtaccccgt ttgtgtaacc gtctcttgga ggttcgtatg 7680
acactagtgg ttcccctcag cttgcgacta gatgttgagg cctaacattt tattagagag 7740
caggctagtt gcttagatac atgatcttca ggccgttatc tgtcagggca agcgaaaatt 7800
ggccatttat gacgaccaat gcccccgcaga agctcccatc tttgccgcca tagacgccgc 7860
gccccccttt tggggtgtag aacatccttt tgccagatgt ggaaaagaag ttcgttgtcc 7920
cattgttggc aatgacgtag tagccggcga aagtgcgaga cccatttgcg ctatatataa 7980
gcctacgatt tccgttgcga ctattgtcgt aattggatga actattatcg tagttgctct 8040
cagagttgtc gtaatttgat ggactattgt cgtaattgct tatggagttg tcgtagttgc 8100
ttggagaaat gtcgtagttg gatggggagt agtcataggg aagacgagct tcatccacta 8160
aaacaattgg caggtcagca agtgcctgcc ccgatgccat cgcaagtacg aggcttagaa 8220
ccaccttcaa cagatcgcgc atagtcttcc ccagctctct aacgcttgag ttaagccgcg 8280
ccgcgaagcg gcgtcggctt gaacgaattg ttagacatta tttgccgact accttggtga 8340
tctcgccttt cacgtagtga acaaattctt ccaactgatc tgcgcgcgag gccaagcgat 8400
cttcttgtcc aagataagcc tgcctagctt caagtatgac gggctgatac tgggccggca 8460
ggcgctccat tgcccagtcg gcagcgacat ccttcggcgc gattttgccg gttactgcgc 8520
tgtaccaaat gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg 8580
gcgagttcca tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg 8640
gatcaaagag ttcctccgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg 8700
tcagcaagat agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt 8760
cattgcgctg ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga 8820
tgtcgtcgtg cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg 8880
aagccgaagt ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta 8940
cggtcaccgt aaccagcaaa tcaatatcac tgtgcggctt caggccgcca tccactgcgg 9000
agccgtacaa atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta 9060
cctctgatag ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactcct 9120
gaattaagcc gcgccgcgaa gcggtgtcgg cttgaatgaa ttgttaggcg tcatcctgtg 9180
ctcccgagaa ccagtaccag tacatcgctg tttcgttcga gacttgaggt ctagttttat 9240
acgtgaacag gtcaatgccg ccgagagtaa agccacattt tgcgtacaaa ctgcaggcag 9300
gtacattgtt cgtttgtgtc tctaatcgta tgccaaggag ctgtctgctt agtgcccact 9360
ttttcgcaaa ttcgatgaga ctgtgcgcga ctcctttgcc tcggtgcgtg tgcgacacaa 9420

caatgtgttc gatagaggct agatcgttcc atgttgagtt gagttcaatc ttcccgacaa 9480
gctcttggtc gatgaatgcg ccatagcaag cagagtcttc atcagagtca tcatccgaga 9540
tgtaatcctt ccggtagggg ctcacacttc tggtagatag ttcaaagcct tggtcggata 9600
ggtgcacatc gaacacttca cgaacaatga aatggttctc agcatccaat gtttccgcca 9660
cctgctcagg gatraccgaa atcttcatat gacgcctaac gcctggcaca gcggatcgca 9720
aacctggcgc ggcttttggc acaaaaggcg tgacaggttt gcgaatccgt tgctgccact 9780
tgttaaccct tttgccagat ttggtaacta taatttatgt tagaggcgaa gtcttgggta 9840
aaaactggcc taaaattgct ggggatttca ggaaagtaaa catcaccttc cggctcgatg 9900
tctattgtag atatatgtag tgtatctact tgatcgggag atctgctgcc tcgcgcgttt 9960
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 10020
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 10080
tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat 10140
gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga 10200
tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg 10260
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 10320
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 10380
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 10440
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 10500
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 10560
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 10620
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 10680
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 10740
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 10800
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 10860
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 10920
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 10980
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 11040
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 11100
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 11160
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 11220
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 11280
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 11340

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 11400
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 11460
agtttgcgca acgttgttgc cattgctgca ggggggggcg gggggggctt ccattgttca 11520
ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc tttcagcacc 11580
tgtcgtttcc tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa 11640
gaacggaaac gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg 11700
ccccgtaacc tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca 11760
tatcacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat 11820
cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg 11880
acactgaata cggggcaacc tcatgtcccc ccccccccgc ccctgcaggc ategtggtgt 11940
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 12000
catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca 12060
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 12120
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 12180
gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg 12240
cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac 12300
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 12360
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 12420
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 12480
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 12540
gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 12600
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 12660
cctttcgtct tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt cgtggagttc 12720
ccgccacaga cccggattga aggcgagatc cagcaactcg cgccagatca tcctgtgacg 12780
gaactttggc gcgtgatgac tggccaggac gtcggccgaa agagcgacaa gcagatcacg 12840
cttttcgaca gcgtcggatt tgcgatcgag gatttttcgg cgctgcgcta cgtccgcgac 12900
cgcgttgagg gatcaagcca cagcagccca ctcgaccttc tagccgaccc agacgagcca 12960
agggatcttt ttggaatgct gctccgtcgt caggctttcc gacgtttggg tggttgaaca 13020
gaagtcatta tcgcacggaa tgccaagcac tcccgagggg aaccctgtgg ttggcatgca 13080
catacaaatg gacgaacgga taaacctttt cacgcccttt taaatatccg attattctaa 13140
taaacgctct tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagtttа 13200
aactgaaggc gggaaacgac aacctgatca tgagcggaga attaagggag tcacgttatg 13260

accccgccg atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt 13320
tgaaggagcc actcagctta attaagtcta actcgagtta ctggtacgta ccaaatccat 13380
ggcacccaac cccatatcga cagaggatgt gaagaacagg taaatcacgc agaagaaccc 13440
atctctgata gcagctatcg attagaacaa cgaatccata ttgggtccgt gggaaatact 13500
tactgcacag gaaggggcg atctgacgag gccccgccac cggcctcgac ccgaggccga 13560
ggccgacgaa gcgccggcga gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg 13620
tgggagggag aggccgcggt ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg 13680
gcgcgggggt cagccgccga gccggcggcg acggaggagc agggcggcgt ggacgcgaac 13740
ttccgatcgg ttggtcagag tgcgcgagtt gggcttagcc aattaggtct caacaatcta 13800
ttgggccgta aaattcatgg gccctggttt gtctaggccc aatatcccgt tcatttcagc 13860
ccacaaatat ttccccagag gattattaag gcccacacgc agcttatagc agatcaagta 13920
cgatgtttcc tgatcgttgg atcggaaacg tacggtcttg atcaggcatg ccgacttcgt 13980
caaagagagg cggcatgacc tgacgcggag ttggttccgg gcaccgtctg gatggtcgta 14040
ccgggaccgg acacgtgtcg cgcctccaac tacatggaca cgtgtcgtgc tgccattggg 14100
ccgtacgcgt ggcggtgacc gcaccggatg ctgcctcgca ccgccttgcc cacgctttat 14160
atagagaggt tttctctcca ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg 14220
agcgaagctt tgcgttctct aatcgcctcg tcaaggtaac taatcaatca cctcgtccta 14280
atcctcgaat ctctcgtggt gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag 14340
cgtaggagga tcccgtgcga gttagtctca atctctcagg gtttcgtgcg attttagggt 14400
gatccacctc ttaatcgagt tacggtttcg tgcgatttta gggtaatcct cttaatctct 14460
cattgattta gggtttcgtg agaatcgagg tagggatctg tgttatttat atcgatctaa 14520
tagatggatt ggttttgaga ttgttctgtc agatgggggat tgtttcgata tattacccta 14580
atgatgtgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 14640
tgtttcgata tattacccta atgatggata ataagagtag ttcacagtta tgttttgatc 14700
ctgccacata gtttgagttt tgtgatcaga tttagtttta cttatttgtg cttagttcgg 14760
atgggattgt tctgatattg ttccaataga tgaatagctc gttaggttaa aatctttagg 14820
ttgagttagg cgacacatag tttatttcct ctggatttgg attggaattg tgttcttagt 14880
ttttttcccc tggatttgga ttggaattgt gtggagctgg gttagagaat tacatctgta 14940
tcgtgtacac ctacttgaac tgtagagctt gggttctaag gtcaatttaa tctgtattgt 15000
atctggctct ttgcctagtt gaactgtagt gctgatgttg tactgtgttt ttttaccegt 15060
tttatttgct ttactcgtgc aaatcaaatc tgtcagatgc tagaactagg tggctttatt 15120
ctgtgttctt acatagatct gttgtcctgt agttacttat gtcagttttg ttattatctg 15180

aagatatttt tggttgttgc ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat 15240
taatgatgct gtccaattgt agtgtagtat gatgtgattg atatgttcat ctattttgag 15300
ctgacagtac cgatatcgta ggatctggtg ccaacttatt ctccagctgc ttttttttac 15360
ctatgttaat tccaatcctt tcttgcctct tccagatcca gataatggcg aacaaacatt 15420
tgtcctctc cctcttcctc gtcctccttg gcctgtcggc cagcttggcc tccgggcaag 15480
tcgctactga tcccttccag tctcgctgta acgagtttca aaacaaaatt gatatcgcaa 15540
acgtcacggt taggtcggtc gcctacgtcg cggctggtca gaatatctct caagcagaag 15600
tcgcttcggt ttgcaaggca agcgtccagg cctccgtgga cctctgtaga gtgacaatga 15660
atattagcac ctcggatagg agccaccttt gggccgaggc ttggctgcct cgcaattata 15720
ctggacggtt cgtctccact gggaacgggg gcctggctgg atgcgtgcag gaaacagact 15780
tgaatttcgc tgccaacttt ggcttcgcaa ccgttggcac taatggaggg cacgatggag 15840
acacggcgaa atactttctg aataactctg aggtcctcgc tgatttcgcc tatcgcagcg 15900
ttcacgaggg cactgtggtc gggaagcaac ttacccagct gttctacgac gaaggttaca 15960
attattcgta ctatctgggc tgcagcaccg gaggcaggca aggctatcag caagtgcagc 16020
gctttccgga tgactacgat ggtgttattg ccggaagtgc agcgatgaac tttatcaatt 16080
tgatttcttg gggcgctttt ctttggaagg ccacgggact cgcagacgac ccagatttta 16140
tctcagcgaa cctctggagc gttatccacc aggagattgt gcggcagtgt gacttggttg 16200
atggtgctct tgacggaatt attgaagatc ctgacttctg cgcccctgtt atcgagagac 16260
tgatttgtga tgggactacc aacggcacct cctgtatcac tggggcacaa gcggctaaag 16320
tgaatagggc cctctcggac ttttatggtc cggatgggac agtctactat ccacgcttga 16380
actacggagg ggaggcagac tctgcgagcc tgtacttcac aggctccatg tattcacgga 16440
ccgaagagtg gtacaagtac gtcgtgtaca acgatactaa ctggaatagt tctcaatgga 16500
cgctcgaaag cgctaagttg gccttggagc agaaccctt taatattcag gcattcgacc 16560
ctaatatcac agcgttcagg gaccgcggtg gaaaactctt gagctaccac ggcacccaag 16620
atccgattat ttcatcgact gacagcaagt tgtattacag aagagttgct aacgccctta 16680
acgctgcccc atccgaactc gacgagtttt atcgcttctt tcagatctct gggatgggtc 16740
actgcggcga cggaaccggg gcttcatacа ttggccaggg atacggcacg tatacctcga 16800
aggccccaca ggtcaacttg cctcggacta tggtggattg ggttgagaat ggaaaagctc 16860
cggaatatat gcctgggaat aagttgaacg ccaacggttc gatcgagtat atgagaaagc 16920
actgtcgtta cccaaaacac aatattcaca cgggcccagg taactacacg gaccctaact 16980
cctggacctg cgtcagcgag aaggacgagc tgtgatcccc gaatttcccc gatcgttcaa 17040
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 17100

tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 17160

ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 17220

acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 17280

atcgggaatt gggtac 17296

<210> 70
<211> 20454
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2096

<400> 70
cgtcgactct agtaacggcc gccagtgtgc tggaattaat tcggcttgtc gaccacccaa 60

cccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catctctgat 120

agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca 180

ggaaggggcc gatctgacga ggccccgcca ccggcctcga cccgaggccg aggccgacga 240

agcgccggcg agtacggcgc cgcggcggcc tctgccccgtg ccctctgcgc gtgggaggga 300

gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg 360

tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 420

gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 480

aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 540

tttcccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 600

ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg tcaaagagag 660

gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 720

gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 780

tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 840

ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 900

ttgcgttctc taatcggctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 960

tctctcgtgg tgccccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 1020

atcccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg tgatccacct 1080

cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 1140

agggtttcgt gagaatcgag gtagggatct gtgttactta tatcgatcta atagatggat 1200

tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 1260

cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 1320

atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 1380

agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 1440
ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 1500
gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag ttttttcccc 1560
ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt atcgtgtaca 1620
cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 1680
tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt tttttacccg ttttatttgc 1740
tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttct 1800
tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 1860
ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 1920
tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctatttttga gctgacagta 1980
ccgatatcgt aggatctggt gccaacttat tctccagctg ctttttttta cctatgttaa 2040
ttccaatcct ttcttgcctc ttccagatcc agataatggc gaacaaacat ttgtccctct 2100
ccctcttcct cgtcctcctt ggcctgtcgg ccagcttggc ctccgggcaa cgaacaagca 2160
ttactctgac atccaacgca tccggtacgt ttgacggtta ctattacgaa ctctggaagg 2220
atactggcaa tacaacaatg acggtctaca ctcaaggtcg cttttcctgc cagtggtcga 2280
acatcaataa cgcgttgttt aggaccggga agaaatacaa ccagaattgg cagtctcttg 2340
gcacaatccg gatcacgtac tctgcgactt acaacccaaa cgggaactcc tacttgtgta 2400
tctatggctg gtctaccaac ccattggtcg agttctacat cgttgagtcc tgggggaact 2460
ggagaccgcc tggtgccacg tccctgggcc aagtgacaat cgatggcggg acctacgaca 2520
tctataggac gacacgcgtc aaccagcctt ccattgtggg gacagccacg ttcgatcagt 2580
actggagcgt gcgcacctct aagcggactt caggaacagt gaccgtgacc gatcacttcc 2640
gcgcctgggc gaaccggggc ctgaacctcg gcacaataga ccaaattaca ttgtgcgtgg 2700
agggttacca aagctctgga tcagccaaca tcaccagaa cacttctct cagggctctt 2760
cttccggcag ttcgggtggc tcatccggct ccacaacgac tactcgcatc gagtgtgaga 2820
acatgtcctt gtccggaccc tacgttagca ggatcaccaa tccctttaat ggtattgcgc 2880
tgtacgccaa cggagacaca gcccgcgcta ccgttaactt ccccgcaagt cgcaactaca 2940
atttccgcct gcggggttgc ggcaacaaca ataatcttgc ccgtgtggac ctgaggatcg 3000
acggacggac cgtcgggacc ttttattacc agggcacata cccctgggag gccccaattg 3060
acaatgttta tgtcagtgcg gggagtcata cagtcgaaat cactgttact gcggataacg 3120
gcacatggga cgtgtatgcc gactacctgg tgatacagtg acctaggtcc ccgaatttcc 3180
ccgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg 3240
cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat 3300

gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 3360
acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcgggtgtcat 3420
ctatgttact agatcgggaa ttggaattca tactaaagct tgcatgcctg caggtcgact 3480
ctagtaacgg ccgccagtgt gctggaatta attcggcttg tcgacccccc aacccccatat 3540
cgacagagga tgtgaagaac aggtaaatca cgcagaagaa cccatctctg atagcagcta 3600
tcgattagaa caacgaatcc atattgggtc cgtgggaaat acttactgca caggaagggg 3660
gcgatctgac gaggcccccgc caccggcctc gacccgaggc cgaggccgac gaagcgccgg 3720
cgagtacggc gccgcggcgg cctctgcccg tgccctctgc gcgtgggagg gagaggccgc 3780
ggtggtgggg gcgcgcgcgc gcgcgcgcgc agctggtgcg gcggcgcggg ggtcagccgc 3840
cgagccggcg gcgacggagg agcagggcgg cgtggacgcg aacttccgat cggttggtca 3900
gagtgcgcga gttgggctta gccaattagg tctcaacaat ctattgggcc gtaaaattca 3960
tgggccctgg tttgtctagg cccaatatcc cgttcatttc agcccacaaa tatttcccca 4020
gaggattatt aaggcccaca cgcagcttat agcagatcaa gtacgatgtt tcctgatcgt 4080
tggatcggaa acgtacggtc ttgatcaggc atgccgactt cgtcaaagag aggcggcatg 4140
acctgacgcg gagttggttc cgggcaccgt ctggatggtc gtaccgggac cggacacgtg 4200
tcgcgcctcc aactacatgg acacgtgtgg tgctgccatt gggccgtacg cgtggcggtg 4260
accgcaccgg atgctgcctc gcaccgcctt gcccacgctt tatatagaga ggttttctct 4320
ccattaatcg catagcgagt cgaatcgacc gaaggggagg gggagcgaag ctttgcgttc 4380
tctaatcgcc tcgtcaaggt aactaatcaa tcacctcgtc ctaatcctcg aatctctcgt 4440
ggtgcccgtc taatctgcgg attttgatgc tcgtggtgga aagcgtagga ggatcccgtg 4500
cgagttagtc tcaatctctc agggtttcgt gcgatttag ggtgatccac ctcttaatcg 4560
agttacggtt tcgtgcgatt ttagggtaat cctcttaatc tctcattgat ttagggtttc 4620
gtgagaatcg aggtagggat ctgtgttatt tatatcgatc taatagatgg attggttttg 4680
agattgttct gtcagatggg gattgtttcg atatattacc ctaatgatgt gtcagatggg 4740
gattgtttcg atatattacc ctaatgatgt gtcagatggg gattgtttcg atatattacc 4800
ctaatgatgg ataataagag tagttcacag ttatgttttg atcctgccac atagtttgag 4860
ttttgtgatc agatttagtt ttacttattt gtgcttagtt cggatgggac tgttctgata 4920
ttgttccaat agatgaatag ctcgttaggt taaaatcttt aggttgagtt aggcgacaca 4980
tagtttattt cctctggatt tggattggaa ttgtgttctt agtttttttc cctggatttt 5040
ggattggaat tgtgtggagc tgggttagag aattacatct gtatcgtgta cacctacttg 5100
aactgtagag cttgggttct aaggtcaatt taatctgtat tgtatctggc tcttcgccta 5160
gttgaactgt agtgctgatg ttgtactgtg ttttttttacc cgttttattt gctttactcg 5220

tgcaaatcaa atctgtcaga tgctagaact aggtggcttt attctgtgtt cttacataga 5280
tctgttgtcc tgtagttact tatgtcagtt ttgttattat ctgaagatat ttttggttgt 5340
tgcttgttga tgtggtgtga gctgtgagca gcgctcttat gattaatgat gctgtccaat 5400
tgtagtgtag tatgatgtga ttgatatgtt catctatttt gagctgacag taccgatatc 5460
gtaggatctg gtgccaactt attctccagc tgcttttttt tacctatgtt aattccaatc 5520
ctttctgcc tcttccagat ccagataatg cagaaactca ttaactcagt gcaaaactat 5580
gcctggggca gcaaaacggc gttgactgaa ctttatggta tggaaaatcc gtccagccag 5640
ccgatggccg agctgtggat gggcgcacat ccgaaaagca gttcacgagt gcagaatgcc 5700
gccggagata tcgtttcact gcgtgatgtg attgagagtg ataaatcgac tctgctcgga 5760
gaggccgttg ccaaacgcct tggcgaactg cctttcctgt tcaaagtatt atgcgcagca 5820
cagccactct ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg ttttgccaaa 5880
gaaaatgccg caggtatccc gatggatgcc gccgagcgta actataaaga tcctaaccac 5940
aagccggagc tggtttttgc gctgacgcct ttccttgcga tgaacgcgtt tcgtgaattt 6000
tccgagattg tctcctact ccagccggtc gcaggtgcac atccggcgat tgctcacttt 6060
ttacaacagc ctgatgccga acgtttaagc gaactgttcg ccagcctgtt gaatatgcag 6120
ggtgaagaaa aatcccgcgc gctggcgatt ttaaaatcgg cctcgatag ccagcagggt 6180
gaaccgtggc aaacgattcg tttaatttct gaattttacc cggaagacag cggtctgttc 6240
tcccccgctat tgctgaatgt ggtgaaattg aaccctggcg aagcgatgtt cctgttcgct 6300
gaaacaccgc acgcttacct gcaaggcgtg gcgctggaag tgatggcaaa ctccgataac 6360
gtgctgcgtg cgggtctgac gcctaaatac attgatattc cggaactggt tgccaatgtg 6420
aaattcgaag ccaaaccggc taaccagttg ttgacccagc cggtgaaaca aggtgcagaa 6480
ctggacttcc cgattccagt ggatgatttt gccttctcgc tgcatgacct tagtgataaa 6540
gaaaccacca ttagccagca gagtgccgcc attttgttct gcgtcgaagg cgatgcaacg 6600
ttgtggaaag gttctcagca gttacagctt aaaccgggtg aatcagcgtt tattgccgcc 6660
aacgaatcac cggtgactgt caaaggccac ggccgtttag cgcgtgttta caacaagctg 6720
taagagctta ctgaaaaaat taacatctct tgctaagctg ggagctctag atccccgaat 6780
ttcccccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt 6840
cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg 6900
taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt 6960
taatacgcga tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg 7020
tcatctatgt tactagatcg ggaattggcg agctcgaatt aattcagtac attaaaaacg 7080
tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc 7140

caccagccag ccaacagctc ccgaccggc agctcggcac aaaatcacca ctcgatacag 7200
gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg cagactttgc 7260
tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg aaacacggat 7320
gatctgcggg agggtagcat gttgattgta acgatgacag agcgttgctg cctgtgatca 7380
aatatcatct cctcgcaga gatccgaatt atcagccttc ttattcattt ctcgcttaac 7440
cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc tggataaagc 7500
cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat cgtcgaccgt 7560
acccccgatga attaattcgg acgtacgttc tgaacacagc tggatacttа cttgggcgat 7620
tgtcatacat gacatcaaca atgtaccccgt ttgtgtaacc gtctcttgga ggttcgtatg 7680
acactagtgg ttcccctcag cttgcgacta gatgttgagg cctaacattt tattagagag 7740
caggctagtt gcttagatac atgatcttca ggccgttatc tgtcagggca agcgaaaatt 7800
ggccatttat gacgaccaat gccccgcaga agctcccatc tttgccgcca tagacgccgc 7860
gccccccttt tggggtgtag aacatccttt tgccagatgt ggaaaagaag ttcgttgtcc 7920
cattgttggc aatgacgtag tagccggcga aagtgcgaga cccatttgcg ctatatataa 7980
gcctacgatt tccgttgcga ctattgtcgt aattggatga actattatcg tagttgctct 8040
cagagttgtc gtaatttgat ggactattgt cgtaattgct tatggagttg tcgtagttgc 8100
ttggagaaat gtcgtagttg gatggggagt agtcataggg aagacgagct tcatccacta 8160
aaacaattgg caggtcagca agtgcctgcc ccgatgccat cgcaagtacg aggcttagaa 8220
ccaccttcaa cagatcgcgc atagtcttcc ccagctctct aacgcttgag ttaagccgcg 8280
ccgcgaagcg gcgtcggctt gaacgaattg ttagacatta tttgccgact accttggtga 8340
tctcgccttt cacgtagtga acaaattctt ccaactgatc tgcgcgcgag gccaagcgat 8400
cttcttgtcc aagataagcc tgcctagctt caagtatgac gggctgatac tgggccggca 8460
ggcgctccat tgcccagtcg gcagcgacat ccttcggcgc gattttgccg gttactgcgc 8520
tgtaccaaat gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg 8580
gcgagttcca tagcgttaag gtttcattta gcggctcaaa tagatcctgt tcaggaaccg 8640
gatcaaagag ttcctccgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg 8700
tcagcaagat agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt 8760
cattgcgctg ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga 8820
tgtcgtcgtg cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg 8880
aagccgaagt ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta 8940
cggtcaccgt aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg 9000
agccgtacaa atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta 9060

cctctgatag ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactcct 9120
gaattaagcc gcgccgcgaa gcggtgtcgg cttgaatgaa ttgttaggcg tcatcctgtg 9180
ctcccgagaa ccagtaccag tacatcgctg tttcgttcga gacttgaggt ctagttttat 9240
acgtgaacag gtcaatgccg ccgagagtaa agccacattt tgcgtacaaa ttgcaggcag 9300
gtacattgtt cgtttgtgtc tctaatcgta tgccaaggag ctgtctgctt agtgcccact 9360
ttttcgcaaa ttcgatgaga ctgtgcgcga ctcctttgcc tcggtgcgtg tgcgacacaa 9420
caatgtgttc gatagaggct agatcgttcc atgttgagtt gagttcaatc ttcccgacaa 9480
gctcttggtc gatgaatgcg ccatagcaag cagagtcttc atcagagtca tcatccgaga 9540
tgtaatcctt ccggtagggg ctcacacttc tggtagatag ttcaaagcct tggtcggata 9600
ggtgcacatc gaacacttca cgaacaatga aatggttctc agcatccaat gtttccgcca 9660
cctgctcagg gatcaccgaa atcttcatat gacgcctaac gcctggcaca gcggatcgca 9720
aacctggcgc ggcttttggc acaaaaggcg tgacaggttt gcgaatccgt tgctgccact 9780
tgttaaccct tttgccagat ttggtaacta taatttatgt tagaggcgaa gtcttgggta 9840
aaaactggcc taaaattgct ggggatttca ggaaagtaaa catcaccttc cggctcgatg 9900
tctattgtag atatatgtag tgtatctact tgatgggggg atctgctgcc tcgcgcgttt 9960
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 10020
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 10080
tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat 10140
gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga 10200
tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg 10260
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 10320
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 10380
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 10440
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 10500
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 10560
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 10620
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 10680
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 10740
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 10800
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 10860
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 10920
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 10980

cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 11040
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 11100
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 11160
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 11220
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 11280
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 11340
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 11400
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 11460
agtttgcgca acgttgttgc cattgctgca ggggggggcg ggggggggtt ccattgttca 11520
ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc tttcagcacc 11580
tgtcgtttcc tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa 11640
gaacggaaac gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg 11700
ccccgtaacc tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca 11760
tatcacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat 11820
cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg 11880
acactgaata cggggcaacc tcatgtcccc cccccccccc ccctgcaggc atcgtggtgt 11940
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 12000
catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca 12060
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 12120
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 12180
gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg 12240
cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac 12300
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 12360
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 12420
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 12480
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 12540
gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 12600
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 12660
cctttcgtct tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt cgtggagttc 12720
ccgccacaga cccggattga aggcgagatc cagcaactcg cgccagatca tcctgtgacg 12780
gaactttggc gcgtgatgac tggccaggac gtcggccgaa agagcgacaa gcagatcacg 12840
cttttcgaca gcgtcggatt tgcgatcgag gatttttcgg cgctgcgcta cgtccgcgac 12900

cgcgttgagg gatcaagcca cagcagccca ctcgaccttc tagccgaccc agacgagcca 12960

agggatcttt ttggaatgct gctccgtcgt caggctttcc gacgtttggg tggttgaaca 13020

gaagtcatta tcgcacggaa tgccaagcac tcccgagggg aaccctgtgg ttggcatgca 13080

catacaaatg gacgaacgga taaacctttt cacgcccttt taaatatccg attatttaaa 13140

taaacgctct tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagttta 13200

aactgaaggc gggaaacgac aacctgatca tgagcggaga attaagggag tcacgttatg 13260

acccccgccg atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt 13320

tgaaggagcc actcagctta attaagtcta actcgagcac ccaaccccat atcgacagag 13380

gatgtgaaga acaggtaaat cacgcagaag aacccatctc tgatagcagc tatcgattag 13440

aacaacgaat ccatattggg tccgtgggaa atacttactg cacaggaagg gggcgatctg 13500

acgaggcccc gccaccggcc tcgacccgag gccgaggccg acgaagcgcc ggcgagtacg 13560

gcgccgcggc ggcctctgcc cgtgccctct gcgcgtggga gggagaggcc gcggtggtgg 13620

gggcgcgcgc gcgcgcgcgc gcagctggtg cggcggcgcg ggggtcagcc gccgagccgg 13680

cggcgacgga ggagcagggc ggcgtggacg cgaacttccg atcggttggt cagagtgcgc 13740

gagttgggct tagccaatta ggtctcaaca atctattggg ccgtaaaatt catgggccct 13800

ggtttgtcta ggccaatat cccgttcatt tcagcccaca aatatttccc cagaggatta 13860

ttaaggccca cacgcagctt atagcagatc aagtacgatg tttcctgatc gttggatcgg 13920

aaacgtacgg tcttgatcag gcatgccgac ttcgtcaaag agaggcggca tgacctgacg 13980

cggagttggt tccgggcacc gtctggatgg tcgtaccggg accggacacg tgtcgcgcct 14040

ccaactacat ggacacgtgt ggtgctgcca ttgggccgta cgcgtggcgg tgaccgcacc 14100

ggatgctgcc tcgcaccgcc ttgcccacgc tttatataga gaggttttct ctccattaat 14160

cgcatagcga gtcgaatcga ccgaagggga gggggagcga agctttgcgt tctctaatcg 14220

cctcgtcaag gtaactaatc aatcacctcg tcctaatcct cgaatctctc gtggtgcccg 14280

tctaatctcg cgattttgat gctcgtggtg gaaagcgtag gaggatcccg tgcgagttag 14340

tctcaatctc tcagggtttc gtgcgatttt agggtgatcc acctccttaat cgagttacgg 14400

tttcgtgcga ttttagggta atcctcttaa tctctcattg atttagggtt tcgtgagaat 14460

cgaggtaggg atctgtgtta tttatatcga tctaatagat ggattggttt tgagattgtt 14520

ctgtcagatg gggattgttt cgatatatta ccctaatgat gtgtcagatg gggattgttt 14580

cgatatatta ccctaatgat gtgtcagatg gggattgttt cgatatatta ccctaatgat 14640

ggataataag agtagttcac agttatgttt tgatcctgcc acatagtttg agtttgtga 14700

tcagatttag ttttacttat ttgtgcttag ttcggatggg attgttctga tattgttcca 14760

atagatgaat agctcgttag gttaaaatct ttaggttgag ttaggcgaca catagtttat 14820

ttcctctgga tttggattgg aattgtgttc ttagtttttt tcccctggat ttggattgga 14880
attgtgtgga gctgggttag agaattacat ctgtatcgtg tacacctact tgaactgtag 14940
agcttgggtt ctaaggtcaa tctaatctgt attgtatctg gctctttgcc tagttgaact 15000
gtagtgctga tgttgtactg tgtttttta cccgttttat ttgctttact cgtgcaaatc 15060
aaatctgtca gatgctagaa ctaggtggct ttattctgtg ttcttacata gatctgttgt 15120
cctgtagtta cttatgtcag ttttgttatt atctgaagat atttttggtt gttgcttgtt 15180
gatgtggtgt gagctgtgag cagcgctctt atgattaatg atgctgtcca attgtagtgt 15240
agtatgatgt gattgatatg ttcatctatt ttgagctgac agtaccgata tcgtaggatc 15300
tggtgccaac ttattctcca gctgcttttt tttacctatg ttaattccaa tcctttcttg 15360
cctcttccag atccagataa tggcgaacaa acatttgtcc ctctcccctc tcctcgtcct 15420
ccttggcctg tcggccagct tggcctccgg gcaagtcgca tccacgcagg gtatctctga 15480
ggacctgtat aatcgcctcg tggaaatggc cacaatttca caagcggctt acgcagatct 15540
ttgtaatatc ccttcgacaa ttatcaaggg agagaaaatc tataacgccc agactgacat 15600
caacggctgg atactgcggg atgacacgag caaggaaatt atcacagtct ttagagggac 15660
cggttccgat acaaatttgc agttggacac gaattacaca ctgacccccc tcgatactct 15720
ccctcaatgc aacgactgtg aggttcacgg tgggtactat attggctgga tctctgttca 15780
agaccaagtc gagtcacttg ttaagcagca agcgtcgcag taccccgact acgcattgac 15840
ggtgacaggg cacagcctgg gtgcctcgat ggcagcgctc accgctgccc agctttctgc 15900
aacctacgat aatgtcaggc tgtacacttt cggagaacca cgctcaggca accaagcgtt 15960
tgcttcgtat atgaacgacg ctttccaggt tagctccccc gagacgacac aatactttcg 16020
ggtgacccat tctaacgacg ggattcctaa cctcccgcca gccgacgaag gttacgcaca 16080
cgggggtgtc gagtactggt cagtggaccc ctacagcgcg cagaatactt tcgtttgcac 16140
gggcgatgag gtccagtgct gtgaagctca agggggtcag ggagtgaatg atgcacacac 16200
aacctatttc ggaatgactt ccgggggttg cacgtggagc gagaaggacg agctgtgatc 16260
cccgaatttc cccgatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt 16320
tgccggtctt gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat 16380
taacatgtaa tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt 16440
atacatttaa tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg 16500
cgcggtgtca tctatgttac tagatcgggа attgccatgg cacccaaccc catatcgaca 16560
gaggatgtga agaacaggta aatcacgcag aagaacccat ctctgatagc agctatcgat 16620
tagaacaacg aatccatatt gggtccgtgg gaaatactta ctgcacagga agggggcgat 16680
ctgacgaggc ccgcccaccg gcctcgaccc gaggccgagg ccgacgaagc gcccggcgagt 16740

acggcgccgc ggcggcctct gcccgtgccc tctgcgcgtg ggagggagag gccgcggtgg 16800
tgggggcgcg cgcgcgcgcg cgcgcagctg gtgcggcggc gcgggggtca gccgccgagc 16860
cggcggcgac ggaggagcag ggcggcgtgg acgcgaactt ccgatcggtt ggtcagagtg 16920
cgcgagttgg gcttagccaa ttaggtctca acaatctatt gggcgtaaaa attcatgggc 16980
cctggttgt ctaggcccaa tatcccgttc atttcagccc acaaatattt ccccagagga 17040
ttattaaggc ccacacgcag cttatagcag atcaagtacg atgtttcctg atcgttggat 17100
cggaaacgta cggtcttgat caggcatgcc gacttcgtca aagaggcgg gcatgacctg 17160
acgcggagtt ggttccgggc accgtctgga tggtcgtacc gggaccggac acgtgtcgcg 17220
cctccaacta catggacacg tgtggtgctg ccattgggcc gtacgcgtgg cggtgaccgc 17280
accggatgct gcctcgcacc gccttgccca cgctttatat agagaggttt tctctccatt 17340
aatcgcatag cgagtcgaat cgaccgaagg ggaggggggag cgaagctttg cgttctctaa 17400
tcgcctcgtc aaggtaacta atcaatcacc tcgtcctaat cctcgaatct ctcgtggtgc 17460
ccgtctaatc tcgcgatttt gatgctcgtg gtggaaagcg taggaggatc ccgtgcgagt 17520
tagtctcaat ctctcagggt ttcgtgcgat tttagggtga tccacctctt aatcgagtta 17580
cggtttcgtg cgattttagg gtaatcctct taatctctca ttgatttagg gtttcgtgag 17640
aatcgaggta gggatctgtg ttatttatat cgatctaata gatggattgg ttttgagatt 17700
gttctgtcag atggggattg tttcgatata ttaccctaat gatgtgtcag atggggattg 17760
tttcgatata ttaccctaat gatgtgtcag atggggattg tttcgatata ttaccctaat 17820
gatggataat aagagtagtt cacagttatg ttttgatcct gccacatagt ttgagttttg 17880
tgatcagatt tagttttact tatttgtgct tagttcggat gggattgttc tgatattgtt 17940
ccaatagatg aatagctcgt taggttaaaa tctttaggtt gagttaggcg acacatagtt 18000
tatttcctct ggatttggat tggaattgtg ttcttagttt ttttcccctg gatttggatt 18060
ggaattgtgt ggagctgggt tagagaatta catctgtatc gtgtacacct acttgaactg 18120
tagagcttgg gttctaaggt caatttaatc tgtattgtat ctggctcttt gcctagttga 18180
actgtagtgc tgatgttgta ctgtgttttt ttacccgttt tatttgcttt actcgtgcaa 18240
atcaaatctg tcagatgcta gaactaggtg gctttattct gtgttcttac atagatctgt 18300
tgtcctgtag ttacttatgt cagttttgtt attatctgaa gatatttttg gttgttgctt 18360
gttgatgtgg tgtgagctgt gagcagcgct cttatgatta atgatgctgt ccaattgtag 18420
tgtagtatga tgtgattgat atgttcatct attttgagct gacagtaccg atatcgtagg 18480
atctggtgcc aacttattct ccagctgctt ttttttacct atgttaattc caatcctttc 18540
ttgcctcttc cagatccaga taatggcgaa caaacatttg tccctctccc tcttcctcgt 18600
cctccttggc ctgtcggcca gcttggcctc cgggcaagtc gctactgatc ccttccagtc 18660

tcgctgtaac gagtttcaaa acaaaattga tatcgcaaac gtcacggtta ggtcggtcgc 18720
ctacgtcgcg gctggtcaga atatctctca agcagaagtc gcttcggttt gcaaggcaag 18780
cgtccaggcc tccgtggacc tctgtagagt gacaatgaat attagcacct cggataggag 18840
ccacctttgg gccgaggctt ggctgcctcg caattatact ggacggttcg tctccactgg 18900
gaacgggggc ctggctggat gcgtgcagga aacagacttg aatttcgctg ccaactttgg 18960
cttcgcaacc gttggcacta atggagggca cgatggagac acggcgaaat actttctgaa 19020
taactctgag gtcctcgctg atttcgccta tcgcagcgtt cacgagggca ctgtggtcgg 19080
gaagcaactt acccagttgt tctacgacga aggttacaat tattcgtact atctgggctg 19140
cagcaccgga ggcaggcaag gctatcagca agtgcagcgc tttccggatg actacgatgg 19200
tgttattgcc ggaagtgcag cgatgaactt tatcaatttg atttcttggg gcgcttttct 19260
ttggaaggcc acgggactcg cagacgaccc agattttatc tcagcgaacc tctggagcgt 19320
tatccaccag gagattgtgc ggcagtgtga cttggttgat ggtgctcttg acggaattat 19380
tgaagatcct gactttctgc ccctgttat cgagagactg atttgtgatg ggactaccaa 19440
cggcacctcc tgtatcactg gggcacaagc ggctaaagtg aatagggccc tctcggactt 19500
ttatggtccg gatgggacag tctactatcc acgcttgaac tacggagggg aggcagactc 19560
tgcgagcctg tacttcacag gctccatgta ttcacggacc gaagagtggt acaagtacgt 19620
cgtgtacaac gatactaact ggaatagttc tcaatggacg ctcgaaagcg ctaagttggc 19680
cttggagcag aaccccttta atattcaggc attcgaccct aatatcacag cgttcaggga 19740
ccgcggtgga aaactcttga gctaccacgg cacccaagat ccgattattt catcgactga 19800
cagcaagttg tattacagaa gagttgctaa cgcccttaac gctgccccat ccgaactcga 19860
cgagttttat cgcttctttc agatctctgg gatgggtcac tgcggcgacg gaaccggggc 19920
ttcatacatt ggccagggat acggcacgta taccttgaag gccccacagg tcaacttgct 19980
tcggactatg gtggattggg ttgagaatgg aaaagctccg gaatatatgc ctgggaataa 20040
gttgaacgcc aacggttcga tcgagtatat gagaaagcac tgtcgttacc caaaacacaa 20100
tattcacacg ggcccaggta actacacgga ccctaactcc tggacctgcg tcagcgagaa 20160
ggacgagctg tgatccccga atttccccga tcgttcaaac atttggcaat aaagtttctt 20220
aagattgaat cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt 20280
taagcatgta ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat 20340
tagagtcccg caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta 20400
ggataaatta tcgcgcgcgg tgtcatctat gttactagat cgggaattgg gtac 20454

<210> 71
<211> 13540

<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2261

<400> 71
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaacccа tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt acctgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggcccа 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggacggga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg cacccgatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttaccegtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tatttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa cacccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggtttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctcccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc cccttttggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg ctttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580

acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500

agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg gggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420

atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttttcacg cccttttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gcccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgact ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220
tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280
atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggctaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaactgtagt gctgatgttg tactgtgttt ttttaccegt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc tttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatgca gatctttgtg aagaccctga ccggcaagac tatcaccctc 12060

gaggtggagt cttctgacac cattgacaat gtcaaggcca agatccagga caaggagggc 12120

atcccccag accagcagcg gctcatcttt gctggcaagc agctggagga cgggcgcacg 12180

cttgcggact acaacatcca gaaggagagc accctccacc tggtgctccg cctcagggga 12240

ggcatgcaaa caagcattac tctgacatcc aacgcatccg gtacgtttga cggttactat 12300

tacgaactct ggaaggatac tggcaataca acaatgacgg tctacactca aggtcgcttt 12360

tcctgccagt ggtcgaacat caataacgcg ttgtttagga ccgggaagaa atacaaccag 12420

aattggcagt ctcttggcac aatccggatc acgtactctg cgacttacaa cccaaacggg 12480

aactcctact tgtgtatcta tggctggtct accaacccat tggtcgagtt ctacatcgtt 12540

gagtcctggg ggaactggag accgcctggt gccacgtccc tgggccaagt gacaatcgat 12600

ggcgggacct acgacatcta taggacgaca cgcgtcaacc agccttccat tgtggggaca 12660

gccacgttcg atcagtactg gagcgtgcgc acctctaagc ggacttcagg aacagtgacc 12720

gtgaccgatc acttccgcgc ctgggcgaac cggggcctga acctcggcac aatagaccaa 12780

attacattgt gcgtggaggg ttaccaaagc tctggatcag ccaacatcac ccagaacacc 12840

ttctctcagg gctcttcttc cggcagttcg ggtggctcat ccggctccac aacgactact 12900

cgcatcgagt gtgagaacat gtccttgtcc ggaccctacg ttagcaggat caccaatccc 12960

tttaatggta ttgcgctgta cgccaacgga gacacagccc gcgctaccgt taacttcccc 13020

gcaagtcgca actacaattt ccgcctgcgg ggttgcggca acaacaataa tcttgcccgt 13080

gtggacctga ggatcgacgg acggaccgtc gggaccttct attaccaggg cacataccgc 13140

tgggaggccc caattgacaa tgtttatgtc agtgcgggga gtcatacagt cgaaatcact 13200

gttactgcgg ataacggcac atgcgacgtg tatgccgact acctggtgat acagtgacct 13260

aggtccccga atttcccccga tcgttcaaac atttggcaat aaagtttctt aagattgaat 13320

cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta 13380

ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg 13440

caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 13500

tcgcgcgcgg tgtcatctat gttactagat cgggaattgg 13540

<210> 72
<211> 14683
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2229

<400> 72
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgcccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatgggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatgggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat actaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccc ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300
atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360
caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480
tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600
ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgcgcg cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tctttaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc cccttttggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160

catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggcgtca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080

ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat accgtcgccg ctttctcccc 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt ttttcgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000

cgacacggaa atgttgaata ctcatactct tccttttcca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttttcacg ccctttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg acccccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atcctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgccccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggccggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggagggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac tcatcgatca cctcgtccta atcctcgaat ctctcgtggt 10920

gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980

gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040

tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100

agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160

ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220

tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaactgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatggcg aacaaacatt tgtccctctc cctcttcctc 12060

gtcctccttg gcctgtcggc cagcttggcc tccgggcaac aaacaagcat tactctgaca 12120

tccaacgcat ccggtacgtt tgacggttac tattacgaac tctggaagga tactggcaat 12180

acaacaatga cggtctacac tcaaggtcgc ttttcctgcc agtggtcgaa catcaataac 12240

gcgttgttta ggaccgggaa gaaatacaac cagaattggc agtctcttgg cacaatccgg 12300

atcacgtact ctgcgactta caacccaaac gggaactcct acttgtgtat ctatggctgg 12360

tctaccaacc cattggtcga gttctacatc gttgagtcct gggggaactg gagaccgcct 12420

ggtgcctgcc tggccgaggg ctcgctcgtc ttggacgcgg ctacgggcca gagggtccct 12480

atcgaaaagg tgcgtccggg gatggaagtt ttctccttgg gacctgatta cagactgtat 12540

cgggtgcccg ttttggaggt ccttgagagc ggggttaggg aagttgtgcg cctcagaact 12600

cggtcaggga gaacgctggt gttgacacca gatcacccgc ttttgacccc cgaaggttgg 12660

aaacctcttt gtgacctccc gcttggaact ccaattgcag tcccccgcaga actgcctgtg 12720

gcgggccact tggccccacc tgaagaacgt gttacgctcc tggctcttct gttgggggat 12780

gggaacacaa agctgtcggg tcggagaggt acacgtccta atgccttctt ctacagcaaa 12840

aaccccgaat tgctcgcggc ttatcgccgg tgtgcagaag ccttgggtgc aaaggtgaaa 12900
gcatacgtcc accegactac gggggtggtt acactegcaa ccctcgctcc acgtcctgga 12960
gctcaagatc ctgtcaaacg cctcgttgtc gaggcgggaa tggttgctaa agccgaagag 13020
aagagggtcc cggaggaggt gtttcgttac cggcgtgagg cgttggccct tttcttgggc 13080
cgtttgttct cgacagacgg ctctgttgaa aagaagagga tctcttattc aagtgccagt 13140
ttgggactgg cccaggatgt cgcacatctc ttgctgcgcc ttggaattac atctcaactc 13200
cgttcgagag ggccacgggc tcacgaggtt cttatatcgg gccgcgagga tattttgcgg 13260
tttgctgaac ttatcggacc ctacctcttg ggggccaaga gggagagact tgcagcgctg 13320
gaagctgagg cccgcaggcg tttgcctgga cagggatggc acttgcggct tgttcttcct 13380
gccgtggcgt acagagtggg cgaggcggaa aggcgctcgg gattttcgtg gagtgaagcc 13440
ggtcggcgcg tcgcagttgc gggatcgtgt ttgtcatctg gactcaacct caaattgccc 13500
agacgctacc tttctcggca ccggttgtcg ctgctcggtg aggcttttgc cgaccctggg 13560
ctggaagcgc tcgcggaagg ccaagtgctc tgggacccta ttgttgctgt cgaaccggcc 13620
ggtaaggcga gaacattcga cttgcgcgtt ccacccttta caaacttcgt gagcgaggac 13680
ctggtggtgc ataacaccgt cccctgggc caagtgacaa tcgatggcgg gacctacgac 13740
atctatagga cgacacgcgt caaccagcct tccattgtgg ggacagccac gttcgatcag 13800
tactggagcg tgcgcacctc taagcggact tcaggaacag tgaccgtgac cgatcacttc 13860
cgcgcctggg cgaaccgggg cctgaacctc ggcacaatag accaaattac attgtgcgtg 13920
gagggttacc aaagctctgg atcagccaac atcacccaga acaccttctc tcagggctct 13980
tcttccggca gttcgggtgg ctcatccggc tccacaacga ctactcgcat cgagtgtgag 14040
aacatgtcct tgtccggacc ctacgttagc aggatcacca atcccttaa tggtattgcg 14100
ctgtacgcca acggagacac agcccgcgct accgttaact tccccgcaag tcgcaactac 14160
aatttccgcc tgcggggttg cggcaacaac aataatcttg cccgtgtgga cctgaggatc 14220
gacggacgga ccgtcgggac cttttattac cagggcacat acccctggga ggccccaact 14280
gacaatgttt atgtcagtgc ggggagtcat acagtcgaaa tcactgttac tgcggataac 14340
ggcacatggg acgtgtatgc cgactacctg gtgatacaga gcgagaagga cgagctgtga 14400
cctaggtccc cgaatttccc cgatcgttca aacatttggc aataaagttt cttaagattg 14460
aatcctgttg ccggtcttgc gatgattatc atataatttc tgttgaatta cgttaagcat 14520
gtaataatta acatgtaatg catgacgtta tttatgagat gggtttttat gattagagtc 14580
ccgcaattat acatttaata cgcgatagaa aacaaaatat agcgcgcaaa ctaggataaa 14640
ttatcgcgcg cggtgtcatc tatgttacta gatcgggaat tgg 14683

<210> 73
<211> 14590
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2233

<400> 73
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tcttctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggagggcga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

nttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgatttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc cgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atcttcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggtttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020

actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080

tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140

acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200

acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260

cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320

tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380

cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440

tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500

cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560

tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620

tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680

gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740

cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800

tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860

gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920

attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980

tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040

cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100

gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5150

catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220

gaccctaccaa ggcaacgcta tgttctcttg ctttgtcagc caagatagcc agatcaatgt 5280

cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340

gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400

cttctacagc gcggagaatc tcgctctctc cagggggaagc cgaagtttcc aaaaggtcgt 5460

tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520

tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggctcta 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttta ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500

agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560

gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620

attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680

agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740

atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800

ccccgtcgtgt agataactac gatacgggag ggcttaccat ctggcccccag tgctgcaatg 7860

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040

gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100

gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160

gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220

aattttcata aatagcgaaa accgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280

aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340

caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400

taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460

gtcccccccc cccccccct gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580

aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640

cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700

tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760

gttgctcttg cccggcgtca acacgggata ataccgcgcc acatagcaga actttaaaag 8820

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940

ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000

cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 9060

agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120

gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180

tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240

acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300

gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360

caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg cccttttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcacccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcacccgatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggagggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatc ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 11220
tgtttcgata tattacccta atgatgtgtc agatggggat cgtttcgata tattacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400
ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640
gaactgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700
aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000
tcttgcctct tccagatcca gataatgcaa acaagcatta ctctgacatc caacgcatcc 12060
ggtacgtttg acggttacaa ttacgaactc tggaaggata ctggcaatac aacaatgacg 12120
gtctacactc aaggtcgctt ttcctgccag tggtcgaaca tcaataacgc gttgtttagg 12180
accgggaaga aatacaacca gaattggcag tctctcggca caatccggat cacgtactct 12240
gcgacttaca acccaaacgg gaactcctac atgtgtatct atggctggtc taccaaccca 12300
ttggtcgagt tctacatcgt tgagtcctgg gggaactgga gaccgcctgg tgccacgtcc 12360
ctgggccaag tgacaatcga tggcgggacc tacgacatct ataggacgac acgcgtcaac 12420
cagccttgcc tggccgaggg ctcgctcgtc ttggacgcgg ctaccgggca gagggtcccc 12480
atcgaaaagg tgcgtccggg gatggaagtt ttctccttgg gacctgatta cagactgtat 12540
cgggtgcccg ttttggaggt ccttgagagc ggggttgggg aagttgtgcg cctccgaact 12600
cggtcaggga gaacgctggt gttgacacca gaacacccgc ttttgacccc cgaaggttgg 12660
aaacctcttt gtgacctccc gcttggaact ccaattgcag tcccctgcaga actgcctgtg 12720
gcgggccact tggccccacc tgaagaacgt gttatgctcc tggctcttct gttgggggat 12780
gggaacacaa agctgtcggg tcggagaggt acacgtccta atgccttctt ctacagcaaa 12840
gactccgaat tgctcgcggc ttatcgccgg tgtgcagaag ccttgggtgc aaaggtgaaa 12900
gcatacgtcc acccgactac gggggtggtt acactcgcaa ccctcgctcc acgtcctgga 12960
gctcaagatc ctgtcaaacg cctcgttgtc gaggcgggaa tggttgctaa agtcgaagag 13020
aagagggtcc cggaggaggt gtttcgttac cggcgtgagg cgttggccct tttcttgggc 13080
cgtttgttct cgacagaacgg ctctgttgaa aagaagagga tctcttattc aagtgccagt 13140
ttgggactgg cccaggatgt cgcacatctc ttgctgcgcc ttggaattac atctcaactc 13200

cgttcgagag ggccacgggc tcacaaggtt cttatatcgg gccgcgagga tattttgcgg 13260

tttgctgaac ttatcggacc ctacctcttg ggggccaaga gggagagact tgcagcgctg 13320

gaagctgagg ccccgcaggcg tttgcctgga cagggatggc acttgcggcc tgttcttcct 13380

gccgtggcgt acagagtgag cgaggctaaa aggcgctcgg gattttcgtg gagtgaagcc 13440

ggtcggcgcg tcgcagttgc gggatcgtgt ttgtcatctg gactcaacct caaattgccc 13500

agacgctacc tttctcggca ccggttgtcg ctgctcggtg aggcttttgc cgaccctggg 13560

ctggaagcgc tcgcggaagg ccaagtgctc tgggacccta ttgttgctgt cgaaccggcc 13620

ggtaaggcga gaacattcga cttgcgcgtc ccaccctttg caaacttcgt gagcgaggac 13680

ctggtggtgc ataactccat tgtggggaca gccacgttcg atcagtactg gagcgtgcgc 13740

acctctaagc ggacttcagg aacagtgacc gtgaccgatc acttccgcgc ctgggcgaac 13800

cggggcctga acctcggcac aatagaccaa attacattgt gcgtggaggg ttaccaaagc 13860

tctggatcag ccaacatcac ccagaacacc ttctctcagg gctcttcttc cggcagttcg 13920

ggtggctcat ccggctccac aacgactact cgcatcgagt gtgagaacat gtccttgtcc 13980

ggaccctacg ttagcaggat caccaatccc tttaatggta ttgcgctgta cgccaacgga 14040

gacacagccc gcgctaccgt taacttcccc gcaagtcgca actacaattt ccgcctgcgg 14100

ggttgcggca acagcaataa tcttgcccgt gtggacctga ggatcgacgg acggaccgtc 14160

gggacctttt attaccaggg cacataccgc tgggaggccc caattgacaa tgtttatgtc 14220

agtgcgggga gtcatacagt cgaaatcact gttactgcgg ataacggcac acgggacgtg 14280

tatgccgact acctggtgat acagtgacct aggtcccga atttccccga tcgttcaaac 14340

atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata 14400

taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 14460

atgagatggg tttttatgat tagagtcccg caattataca tttaatacgc gatagaaaac 14520

aaaatatagc gcgcaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat 14580

cgggaattgg 14590

<210> 74
<211> 14662
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2234

<400> 74
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaacccа tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg cctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgactttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgatttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatatttttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgcggcaaac gattcgttta 2760

atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg accttccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggtttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020

actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccgttttggg gtgtagaaca 4440
tccttttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgcccgga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940

gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg tagggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
ccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040

gctgcagggg gggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100

gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160

gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220

aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280

aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340

caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400

taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460

gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580

aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640

cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700

tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760

gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940

ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000

cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060

agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120

gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180

tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240

acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300

gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360

caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420

atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480

agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540

cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600

aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660

ccttttcacg cccttttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720

acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcgggа aacgacaacc 9780

tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agctctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaagggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt ccatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggtttcgaga 11160
ttgttctgtc agatggggat tgtttcgata tattaccctа atgatgtgtc agatggggat 11220
tgtttcgata tattaccctа atgatgtgtc agatggggat tgtttcgata tattaccctа 11280
atgatggata acaagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tcrgatattg 11400
ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtc 11640
gaactgcagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ctactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000
tcttgcctct tccagatcca gataatggcg aacaaacatt tgtccctctc cctcttcctc 12060
gtcctccttg gcctgtcggc cagcttggcc tccgggcaac aaacaagcat tactctgaca 12120
tccaacgcat ccggtacgtt tgacggttac aattacgaac tctggaagga tactggcaat 12180
acaacaatga cggtctacac tcaaggtcgc ttttcctgcc agtggtcgaa catcaataac 12240
gcgttgttta ggaccgggaa gaaatacaac cagaattggc agtctctcgg cacaatccgg 12300
atcacgtact ctgcgacttа caacccaaac gggaactcct acatgtgtat ctatggctgg 12360
tctaccaacc cattggtcga gttctacatc gttgagtcct gggggaactg gagaccgcct 12420
ggtgccacgt ccctgggcca agtgacaatc gatggcggga cctacgacat ctataggacg 12480
acacgcgtca accagccttg cctggccgag ggctcgctcg tcttggacgc ggctaccggg 12540
cagagggtcc ctatcgaaaa ggtgcgtccg gggatggaag ttttctcctt gggacctgat 12600
tacagactgt atcgggtgcc cgttttggag gtccttgaga gcggggttgg ggaagttgtg 12660
cgcctccgaa ctcggtcagg gagaacgctg gtgttgacac cagaacaccc gcttttgacc 12720
cccgaaggtt ggaaacctct ttgtgacctc ccgcttggaa ctccaattgc agtccccgca 12780
gaactgcctg tggcgggcca cttggcccca cctgaagaac gtgttatgct cctggctctt 12840
ctgttggggg atgggaacac aaagctgtcg ggtcggagag gtacacgtcc taatgccttc 12900
ttctacagca aagactccga attgctcgcg gcttatcgcc ggtgtgcaga agccttgggt 12960
gcaaaggtga aagcatacgt ccacccgact acgggggtgg ttacactcgc aaccctcgct 13020
ccacgtcctg gagctcaaga tcctgtcaaa cgcctcgttg tcgaggcggg aatggttgct 13080
aaagtcgaag agaagagggt cccggaggag gtgtttcgtt acccgcgtga ggcgttggcc 13140
cttttcttgg gccgtttgtt ctcgacagac ggctctgttg aaaagaagag gatctcttac 13200
tcaagtgcca gtttgggact ggccccaggat gtcgcacatc tcttgctgcg ccttggaatt 13260
acatctcaac tccgttcgag agggccacgg gctcacaagg ttcttatatc gggccgcgag 13320
gatattttgc ggtttgctga acttatcgga cccтaccтct tgggggccaa gagggagaga 13380
cttgcagcgc tggaagctga ggcccgcagg cgtttgcctg gacagggatg gcacttgcgg 13440
cctgttcttc ctgccgtggc gtacagagtg agcgaggcta aaaggcgctc gggattttcg 13500
tggagtgaag ccggtcggcg cgtcgcagtt gcgggatcgt gtttgtcatc tggactcaac 13560
ctcaaattgc ccagacgcta cctttctcgg caccggttgt cgctgctcgg tgaggctttt 13620

gccgaccctg ggctggaagc gctcgcggaa ggccaagtgc tctgggaccc tattgttgct 13680

gtcgaaccgg ccggtaaggc gagaacattc gacttgcgcg tcccaccctt tgcaaacttc 13740

gtgagcgagg acctggtggt gcataactcc attgtgggga cagccacgtt cgatcagtac 13800

tggagcgtgc gcacctctaa gcggacttca ggaacagtga ccgtgaccga tcactttccgc 13860

gcctgggcga accggggcct gaacctcggc acaatagacc aaattacatt gtgcgtggag 13920

ggttaccaaa gctctggatc agccaacatc acccagaaca ccttctctca gggctcttct 13980

tccggcagtt cgggtggctc atccggctcc acaacgacta ctcgcatcga gtgtgagaac 14040

atgtccttgt ccggacccta cgttagcagg atcaccaatc cctttaatgg tattgcgctg 14100

tacgccaacg gagacacagc ccgcgctacc gttaacttcc ccgcaagtcg caactacaat 14160

ttccgcttgc ggggttgcgg caacagcaat aatcttgccc gtgtggacct gaggatcgac 14220

ggacggaccg tcgggacctt ttattaccag ggcacatacc cctgggaggc cccaattgac 14280

aatgtttatg tcagtgcggg gagtcataca gtcgaaatca ctgttactgc ggataacggc 14340

acacgggacg tgtatgccga ctacctggtg atacagtgac ctaggtcccc gaatttcccc 14400

gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg 14460

atgattatca tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc 14520

atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 14580

gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct 14640

atgttactag atcgggaatt gg 14662

<210> 75
<211> 17309
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2242

<400> 75
cgtcgacttct agtaacggcc gccagtgtgc tggaattaat tcggcttgtc gaccacccaa 60

ccccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catcttctgat 120

agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca 180

ggaaggggcc gatctgacga ggcccccgcca ccggcctcga cccgaggccg aggccgacga 240

agcgccggcg agtacggcgc cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga 300

gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcgggcg 360

tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 420

gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 480

aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 540

tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 600
ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgactttg tcaaagagag 660
gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 720
gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 780
tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 840
ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 900
ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 960
tctctcgtgg tgcccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 1020
atcccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg tgatccacct 1080
cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 1140
agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta atagatggat 1200
tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 1260
cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 1320
atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 1380
agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 1440
ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 1500
gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag ttttttcccc 1560
ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt atcgtgtaca 1620
cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 1680
tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt tttttacccg ttttatttgc 1740
tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttct 1800
tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 1860
ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 1920
tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga gctgacagta 1980
ccgatatcgt aggatctggt gccaacttat tctccagctg cttttttta cctatgttaa 2040
ttccaatcct ttcttgcctc ttccagatgc agatctttgt gaagacccctg accggcaaga 2100
ctatcaccct cgaggtggag tcttctgaca ccattgacaa tgtcaaggcc aagatccagg 2160
acaaggaggg catccccccca gaccagcagc ggctcatctt tgctggcaag cagctggagg 2220
acgggcgcac gcttgcggac tacaacatcc agaaggagag cacccttcac ctggtgctcc 2280
gcctcagggg aggcatgcaa acaagcatta ctctgacatc caacgcatcc ggtacgtttg 2340
acggttacta ttacgaactc tggaaggata ctggcaatac aacaatgacg gtctacactc 2400
aaggtcgctt ttcctgccag tggtcgaaca tcaataacgc gttgtttagg accgggaaga 2460

aatacaacca gaattggcag tctcttggca caatccggat cacgtactct gcgacttaca 2520
acccaaacgg gaactcctac ttgtgtatct atggctggtc taccaacccа ttggtcgagt 2580
tctacatcgt tgagtcctgg gggaactgga gaccgcctgg tgccacgtcc ctgggccaag 2640
tgacaatcga tggcgggacc tacgacatct ataggacgac acgcgtcaac cagccttcca 2700
ttgtggggac agccacgttc gatcagtact ggagcgtgcg cacctctaag cggacttcag 2760
gaacagtgac cgtgacgcgat cacttccgcg cctgggcgaa ccggggcctg aacctggcа 2820
caatagacca aattacattg tgcgtggagg gttaccaaag ctctggatca gccaacatca 2880
cccagaacac cttctctcag ggctcttctt ccggcagttc gggtggctca tccggctcca 2940
caacgactac tcgcatcgag tgtgagaaca tgtccttgtc cggaccctac gttagcagga 3000
tcaccaatcc ctttaatggt attgcgctgt acgccaacgg agacacagcc cgcgctaccg 3060
ttaacttccc cgcaagtcgc aactacaatt tccgcctgcg gggttgcggc aacaacaata 3120
atcttgcccg tgtggacctg aggatcgacg gacggaccgt cgggaccttt tattaccagg 3180
gcacataccc ctgggaggcc ccaattgaca atgtttatgt cagtgcgggg agtcatacag 3240
tcgaaatcac tgttactgcg gataacggca catgggacgt gtatgccgac tacctggtga 3300
tacagtgacc taggtccccg aatttccccg atcgttcaaa catttggcaa taaagtttct 3360
taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg ttgaattacg 3420
ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga 3480
ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact 3540
aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaattg gaattcatac 3600
taaagcttgc atgcctgcag gtcgactcta gtaacggccg ccagtgtgct ggaattaatt 3660
cggcttgtcg accacccaac cccatatcga cagaggatgt gaagaacagg taaatcacgc 3720
agaagaaccc atctctgata gcagctatcg attagaacaa cgaatccata ttgggtccgt 3780
gggaaatact tactgcacag gaaggggggcg atctgacgag gccccgccac cggcctcgac 3840
ccgaggccga ggccgacgaa gcgccggcga gtacggcgcc gcggcggcct ctgcccgtgc 3900
cctctgcgcg tgggagggag aggccgcggt ggtgggggcg cgcgcgcgcg cgcgcgcagc 3960
tggtgcggcg gcgcgggggt cagccgccga gccggcggcg acggaggagc agggcggcgt 4020
ggacgcgaac ttccgatcgg ttggtcagag tgcgcgagtt gggcttagcc aattaggtct 4080
caacaatcta ttgggccgta aaattcatgg gccctggttt gtctaggccc aatatcccgt 4140
tcatttcagc ccacaaatat ttccccagag gattattaag gcccacacgc agcttatagc 4200
agatcaagta cgatgtttcc tgatcgttgg atcggaaacg tacggtcttg atcaggcatg 4260
ccgacttcgt caaagagagg cggcatgacc tgacgcggag ttggttccgg gcaccgtctg 4320
gatggtcgta ccgggaccgg acacgtgtcg cgcctccaac tacatggaca cgtgtggtgc 4380

tgccattggg ccgtacgcgt ggcggtgacc gcaccggatg ctgcctcgca ccgccttgcc 4440
cacgctttat atagagaggt tttctctcca ttaatcgcat agcgagtcga atcgaccgaa 4500
ggggaggggg agcgaagctt tgcgttctct aatcgcctcg tcaaggtaac taatcaatca 4560
cctcgtccta atcctcgaat ctctcgtggt gcccgtctaa tctcgcgatt ttgatgctcg 4620
tggtggaaag cgtaggagga tcccgtgcga gttagtctca atctctcagg gtttcgtgcg 4680
attttagggt gatccacctc ttaatcgagt tacggtttcg tgcgatttta gggtaatcct 4740
cttaatctct cattgattta gggtttcgtg agaatcgagg tagggatctg tgttatttat 4800
atcgatctaa tagatggatt ggttttgaga ttgttctgtc agatggggat tgtttcgata 4860
tattacccta atgatgtgtc agatggggat tgtttcgata tattacccta atgatgtgtc 4920
agatggggat tgtttcgata tattacccta atgatggata ataagagtag ttcacagtta 4980
tgttttgatc ctgccacata gtttgagttt tgtgatcaga tttagtttta cttatttgtg 5040
cttagttcgg atgggattgt tctgatattg ttccaataga tgaatagctc gttaggttaa 5100
aatctttagg ttgagttagg cgacacatag tttatttcct ctggatttgg attggaattg 5160
tgttcttagt ttttttcccc tggatttgga ttggaattgt gtggagctgg gttagagaat 5220
tacatctgta tcgtgtacac ctacttgaac tgtagagctt gggttctaag gtcaatttaa 5280
tctgtattgt atctggctct ttgcctagtt gaactgtagt gctgatgttg tactgtgttt 5340
ttttacccgt tttatttgct ttactcgtgc aaatcaaatc tgtcagatgc tagaactagg 5400
tggctttatt ctgtgttctt acatagatct gttgtcctgt agttacttat gtcagttttg 5460
ttattatctg aagatatttt tggttgttgc ttgttgatgt ggtgtgagct gtgagcagcg 5520
ctcttatgat taatgatgct gtccaattgt agtgtagtat gatgtgattg atatgttcat 5580
ctattttgag ctgacagtac cgatatcgta ggatctggtg ccaacttatt ctccagctgc 5640
ttttttttac ctatgttaat tccaatcctt tcttgcctct tccagatcca gataatgcag 5700
aaactcatta actcagtgca aaactatgcc tggggcagca aaacggcgtt gactgaactt 5760
tatggtatgg aaaatccgtc cagccagccg atggccgagc tgtggatggg cgcacatccg 5820
aaaagcagtt cacgagtgca gaatgccgcc ggagatatcg tttcactgcg tgatgtgatt 5880
gagagtgata aatcgactct gctcggagag gccgttgcca aacgctttgg cgaactgcct 5940
ttcctgttca aagtattatg cgcagcacag ccactctcca ttcaggttca tccaaacaaa 6000
cacaattctg aaatcggttt tgccaaagaa aatgccgcag gtatcccgat ggatgccgcc 6060
gagcgtaact ataaagatcc taaccacaag ccggagctgg tttttgcgct gacgcctttc 6120
cttgcgatga acgcgtttcg tgaattttcc gagattgtct ccctactcca gccggtcgca 6180
ggtgcacatc cggcgattgc tcactttttа caacagcctg atgccgaacg tttaagcgaa 6240
ctgttcgcca gcctgttgaa tatgcagggt gaagaaaaat cccgcgcgct ggcgatttta 6300

aaatcggccc tcgatagcca gcagggtgaa ccgtggcaaa cgattcgttt aatttctgaa 6360
ttttacccgg aagacagcgg tctgttctcc ccgctattgc tgaatgtggt gaaattgaac 6420
cctggcgaag cgatgttcct gttcgctgaa acaccgcacg cttacctgca aggcgtgggg 6480
ctggaagtga tggcaaactc cgataacgtg ctgcgtgcgg gtctgacgcc taaatacatt 6540
gatattccgg aactggttgc caatgtgaaa ttcgaagcca aaccggctaa ccagttgttg 6600
acccagccgg tgaaacaagg tgcagaactg gacttcccga ttccagtgga tgattttgcc 6660
ttctcgctgc atgaccttag tgataaagaa accaccatta gccagcagag tgccgccatt 6720
ttgttctgcg tcgaaggcga tgcaacgttg tggaaaggtt ctcagcagtt acagcttaaa 6780
ccgggtgaat cagcgtttat tgccgccaac gaatcaccgg tgactgtcaa aggccacggc 6840
cgtttagcgc gtgtttacaa caagctgtaa gagcttactg aaaaaattaa catcttcttgc 6900
taagctggga gctctagatc cccgaatttc ccccgatcgtt caaacatttg gcaataaagt 6960
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 7020
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggtttt 7080
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 7140
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcggga attggcgagc 7200
tcgaattaat tcagtacatt aaaaacgtcc gcaatgtgtt attaagttgt ctaagcgtca 7260
atttgtttac accacaatat atcctgccac cagccagcca acagctcccc gaccggcagc 7320
tcggcacaaa atcaccactc gatacaggca gcccatcagt ccgggacggc gtcagcgggg 7380
gagccgttgt aaggcggcag actttgctca tgttaccgat gctattcgga agaacggcaa 7440
ctaagctgcc gggtttgaaa cacggatgat ctcgcggagg gtagcatgtt gattgtaacg 7500
atgacagagc gttgctgcct gtgatcaaat atcatctccc tcgcagagat ccgaattatc 7560
agccttctta ttcatttctc gcttaaccgt gacaggctgt cgatcttgag aactatgccg 7620
acataatagg aaatcgctgg ataaagccgc tgaggaagct gagtggcgct atttctttag 7680
aagtgaacgt tgacgatcgt cgaccgtacc ccgatgaatt aattcggacg tacgttctga 7740
acacagctgg atacttactt gggcgattgt catacatgac atcaacaatg tacccgtttg 7800
tgtaaccgtc tcttggaggt tcgtatgaca ctagtggttc ccctcagctt gcgactagat 7860
gttgaggcct aacattttat tagagagcag gctagttgct tagatacatg atcttcaggc 7920
cgttatctgt cagggcaagc gaaaattggc catttatgac gaccaatgcc ccgcagaagc 7980
tcccatcttt gccgccatag acgccgcgcc cccttttgg ggtgtagaac atcctttgc 8040
cagatgtgga aaagaagttc gttgtcccat tgttggcaat gacgtagtag ccggcgaaag 8100
tgcgagaccc atttgcgcta tatataagcc tacgatttcc gttgcgacta ttgtcgtaat 8160
tggatgaact attatcgtag ttgctctcag agttgtcgta atttgatgga ctattgtcgc 8220

aattgcttat ggagttgtcg tagttgcttg gagaaacgtc gtagttggat ggggagtagt 8280
catagggaag acgagcttca tccactaaaa caattggcag gtcagcaagt gcctgccccg 8340
atgccatcgc aagtacgagg cttagaacca ccttcaacag atcgcgcata gtcttcccca 8400
gctctctaac gcttgagtta agccgcgccg cgaagcggcg tcggcttgaa cgaattgtta 8460
gacattattt gccgactacc ttggtgatct cgcctttcac gtagtgaaca aattcttcca 8520
actgatctgc gcgcgaggcc aagcgatctt cttgtccaag ataagcctgc ctagcttcaa 8580
gtatgacggg ctgatactgg gccggcaggc gctccattgc ccagtcggca gcgacatcct 8640
tcggcgcgat tttgccggtt actgcgctgt accaaatgcg ggacaacgta agcactacat 8700
ttcgctcatc gccagcccag tcgggcggcg agttccatag cgttaaggtt tcatttagcg 8760
cctcaaatag atcctgttca ggaaccggat caaagagttc ctccgccgct ggacctacca 8820
aggcaacgct atgttctctt gcttttgtca gcaagatagc cagatcaatg tcgatcgtgg 8880
ctggctcgaa gatacctgca agaatgtcat tgcgctgcca ttctccaaat tgcagttcgc 8940
gcttagctgg ataacgccac ggaatgatgt cgtcgtgcac aacaatggtg acttctacag 9000
cgcggagaat ctcgctctct ccaggggaag ccgaagtttc caaaaggtcg ttgatcaaag 9060
ctcgccgcgt tgtttcatca agccttacgg tcaccgtaac cagcaaatca atatcactgt 9120
gtggcttcag gccgccatcc actgcggagc cgtacaaatg tacggccagc aacgtcggtt 9180
cgagatggcg ctcgatgacg ccaactacct ctgatagttg agtcgatact tcggcgatca 9240
ccgcttccct catgatgttt aactcctgaa ttaagccgcg ccgcgaagcg gtgtcggctt 9300
gaatgaattg ttaggcgtca tcctgtgctc ccgagaacca gtaccagtac atcgctgttt 9360
cgttcgagac ttgaggtcta gttttatacg tgaacaggtc aatgccgccg agagtaaagc 9420
cacattttgc gtacaaattg caggcaggta cattgttcgt ttgtgtctct aatcgtatgc 9480
caaggagctg tctgcttagt gcccactttt tcgcaaattc gatgagactg tgcgcgactc 9540
ctttgcctcg gtgcgtgtgc gacacaacaa tgtgttcgat agaggctaga tcgttccatg 9600
ttgagttgag ttcaatcttc ccgacaagct cttggtcgat gaatgcgcca tagcaagcag 9660
agtcttcatc agagtcatca tccgagatgt aatccttccg gtaggggctc acacttctgg 9720
tagatagttc aaagccttgg tcggataggt gcacatcgaa cacttcacga acaatgaaat 9780
ggttctcagc atccaatgtt tccgccacct gctcagggat caccgaaatc ttcatatgac 9840
gcctaacgcc tggcacagcg gatcgcaaac ctggcgcggc ttttggcaca aaaggcgtga 9900
caggtttgcg aatccgttgc tgccacttgt taaccctttt gccagatttg gtaactataa 9960
tttatgttag aggcgaagtc ttgggtaaaa actggcctaa aattgctggg gatttcagga 10020
aagtaaacat cacgttcggg ctcgatgtct attgtagata tatgtagtgt atctacttga 10080
tcgggggatc tgctgcctcg cgcgtttcgg tgatgacggt gaaaacctct gacacatgca 10140

gctcccggag acggtcacag cttgtctgta agcggatgcc gggagcagac aagcccgtca 10200

gggcgcgtca gcgggtgttg gcgggtgtcg gggcgcagcc atgacccagt cacgtagcga 10260

tagcggagtg tatactggct taactatgcg gcatcagagc agattgtact gagagtgcac 10320

catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgctct 10380

tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca 10440

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 10500

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 10560

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 10620

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 10680

tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 10740

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 10800

aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac 10860

tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 10920

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 10980

aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 11040

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 11100

tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 11160

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 11220

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 11280

tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 11340

gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 11400

tagataacta cgatacggga gggcttacca tctggccccа gtgctgcaat gataccgcga 11460

gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 11520

cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 11580

gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctgcaggg 11640

gggggggggg ggggcttcca ttgttcattc cacggacaaa aacagagaaa ggaaacgaca 11700

gaggccaaaa agctcgcttt cagcacctgt cgtttccttt cttttcagag ggtattttaa 11760

ataaaaacat taagttatga cgaagaagaa cggaaacgcc ttaaaccgga aaatttttcat 11820

aaatagcgaa aacccgcgag gtcgccgccc cgtaacctgt cggatcaccg gaaaggaccc 11880

gtaaagtgat aatgattatc atctacatat cacaacgtgc gtggaggcca tcaaaccacg 11940

tcaaataatc aattatgacg caggtatcgt attaattgat ctgcatcaac ttaacgtaaa 12000

aacaacttca gacaatacaa atcagcgaca ctgaatacgg ggcaacctca tgtccccccc 12060

ccccccccc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct 12120
ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa aaagcggtta 12180
gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg 12240
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 12300
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 12360
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 12420
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 12480
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 12540
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 12600
aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt 12660
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 12720
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 12780
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattggtc gacgatcttg 12840
ctgcgttcgg atattttcgt ggagttcccg ccacagaccc ggattgaagg cgagatccag 12900
caactcgcgc cagatcatcc tgtgacggaa ctttggcgcg tgatgactgg ccaggacgtc 12960
ggccgaaaga gcgacaagca gatcacgctt ttcgacagcg tcggatttgc gatcgaggat 13020
ttttcggcgc tgcgctacgt ccgcgaccgc gttgagggat caagccacag cagcccactc 13080
gaccttctag ccgacccaga cgagccaagg gatcttttg gaatgctgct ccgtcgtcag 13140
gctttccgac gtttgggtgg ttgaacagaa gtcattatcg cacggaatgc caagcactcc 13200
cgaggggaac cctgtggttg gcatgcacat acaaatggac gaacggataa agcttttcac 13260
gccctttaa atatccgatt attctaataa acgctctttt ctcttaggtt tacccgccaa 13320
tatatcctgt caaacactga tagtttaaac tgaaggcggg aaacgacaac ctgatcatga 13380
gcggagaatt aagggagtca cgttatgacc cccgccgatg acgcgggaca agccgtttta 13440
cgcttggaac tgacagaacc gcaacgttga aggagccact cagcttaatt aagtctaact 13500
cgagttactg gtacgtacca aatccatgga atcaaggtac cgtcgactct agtaacggcc 13560
gccagtgtgc tggaattaat tcgggttgtc gacccacccaa ccccatatcg acagaggatg 13620
tgaagaacag gtaaatcacg cagaagaacc catctctgat agcagctatc gattagaaca 13680
acgaatccat attgggtccg tgggaaatac ttactgcaca ggaaggggcc gatctgacga 13740
ggccccgcca ccggcctcga cccgaggccg aggccgacga agcgccggcg agtacggcgc 13800
cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga gaggccgcgg tggtgggggc 13860
gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg tcagccgccg agccggcggc 13920
gacggaggag cagggcggcg tggacgcgaa cttccgatcg gttggtcaga gtgcgcgagt 13980

tgggcttagc caattaggtc tcaacaatct attgggccgt aaaattcatg ggcccctggtt 14040

tgtctaggcc caatatcccg ttcatttcag cccacaaata tttccccaga ggattattaa 14100

ggcccacacg cagcttatag cagatcaagt acgatgtttc ctgatcgttg gatcggaaac 14160

gtacggtctt gatcaggcat gccgactteg tcaaagagag gcggcatgac ctgacgcgga 14220

gttggttccg ggcacegtct ggatggtcgt accgggaccg gacacgtgtc gcgcctccaa 14280

ctacatggac acgtgtggtg ctgccattgg gccgtacgcg tggcggtgac cgcaccggat 14340

gctgcctcgc accgccttgc ccacgcttta tatagagagg ttttctctcc attaatcgca 14400

tagcgagtcg aatcgaccga aggggagggg gagcgaagct ttgcgttccc taatcgcctc 14460

gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa tctctcgtgg tgcccgtcta 14520

atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg atcccgtgcg agttagtctc 14580

aatctctcag ggtttcgtgc gattttaggg tgatccacct cttaatcgag ttacggtttc 14640

gtgcgatttt agggtaatcc tcttaatctc tcattgattt agggtttcgt gagaatcgag 14700

gtagggatct gtgttattta tatcgatcta atagatggat tggttttgag attgttctgt 14760

cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 14820

atattaccct aatgatgtgt cagatgggga ttgtttcgat atattaccct aatgatggat 14880

aataagagta gttcacagtt atgttttgat cctgccacat agtttgagtt tcgtgatcag 14940

atttagtttt acttatttgt gcttagttcg gatgggattg ttctgatatt gttccaatag 15000

acgaatagct cgttaggtta aaatctttag gttgagttag gcgacacata gtttatttcc 15060

tctggatttg gattggaatt gtgttcttag ttttttccc ctggatttgg attggaattg 15120

tgtggagctg ggttagagaa ttacatctgt atcgtgtaca cctacttgaa ctgtagagct 15180

tgggttctaa ggtcaattta atctgtattg tatctggctc tttgcctagt tgaactgtag 15240

tgctgatgtt gtactgtgtt ttttaccccg ttttatttgc tttactcgtg caaatcaaat 15300

ctgtcagatg ctagaactag gtggctttat tctgtgttct tacatagatc tgttgtcctg 15360

tagttactta tgtcagtttt gttattatct gaagatattt ttggttgttg cttgttgatg 15420

tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc tgtccaattg tagtgtagta 15480

tgatgtgatt gatatgttca tctatttttga gctgacagta ccgatatcgt aggatctggt 15540

gccaacttat tctccagctg cttttttta cctatgttaa ttccaatcct ttcttgcctc 15600

ttccagatcc agataatggg cttcgtgctc ttctcccagc tgccttcctt ccttcttgtc 15660

tccaccctgc tcttgttcct cgtgatctcc cactcctgcc gcgccgctta cgactacaag 15720

caggtgttgc gggactcgct actattctat gaggcccaga gatccggccg gctcccagcc 15780

gaccagaagg tcacgtggag gaaggatagc gcgctgaatg accagggtga ccagggacaa 15840

gacttgaccg gcggctactt tgacgctggg gacttcgtca agttcgggtt ccccatggct 15900

tataccgcaa ccgtgctggc atggggcctc atagattttg aggccggcta cagcagtgcc 15960

ggggccttgg atgatggacg gaaggctgtc aaatgggcca ccgactattt cataaaggcc 16020

cacacaagtc aaaatgagtt ctatggtcag gtcggccagg gtgacgccga tcacgctttc 16080

tggggaagac cagaggatat gacgatggcg cgcccggcgt acaagataga cacctcaagg 16140

cctggctctg atctggcagg cgagacagcg gctgctcttg ccgctgcttc aatcgtgttc 16200

cggaacgtcg atggcactta ctcaaataac ctgttaacac acgctcgcca gctattcgac 16260

ttcgcgaaca actaccgggg aaagtatagt gactctatta ctgacgcaag aaatttctac 16320

gcaagcgcag actacagaga cgagttggtt tgggctgctg cgtggttata cagagcgacc 16380

aacgacaaca cctacctcaa cactgctgag tcactgtacg atgagtttgg gctacagaac 16440

tgggggggag gcctgaactg ggatagcaag gtgtctggcg tgcaggtgtt gttggccaag 16500

cttaccaata agcaggccta caaggacacg gtgcagtctt acgtcaatta cctaattaat 16560

aaccagcaga agactcccaa gggcctcctc tacatcgaca tgtggggcac ccttcgccac 16620

gctgccaacg ccgcattcat catgctcgaa gccgccgagc tgggcttgtc cgcctcctct 16680

tatagacagt tcgcgcaaac gcaaatcgac tacgccctgg gcgatggtgg ccgctccttt 16740

gtgtgcgggt tcgggagtaa tcctcctacg agaccgcacc acagatcctc gtcgtgcccg 16800

ccagctcccg ctacttgcga ctggaataca ttcaactcac ctgacccaaa ctaccacgtc 16860

ctctctgggg ccctagtggg cggacctgat cagaatgaca actacgtcga tgaccgttca 16920

gactatgttc acaacgaagt cgccactgat tacaacgcgg gtttccagtc cgcgttagct 16980

gctttggtgg ccttggttta cagcgagaag gacgagctgt gacctaggtc ccgaatttcc 17040

cccgatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 17100

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 17160

tgcatgacgt tatttatgag atgggtttt atgattagag tcccgcaatt atacatttaa 17220

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 17280

tctatgttac tagatcgggg attgggtac 17309

<210> 76
<211> 13042
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2252

<400> 76
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttttcccc ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatatttct ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

atcaatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagctatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttaccccgga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg actttccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggtttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020

```
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta   4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt   4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt   4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg   4260
cgactagatg ttgaggccta acatttttatt agagagcagg ctagttgctt agatacatga   4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc   4380
cgcagaagct cccatcttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca   4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc   4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat   4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac   4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg   4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg   4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag   4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac   4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa   4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc   4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag   5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa   5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt   5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg   5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt   5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt   5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga   5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt   5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa   5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca   5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt   5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg   5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca   5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga   5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta   5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt   5940
```

gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaaccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt ttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
ccccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg gggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9340
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg cccttttaaa taccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780

tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accaccccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcgg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440
gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500
gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560
atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620
tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680
cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740
gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800
ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg agcgaagctt tgcgttctct 10860
aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920
gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980
gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040
tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100
agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160
ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220
tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280
atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340
tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400
ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460
tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520
ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580
tgtagagctt gggttctaag gtcaatttaa tcrgtattgt atctggccct ttgcctagtt 11640
gaactgtagt gctgatgttg tactgtgttt ttttacccgt trtatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatgtgc gattggctgt tccctgatgg agataatgga 12060

aaggagcccg aaccagaacc agaaccaacg gtagagctct gcggaagatg ggatgcaaga 12120

gacgtagcgg gaggaagata tagagttatt aacaatgtgt gggggccga aacagcacag 12180

tgtatagaag taggtcttga aacaggtaat tttacgataa caagagccga tcatgacaat 12240

ggtaataatg ttgcagcata tccagcaatt tacttcggat gtcattgggc accagcgaga 12300

gcaataaggg attgtgccgc gcgtgcggga gcggttagga gagcacacga attggatgtt 12360

acaccaatta ccacgggaag atggaacgca gcttacgata tatggtttag tccggtaaca 12420

aactctggta acgggtactc gggaggcgcc gaacttatga tatggctgaa ttggaatggt 12480

ggcgtaatgc caggaggatc acgggtagca actgtcgaat tggctggagc gacatgggaa 12540

gtgtggtatg cagattggga ttggaattac atcgcatata gaagaacgac tccgacaacc 12600

tcagtgagtg aacttgactt gaaagccttt attgatgatg cagtagcgag aggatacata 12660

aggccagaat ggtatctgca tgcagtggaa acgggatttg aattgtggga agggggggct 12720

gggttgagga cagcagattt tagcgtaact gtacagtgac ctaggtcccc gaatttcccc 12780

gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg 12840

atgattatca tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc 12900

atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 12960

gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct 13020

atgttactag atcgggaatt gg 13042

<210> 77
<211> 13120
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтезична послідовність, вектор, pAG2253

<400> 77
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggccca 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggcgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgatttcag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt tttttccccc ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220
gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280
gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340
gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400
ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460
gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520
acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctaccccag 2580
ccggtcgcag gtgcacatcc ggcgattgct cacttttac aacagcctga tgccgaacgt 2640
ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700
gcgattttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760
atttctgaat tttacccgga agacagcggt ctgttctccc cgctattgct gaatgtggcg 2820
aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa cacccacgc ttacctgcaa 2880
ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940
aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000
cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060
gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120
gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180
cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240
ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300
atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360
caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480
tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600
ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgcgga catattagga atcgcgtgga taaagccgct gaggaagctg agtggcgcta 4080

tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttttggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgcttcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctgct caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tgccgcgttt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgccgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000

cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggcg ggggttccat tgttcattcc acggacaaaa acagaggaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttcacg ccctttaaa tatccgatta tttaataaa cgatcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcgggaa aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840

gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900

agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960

gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020

cagaggatgt gaagaacagg tagatcacgc agaagaaccc atctctgata gcagctatcg 10080

attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggtcg 10140

atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200

gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260

ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcggggtt cagccgccga 10320

gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380

tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440

gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500

gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560

atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620

tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680

cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740

gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttcttccca 10800

ttaatcgcat agcgagtcga atcgaccgaa ggggagggcg agcgaagctt tgcgttcttt 10860

aatcgccttg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920

gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980

gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040

tacggtttcg tgcgatttta gggtaatcct cttaatcttt cattgattta gggtttcgtg 11100

agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160

ttgttctgtc agatggggat tgtttcgata ttatacccta atgatgtgtc agatggggat 11220

tgtttcgata ttatacccta atgatgtgtc agatggggat tgtttcgata ttatacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaacrgtagt gctgatgttg tactgtgttt ttttacccgt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

```
gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc    11820
ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt    11880
agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta    11940
ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt    12000
tcttgcctct tccagatcca gataatggcg aacaaacatt tgtccctctc cctcttcctc    12060
gtcctccttg gcctgtcggc cagcttggcc tccgggcaac aggtgtgcga ttggctgttc    12120
cctgatggag ataatggaaa ggagcccgaa ccagaaccag aaccaacggt agagctctgc    12180
ggaagatggg atgcaagaga cgtagcggga ggaagatata gagttattaa caatgtgtgg    12240
ggggccgaaa cagcacagtg tatagaagta ggtcttgaaa caggtaattt tacgataaca    12300
agagccgatc atgacaatgg taataatgtt gcagcatatc cagcaattta cttcggatgt    12360
cattgggcac cagcgagagc aataagggat tgtgccgcgc gtgcgggagc ggttaggaga    12420
gcacacgaat tggatgttac accaattacc acgggaagat ggaacgcagc ttacgatata    12480
tggtttagtc cggtaacaaa ctctggtaac gggtactcgg gaggcgccga acttatgata    12540
tggctgaatt ggaatggtgg cgtaatgcca ggaggatcac gggtagcaac tgtcgaattg    12600
gctggagcga catgggaagt gtggtatgca gattgggatt ggaattacat cgcatataga    12660
agaacgactc cgacaacctc agtgagtgaa cttgacttga aagcctttat tgatgatgca    12720
gtagcgagag gatacataag gccagaatgg tatctgcatg cagtggaaac gggatttgaa    12780
ttgtgggaag gggggctgg gttgaggaca gcagatttta gcgtaactgt acagtgacct    12840
aggtccccga atttccccga tcgttcaaac atttggcaat aaagtttctt aagattgaat    12900
cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta    12960
ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg    13020
caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta    13080
tcgcgcgcgg tgtcatctat gttactagat cgggaattgg                          13120
```

<210> 78
<211> 17111
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2309

<400> 78

```
cgtcgactct agtaacggcc gccagtgtgc tggaattaat ccggcttgtc gaccacccaa       60
ccccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catctctgat      120
agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca      180
ggaagggggc gatctgacga ggccccgcca ccggcctcga cccgaggccg aggccgacga      240
```

agcgccggcg agtacggcgc cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga 300

gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg 360

tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 420

gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 480

aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 540

tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 600

ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg tcaaagagag 660

gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 720

gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 780

tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 840

ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 900

ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 960

tctctcgtgg tgcccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 1020

atccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg tgatccacct 1080

cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 1140

agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta atagatggat 1200

tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 1260

cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 1320

atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 1380

agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 1440

ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 1500

gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag ttttttccc 1560

ctggatttgg attggaattg tgtggagctg ggttagagaa ttacacctgt atcgtgtaca 1620

cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 1680

tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt ttttttaccg ttttatttgc 1740

tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttcc 1800

tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 1860

ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 1920

tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga gctgacagta 1980

ccgatatcgt aggatctggt gccaacttat tctccagctg cttttttta cctatgttaa 2040

ttccaatcct ttcttgcctc ttccagatcc agataatgca aacaagcatt actctgacat 2100

ccaacgcatc cggtacgttt gacggttact attacgaacc ctggaaggat actggcaata 2160

caacaatgac ggtctacact caaggtcgct tttcctgcca gtggtcgaac atcaataacg 2220
cgttgtttag gaccgggaag aaatacaacc agaattggca gtctcttggc acaatccgga 2280
tcacgtactc tgcgacttac aacccaaacg ggaactccta cttgtgtatc tatggctggt 2340
ctaccaaccc attggtcgag ttctacatcg ttgagtcctg ggggaactgg agaccgcctg 2400
gtgccacgtc cctgggccaa gtgacaatcg atggcgggac ctacgacatc tataggacga 2460
cacgcgtcaa ccagccttcc attgtgggga cagccacgtt cgatcagtac tggagcgtgc 2520
gcacctctaa gcggacttca ggaacagtga ccgtgaccga tcacttccgc gcctgggcga 2580
accggggcct gaacctcggc acaatagacc aaattacatt gtgcgtggag ggttaccaaa 2640
gctctggatc agccaacatc acccagaaca ccttctctca gggctcttct tccggcagtt 2700
cgggtggctc atccggctcc acaacgacta ctcgcatcga gtgtgagaac atgtccttgt 2760
ccggacccta cgttagcagg atcaccaatc cctttaatgg tattgcgctg tacgccaacg 2820
gagacacagc ccgcgctacc gttaacttcc ccgcaagtcg caactacaat ttccgcctgc 2880
ggggttgcgg caacaacaat aatcttgccc gtgtggacct gaggatcgac ggacggaccg 2940
tcgggacctt ttattaccag ggcacatacc cctgggaggc cccaattgac aatgtttatg 3000
tcagtgcggg gagtcataca gtcgaaatca ctgttactgc ggataacggc acatgggacg 3060
tgtatgccga ctacctggtg atacagtgac ctaggtcccc gaatttcccc gatcgttcaa 3120
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 3180
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 3240
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 3300
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 3360
atcgggaatt ggaattcata ctaaagcttg catgcctgca ggtcgactct agtaacggcc 3420
gccagtgtgc tggaattaat tcggcttgtc gaccacccaa ccccatatcg acagaggatg 3480
tgaagaacag gtaaatcacg cagaagaacc catctctgat agcagctatc gattagaaca 3540
acgaatccat attgggtccg tgggaaatac ttactgcaca ggaagggggc gatctgacga 3600
ggccccgcca ccggcctcga cccgaggccg aggccgacga agcgccggcg agtacggcgc 3660
cgcggcggcc tctgccgtg ccctctgcgc gtgggaggga gaggccgcgg tggtgggggc 3720
gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg tcagccgccg agccggcggc 3780
gacggaggag cagggcggcg tggacgcgaa cttccgatcg gttggtcaga gtgcgcgagt 3840
tgggcttagc caattaggtc tcaacaatct attgggccgt aaaattcatg ggcctggttt 3900
tgtctaggcc caatatcccg ttcatttcag cccacaaata tttccccaga ggattattaa 3960
ggcccacacg cagcttatag cagatcaagt acgatgtttc ctgatcgttg gatcggaaac 4020
gtacggtctt gatcaggcat gccgacttcg tcaaagagag gcggcatgac ctgacgcgga 4080

gttggttccg ggcaccgtct ggatggtcgt accgggaccg gacacgtgtc gcgcctccaa 4140
ctacatggac acgtgtgggg ctgccattgg gccgtacgcg tggcggtgac cgcaccggat 4200
gctgcctcgc accgccttgc ccacgcttta tatagagagg ttttctctcc attaatcgca 4260
tagcgagtcg aatcgaccga aggggagggg gagcgaagct ttgcgttctc taatcgcctc 4320
gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa tctctcgtgg tgcccgtcta 4380
atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg atcccgtgcg agttagtctc 4440
aatctctcag ggtttcgtgc gattttaggg tgatccacct cttaatcgag ttacggtttc 4500
gtgcgatttt agggtaatcc tcttaatctc tcattgattt agggtttcgt gagaatcgag 4560
gtagggatct gtgttattta tatcgatcta atagatggat tggttttgag attgttctgt 4620
cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 4680
atattaccct aatgatgtgt cagatgggga ttgtttcgat atattaccct aatgatggat 4740
aataagagta gttcacagtt atgttttgat cctgccacat agtttgagtt ttgtgatcag 4800
atttagtttt acttatttgt gcttagttcg gatgggattg ttctgatatt gttccaatag 4860
atgaatagct cgttaggtta aaatctttag gttgagttag gcgacacata gtttatttcc 4920
tctggatttg gattggaatt gtgttcttag ttttttcccc ctggatttgg attggaattg 4980
tgtggagctg ggttagagaa ttacatctgt atcgtgtaca cctacttgaa ctgtagagct 5040
tgggttctaa ggtcaattta atctgtattg tatctggctc tttgcctagt tgaactgtag 5100
tgctgatgtt gtactgtgtt tttttacccg ttttatttgc tttactcgtg caaatcaaat 5160
ctgtcagatg ctagaactag gtggctttat tctgtgttct tacatagatc tgttgtcctg 5220
tagttactta tgtcagtttt gttattatct gaagatattt ttggttgttg cttgttgatg 5280
tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc tgtccaattg tagtgtagta 5340
tgatgtgatt gatatgttca tctattttga gctgacagta ccgatatcgt aggatctggt 5400
gccaacttat tctccagctg ctttttttta cctatgttaa ttccaatcct ttcttgcctc 5460
ttccagatcc agataatgca gaaactcatt aactcagtgc aaaactatgc ctggggcagc 5520
aaaacggcgt tgactgaact ttatggtatg gaaaatccgt ccagccagcc gatggccgag 5580
ctgtggatgg gcgcacatcc gaaaagcagt tcacgagtgc agaatgccgc cggagatatc 5640
gtttcactgc gtgatgtgat tgagagtgat aaatcgactc tgctcggaga ggccgttgcc 5700
aaacgctttg gcgaactgcc tttcctgttc aaagtattat gcgcagcaca gccactctcc 5760
attcaggttc atccaaacaa acacaattct gaaatcggtt ttgccaaaga aaatgccgca 5820
ggtatcccga tggatgccgc cgagcgtaac tataaagatc ctaaccacaa gccggagctg 5880
gtttttgcgc tgacgccttt ccttgcgatg aacgcgtttc gtgaattttc cgagattgtc 5940
tccctactcc agccggtcgc aggtgcacat ccggcgattg ctcacttttt acaacagcct 6000

gatgccgaac gtttaagcga actgttcgcc agcctgttga atatgcaggg tgaagaaaaa 6060
tcccgcgcgc tggcgatttt aaaatcggcc ctcgatagcc agcagggtga accgtggcaa 6120
acgattcgtt taatttctga attttacccg gaagacagcg gtctgttctc cccgctattg 6180
ctgaatgtgg tgaaattgaa ccctggcgaa gcgatgttcc tgttcgctga aacaccgcac 6240
gcttacctgc aaggcgtggc ggtggaagtg atggcaaact ccgataacgt gctgcgtgcg 6300
ggtctgacgc ctaaatacat tgatattccg gaactggttg ccaatgtgaa attcgaagcc 6360
aaaccggcta accagttgtt gacccagccg gtgaaacaag gtgcagaact ggactttccg 6420
attccagtgg atgattttgc cttctgctg catgacctta gtgataaaga aaccaccatt 6480
agccagcaga gtgccgccat tttgttctgc gtcgaaggcg atgcaacgtt gtggaaaggt 6540
tctcagcagt tacagcttaa accgggtgaa tcagcgttta ttgccgccaa cgaatcaccg 6600
gtgactgtca aaggccacgg ccgtttagcg cgtgtttaca acaagctgta agagcttact 6660
gaaaaaatta acatctcttg ctaagctggg agctctagat ccccgaattt ccccgatcgt 6720
tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 6780
atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 6840
ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata 6900
gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 6960
ctagatcggg aattggcgag cttgaattaa ttcagtacat taaaaacgtc cgcaatgtgt 7020
tattaagttg tctaagcgtc aatttgttta caccacaata tatcctgcca ccagccagcc 7080
aacagctccc cgaccggcag ctcggcacaa aatcaccact cgatacaggc agcccatcag 7140
tccgggacgg cgtcagcggg agagccgttg taaggcggca gactttgctc atgttaccga 7200
tgctattcgg aagaacggca actaagctgc cgggtttgaa acacggatga tctcgcggag 7260
ggtagcatgt tgattgtaac gatgacagag cgttgctgcc tgtgatcaaa tatcatctcc 7320
ctcgcagaga tccgaattat cagccttctt attcatttct cgcttaaccg tgacaggctg 7380
tcgatcttga gaactatgcc gacataataq gaaatcgctg gataaagccg ctgaggaagc 7440
tgagtggcgc tatttcttta gaagtgaacg ttgacgatcg tcgaccgtac ccccgatgaat 7500
taattcggac gtacgttctg aacacagctg gatacttact tgggcgattg tcatacatga 7560
catcaacaat gtacccgttt gtgtaaccgt cctttggagg ttcgtatgac actagtggtt 7620
cccctcagct tgcgactaga tgttgaggcc taacattttta ttagagagca ggctagttgc 7680
ttagatacat gatcttcagg ccgttatctg tcagggcaag cgaaaattgg ccatttatga 7740
cgaccaatgc cccgcagaag ctcccatctt tgccgccata gacgccgcgc ccccttttg 7800
gggtgtagaa catcctttg ccagatgtgg aaaagaagtt cgttgtccca ttgttggcaa 7860
tgacgtagta gccggcgaaa gtgcgagacc catttgcgct atatataagc ctacgattte 7920

cgttgcgact attgtcgtaa ttggatgaac tattatcgta gttgctctca gagttgtcgt 7980
aatttgatgg acattgtcg taattgctta tggagttgtc gtagttgctt ggagaaatgt 8040
cgtagttgga tggggagtag tcatagggaa gacgagcttc atccactaaa acaattggca 8100
ggtcagcaag tgcctgcccc gatgccatcg caagtacgag gcttagaacc accttcaaca 8160
gatcgcgcat agtcttcccc agctctctaa cgcttgagtt aagccgcgcc gcgaagcggc 8220
gtcggcttga acgaattgtt agacattatt tgccgactac cttggtgatc tcgccttcca 8280
cgtagtgaac aaattcttcc aactgatctg cgcgcgaggc caagcgatct tcttgtccaa 8340
gataagcctg cctagcttca gtatgacgg gctgatactg ggccggcagg cgctccattg 8400
cccagtcggc agcgacatcc ttcggcgcga tttgccggt tactgcgctg taccaaatgc 8460
gggacaacgt aagcactaca tttcgctcat cgccagccca gtcgggcggc gagttccata 8520
gcgttaaggt ttcatttagc gcctcaaata gatcctgttc aggaacogga tcaaagagtt 8580
cctccgccgc tggacctacc aaggcaacgc tatgttctct tgcttttgtc agcaagatag 8640
ccagatcaat gtcgatcgtg gctggctcga agatacctgc aagaatgtca ttgcgctgcc 8700
attctccaaa ttgcagttcg cgcttagctg gataacgcca cggaatgatg tcgtcgtgca 8760
caacaatggt gacttctaca gcgcggagaa tctcgctctc tccagggaa gccgaagttt 8820
ccaaaaggtc gttgatcaaa gctcgccgcg ttgtttcatc aagccttacg gtcaccgtaa 8880
ccagcaaatc aatatcactg tgtggcttca ggccgccatc cactgcggag ccgtacaaat 8940
gtacggccag caacgtcggt tcgagatggc gctcgatgac gccaactacc tctgatagtt 9000
gagtcgatac ttcggcgatc accgcttccc tcatgatgtt taactcctga attaagccgc 9060
gccgcgaagc ggtgtcggct tgaatgaatt gttaggcgtc atcctgtgct cccgagaacc 9120
agtaccagta catcgctgtt tcgttcgaga cttgaggtct agttttatac gtgaacaggt 9180
caatgccgcc gagagtaaag ccacattttg cgtacaaatt gcaggcaggt acattgttcg 9240
tttgtgtctc taatcgtatg ccaaggagct gtctgcttag tgcccacttt ttcgcaaatt 9300
cgatgagact gtgcgcgact cctttgcctc ggtgcgtgtg cgacacaaca atgtgttcga 9360
tagaggctag atcgttccat gttgagttga gttcaatctt cccgacaagc tcttggtcga 9420
tgaatgcgcc atagcaagca gagtcttcat cagagtcatc atccgagatg taatccttcc 9480
ggtagggggt cacacttctg gtagatagtt caaagccttg gtcggatagg tgcacatcga 9540
acacttcacg aacaatgaaa tggttctcag catccaatgt ttccgccacc tgctcaggga 9600
tcaccgaaat cttcatatga cgcctaacgc ctggcacagc ggatcgcaaa cctggcgcgg 9660
cttttggcac aaaaggcgtg acaggtttgc gaatccgttg ctgccacttg ttaacccttt 9720
tgccagattt ggtaactata atttatgtta gaggcgaagt cttgggtaaa aactggccta 9780
aaattgctgg ggatttcagg aaagtaaaca tcaccttccg gctcgatgtc tattgtagat 9840

atatgtagtg tatctacttg atcgggggat ctgctgcctc gcgcgtttcg gtgatgacgg 9900
tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc 9960
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc 10020
catgacccag tcacgtagcg atagcggagt gtatactggc ttaactatgc ggcatcagag 10080
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 10140
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 10200
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 10260
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 10320
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 10380
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 10440
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 10500
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 10560
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 10620
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 10680
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 10740
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc 10800
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 10860
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 10920
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 10980
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta 11040
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 11100
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 11160
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 11220
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac 11280
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 11340
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 11400
gttgttgcca ttgctgcagg gggggggggg gggggttcc attgttcatt ccacggacaa 11460
aaacagagaa aggaaacgac agaggccaaa aagctcgctt tcagcacctg tcgtttcctt 11520
tcttttcaga gggtatttta aataaaaaca ttaagttatg acgaagaaga acggaaacgc 11580
cttaaaccgg aaaattttca taaatagcga aaacccgcga ggtcgccgcc ccgtaacctg 11640
tcggatcacc ggaaaggacc cgtaaagtga taatgattat catctacata tcacaacgtg 11700
cgtggaggcc atcaaaccac gtcaaataat caattatgac gcaggtatcg tattaattga 11760

tctgcatcaa cttaacgtaa aaacaacttc agacaataca aatcagcgac actgaatacg 11820

gggcaacctc atgtcccccc cccccccccc ctgcaggcat cgtggtgtca cgctcgtcgt 11880

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatcccccа 11940

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 12000

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 12060

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 12120

tgcggcgacc gagttgctct tgcccggcgt caacacggga taataccgcg ccacatagca 12180

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 12240

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 12300

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 12360

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 12420

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 12480

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 12540

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc 12600

aagaattggt cgacgatctt gctgcgttcg gatattttcg tggagttccc gccacagacc 12660

cggattgaag gcgagatcca gcaactcgcg ccagatcatc ctgtgacgga actttggcgc 12720

gtgatgactg gccaggacgt cggccgaaag agcgacaagc agatcacgct tttcgacagc 12780

gtcggatttg cgatcgagga ttttcggcgc ctgcgctacg tccgcgaccg cgttgaggga 12840

tcaagccaca gcagcccact cgaccttcta gccgacccag acgagccaag ggatcttttt 12900

ggaatgctgc tccgtcgtca ggctttccga cgtttcggtg gttgaacaga agtcattatc 12960

gcacggaatg ccaagcactc ccgaggggaa ccctgtggtt ggcatgcaca tacaaatgga 13020

cgaacggata aacctttcca cgcccttta aatatccgat tattctaata aacgctcttt 13080

tctcttaggt ttacccgcca atatatcctg tcaaacactg atagtttaaa ctgaaggcgg 13140

gaaacgacaa ccctgatcatg agcggagaat taagggagtc acgttatgac ccccgccgat 13200

gacgcgggac aagccgtttt acgtttggaa ctgacagaac cgcaacgttg aaggagccac 13260

tcagcttaat taagtctaac tcgagttact ggtacgtacc aaatccatgg aatcaaggta 13320

ccgtcgactc tagtaacggc cgccagtgtg ctggaattaa ttcggcttgt cgaccaccca 13380

accccatatc gacagaggat gtgaagaaca ggtaaatcac gcagaagaac ccatctctga 13440

tagcagctat cgattagaac aacgaatcca tattgggtcc gtgggaaata cttactgcac 13500

aggaaggggg cgatctgacg aggccccgcc accggcctcg accgaggcc gaggccgacg 13560

aagcgccggc gagtacggcg ccgcggcggc ctctgcccgt gccctctgcg cgtgggaggg 13620

agaggccgcg gtggtggggg cgcgcgcgcg cgcgcgcgca gctggtgcgg cggcgcgggg 13680

gtcagccgcc gagccggcgg cgacggagga gcagggcggc gtggacgcga acttccgatc 13740

ggttggtcag agtgcgcgag ttgggcttag ccaattaggt ctcaacaatc tattgggccg 13800

taaaattcat gggccctggt ttgtctaggc ccaatatccc gttcatttca gcccacaaat 13860

atttccccag aggattatta aggcccacac gcagcttata gcagatcaag tacgatgttt 13920

cctgatcgtt ggatcggaaa cgtacggtct tgatcaggca tgccgacttc gtcaaagaga 13980

ggcggcatga cctgacgcgg agttggttcc gggcaccgtc tggatggtcg taccgggacc 14040

ggacacgtgt cgcgcctcca actacatgga cacgtgtggt gctgccattg ggccgtacgc 14100

gtggcggtga ccgcaccgga tgctgcctcg cacccgcttg cccacgcttt atatagagag 14160

gttttctctc cattaatcgc atagcgagtc gaatcgaccg aagggagggg ggagcgaagc 14220

tttgcgttct ctaatcgcct cgtcaaggta actaatcaat cacctcgtcc taatcctcga 14280

atctctcgtg gtgcccgtct aatctcgcga ttttgatgct cgtggtggaa agcgtaggag 14340

gatcccgtgc gagttagtct caatctctca gggtttcgtg cgattttagg gtgatccacc 14400

tcttaatcga gttacggttt cgtgcgattt tagggtaatc ctcttaatct ctcattgatt 14460

tagggtttcg tgagaatcga ggtagggatc tgtgttattt atatcgatct aatagatgga 14520

ttggttttga gattgttctg tcagatgggg attgtttcga tatattaccc taatgatgtg 14580

tcagatgggg attgtttcga tatattaccc taatgatgtg tcagatgggg attgtttcga 14640

tatattaccc taatgatgga taataagagt agttcacagt tatgttttga tcctgccaca 14700

tagtttgagt tttgtgatca gatttagttt tacttatttg tgcttagttc ggatgggatt 14760

gttctgatat tgttccaata gatgaatagc tcgttaggtt aaaatcttta ggttgagtta 14820

ggcgacacat agtttatttc ctctggattt ggattggaat tgtgttctta gttttttttc 14880

cctggatttg gattggaatt gtgtggagct gggttagaga attacatctg tatcgtgtac 14940

acctacttga actgtagagc ttgggttcta aggtcaattt aatctgtatt gtatctggct 15000

ctttgcctag ttgaactgta gtgctgatgt tgtactgtgt tttttaccc gttttatttg 15060

ctttactcgt gcaaatcaaa tctgtcagat gctagaacta ggtggcttta ttctgtgttc 15120

ttacatagat ctgttgtcct gtagttactt atgtcagttt tgttattatc tgaagatatt 15180

tttggttgtt gcttgttgat gtggtgtgag ctgtgagcag cgctcttatg attaatgatg 15240

ctgtccaatt gtagtgtagt atgatgtgat tgatatgttc atctattttg agctgacagt 15300

accgatatcg taggatctgg tgccaactta ttctccagct gcttttttt acctatgtta 15360

attccaatcc tttcttgcct cttccagatc cagataatgg cccacgcccg cgtcctcctc 15420

ctggcgctcg ccgtcctggc caccgccgcc gtcgccgtcg cctcctcctc ctccttcgcc 15480

gactccaacc cgatccgccc ggtgaccgac cgcgccgcct ccacgcttta cgactacaag 15540

caggtgttgc gggactcgct actattctat gaggcccaga gatccggccg gctcccagcc 15600

gaccagaagg tcacgtggag gaaggatagc gcgctgaatg accagggtga ccagggacaa 15660

gacttgaccg gcggctactt tgacgctggg gacttcgtca agttcgggtt ccccatggct 15720

tataccgcaa ccgtgctggc atggggcctc atagattttg aggccggcta cagcagtgcc 15780

ggggccttgg atgatggacg gaaggctgtc aaatgggcca ccgactattt cataaaggcc 15840

cacacaagtc aaaatgagtt ctatggtcag gtcggccagg gtgacgccga tcacgctttc 15900

tggggaagac cagaggatat gacgatggcg cgcccggcgt acaagataga cacctcaagg 15960

cctggctctg atctggcagg cgagacagcg gctgctcttg ccgctgcttc aatcgtgttc 16020

cggaacgtcg atggcactta ctcaaataac ctgttaacac acgctcgcca gctattcgac 16080

ttcgcgaaca actaccgggg aaagtatagt gactctatta ctgacgcaag aaatttctac 16140

gcaagcgcag actacagaga cgagttggtt tgggctgctg cgtggttata cagagcgacc 16200

aacgacaaca cctacctcaa cactgctgag tcactgtacg atgagtttgg gctacagaac 16260

tgggggggcg gcctgaactg ggatagcaag gtgtctggcg tgcaggtgtt gttggccaag 16320

cttaccaata agcaggccta caaggacacg gtgcagtctt acgtcaatta cctaattaat 16380

aaccagcaga agactcccaa gggcctcctc tacatcgaca tgtggggcac ccttcgccac 16440

gctgccaacg ccgcattcat catgctcgaa gccgccgagc tgggcttgtc cgcctcctct 16500

tatagacagt tcgcgcaaac gcaaatcgac tacgccctgg gcgatggtgg ccgctccttt 16560

gtgtgcgggt tcgggagtaa tcctcctacg agaccgcacc acagatcctc gtcgtgcccg 16620

ccagctcccg ctacttgcga ctggaatacg ttcaacttcac ctgacccaaa ctaccacgtc 16680

ctctctgggg ccctagtggg cggacctgat cagaatgaca actacgtcga tgaccgttca 16740

gactatgttc acaacgaagt cgccactgat tacaacgcgg gtttccagtc cgcgttagct 16800

gctttggtgg ccttggttta ctgacctagg tccccgaatt tccccgatcg ttcaaacatt 16860

tggcaataaa gtttcttaag attgaatcct gttgccggtc ttgcgatgat tatcatataa 16920

tttctgttga attacgttaa gcatgtaata attaacatgt aatgcatgac gttatttatg 16980

agatgggttt ttatgattag agtcccgcaa ttatacattt aatacgcgat agaaaacaaa 17040

atatagcgcg caaactagga taaattatcg cgcgcggtgt catctatgtt actagatcgg 17100

gaattgggta c 17111

<210> 79
<211> 12892
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2310

<400> 79
ctaggtcccc gaatttcccc gatcgttcaa acatttggca ataaagtttc ttaagattga 60

atcctgttgc cggtcttgcg atgattatca tataatttct gttgaattac gttaagcatg 120

taataattaa catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc 180

cgcaattata catttaatac gcgatagaaa acaaaatata gcgcgcaaac taggataaat 240

tatcgcgcgc ggtgtcatct atgttactag atcgggaatt ggaattcata ctaaagcttg 300

catgcctgca ggtcgactct agtaacggcc gccagtgtgc tggaattaat tcggcttgtc 360

gaccacccaa ccccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc 420

catctctgat agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac 480

ttactgcaca ggaagggggc gatctgacga ggccccgcca ccggcctcga cccgaggccg 540

aggccgacga agcgccggcg agtacggcgc cgcggcggcc tctgccccgtg ccctctgcgc 600

gtgggagggga gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc 660

ggcgcggggg tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa 720

cttccgatcg gttggtcaga gtgcgccgagt tgggcttagc caattaggtc tcaacaatct 780

attgggccgt aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag 840

cccacaaata tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt 900

acgatgtttc ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg 960

tcaaagagag gcggcatgac ctgacgcgga gttggttccg ggcacccgtct ggatggtcgt 1020

accgggaccg gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg 1080

gccgtacgcg tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta 1140

tatagagagg ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg 1200

gagcgaagct ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc accicgtcct 1260

aatcctcgaa tctctcgtgg tgcccgtcta atctcgcgat tttgatgctc gtggtggaaa 1320

gcgtaggagg atccccgtgcg agttagtctc aatctctcag ggtttcgtgc gattttaggg 1380

cgatccacct cttaatcgag ttacggtttc gcgcgatttt agggtaatcc tcttaatctc 1440

tcattgattt agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta 1500

atagatggat tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct 1560

aatgatgtgt cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga 1620

ttgtttcgat atattaccct aatgatggat aataagagta gttcacagtt atgttttgat 1680

cctgccacat agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg 1740

gatgggattg ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag 1800

gttgagttag gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag 1860

ttttttcccc ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt 1920

atcgtgtaca cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgcattg 1980

tatctggctc tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt tttttacccg 2040

ttttatttgc tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat 2100

tctgtgttct tacatagatc tgttgtcctg tagttactca tgtcagtttt gttattatcc 2160

gaagatattt ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga 2220

ttaatgatgc tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga 2280

gctgacagta ccgatatcgt aggatctggt gccaacttat tctccagctg ctttttttta 2340

cctatgttaa ttccaatcct ttcttgcctc ttccagatcc agataatgca gaaactcatt 2400

aactcagtgc aaaactatgc ctggggcagc aaaacggcgt tgactgaact ttatggtatg 2460

gaaaatccgt ccagccagcc gatggccgag ctgtggatgg gcgcacatcc gaaaagcagt 2520

tcacgagtgc agaatgccgc cggagatatc gtttcactgc gtgatgtgat tgagagtgat 2580

aaatcgactc tgctcggaga ggccgttgcc aaacgctttg gcgaactgcc tttcctgttc 2640

aaagtattat gcgcagcaca gccactctcc attcaggttc atccaaacaa acacaattct 2700

gaaatcggtt ttgccaaaga aaatgccgca ggtatcccga tggatgccgc cgagcgtaac 2760

tataaagatc ctaaccacaa gccggagctg gtttttgcgc tgacgccttt ccttgcgatg 2820

aacgcgtttc gtgaatttta cgagattgtc tccctactcc agccggtcgc aggtgcacat 2880

ccggcgattg ctcacttttt acaacagcct gatgccgaac gtttaagcga actgttcgcc 2940

agcctgttga atatgcaggg tgaagaaaaa tcccgcgcgc tggcgatttt aaaatcggcc 3000

ctcgatagcc agcagggtga acgtggcaa acgattcgtt taatttctga attttacccg 3060

gaagacagcg gtctgttctc cccgctattg ctgaatgtgg tgaaattgaa ccctggcgaa 3120

gcgatgttcc tgttcgctga aacaccgcac gcttacctgc aaggcgtggc gctggaagtg 3180

atggcaaact ccgataacgt gctgcgtgcg ggtctgacgc ctaaatacat tgatattccg 3240

gaactggttg ccaatgtgaa attcgaagcc aaaacggcta accagttgtt gaccccagcg 3300

gtgaaacaag gtgcagaact ggacttcccg attccagtgg atgattttgc cttctcgctg 3360

catgaccttа gtgataaaga accaccatt agccagcaga gtgccgccat tttgttctgc 3420

gtcgaaggcg atgcaacgtt gtggaaaggt tctcagcagt tacagcttaa acccgggtgaa 3480

tcagcgttta ttgccgccaa cgaatcaccg gtgactgtca aaggccacgg ccgtttagcg 3540

cgtgtttaca acaagctgta agagcttact gaaaaaatta acatctcttg ctaagctggg 3600

agctctagat ccccgaattt ccccgatcgt tcaaacattt ggcaataaag tttcttaaga 3660

ttgaatcctg ttgccggtct tgcgatgatt atcatataat ttctgttgaa ttacgttaag 3720

catgtaataa ttaacatgta atgcatgacg ttatttatga gatgggtttt tatgattaga 3780

gtcccgcaat tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat 3840

aaattatcgc gcgcggtgtc atctatgtta ctagatcggg aattggcgag ctcgaattaa 3900

ttcagtacat taaaaacgtc cgcaatgtgt tattaagttg tctaagcgtc aatttgttta 3960
caccacaata tatcctgcca ccagccagcc aacagctccc cgaccggcag ctcggcacaa 4020
aatcaccact cgatacaggc agcccatcag tccgggacgg cgtcagcggg agagccgttg 4080
taaggcggca gactttgctc atgttaccga tgctattcgg aagaacggca actaagctgc 4140
cgggtttgaa acacggatga tctcgcggag ggtagcatgt tgattgtaac gatgacagag 4200
cgttgctgcc tgtgatcaaa tatcatctcc ctcgcagaga tccgaattat cagccttctt 4260
attcatttct cgcttaaccg tgacaggctg tcgatcttga gaactatgcc gacataatag 4320
gaaatcgctg gataaagccg ctgaggaagc tgagtggcgc tatttcttta gaagtgaacg 4380
ttgacgatcg tcgaccgtac ccgatgaat taattcggac gtacgttctg aacacagctg 4440
gatacttact tgggcgattg tcatacatga catcaacaat gtacccgttt gtgtaaccgt 4500
cttcttggagg ttcgtatgac actagtggtt ccctcagct tgcgactaga tgttgaggcc 4560
taacatttta ttagagagca ggctagttgc ttagatacat gatcttcagg ccgttatctg 4620
tcagggcaag cgaaaattgg ccatttatga cgaccaatgc ccgcagaag ctcccatctt 4680
tgccgccata gacgccgcgc ccccttttg gggtgtagaa catcctttg ccagatgtgg 4740
aaaagaagtt cgttgtccca ttgttggcaa tgacgtagta gccggcgaaa gtgcgagacc 4800
catttgcgct atatataagc ctacgatttc cgttgcgact attgtcgtaa ttggatgaac 4860
tattatcgta gttgctctca gagttgtcgt aatttgatgg actattgtcg taattgctta 4920
tggagttgtc gtagttgctt ggagaaatgt cgtagttgga tggggagtag tcatagggaa 4980
gacgagcttc atccactaaa acaattggca ggtcagcaag tgcctgcccc gatgccatcg 5040
caagtacgag gcttagaacc accttcaaca gatcgcgcat agtcttcccc agcctctaa 5100
cgcttgagtt aagccgcgcc gcgaagcggc gtcggcttga acgaattgtt agacattatt 5160
tgccgactac cttggtgatc tcgcctttca cgtagtgaac aaattcttcc aactgatctg 5220
cgcgcgaggc caagcgatct tcttgtccaa gataagcctg cctagcttca agtatgacgg 5280
gctgatactg ggccggcagg cgctccattg cccagtcggc agcgacatcc ttcggcgcga 5340
ttttgccggt tactgcgctg taccaaatgc gggacaacgt aagcactaca tttcgctcat 5400
cgccagccca gtcgggcggc gagttccata gcgttaaggt ttcatttagc gcctcaaata 5460
gatcctgttc aggaaccgga tcaaagagtt cctccgccgc tggacctacc aaggcaacgc 5520
tatgttctct tgcttttgtc agcaagatag ccagatcaat gtcgatcgtg gctggctcga 5580
agatacctgc aagaatgtca ttgcgctgcc attctccaaa ttgcagttcg cgcttagctg 5640
gataacgcca cggaatgatg tcgtcgtgca caacaatggt gacttctaca gcgcggagaa 5700
tctcgctctc tccagggaa gccgaagttt ccaaaaggtc gttgatcaaa gctcgccgcg 5760
ttgtttcatc aagccttacg gtcaccgtaa ccagcaaatc aatatcactg tgtggcttca 5820

ggcegccatc cactgcggag ccgtacaaat gtacggccag caacgtcggt tcgagatggc 5880
gctcgatgac gccaactacc tctgatagtt gagtcgatac ttcggcgatc accgcttccc 5940
tcatgatgct taactcctga attaagccgc gccgcgaagc ggtgtcgget tgaatgaatt 6000
gttaggcgtc atcctgtgct cccgagaacc agtaccagta catcgctgtt tcgttcgaga 6060
cttgaggtct agttttatac gtgaacaggt caatgccgcc gagagtaaag ccacattttg 6120
cgtacaaatt gcaggcaggt acattgttcg tttgtgtctc taatcgtatg ccaaggagct 6180
gtctgcttag tgcccacttt ttcgcaaatt cgatgagact gtgcgcgact cctttgcctc 6240
ggtgcgtgtg cgacacaaca atgcgttcga tagaggctag atcgttccat gttgagttga 6300
gttcaatctt cccgacaagc tcttggtcga tgaatgcgcc atagcaagca gagtcttcat 6360
cagagtcatc atccgagatg taatccttcc ggtaggggct cacacttctg gtagatagtt 6420
caaagccttg gtcggatagg tgcacatcga acacttcacg aacaatgaaa tggttctcag 6480
catccaatgt ttccgccacc tgctcaggga tcaccgaaat cttcatatga cgcctaacgc 6540
ctggcacagc ggatcgcaaa cctggcgcgg cttttggcac aaaaggcgtg acaggtttgc 6600
gaatccgttg ctgccacttg ttaacccttt tgccagattt ggtaactata atttatgtta 6660
gaggcgaagt cttgggtaaa aactggccta aaattgctgg ggatttcagg aaagtaaaca 6720
tcaccttccg gctcgatgtc tattgtagat atatgtagtg tatctacttg atcgggggat 6780
ctgctgcctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctcccgga 6840
gacggtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc agggcgcgtc 6900
agcgggtgtt ggcgggtgtc ggggcgcagc catgacccag tcacgtagcg atagcggagt 6960
gtatactggc ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg 7020
tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc 7080
tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 7140
aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 7200
aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 7260
ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 7320
acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 7380
ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 7440
tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 7500
tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 7560
gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 7620
agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 7680
tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 7740

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt 7800
tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 7860
acgggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 7920
tcaaaaagga tcttcaccta gatccttta aattaaaaat gaagttttaa atcaatctaa 7980
agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 8040
tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 8100
acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 8160
tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 8220
ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccgggg agctagagta 8280
agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctgcagg gggggggggg 8340
gggggttcc attgttcatt ccacggacaa aaacagagaa aggaaacgac agaggccaaa 8400
aagctcgctt tcagcacctg tcgtttcctt tcttttcaga gggtatttta aataaaaaca 8460
ttaagttatg acgaagaaga acggaaacgc cttaaaccgg aaaattttca taaatagcga 8520
aaaccgcgca ggtcgccgcc ccgtaacctg tcggatcacc ggaaaggacc cgtaaagtga 8580
taatgattat catctacata tcacaacgtg cgtggaggcc atcaaaccac gtcaaataat 8640
caattatgac gcaggtatcg tattaattga tctgcatcaa cttaacgtaa aaacaacttc 8700
agacaatacа aatcagcgac actgaatacg gggcaacctc atgtcccccc ccccccccc 8760
ctgcaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 8820
aacgatcaag gcgagttaca tgatcccccа tgttgtgcaa aaaagcggtt agctccttcg 8880
gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 8940
cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 9000
actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 9060
caacacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac 9120
gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 9180
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 9240
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 9300
tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 9360
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 9420
cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa 9480
ataggcgtat cacgaggccc tttcgtcttc aagaattggt cgacgatctt gctgcgttcg 9540
gatattttcg cggagttccc gccacagacc cggattgaag gcgagatcca gcaactcgcg 9600
ccagatcatc ctgtgacgga actttggcgc gtgatgactg gccaggacgt cggccgaaag 9660

agcgacaagc agatcacgct tttcgacagc gtcggatttg cgatcgagga tttttcggcg 9720
ctgcgctacg tccgcgaccg cgttgaggga tcaagccaca gcagcccact cgaccttcta 9780
gccgacccag acgagccaag ggatcttttt ggaatgctgc tccgtcgtca ggctttccga 9840
cgtttgggtg gttgaacaga agtcattatc gcacggaatg ccaagcactc ccgagggaa 9900
cctgtggttt ggcatgcaca tacaaatgga cgaacggata aacctttcca cgcccttta 9960
aatatccgat tattctaata aacgctcttt tctcttaggt ttacccgcca atatatcctg 10020
tcaaacactg atagtttaaa ctgaaggcgg gaaacgacaa tctgatcatg agcggagaat 10080
taagggagtc acgttatgac ccccgccgat gacgcgggac aagccgtttt acgtttggaa 10140
ctgacagaac cgcaacgttg aaggagccac tcagcttaat taagtctaac tcgagttact 10200
ggtacgtacc aaatccatgg aatcaaggta ccgtcgactc tagtaacggc cgccagtgtg 10260
ctggaattaa ttcggcttgt cgaccaccca accccatatc gacagaggat gtgaagaaca 10320
ggtaaatcac gcagaagaac ccatctctga tagcagctat cgattagaac aacgaatcca 10380
tattgggtcc gtgggaaata cttactgcac aggaaggggg cgatctgacg aggccccgcc 10440
accggcctcg acccgaggcc gaggccgacg aagcgccggc gagtacggcg ccgcggcggc 10500
ctctgcccgt gccctctgcg cgtgggaggg agaggccgcg gtggtggggg cgcgcgcgcg 10560
cgcgcgcgca gctggtgcgg cggcgcgggg gtcagccgcc gagccggcgg cgacggagga 10620
gcagggcggc gtggacgcga acttccgatc ggttggtcag agtgcgcgag ttgggcttag 10680
ccaattaggt ctcaacaatc tattgggccg taaaattcat gggcctggt ttgtctaggc 10740
ccaatatccc gttcatttca gcccacaaat atttccccag aggattatta aggcccacac 10800
gcagcttata gcagatcaag tacgatgttt cctgatcgtt ggatcggaaa cgtacggtct 10860
tgatcaggca tgccgacttc gtcaaagaga ggcggcatga cctgacgcgg agttggttcc 10920
gggcaccgtc tggatggtcg taccgggacc ggacacgtgt cgcgcctcca actacatgga 10980
cacgtgtggt gctgccattg ggccgtacgc gtggcggtga ccgcaccgga tgctgcctcg 11040
cacccgcttg cccacgcttt atatagagag gttttctctc cattaatcgc atagcgagtc 11100
gaatcgaccg aagggagggg ggagcgaagc tttgcgttct ctaatcgcct cgtcaaggta 11160
actaatcaat cacctcgtcc taatcctcga atctctcgtg gtgcccgtct aatctcgcga 11220
ttttgatgct cgtggtggaa agcgtaggag gatcccgtgc gagttagtct caatctctca 11280
gggtttcgtg cgattttagg gtgatccacc tcttaatcga gttacggttt cgtgcgattt 11340
tagggtaatc ctcttaatct ctcattgatt tagggtttcg tgagaatcga ggtagggatc 11400
tgtgttattt atatcgatct aatagatgga ttggttttga gattgttctg tcagatgggg 11460
attgtttcga tatattaccc taatgatgtg tcagatgggg attgtttcga tatattaccc 11520
taatgatgtg tcagatgggg attgtttcga tatattaccc taatgatgga taataagagt 11580

agttcacagt tatgttttga tcctgccaca tagtttgagt tttgtgatca gatttagttt 11640
tacttatttg tgcttagttc ggatgggatt gttctgatat tgttccaata gatgaatagc 11700
tcgttaggtt aaaatcttta ggttgagcta ggcgacacat agtttatttc ctctggattt 11760
ggattggaat tgtgttctta gttttttttc cctggatttg gattggaatt gtgtggagct 11820
gggttagaga attacatctg tatcgtgtac acctacttga actgtagagc ttgggttcta 11880
aggtcaattt aatctgtatt gtatctggct ctttgcctag ttgaactgta gtgctgatgt 11940
tgtactgtgt ttttttaccc gttttatttg ctttactcgt gcaaatcaaa tctgtcagat 12000
gctagaacta ggtggcttta ttctgtgttc ttacatagat ctgttgtcct gtagttactt 12060
atgtcagttt tgttattatc tgaagatatt tttggttgtt gcttgttgat gtggtgtgag 12120
ctgtgagcag cgctcttatg attaatgatg ctgtccaatt gtagtgtagt atgatgtgat 12180
tgatatgttc atctattttg agctgacagt accgatatcg taggatctgg tgccaactta 12240
ttctccagct gtttttttct acctatgtta attccaatcc tttcttgcct cttccagatc 12300
cagataatgg cccaaacatg tctgaccagc ccacaaacag gattccacaa cggcttcttt 12360
tactcgtttt ggaaagattc tcctggtact gttaactttt gccttcttga aggggaaga 12420
tatacctcca attggtctgg gataaacaat tgggtcggtg gcaagggctg gcagacggt 12480
agtaggcgca acatcacgta ttccgggagc ttcaatacac cagggaatgg atacctggct 12540
ctctacggtt ggaccaccaa cccgttggtt gagtactatg ttgtcgatag ctggggctcc 12600
tggcgccctc ccggaagtga tgggacattt cttggcacag tgaactcaga cggcggcacg 12660
tatgacatct acagggcgca aagagtcaac gcaccgagca tcattggcaa tgccaccttc 12720
taccagtatt ggtccgtgcg gcagtccaag cgtgtcggtg ggacaatcac gactggcaac 12780
cacttcgacg cgtgggccag cgtgggcctc aacctcggga ctcataacta ccagataatg 12840
gctactgagg gttaccagtc gtcaggatct tcagacatta cggtgtcatg ac 12892

<210> 80
<211> 21051
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2339

<400> 80
catggaatca aggtaccgtc gactctagta acggccgcca gtgtgctgga attaattcgg 60
cttgtcgacc acccaacccc atatcgacag aggatgtgaa gaacaggtaa atcacgcaga 120
agaacccatc tctgatagca gctatcgatt agaacaacga atccatattg ggtccgtggg 180
aaatacttac tgcacaggaa gggggcgatc tgacgaggcc ccgccaccgg cctcgacccg 240
aggccgaggc cgacgaagcg ccggcgagta cggcgccgcg gcggcctctg cccgtgccct 300

ctgcgcgtgg gagggagagg ccgcggtggt gggggcgcgc gcgcgcgcgc gcgcagctgg 360

tgcggcggcg cgggggtcag ccgccgagcc ggcggcgacg gaggagcagg gcggcgtgga 420

cgcgaacttc cgatcggttg gtcagagtgc gcgagttggg cttagccaat taggtctcaa 480

caatctattg ggccgtaaaa ttcatgggcc ctggtttgtc taggcccaat atcccgttca 540

tttcagccca caaatatttc cccagaggat tattaaggcc cacacgcagc ttatagcaga 600

tcaagtacga tgtttcctga tcgttggatc ggaaacgtac ggtcttgatc aggcatgccg 660

acttcgtcaa agagaggcgg catgacctga cgcggagttg gttccgggca ccgtctggat 720

ggtcgtaccg ggaccggaca cgtgtcgcgc ctccaactac atggacacgt gtggtgctgc 780

cattgggccg tacgcgtggc ggtgaccgca ccggatgctg cctcgcaccg ccttgcccac 840

gctttatata gagaggtttt ctctccatta atcgcatagc gagtcgaatc gaccgaaggg 900

gaggggagc gaagctttgc gttctctaat cgcctcgtca aggtaactaa tcaatcacct 960

cgtcctaatc ctcgaatctc tcgtggtgcc cgtctaatct cgcgattttg atgctcgtgg 1020

tggaaagcgt aggaggatcc cgtgcgagtt agtctcaatc tctcagggtt tcgtgcgatt 1080

ttagggtgat ccacctctta atcgagttac ggtttcgtgc gattttaggg taatcctctt 1140

aatctctcat tgatttaggg tttcgtgaga atcgaggtag ggatctgtgt tatttatatc 1200

gatctaatag atggattggt tttgagattg ttctgtcaga tggggattgt ttcgatatat 1260

taccctaatg atgtgtcaga tggggattgt ttcgatatat taccctaatg atgtgtcaga 1320

tggggattgt ttcgatatat taccctaatg atggataata agagtagttc acagttatgt 1380

tttgatcctg ccacatagtt tgagttttgt gatcagattt agttttactt atttgtgctt 1440

agttcggatg ggattgttct gatattgttc caatagatga atagctcgtt aggttaaaat 1500

ctttaggttg agttaggcga cacatagttt atttcctctg gatttggatt ggaattgtgt 1560

tcttagtttt tttccccttgg atttggattg gaattgtgtg gagctgggtt agagaattac 1620

atctgtatcg tgtacacctа cttgaactgt agagcttggg ttctaaggtc aatttaatct 1680

gtattgtatc tggctctttg cctagttgaa ctgtagtgct gatgttgtac cgtgtttttt 1740

tacccgtttt atttgcttta ctcgtgcaaa tcaaatctgt cagatgctag aactaggtgg 1800

ctttattctg tgttcttaca tagatctgtt gtcctgtagt tacttatgtc agttttgtta 1860

ctatctgaag atatttttgg ttgttgcttg ttgatgtggt gtgagctgtg agcagcgctc 1920

ttatgattaa tgatgctgtc caattgtagt gtagtatgat gtgattgata tgttcatcta 1980

ttttgagctg acagtaccga tatcgtagga tctggtgcca acttattctc cagctgcttt 2040

tttttaccta tgttaattcc aatcctttct tgcctcttcc agatccagat aatggcgaac 2100

aaacatttgt cctctccct cttcctcgtc ctccttggcc tgtcggccag cttggcctcc 2160

gggcaacagg tgtgcgattg gctgttccct gatggagata atggaaagga gcccgaacca 2220

gaaccagaac caacggtaga gctctgcgga agatgggatg caagagacgt agcgggagga 2280

agatatagag ttattaacaa tgtgtggggg gccgaaacag cacagtgtat agaagtaggt 2340

cttgaaacag gtaattttac gataacaaga gccgatcatg acaatggtaa taatgttgca 2400

gcatatccag caatttactt cggatgtcat tgggcaccag cgagagcaat aagggattgt 2460

gccgcgcgtg cgggagcggt taggagagca cacgaattgg atgttacacc aattaccacg 2520

ggaagatgga acgcagctta cgatatatgg tttagtccgg taacaaactc tggtaacggg 2580

tactcgggag gcgccgaact tatgatatgg ctgaattgga atggtggcgt aatgccagga 2640

ggatcacggg tagcaactgt cgaattggct ggagcgacat gggaagtgtg gtatgcagat 2700

tgggattgga attacatcgc atatagaaga acgactccga caacctcagt gagtgaactt 2760

gacttgaaag cctttattga tgatgcagta gcgagaggat acataaggcc agaatggtat 2820

ctgcatgcag tggaaacggg atttgaattg tgggaagggg gggctgggtt gaggacagca 2880

gattttagcg taactgtaca gtgacctagg tccccgaatt tccccgatcg ttcaaacatt 2940

tggcaataaa gtttcttaag attgaatcct gttgccggtc ttgcgatgat tatcatataa 3000

ttttctgttga attacgttaa gcatgtaata attaacatgt aatgcatgac gttatttatg 3060

agatgggttt ttatgattag agtcccgcaa ttatacattt aatacgcgat agaaaacaaa 3120

atatagcgcg caaactagga taaattatcg cgcgcggtgt catctatgtt actagatcgg 3180

gaattgggta ccgtcgactc tagtaacggc cgccagtgtg ctggaattaa ttcggcttgt 3240

cgaccacccg acccccatatc gacagaggat gtgaagaaca ggtaaatcac gcagaagaac 3300

ccatctctga tagcagctat cgattagaac aacgaatcca tattgggtcc gtgggaaata 3360

cttactgcac aggaaggggg cgatctgacg aggccccgcc accggcctcg acccgaggcc 3420

gaggccgacg aagcgccggc gagtacggcg ccgcggcggc ctctgccccgt gccctctgcg 3480

cgtgggaggg agaggccgcg gtggtggggg cgcgcgcgcg cgcgcgcgca gctggtgcgg 3540

cggcgcgggg gtcagccgcc gsgccggcgg cgacggagga gcagggcggc gtggacgcga 3600

acttccgatc ggttggtcag agtgcgcgag ttgggcttag ccaattaggt ctcaacaatc 3660

tattgggccg taaaattcat gggccctggt ttgtctaggc ccaatatccc gttcatttca 3720

gcccacaaat atttccccag aggattatta aggcccacac gcagcttata gcagatcaag 3780

tacgatgttt cctgatcgtt ggatcggaaa cgtacggtct tgatcaggca tgccgacttc 3840

gtcaaagaga ggcggcatga cctgacgcgg agttggttcc gggcaccgtc tggatggtcg 3900

tacgggacc ggacacgtgt cgcgcctcca actacatgga cacgtgtggt gctgccattg 3960

ggccgtacgc gtggcggtga ccgcatcgga tgctgcctcg cacccgcttg cccacgcttt 4020

atatagagag gttttctctc cattaatcgc atagcgagtc gaatcgaccg aaggggaggg 4080

ggagcgaagc tttgcgttct ctaatcgcct cgtcaaggta actaatcaat cacctcgtcc 4140

taatcctcga atctctcgtg gtgcccgtct aatctcgcga ttttgatgct cgtggtggaa 4200
agcgtaggag gatcccgtgc gagttagtct caatctctca gggtttcgtg cgattttagg 4260
gtgatccacc tcttaatcga gttacggttt cgtgcgattt tagggtaatc ctcttaatct 4320
ctcattgatt tagggtttcg tgagaatcga ggtagggatc tgtgttattt atatcgatct 4380
aatagatgga ttggttttga gattgttctg tcagatgggg attgtttcga tatattaccc 4440
taatgatgtg tcagatgggg attgtttcga tatattaccc taatgatgtg tcagatgggg 4500
attgtttcga tatattaccc taatgatgga taataagagt agttcacagt tatgttttga 4560
tcctgccaca tagtttgagt tttgtgatca gatttagttt tacttatttg tgcttagttc 4620
ggatgggatt gttctgatat tgttccaata gatgaatagc tcgttaggtt aaaatcttta 4680
ggttgagtta ggcgacacat agtttatttc ctctggattt ggattggaat tgtgttctta 4740
gttttttcc cctggatttg gattggaatt gtgtggagct gggttagaga attacatctg 4800
tatcgtgtac acctacttga actgtagagc ttgggttcta aggtcaattt aatctgtatt 4860
gtatctggct ctttgcctag ttgaactgta gtgctgatgt tgtactgtgt ttttttaccc 4920
gttttatttg ctttactcgt gcaaatcaaa tctgtcagat gctagaacta ggtggcttta 4980
ttctgtgttc ttacatagat ctgttgtcct gtagttactt atgtcagttt tgttattatc 5040
tgaagatatt tttggttgtt gcttgttgat gtggtgtgag ctgtgagcag cgctcttatg 5100
attaatgatg ctgtccaatt gtagtgtagt atgatgtgat tgatatgttc atctatttta 5160
agctgacagt accgatatcg taggatccgg tgccaacttа ttctccagct gcttttttt 5220
acctatgtta attccaatcc tttcttgcct cttccagatc cagataatgg cccaaacatg 5280
tctgaccagc ccacaaacag gattccacaa cggcttcttt tactcgtttt ggaaagattc 5340
tcctggtact gttaactttt gccttcttga aggggaaga tatacctcca attggtctgg 5400
gataaacaat tgggtcggtg gcaagggctg gcagaccggt agtaggcgca acatcacgta 5460
ttccgggagc ttcaatacac cagggaatgg ataccggcct ctctacggtt ggaccaccaa 5520
cccgttggtt gagtactatg ttgtcgatag ctggggctcc tggcgccctc ccggaagtga 5580
tgggacattt cttggcacag tgaactcaga cggcggcacg tatgacatct acagggcgca 5640
aagagtcaac gcaccgagca tcattggcaa tgccaccttc taccagtatt ggtccgtgcg 5700
gcagtccaag cgtgtcggtg ggacaatcac gactggcaac cacttcgacg cgtgggccag 5760
cgtgggcctc aacctcggga ctcataacta ccagataatg gctactgagg gttaccagtc 5820
gtcaggatct tcagacatta cggtgtcatg acctaggtcc ccgaatttcc ccgatcgttc 5880
aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat 5940
catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt 6000
atttatgaga tgggtttta tgattagagt cccgcaatta tacatttaat acgcgataga 6060

aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact 6120
agatcgggaa ttggaattca tactaaagct tgcatgcctg caggtcgact ctagtaacgg 6180
ccgccagtgt gctggaatta attcggcttg tcgaccaccc aaccccatat cgacagagga 6240
tgtgaagaac aggtaaatca cgcagaagaa cccatcttcg atagcagcta tcgattagaa 6300
caacgaatcc atattgggtc cgtgggaaat acttactgca caggaagggg gcgatctgac 6360
gaggccccgc caccggcctc gaccgaggc cgaggccgac gaagcgccgg cgagtacggc 6420
gccgcggcgg cctctgcccg tgccctctgc gcgtgggagg gagaggccgc ggtggcgggg 6480
gcgcgcgcgc gcgcgcgcgc agctggtgcg gcggcgcggg ggtcagccgc cgagccggcg 6540
gcgacggagg agcagggcgg cgtggacgcg aacttccgat cggttggtca gagtgcgcga 6600
gttgggctta gccaattagg tctcaacaat ctattgggcc gtaaaattca tgggcctgg 6660
tttgtctagg cccaatatcc cgttcatttc agcccacaaa tatttcccca gaggattatt 6720
aaggcccaca cgcagcttat agcagatcaa gtacgatgtt tcctgatcgt tggatcggaa 6780
acgtacggtc ttgatcaggc atgccgactt cgtcaaagag aggcggcatg acctgacgcg 6840
gagttggttc cgggcaccgt ctggatggtc gtaccgggac cggacacgtg tcgcgcctcc 6900
aactacatgg acacgtgtgg tgctgccatt gggccgtacg cgtggcggtg accgcaccgg 6960
atgctgcctc gcaccgcctt gcccacgctt tatatagaga ggttttctct ccattaatcg 7020
catagcgagt cgaatcgacc gaaggggagg gggagcgaag ctttgcgttc tctaatcgcc 7080
tcgtcaaggc aactaatcaa tcacctcgtc ctaatcctcg aatctctcgt ggtgcccgtc 7140
taatctcgcg attttgatgc tcgtggtgga aagcgtagga ggatcccgtg cgagttagtc 7200
tcaatctctc agggtttcgt gcgattttag ggtgatccac ctcttaatcg agttacggtt 7260
tcgtgcgatt ttagggtaat cctcttaatc tctcattgat ttagggtttc gtgagaatcg 7320
aggtagggat ctgtgttatt tatatcgatc taatagatgg attggttttg agattgttct 7380
gtcagatggg gattgtttcg atatattacc ctaatgatgt gtcagatggg gattgtttcg 7440
atatattacc ctaatgatgt gtcagatggg gattgtttcg atatattacc ctaatgatgg 7500
ataataagag tagttcacag ttatgttttg atcctgccac atagtttgag ttttgtgatc 7560
agatttagtt ttacttattt gtgcttagtt cggatgggat tgttctgata ttgttccaat 7620
agatgaatag ctcgttaggt taaaatcttt aggttgagtt aggcgacaca tagtttattt 7680
cctctggatt tggattggaa ttgtgttctt agtttttttc ccctggattt ggattggaat 7740
tgtgtggagc tgggttagag aattacatct gtatcgtgta cacctacttg aactgtagag 7800
cttgggttct aaggtcaatt taatctgtat tgtatctggc tctttgccta gttgaactgt 7860
agtgctgatg ttgtactgtg ttttttttacc cgttttattt gctttactcg tgcaaatcaa 7920
atctgtcaga tgctagaact aggtggcttt attctgtgtt cttacataga tctgttgtcc 7980

tgtagttact tatgtcagtt ttgttattat ctgaagatat ttttggttgt tgcttgttga 8040
tgtggtgtga gctgtgagca gcgctcttat gattaatgat gctgtccaat tgtagtgtag 8100
tatgatgtga ttgatatgtt catctatttt gagctgacag taccgatatc gtaggatctg 8160
gtgccaactt attctccagc tgcttttttt tacctatgtt aattccaatc ctttcttgcc 8220
tcttccagat ccagataatg cagaaactca ttaactcagt gcaaaactat gcctgtgggca 8280
gcaaaacggc gttgactgaa ctttatggta tggaaaatcc gtccagccag ccgatggccg 8340
agctgtggat gggcgcacat ccgaaaagca gttcacgagt gcagaatgcc gccggagata 8400
tcgtttcact gcgtgatgtg attgagagtg ataaatcgac tctgcttcgga gaggccgttg 8460
ccaaacgctt tggcgaactg cctttcctgt tcaaagtatt atgcgcagca cagccactct 8520
ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg ttttgccaaa gaaaatgccg 8580
caggtatccc gatggatgcc gccgagcgta actataaaga tcctaaccac aagccggagc 8640
tggtttttgc gctgacgcct ttccttgcga tgaacgcgtt tcgtgaattt tccgagattg 8700
tctccctact ccagccggtc gcaggtgcac atccggcgat tgctcacttt ttacaacagc 8760
ctgatgccga acgtttaagc gaactgttcg ccagcctgtt gaatatgcag ggtgaagaaa 8820
aatcccgcgc gctggcgatt ttaaaatcgg cctcgatag ccagcagggt gaaccgtggc 8880
aaacgattcg tttaatttct gaattttacc cggaagacag cggtctgttc tccccgctat 8940
tgctgaatgt ggtgaaattg aaccctggcg aagcgatgtt cctgttcgct gaaacaccgc 9000
acgcttacct gcaaggcgtg gcgctggaag tgatggcaaa ctccgataac gtgctgcgtg 9060
cgggtctgac gcctaaatac attgatattc cggaactggt tgccaatgtg aaattcgaag 9120
ccaaaccggc taaccagttg ttgacccagc cggtgaaaca aggtgcagaa ctggacttcc 9180
cgattccagt ggatgatttt gccttctcgc tgcatgacct tagtgataaa gaaaccacca 9240
ttagccagca gagtgccgcc attttgttct gcgtcgaagg cgatgcaacg ttgtggaaag 9300
gtttctcagca gttacagctt aaaccgggtg aatcagcgtt tattgccgcc aacgaatcac 9360
cggtgactgt caaaggccac ggccgtttag cgcgtgttta caacaagctg taagagctta 9420
ctgaaaaaat taacatctct tgctaagctg ggagctctag atcccgaat ttcccgatc 9480
gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt cttgcgatga 9540
ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg taatgcatga 9600
cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt taatacgcga 9660
tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg tcatctatgt 9720
tactagatcg ggaattggcg agctcgaatt aattcagtac attaaaaacg tccgcaatgt 9780
gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc caccagccag 9840
ccaacagctc cccgaccggc agctcggcac aaaatcacca ctcgatacag gcagcccatc 9900

agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg cagactttgc tcatgttacc 9960

gatgctattc ggaagaacgg caactaagct gccgggtttg aaacacggat gatctcgcgg 10020

agggtagcat gttgattgta acgatgacag agcgttgctg cctgtgatca aatatcatct 10080

ccctcgcaga gatccgaatt atcagccttc ttattcattt ctcgcttaac cgtgacaggc 10140

tgtcgatctt gagaactatg ccgacataat aggaaatcgc tggataaagc cgctgaggaa 10200

gctgagtggc gctatttctt tagaagtgaa cgttgacgat cgtcgaccgt acccccgatga 10260

attaattcgg acgtacgttc tgaacacagc tggatactta cttgggcgat tgtcatacat 10320

gacatcaaca atgtacccgt ttgtgtaacc gtctcttgga ggttcgtatg acactagtgg 10380

ttcccctcag cttgcgacta gatgttgagg cctaacattt tattagagag caggctagtt 10440

gcttagatac atgatcttca ggccgttatc tgtcagggca agcgaaaatt ggccatttat 10500

gacgaccaat gccccgcaga agctcccatc tttgccgcca tagacgccgc gccccccttt 10560

tggggtgtag aacatccttt tgccagatgt ggaaaagaag ttcgttgtcc cattgttggc 10620

aatgacgtag tagccggcga aagtgcgaga cccatttgcg ctatatataa gcctacgatt 10680

tccgttgcga ctattgtcgt aattggatga actattatcg tagttgctct cagagttgtc 10740

gtaatttgat ggactattgt cgtaattgct tatggagttg tcgtagttgc ttggagaaat 10800

gtcgtagttg gatggggagt agtcataggg aagacgagct tcatccacta aaacaattgg 10860

caggtcagca agtgcctgcc ccgatgccat cgcaagtacg aggcttagaa ccaccttcaa 10920

cagatcgcgc atagtcttcc ccagctctct aacgcttgag ttaagccgcg ccgcgaagcg 10980

gcgtcggctt gaacgaattg ttagacatta tttgccgact accttggtga tctcgccttt 11040

cacgtagtga acaaattctt ccaactgatc tgcgcgcgag gccaagcgat cttcttgtcc 11100

aagataagcc tgcctagctt caagtatgac gggctgatac tgggccggca ggcgctccat 11160

tgcccagtcg gcagcgacat ccttcggcgc gattttgccg gttactgcgc tgtaccaaat 11220

gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg gcgagttcca 11280

tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag 11340

ttcctccgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg tcagcaagat 11400

agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt cattgcgctg 11460

ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga tgtcgtcgtg 11520

cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg aagccgaagt 11580

ttccaaaagg tcgttgatca aagctgccgc cgttgtttca tcaagcctta cggtcaccgt 11640

aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg agccgtacaa 11700

atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta cctctgatag 11760

ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactcct gaattaagcc 11820

gcgccgcgaa gcggtgtcgg cttgaatgaa ttgttaggcg tcatcctgtg ctcccgagaa 11880
ccagtaccag tacatcgctg tttcgttcga gacttgaggt ctagttttat acgtgaacag 11940
gtcaatgccg ccgagagtaa agccacattt tgcgtacaaa ttgcaggcag gtacattgtt 12000
cgtttgtgtc tctaatcgta tgccaaggag ctgtctgctt agtgcccact ttttcgcaaa 12060
ttcgatgaga ctgtgcgcga ctcctttgcc tcggtgcgtg tgcgacacaa caatgtgttc 12120
gatagaggct agatcgttcc atgttgagtt gagttcaatc ttcccgacaa gctcttggtc 12180
gatgaatgcg ccatagcaag cagagtcttc atcagagtca tcatccgaga tgtaatcctt 12240
ccggtagggg ctcacacttc tggtagatag ttcaaagcct tggtcggata ggtgcacatc 12300
gaacacttca cgaacaatga aatggttctc agcatccaat gtttccgcca cctgctcagg 12360
gatcaccgaa atcttcatat gacgcctaac gcctggcaca gcggatcgca aacctggcgc 12420
ggcttttggc acaaaaggcg tgacaggttt gcgaatccgt tgctgccact tgttaaccct 12480
tttgccagat ttggtaacta taatttatgt tagaggcgaa gtcttgggta aaaactggcc 12540
taaaattgct ggggatttca ggaaagtaaa catcaccttc cggctcgatg tctattgtag 12600
atatatgtag tgtatctact tgatcggggg atctgctgcc tcgcgcgttt cggtgatgac 12660
ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct gtaagcggat 12720
gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg tcggggcgca 12780
gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat gcggcatcag 12840
agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga tgcgtaagga 12900
gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 12960
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat 13020
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 13080
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa 13140
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 13200
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 13260
ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca 13320
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg 13380
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat 13440
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 13500
cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct 13560
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 13620
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 13680
aaggatctca agaagatcct ttgatctttt ctacgggctc tgacgctcag tggaacgaaa 13740

actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 13800
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca 13860
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 13920
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 13980
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 14040
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 14100
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 14160
acgttgttgc cattgctgca gggggggggg gggggggttt ccattgttca ttccacggac 14220
aaaaacagag aaaggaaacg acagaggcca aaaagctcgc tttcagcacc tgtcgtttcc 14280
tttctttcca gagggtattt taaataaaaa cattaagtta tgacgaagaa gaacggaaac 14340
gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg ccccgtaacc 14400
tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca tatcacaacg 14460
tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat cgtattaatt 14520
gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg acactgaata 14580
cggggcaacc tcatgtcccc cccccccccc ccctgcaggc atcgtggtgt cacgctcgtc 14640
gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 14700
catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 14760
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 14820
atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 14880
tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag 14940
cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 15000
cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 15060
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 15120
aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 15180
ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 15240
aaataaacaa atagggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 15300
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct 15360
tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt cgtggagttc ccgccacaga 15420
cccggattga aggcgagatc cagcaactcg cgccagatca tcctgtgacg gaactttggc 15480
gcgtgatgac tggccaggac gtcggccgaa agagcgacaa gcagatcacg cttttcgaca 15540
gcgtcggatt cgcgatcgag gatttttcgg cgctgcgcta cgtccgcgac cgcgttgagg 15600
gatcaagcca cagcagccca ctcgaccttc tagccgaccc agacgagcca agggatcttt 15660

ttggaatgct gctccgtcgt caggctttcc gacgtttggg tggttgaaca gaagtcatta 15720
tcgcacggaa tgccaagcac tcccgagggg aaccctgtgg ttggcatgca catacaaatg 15780
gacgaacgga caaacctttt cacgcccttt taaatatccg attattctaa taaacgctct 15840
tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc 15900
gggaaacgac aacctgatca tgagcggaga attaagggag tcacgttatg acccccgccg 15960
atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc 16020
actcagctta attaagtcta actcgagtta ctggtacgta ccaaatccat ggaatcaagg 16080
taccgtcgac tctagtaacg gccgccagtg tgctggaatt aattcggctt gtcgaccacc 16140
caacccccata tcgacagagg atgtgaagaa caggtaaatc acgcagaaga acccatctct 16200
gatagcagct atcgattaga acaacgaatc catattgggt ccgtgggaaa tacttactgc 16260
acaggaaggg ggcgatctga cgaggccccg ccaccggcct cgacccgagg ccgaggccga 16320
cgaagcgccg gcgagtacgg cgccgcggcg gcctctgccc gtgccctctg cgcgtgggag 16380
ggagaggccg cggtggtggg ggcgcgcgcg cgcgcgcgcg cagctggtgc ggcggcgcgg 16440
gggtcagccg ccgagccggc ggcgacggag gagcagggcg gcgtggacgc gaacttccga 16500
tcggttggtc agagtgcgcg agttgggctt agccaattag gtctcaacaa tctattgggc 16560
cgtaaaattc atgggccctg gtttgtctag gcccaatatc ccgttcattt cagcccacaa 16620
atatttcccc agaggattat taaggcccac acgcagctta tagcagatca agtacgatgt 16680
ttcctgatcg ttggatcgga aacgtacggt cttgatcagg catgccgact tcgtcaaaga 16740
gaggcggcat gacctgacgc ggagttggtt ccgggcaccg tctggatggt cgtaccgggа 16800
ccggacacgt gtcgcgcctc caactacatg gacacgtgtg gtgctgccat tgggccgtac 16860
gcgtggcggt gaccgcaccg gatgctgcct cgcaccgcct tgcccacgct ttatatagag 16920
aggttttctc tccattaatc gcatagcgag tcgaatcgac cgaaggggag ggggagcgaa 16980
gctttgcgtt ctctaatcgc ctcgtcaagg caactaatca atcacctcgt cctaatcctc 17040
gaatctctcg tggtgcccgt ctaatctcgc gattttgatg ctcgtggtgg aaagcgtagg 17100
aggatcccgt gcgagttagt ctcaatctct cagggtttcg tgcgatttta gggtgatcca 17160
cctcttaatc gagttacggt ttcgtgcgat tttagggtaa tcttcttaat ctctcattga 17220
tttagggttt cgtgagaatc gaggtaggga tctgtgttat ttatatcgat ctaatagatg 17280
gattggtttt gagattgttc tgtcagatgg ggattgtttc gatatattac cctaatgatg 17340
tgtcagatgg ggattgttcc gatatattac cctaatgatg tgtcagatgg ggattgtttc 17400
gatatattac cctaatgatg gataataaga gtagttcaca gttatgtttt gatcctgcca 17460
catagtttga gttttgtgat cagatttagt tttacttatt tgtgcttagt tcggatggga 17520
ttgttctgat attgttccaa tagatgaata gctcgttagg ttaaaatctt taggttgagt 17580

taggcgacac atagtttatt tcctctggat ttggattgga attgtgttct tagtttttt 17640
ccccctggatt tggattggaa ttgtgtggag ctgggttaga gaattacatc tgtatcgtgt 17700
acacctactt gaactgtaga gcttgggttc taaggtcaat ttaatctgta ttgtatctgg 17760
ctctttgcct agttgaactg tagtgctgat gttgtactgt gttttttac cggttttatt 17820
tgctttactc gtgcaaatca aatctgtcag atgctagaac taggtggctt tattctgtgt 17880
tcttacatag atctgttgtc ctgtagttac ttatgtcagt tttgttatta tctgaagata 17940
tttttggttg ttgcttgttg atgtggtgtg agctgtgagc agcgctctta tgattaatga 18000
tgctgtccaa ttgtagtgta gtatgatgtg attgatatgt tcatctattt tgagctgaca 18060
gtaccgatat cgtaggatct ggtgccaact tattctccag ctgctttttt ttacctatgt 18120
taattccaat cctttcttgc ctcttccaga tccagataat gctggaggac aagtctccca 18180
aactgcctga ttataagaac gaccttctgt acgaacgcac attcgacgag gggctctgct 18240
tcccgtggca cacgtgcgaa gattcaggag ggaaatgcga ttttgccgtg gtcgacgttc 18300
caggcgagcc tgggaacaag gcgttcaggc tcactgttat cgataagggt cagaacaagt 18360
ggtcggtcca aatgagacac cggggtatca cgttggagca ggggcacaca tacaccgttc 18420
ggtttactat ctggagcgac aagagctgcc gcgtgtatgc caaaatcggc caaatgggtg 18480
aaccctacac ggagtactgg aacaataact ggaatccgtt caacctcact ccggggcaga 18540
aattgacggt ggaacagaac tttactatga attatcccac ggacgacacg tgtgagttta 18600
ccttccactt gggaggggaa ctggcagccg ggaccccttа ctacgtgtac ctcgacgacg 18660
tttctcttta cgatccccgc tttgtcaagc cagtggaata cgtcctgcct caaccggatg 18720
tcagggttaa tcaagttgga tacctccctt ttgctaagaa atatgctact gtcgtgtcat 18780
cgagcacgtc cccattgaag tggcaacttc tgaatagtgc aaaccaagtt gtcttggagg 18840
gcaatacaat ccccaaggga ctggacaaag attcacaaga ctacgttcat tggatcgatt 18900
tctcgaactt taagaccgaa ggcaaggggt actatttcaa gttgcccact gtgaactccg 18960
atactaacta ctcccacccg tttgatattt ctgcagatat ctattcaaag atgaagttcg 19020
acgcgctcgc tttcttttac cataaaaggt cgggaatacc aatcgagatg ccctacgccg 19080
ggggagagca gtggacaagg ccgcagggc acattggtgt cgcgccgaac aagggcgaca 19140
cgaatgtgcc aacttggccc caggatgacg aatatgctgg acgccccag aaatactata 19200
cgaaagacgt gaccggcggg tggtacgatg ccggtgacca cggcaagtac gtcgtgaacg 19260
ggggtatcgc agtttggacc cttatgaata tgtacgagag agcaaagatt agaggaatcg 19320
ctaaccaggg tgcctacaaa gatggaggaa tgaatatccc ggaaaggaat aacggctatc 19380
ctgatattct ggacgaggcc agatgggaga tcgaattttt taagaagatg caagtcactg 19440
agaaagaaga tccgtcgatt gcaggtatgg tgcaccacaa gatccacgat ttcaggtgga 19500

cggcgctcgg aatgttgcct cacgaggacc cccagccacg ctaccttcgg cccgtcagca 19560

cagcggcaac cctgaatttc gcagcgaccc tcgctcagtc tgccagattg tggaaggatt 19620

acgaccccgac ttttgcagcg gactgccttg agaaagctga aattgcctgg caagcagcac 19680

tcaaacaccc ggacatctac gctgagtaca cgccaggaag cggtgggccg ggtggaggtc 19740

cttataatga cgattatgtc ggggacgagt tctactgggc cgcttgtgaa ctctatgtga 19800

caaccggtaa ggatgagtac aagaattact tgatgaatag tccgcactat ctggaaatgc 19860

cagcgaagat gggcgagaac ggagggggcta acggcgagga caacggtctc tggggctgct 19920

ttacttgggg aacgacacag gggttgggta caattaccct tgccctcgtt gaaaacggcc 19980

tccctteggc ggatattcaa aaggcccgca acaatatcgc taaagccgca gataagtggc 20040

ttgagaatat tgaagaacaa ggttacccgcc tgcctatcaa acaagcggag gatgaacggg 20100

gcggataccc gtggggtagt aattctttca ttctcaacca gatgatcgtc atgggctacg 20160

cttacgactt cacgggaaac agcaagtatc ttgacgggat gcaggacggc atgtcctccc 20220

tgctcggtag aaacggactt gatcaatcgt acgttactgg gtacgggggag aggccacttc 20280

agaaccccca cgaccgcttt tggacccctc aaacttcgaa gaaattcccg gccccacccc 20340

ctggtattat cgcaggcggg ccgaatagcc ggtttgaaga tccaacgatc actgcagcgg 20400

ttaagaagga tacacccccg cagaagtgct atattgacca caccgattcc tggtctacta 20460

acgagatcac gattaattgg aacgcccccct tcgcgtgggt cacagcgtat ctggacgaaa 20520

ttgacttgat tacccccacc ggcggagtgg acccctgaaga gccggaagtt atctacggtg 20580

attgtaacgg cgacggaaag gttaattcga ccgatgctgt ggcccttaaa aggtatatcc 20640

tccgcagcgg tatctcgatc aacacggaca acgcggacgt taatgcagat ggtcgcgtga 20700

atagcactga cctcgctatt ttgaagcgct atattttgaa ggagatcgat gttcttcctc 20760

acaagtgacc taggtccccg aatttccccg atcgttcaaa catttggcaa taaagtttct 20820

taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg ttgaattacg 20880

ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga 20940

ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact 21000

aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaattg c 21051

<210> 81
<211> 17129
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2342

<400> 81
cgtcgacttt agtaacggcc gccagtgtgc tggaattaat tcggcttgtc gaccacccaa 60

ccccatatcg acagaggatg tgaagaacag gtaaatcacg cagaagaacc catctctgat 120

agcagctatc gattagaaca acgaatccat attgggtccg tgggaaatac ttactgcaca 180

ggaaggggge gatctgacga ggccccgcca ccggcctcga cccgaggccg aggccgacga 240

agcgccgggg agtacggcgc cgcggcggcc tctgcccgtg ccctctgcgc gtgggaggga 300

gaggccgcgg tggtgggggc gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcgggag 360

tcagccgccg agccggcggc gacggaggag cagggcggcg tggacgcgaa cttccgatcg 420

gttggtcaga gtgcgcgagt tgggcttagc caattaggtc tcaacaatct attgggccgt 480

aaaattcatg ggccctggtt tgtctaggcc caatatcccg ttcatttcag cccacaaata 540

tttccccaga ggattattaa ggcccacacg cagcttatag cagatcaagt acgatgtttc 600

ctgatcgttg gatcggaaac gtacggtctt gatcaggcat gccgacttcg tcaaagagag 660

gcggcatgac ctgacgcgga gttggttccg ggcaccgtct ggatggtcgt accgggaccg 720

gacacgtgtc gcgcctccaa ctacatggac acgtgtggtg ctgccattgg gccgtacgcg 780

tggcggtgac cgcaccggat gctgcctcgc accgccttgc ccacgcttta tatagagagg 840

ttttctctcc attaatcgca tagcgagtcg aatcgaccga aggggagggg gagcgaagct 900

ttgcgttctc taatcgcctc gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa 960

tctctcgtgg tgcccgtcta atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg 1020

atcccgtgcg agttagtctc aatctctcag ggtttcgtgc gatttttaggg tgatccacct 1080

cttaatcgag ttacggtttc gtgcgatttt agggtaatcc tcttaatctc tcattgattt 1140

agggtttcgt gagaatcgag gtagggatct gtgttattta tatcgatcta atagatggat 1200

tggttttgag attgttctgt cagatgggga ttgtttcgat atattaccct aatgatgtgt 1260

cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 1320

atattaccct aatgatggat aataagagta gttcacagtt atgttttgat cctgccacat 1380

agtttgagtt ttgtgatcag atttagtttt acttatttgt gcttagttcg gatgggattg 1440

ttctgatatt gttccaatag atgaatagct cgttaggtta aaatctttag gttgagttag 1500

gcgacacata gtttatttcc tctggatttg gattggaatt gtgttcttag ttttttccc 1560

ctggatttgg attggaattg tgtggagctg ggttagagaa ttacatctgt atcgtgtaca 1620

cctacttgaa ctgtagagct tgggttctaa ggtcaattta atctgtattg tatctggctc 1680

tttgcctagt tgaactgtag tgctgatgtt gtactgtgtt tttttacccg ttttatttgc 1740

tttactcgtg caaatcaaat ctgtcagatg ctagaactag gtggctttat tctgtgttct 1800

tacatagatc tgttgtcctg tagttactta tgtcagtttt gttattatct gaagatattt 1860

ttggttgttg cttgttgatg tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc 1920

tgtccaattg tagtgtagta tgatgtgatt gatatgttca tctattttga gctgacagta 1980

ccgatatcgt aggatctggt gccaacttat tctccagctg ctttttttta cctatgttaa 2040

ttccaatcct ttcttgcctc ttccagatcc agataatgca aacaagcatt actctgacat 2100

ccaacgcatc cggtacgttt gacggttact attacgaact ctggaaggat actggcaata 2160

caacaatgac ggtctacact caaggtcgct tttcctgcca gtggtcgaac atcaataacg 2220

cgttgtttag gaccgggaag aaatacaacc agaattggca gtctcttggc acaatccgga 2280

tcacgtactc tgcgacttac aacccaaacg ggaactccta cttgtgtatc tatggctggt 2340

ctaccaaccc attggtcgag ttctacatcg ttgagtcctg ggggaactgg agaccgcctg 2400

gtgccacgtc cctgggccaa gtgacaatcg atggcgggac ctacgacatc tataggacga 2460

cacgcgtcaa ccagccttcc attgtgggga cagccacgtt cgatcagtac tggagcgtgc 2520

gcacctctaa gcggacttca ggaacagtga ccgtgaccga tcacttccgc gcctgggcga 2580

accggggcct gaacctcggc acaatagacc aaattacatt gtgcgtggag ggttaccaaa 2640

gcttctggatc agccaacatc acccagaaca ccttctctca gggctcttct tccggcagtt 2700

cgggtggctc atccggctcc acaacgacta ctcgcatcga gtgtgagaac atgtccttgt 2760

ccggacccta cgttagcagg atcaccaatc cctttaatgg tattgcgctg tacgccaacg 2820

gagacacagc ccgcgctacc gttaacttcc ccgcaagtcg caactacaat ttccgcctgc 2880

ggggttgcgg caacaacaat aatcttgccc gtgtggacct gaggatcgac ggacggaccg 2940

tcgggacctt ttattaccag ggcacatacc cctgggaggc cccaattgac aatgtttatg 3000

tcagtgcggg gagtcataca gtcgaaatca ctgttactgc ggataacggc acatgggacg 3060

tgtatgccga ctacctggtg atacagtgac ctaggtcccc gaatttcccc gatcgttcaa 3120

acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 3180

tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 3240

ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 3300

acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 3360

atcgggaatt ggaattccata ctaaagcttg catgcctgca ggtcgactct agtaacggcc 3420

gccagtgtgc tggaattaat tcggcttgtc gaccacccaa ccccatatcg acagaggatg 3480

tgaagaacag gtaaatcacg cagaagaacc catctctgat agcagctatc gattagaaca 3540

acgaatccat attgggtccg tgggaaatac ttactgcaca ggaagggggc gatctgacga 3600

ggccccgcca ccggcctcga cccgaggccg aggccgacga agcgccggcg agtacggcgc 3660

cgcggcggcc tctgcccgtg ccctctgcgc gtgggagggga gaggccgcgg tggtgggggc 3720

gcgcgcgcgc gcgcgcgcag ctggtgcggc ggcgcggggg tcagccgccg agccggcggc 3780

gacggaggag cagggcggcg tggacgcgaa cttccgatcg gttggtcaga gtgcgcgagt 3840

tgggcttagc caattaggtc tcaacaatct attgggccgt aaaattcatg ggccctggtt 3900

tgtctaggcc caatatcccg ttcatttcag cccacaaata tttcccaga ggattattaa 3960
ggcccacacg cagcttatag cagatcaagt acgatgtttc ctgatcgttg gatcggaaac 4020
gtacggtctt gatcaggcat gccgacttcg tcaaagagag gcggcatgac ctgacgcgga 4080
gttggttccg ggcaccgtct ggatggtcgt accgggaccg gacacgtgtc gcgcctccaa 4140
ctacatggac acgtgtggtg ctgccattgg gccgtacgcg tggcggtgac cgcaccggat 4200
gctgcctcgc accgccttgc ccacgcttta tatagagagg ttttctctcc attaatcgca 4260
tagcgagtcg aatcgaccga aggggagggg gagcgaagct ttgcgttctc taatcgcctc 4320
gtcaaggtaa ctaatcaatc acctcgtcct aatcctcgaa tctctcgtgg tgcccgtcta 4380
atctcgcgat tttgatgctc gtggtggaaa gcgtaggagg atcccgtgcg agttagtctc 4440
aatctctcag ggttttcgtgc gattttaggg tgatccacct cttaatcgag ttacggtttc 4500
gtgcgatttt agggtaatcc tcttaatctc tcattgattt agggtttcgt gagaatcgag 4560
gtagggatct gtgttattta tatcgatcta atagatggat tggttttgag attgttctgt 4620
cagatgggga ttgtttcgat atattaccct aatgatgtgt cagatgggga ttgtttcgat 4680
atattaccct aatgatgtgt cagatgggga ttgtttcgat atattaccct aatgatggat 4740
aataagagta gttcacagtt atgttttgat cctgccacat agtttgagtt ttgtgatcag 4800
atttagtttt acttatttgt gcttagttcg gatgggattg ttctgatatt gttccaatag 4860
atgaatagct cgttaggtta aaatctttag gttgagttag gcgacacata gtttatttcc 4920
tctggatttg gattggaatt gtgttcttag ttttttccc ctggatttgg attggaattg 4980
tgtggagctg ggttagagaa ttacatctgt atcgtgtaca cctacttgaa ctgtagagct 5040
tgggttctaa ggtcaattta atctgtattg tatctggctc tttgcctagt tgaactgtag 5100
tgctgatgtt gtactgtgtt tttttacccg ttttatttgc tttactcgtg caaatcaaat 5160
ctgtcagatg ctagaactag gtggctttat tctgtgttct tacatagatc tgttgtcctg 5220
tagttacttc tgtcagtttt gttattatct gaagatattt ttggttgttg cttgttgatg 5280
tggtgtgagc tgtgagcagc gctcttatga ttaatgatgc tgtccaattg tagtgtagta 5340
tgatgtgatt gatatgttca tctattttga gctgacagta ccgatatcgt aggatctggt 5400
gccaacttat tctccagctg ctttttttta cctatgttaa ttccaatcct ttcttgcctc 5460
ttccagatcc agataatgca gaaactcatt aactcagtgc aaaactatgc ctggggcagc 5520
aaaacggcgt tgactgaact ttatggtatg gaaaatccgt ccagccagcc gatggccgag 5580
ctgtggatgg gcgcacatcc gaaaagcagt tcacgagtgc agaatgccgc cggagatatc 5640
gtttcactgc gtgatgtgat tgagagtgat aaatcgactc tgctcggaga ggccgttgcc 5700
aaacgctttg gcgaactgcc tttcctgttc aaagtattat gcgcagcaca gccacttctc 5760
attcaggttc atccaaacaa acacaattct gaaatcggtt ttgccaaaga aaatgccgca 5820

ggtatcccga tggatgccgc cgagcgtaac tataaagatc ctaaccacaa gccggagctg 5880

gttttgcgc tgacgccttt ccttgcgatg aacgcgtttc gtgaattttc cgagattgtc 5940

tccctactcc agccggtcgc aggtgcacat ccggcgattg ctcacttttt acaacagcct 6000

gatgccgaac gtttaagcga actgttcgcc agcctgttga atatgcaggg tgaagaaaaa 6060

tcccgcgcgc tggcgatttt aaaatcggcc ctcgatagcc agcagggtga accgtggcaa 6120

acgattcgtt taatttctga attttacccg gaagacagcg gtctgttctc cccgctattg 6180

ctgaatgtgg tgaaattgaa ccctggcgaa gcgatgttcc tgttcgctga aacacgcac 6240

gcttacctgc aaggcgtggc gctggaagtg atggcaaact ccgataacgt gctgcgtgcg 6300

ggtctgacgc ctaaatacat tgatattccg gaactggttg ccaatgtgaa attcgaagcc 6360

aaaccggcta accagttgtt gacccagccg gtgaaacaag gtgcagaact ggacttcccg 6420

attccagtgg atgattttgc cttctcgctg catgaccttа gtgataaaga aaccaccatt 6480

agccagcaga gtgccgccat tttgttctgc gtcgaaggcg atgcaacgtt gtggaaaggt 6540

tctcagcagt tacagcttaa accgggtgaa tcagcgttta ttgccgccaa cgaatcaccg 6600

gtgactgtca aaggccacgg ccgtttagcg cgtgtttaca acaagctgta agagcttact 6660

gaaaaaatta acatctcttg ctaagctggg agctctagat ccccgaattt ccccgatcgt 6720

tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 6780

atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 6840

ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata 6900

gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 6960

ctagatcggg aattggcgag ctcgaattaa ttcagtacat taaaaacgtc cgcaatgtgt 7020

tattaagttg tctaagcgtc aatttgttta caccacaata tatcctgcca ccagccagcc 7080

aacagctccc cgaccggcag ctcggcacaa aatcaccact cgatacaggc agcccatcag 7140

tccgggacgg cgtcagcggg agagccgttg taaggcggca gactttgctc atgttaccga 7200

tgctattcgg aagaacggca actaagctgc cgggtttgaa acacggatga tcтcgcggag 7260

ggtagcatgt tgattgtaac gatgacagag cgttgctgcc tgtgatcaaa tatcatctcc 7320

ctcgcagaga tccgaattat cagccttctt attcatttct cgcttaaccg tgacaggctg 7380

tcgatcttga gaactatgcc gacataatag gaaatcgctg gataaagccg ctgaggaagc 7440

tgagtggcgc tatttcttta gaagtgaacg ttgacgatcg tcgaccgtac ccccgatgaat 7500

taattcggac gtacgttctg aacacagctg gatacttact tgggcgattg tcatacatga 7560

catcaacaat gtaccccgttt gtgtaaccgt ctcttggagg ttcgtatgac actagtggtt 7620

cccctcagct tgcgactaga tgttgaggcc taacatttta ttagagagca ggctagttgc 7680

ttagatacat gatcttcagg ccgttatctg tcagggcaag cgaaaattgg ccatttatga 7740

cgaccaatgc ccgcagaaag ctcccatctt tgccgccata gacgccgcgc cccctttttg 7800
gggtgtagaa catcctttg ccagatgtgg aaaagaagtt cgttgtccca ttgttggcaa 7860
tgacgtagta gccggcgaaa gtgcgagacc catttgcgct atatataagc ctacgatttc 7920
cgttgcgact attgtcgtaa ttggatgaac tattatcgta gttgctctca gagttgtcgt 7980
aatttgatgg actattgtcg taattgctta tggagttgtc gtagttgctt ggagaaatgt 8040
cgtagttgga tggggagtag tcataggaaa gacgagcttc atccactaaa acaattggca 8100
ggtcagcaag tgcctgcccc gatgccatcg caagtacgag gcttagaacc accttcaaca 8160
gatcgcgcat agtcttcccc agctctctaa cgcttgagtt aagccgcgcc gcgaagcggc 8220
gtcggcttga acgaattgtt agacattatt tgccgactac cttggtgatc tcgcctttca 8280
cgtagtgaac aaattcttcc aactgatctg cgcgcgaggc caagcgatct tcttgtccaa 8340
gataagcctg cctagcttca agtatgacgg gctgatactg ggccggcagg cgctccattg 8400
cccagtcggc agcgacatcc ttcggcgcga ttttgccggt tactgcgctg taccaaatgc 8460
gggacaacgt aagcactaca tttcgctcat cgccagccca gtcgggcggc gagttccata 8520
gcgttaaggt ttcatttagc gcctcaaata gatcctgttc aggaaccgga tcaaagagtt 8580
cctccgccgc tggacctacc aaggcaacgc tatgttctct tgcttttgtc agcaagatag 8640
ccagatcaat gtcgatcgtg gctggctcga agatacctgc aagaatgtca ttgcgctgcc 8700
attctccaaa ttgcagttcg cgcttagctg gataacgcca cggaatgatg tcgtcgtgca 8760
caacaatggt gacttctaca gcgcggagaa tctcgctctc tccagggggaa gccgaagttt 8820
ccaaaaggtc gttgatcaaa gctcgccgcg ttgtttcatc aagccttacg gtcaccgtaa 8880
ccagcaaatc aatatcactg tgtggcttca ggccgccatc cactgcggag ccgtacaaat 8940
gtacggccag caacgtcggt tcgagatggc gctcgatgac gccaactacc tctgatagtt 9000
gagtcgatac ttcggcgatc accgcttccc tcatgatgtt taactcctga attaagccgc 9060
gccgcgaagc ggtgtcggct tgaatgaatt gttaggcgtc atcctgtgct cccgagaacc 9120
agtaccagta catcgctgtt tcgttcgaga cttgaggtct agttttatac gtgaacaggt 9180
caatgccgcc gagagtaaag ccacattttg cgtacaaatt gcaggcaggt acattgttcg 9240
tttgtgtctc taatcgtatg ccaaggagct gtctgcttag tgcccacttt ttcgcaaatt 9300
cgatgagact gtgcgcgact cctttgcctc ggtgcgtgtg cgacacaaca atgtgttcga 9360
tagaggctag atcgttccat gttgagttga gttccaatct ccccgacaagc tcttggtcga 9420
tgaatgcgcc atagcaagca gagtcttcat cagagtcatc atccgagatg taatccttcc 9480
ggtaggggct cacacttctg gtagatagtt caaagccttg gtcggatagg tgcacatcga 9540
acacttcacg aacaatgaaa tggttctcag catccaatgt ttccgccacc tgctcaggga 9600
tcaccgaaat cttcatatga cgcctaacgc ctggcacagc ggatcgcaaa cctggcgcgg 9660

cttttggcac aaaaggcgtg acaggtttgc gaatccgttg ctgccacttg ttaaccettt 9720
tgccagattt ggtaactata atttatgtta gaggcgaagt cttgggtaaa aactggccta 9780
aaattgctgg ggatttcagg aaagtaaaca tcaccttccg gctcgatgtc tattgtagat 9840
atatgtagtg tatctacttg atcgggggat ctgctgcctc gcgcgtttcg gtgatgacgg 9900
tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc 9960
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc 10020
catgacccag tcacgtagcg atagcggagt gtatactggc ttaactatgc ggcatcagag 10080
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 10140
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 10200
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 10260
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 10320
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 10380
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 10440
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 10500
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 10560
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 10620
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 10680
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 10740
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc 10800
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 10860
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 10920
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 10980
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta 11040
aatttaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 11100
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 11160
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 11220
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac 11280
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 11340
tctattaatt gttgccgggg agctagagta agtagttcgc cagttaatag tttgcgcaac 11400
gttgttgcca ttgctgcagg gggggggggg gggggttcc attgttcatt ccacggacaa 11460
aaacagagaa aggaaacgac agaggccaaa aagctcgctt tcagcacctg tcgtttcctt 11520
tcttttcaga gggtatttta aataaaaaca ttaagttatg acgaagaaga acggaaacgc 11580

cttaaacggg aaaatttttca taaatagcga aaacccgcga ggtcgccgcc ccgtaacctg 11640

tcggatcacc ggaaaggacc cgtaaagtga taatgattat catctacata tcacaacgtg 11700

cgtggaggcc atcaaaccac gtcaaataat caattatgac gcaggtatcg tattaattga 11760

tctgcatcaa cttaacgtaa aaacaacttc agacaataca aatcagcgac actgaatacg 11820

gggcaacctc atgtcccccc cccccccccc ctgcaggcat cgtggtgtca cgctcgtcgt 11880

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatcccccc 11940

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 12000

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 12060

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 12120

tgcggcgacc gagttgctct tgcccggcgt caacacggga taataccgcg ccacatagca 12180

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 12240

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 12300

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 12360

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 12420

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 12480

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 12540

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc 12600

aagaattggt cgacgatctt gctgcgttcg gatattttcg tggagttccc gccacagacc 12660

cggattgaag gcgagatcca gcaactcgcg ccagatcatc ctgtgacgga actttggcgc 12720

gtgatgactg gccaggacgt cggccgaaag agcgacaagc agatcacgct tttcgacagc 12780

gtcggatttg cgatcgagga ttttcggcgc ctgcgctacg tccgcgaccg cgttgaggga 12840

tcaagccaca gcagcccact cgaccttcta gccgacccag acgagccaag ggatcttttt 12900

ggaatgctgc tccgtcgtca ggctttccga cgtttgggtg gttgaacaga agtcattatc 12960

gcacggaatg ccaagcactc ccgagggaaa ccctgcggtt ggcatgcaca tacaaatgga 13020

cgaacggata aaccttttca cgcccttta aatatccgat tattctaata aacgctcttt 13080

tctcttaggt ttacccgcca atatatcctg tcaaacactg atagtttaaa ctgaaggcgg 13140

gaaacgacaa cctgatcatg agcggagaat taagggagtc acgttatgac ccccgccgat 13200

gacgcgggac aagccgtttt acgtttggaa ctgacagaac cgcaacgttg aaggagccac 13260

tcagcttaat taagtctaac tcgagttact ggtacgtacc aaatccatgg aatcaaggta 13320

ccgtcgactc tagtaacggc cgccagtgtg ctggaattaa ttcggcttgt cgaccaccca 13380

acccccatatc gacagaggat gtgaagaaca ggtaaatcac gcagaagaac ccatctctga 13440

cagcagctat cgattagaac aacgaatcca tattgggtcc gtgggaaata cttactgcac 13500

aggaaggggg cgatctgacg aggccccgcc accggcctcg accccgaggcc gaggccgacg 13560
aagcgccggc gagtacggcg ccgcggcggc ctctgccccgt gccctctgcg cgtgggaggg 13620
agaggccgcg gtggtggggg cgcgcgcgcg cgcgcgcgca gctggtgcgg cggcgcgggg 13680
gtcagccgcc gagccggcgg cgacggagga gcagggcggc gtggacgcga acttccgatc 13740
ggttggtcag agtgcgcgag ttgggcttag ccaattaggt ctcaacaatc tattgggccg 13800
taaaattcat gggccctggt ttgtctaggc ccaatatccc gttcatttca gccccacaaat 13860
atttccccag aggattatta aggcccacac gcagcttata gcagatcaag tacgatgttt 13920
cctgatcgtt ggatcggaaa cgtacggtct tgatcaggca tgccgacttc gtcaaagaga 13980
ggcggcatga cctgacgcgg agttggttcc gggcaccgtc tggatggtcg taccgggacc 14040
ggacacgtgt cgcgcctcca actacatgga cacgtgtggt gctgccattg ggccgtacgc 14100
gtggcggtga ccgcaccgga tgctgcctcg caccgccttg cccacgcttt atatagagag 14160
gttttctctc cattaatcgc atagcgagtc gaatcgaccg aaggggaggg ggagcgaagc 14220
tttgcgttct ctaatcgcct cgtcaaggta actaatcaat cacctcgtcc taatcctcga 14280
atctctcgtg gtgcccgtct aatctcgcga ttttgatgct cgtggtggaa agcgtaggag 14340
gatcccgtgc gagttagtct caatctctca gggtttcgtg cgattttagg gtgatccacc 14400
tcttaatcga gttacggttt cgtgcgattt tagggtaatc ctcttaatct ctcattgatt 14460
tagggtttcg tgagaatcga ggtagggatc tgtgttattt atatcgatct aatagatgga 14520
ttggttttga gattgttctg tcagatgggg attgtttcga tatattaccc taatgatgtg 14580
tcagatgggg attgtttcga tatattaccc taatgatgtg tcagatgggg attgtttcga 14640
tatattaccc taatgatgga taataagagt agttcacagt tatgttttga tcctgccaca 14700
tagtttgagt tttgtgatca gatttagttt tacttatttg tgcttagttc ggatgggatt 14760
gttctgatat tgttccaata gatgaatagc tcgttaggtt aaaatcttta ggttgagtta 14820
ggcgacacat agtttatttc ctctggattt ggattggaat tgtgttctta gttttttccc 14880
cctggatttg gattggaatt gtgtggagct gggttagaga attacatctg tatcgtgtac 14940
acctacttga actgtagagc ttgggttcta aggtcaattt aatctgtatt gtatctggct 15000
ctttgcctag ttgaactgta gtgctgatgt tgtactgtgt ttttttaccc gttttatttg 15060
ctttactcgt gcaaatcaaa tctgtcagat gctagaacta ggtggcttta ttctgtgttc 15120
ttacatagat ctgttgtcct gtagttactt atgtcagttt tgttattatc tgaagatatt 15180
tttggttgtt gcttgttgat gtggtgtgag ctgtgagcag cgatcttatg attaatgatg 15240
ctgtccaatt gtagtgtagt atgatgtgat tgatatgttc atctattttg agctgacagt 15300
accgatatcg taggatctgg tgccaactta ttctccagct gcttttttct acctatgtta 15360
attccaatcc tttcttgcct cttccagatc cagataatgg cccacgcccg cgtcctccta 15420

ctggcgctcg ccgtcctggc caccgccgcc gtcgccgtcg cctcctcctc ctccttcgcc 15480

gactccaacc cgatccgccc ggtgaccgac cgcgccgcct ccaccgctta cgactacaag 15540

caggtgttgc gggactcgct actattctat gaggcccaga gatccggccg gctcccagcc 15600

gaccagaagg tcacgtggag gaaggatagc gcgctgaatg accagggtga ccagggacaa 15660

gacttgaccg gcggctactt tgacgctggg gacttcgtca agttcgggtt ccccatggct 15720

tataccgcaa ccgtgctggc atggggcctc atagattttg aggccggcta cagcagtgcc 15780

ggggccttgg atgatggacg gaaggctgtc aaatgggcca ccgactattt cataaaggcc 15840

cacacaagtc aaaatgagtt ctatggtcag gtcggccagg gtgacgccga tcacgctttc 15900

tggggaagac cagaggatat gacgatggcg cgcccggcgt acaagataga cacctcaagg 15960

cctggctctg atctggcagg cgagacagcg gctgctcttg ccgctgcttc aatcgtgttc 16020

cggaacgtcg atggcactta ctcaaataac ctgttaacac acgctcgcca gctattcgac 16080

ttcgcgaaca actaccgggg aaagtatagt gactctatta ctgacgcaag aaatttctac 16140

gcaagcgcag actacagaga cgagttggtt tgggctgctg cgtggttata cagagcgacc 16200

aacgacaaca cctacctcaa cactgctgag tcactgtacg atgagtttgg gctacagaac 16260

tgggggggcg gcctgaactg ggatagcaag gtgtctggcg tgcaggtgtt gttggccaag 16320

cttaccaata agcaggccta caaggacacg gtgcagtctt acgtcaatta cctaattaat 16380

aaccagcaga agactcccaa gggcctcctc tacatcgaca tgtgggcac ccttcgccac 16440

gctgccaacg ccgcattcat catgctcgaa gccgccgagc tgggcttgtc cgcctcctct 16500

tatagacagt tcgcgcaaac gcaaatcgac tacgccctgg gcgatggtgg ccgctccttt 16560

gtgtgcgggt tcgggagtaa tcctcctacg agaccgcacc acagatcctc gtcgtgcccg 16620

ccagctcccg ctacttgcga ctggaataca ttcaactcac ctgacccaaa ctaccacgtc 16680

ctctctgggg ccctagtggg cggacctgat cagaatgaca actacgtcga tgaccgttca 16740

gactatgttc acaacgaagt cgccactgat tacaacgcgg gtttccagtc cgcgttagct 16800

gctttggtgg ccttggttta cagcgagaag gacgagctgt gacctaggtc cccgaatttc 16860

cccgatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 16920

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 16980

tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 17040

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 17100

tctatgttac tagatcgggа attgggtac 17129

<210> 82
<211> 21841
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2345

<400> 82
tcgagttact ggtacgtacc aaatccatgg aatcaaggta ccgtcgactc tagtaacggc 60

cgccagtgtg ctggaattaa ttcggcttgt cgaccaccca accccatatc gacagaggat 120

gtgaagaaca ggtaaatcac gcagaagaac ccatctctga tagcagctat cgattagaac 180

aacgaatcca tattgggtcc gtgggaaata cttactgcac aggaaggggg cgatctgacg 240

aggccccgcc accggcctcg accgaggcc gaggccgacg aagcgccggc gagtacggcg 300

ccgcggcggc ctctgccgct gccctctgcg cgtgggaggg agaggccgcg gtggtggggg 360

cgcgcgcgcg cgcgcgcgca gctggtgcgg cggcgcgggg gtcagccgcc gagccggcgg 420

cgacggagga gcagggcggc gtggacgcga acttccgatc ggttggtcag agtgcgcgag 480

ttgggcttag ccaattaggt ctcaacaatc tattgggccg taaaattcat gggcccctggt 540

ttgtctaggc ccaatatccc gttcatttca gcccacaaat atttccccag aggattatta 600

aggcccacac gcagcttata gcagatcaag tacgatgttt cctgatcgtt ggatcggaaa 660

cgtacggtct tgatcaggca tgccgacttc gtcaaagaga ggcggcatga cctgacgcgg 720

agttggttcc gggcaccgtc tggatggtcg taccgggacc ggacacgtgt cgcgcctcca 780

actacatgga cacgtgtggt gctgccattg ggccgtacgc gtggcggtga ccgcaccgga 840

tgctgcctcg caccgccttg cccacgcttt atatagagag gttttctctc cattaatcgc 900

atagcgagtc gaatcgaccg aaggggaggg ggagcgaagc tttgcgttct ctaatcgcct 960

cgtcaaggta actaatcaat cacctcgtcc taatcctcga atctctcgtg gtgcccgtct 1020

aatctcgcga ttttgatgct cgtggtggaa agcgtaggag gatcccgtgc gagttagtct 1080

caatctctca gggtttcgtg cgattttagg gtgatccacc tcttaatcga gttacggttt 1140

cgtgcgattt tagggtaatc ctcttaatct ctcattgatt tagggtttcg tgagaatcga 1200

ggtagggatc tgtgttattt atatcgatct aatagatgga ttggttttga gattgttctg 1260

tcagatgggg attgtttcga tatattaccc taatgatgtg tcagatgggg attgtttcga 1320

tatattaccc taatgatgtg tcagatgggg attgtttcga tatattaccc taatgatgga 1380

taataagagt agttcacagt tatgttttga tcctgccaca tagtttgagt tttgtgatca 1440

gatttagttt tacttatttg tgcttagttc ggatgggatt gttctgatat tgttccaata 1500

gatgaatagc tcgttaggtt aaaatcttta ggttgagtta ggcgacacat agtttatttc 1560

ctctggattt ggattggaat tgtgttctta gttttttttcc cctggatttg gattggaatt 1620

gtgtggagct gggttagaga attacatctg tatcgtgtac acctacttga actgtagagc 1680

ttgggttcta aggtcaattt aatctgtatt gtatctggct ctttgcctag ttgaactgta 1740

gtgctgatgt tgtactgtgt ttttttaccc gttttatttg ctttactcgt gcaaatcaaa 1800

tctgtcagat gctagaacta ggtggcttta ttctgtgttc ttacatagat ctgttgtcct 1860
gtagttactt atgtcagttt tgttattatc tgaagatatt tttggttgtt gcttgttgat 1920
gtggtgtgag ctgtgagcag cgctcttatg attaatgatg ctgtccaatt gtagtgtagt 1980
atgatgtgat tgatatgttc atctattttg agctgacagt accgatatcg taggatctgg 2040
tgccaactta ttctccagct gtttttttt acctatgtta attccaatcc tttcttgcct 2100
cttccagatc cagataatgg cccacgcccg cgtcctcctc ctggcgctcg ccgtcctggc 2160
cacccgccgcc gtcgccgtcg cctcctcctc ctccttcgcc gactccaacc cgatccgccc 2220
ggtgaccgac cgcgccgcct ccaccgctta cgactacaag caggtgttgc gggactcgct 2280
actattctat gaggcccaga gatccggccg gctcccagcc gaccagaagg tcacgtggag 2340
gaaggatagc gcgctgaatg accagggtga ccagggacaa gacttgaccg gcggctactt 2400
tgacgctggg gacttcgtca agttcgggtt ccccatggct tataccgcaa ccgtgctggc 2460
atggggcctc atagattttg aggccggcta cagcagtgcc ggggccttgg atgatggacg 2520
gaaggctgtc aaatgggcca ccgactattt cataaaggcc cacacaagtc aaaatgagtt 2580
ctatggtcag gtcggccagg gtgacgccga tcacgctttc tggggaagac cagaggatat 2640
gacgatggcg cgcccggcgt acaagataga cacctcaagg cctggctctg atctggcagg 2700
cgagacagcg gctgctcttg ccgctgcttc aatcgtgttc cggaacgtcg atggcactta 2760
ctcaaataac ctgttaacac acgctcgcca gctattcgac ttcgcgaaca actaccgggg 2820
aaagtatagt gactctatta ctgacgcaag aaatttctac gcaagcgcag actacagaga 2880
cgagttggtt tgggctgctg cgtggttata cagagcgacc aacgacaaca cctacctcaa 2940
cactgctgag tcactgtacg atgagtttgg gctacagaac tgggggggggg gcctgaactg 3000
ggatagcaag gtgtctggcg tgcaggtgtt gttggccaag cttaccaata agcaggccta 3060
caaggacacg gtgcagtctt acgtcaatta cctaattaat aaccagcaga agactcccaa 3120
gggcctcctc tacatcgaca tgtggggcac ccttcgccac gctgccaacg ccgcattcat 3180
catgctcgaa gccgccgagc tgggcttgtc cgcctcctct tatagacagt tcgcgcaaac 3240
gcaaatcgac tacgccctgg gcgatggtgg ccgctccttt gtgtgcgggt tcgggagtaa 3300
tcctcctacg agaccgcacc acagatcctc gtcgtgcccg ccagctcccg ctacttgcga 3360
ctggaataca ttcaactcac ctgacccaaa ctaccacgtc ctctctgggg ccctagtggg 3420
cggacctgat cagaatgaca actacgtcga tgaccgttca gactatgttc acaacgaagt 3480
cgccactgat tacaacgcgg gtttccagtc cgcgttagct gcttggtgtg ccttggttta 3540
cagcgagaag gacgagctgt gacctaggtc ccgaatttc cccgatcgtt caaacatttg 3600
gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt 3660
tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag 3720

atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat 3780
atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcggga 3840
attgggtacc gtcgactcta gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg 3900
accacccaac cccatatcga cagaggatgt gaagaacagg taaatcacgc agaagaaccc 3960
atctctgata gcagctatcg attagaacaa cgaatccata ttgggtccgt gggaaatact 4020
tactgcacag gaaggggggcg atctgacgag gccccgccac cggcctcgac ccgaggccga 4080
ggccgacgaa gcgccggcga gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg 4140
tgggagggag aggccgcggt ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg 4200
gcgcgggggt cagccgccga gccggccggcg acggaggagc agggcggcgt ggacgcgaac 4260
ttccgatcgg ttggtcagag tgcgcgagtt gggcttagcc aattaggtct caacaatcta 4320
ttgggccgta aaattcatgg gccctggttt gtctaggccc aatatcccgt tcatttcagc 4380
ccacaaatat ttccccagag gattattaag gcccacacgc agcttatagc agatcaagta 4440
cgatgtttcc tgatcgttgg atcggaaacg tacggtcttg atcaggcatg ccgacttcgt 4500
caaagagagg cggcatgacc tgacgcggag ttggttccgg gcaccgtctg gatggtcgta 4560
ccgggaccgg acacgtgtcg cgcctccaac tacatggaca cgtgtggtgc tgccattggg 4620
ccgtacgcgt ggcggtgacc gcaccggatg ctgcctcgca ccgccttgcc cacgctttat 4680
atagagaggt tttctctcca ttaatcgcat agcgagtcga atcgaccgaa ggggaggggg 4740
agccgaagctt tgcgttctct aatcgcctcg tcaaggtaac taatcaatca cctcgtccta 4800
atcctcgaat ctctcgtggt gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag 4860
cgtaggagga tcccgtgcga gttagtctca atctctcagg gtttcgtgcg attttagggt 4920
gatccacctc ttaatcgagt tacggtttcg tgcgatttta gggtaatcct cttaatctct 4980
cattgattta gggtttcgtg agaatcgagg tagggatctg tgttatttat atcgatctaa 5040
tagatggatt ggttttgaga ttgttctgtc agatggggat tgtttcgata tattacccta 5100
atgatgtgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 5160
tgtttcgata tattacccta atgatggata ataagagtag ttcacagtta tgttttgatc 5220
ctgccacata gtttgagttt tgtgatcaga tttagtttta cttatttgtg cttagttcgg 5280
atgggattgt tctgatattg ttccaataga tgaatagctc gttaggttaa aatctttagg 5340
ttgagttagg cgacacatag tttatttcct ctggatttgg attggaattg tgttcttagt 5400
ttttttcccc tggatttgga ttggaattgt gtggagctgg gttagagaat tacatctgta 5460
tcgtgtacac ctacttgaac tgtagagctt gggttctaag gtcaatttaa tctgtattgt 5520
atctggctct ttgcctagtt gaactgtagt gctgatgttg tactgtgttt ttttacccgt 5580
tttatttgct ttactcgtgc aaatcaaatc tgtcagatgc tagaactagg tggctttatt 5640

ctgtgttctt acatagatct gttgtcctgt agttacttat gtcagtttg ttattatctg 5700

aagatatttt tggttgttgc ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat 5760

taatgatgct gtccaattgt agtgtagtat gatgtgattg atatgttcat ctattttgag 5820

ctgacagtac cgatatcgta ggatctggtg ccaacttatt ctccagctgc ttttttttac 5880

ctatgttaat tccaatcctt tcttgcctct tccagatcca gataatggcg aacaaacatt 5940

tgtccctctc cctcttcctc gtcctccttg gcctgtcggc cagcttggcc tccgggcaag 6000

tcttcccagc tggaaacgca acggaattgg agaaaagaca aaccacccct aactctgagg 6060

gctggcatga cggatactac tactcttggt ggagcgatgg tggtgcacag gccacctata 6120

caaacctcga aggcggcact tatgagattt catggggtga cggtggcaac cttgtcggcg 6180

gaaaggggtg gaaccccgga cttaacgcca gggcaatcca cttcgaaggg gtgtaccagc 6240

ccaatggcaa ctcatacctg gccgtctacg ggtggacgcg caatccgctg gttgagtact 6300

atatcgtgga gaatttcgga acttatgacc ctagctccgg tgccacggac ctcgggacag 6360

tcgagtgtga cggaagcatc tacaggctgg gtaaaactac ccgcgttaat gctccatcga 6420

tcgacggcac gcaaacattt gatcaatact ggtccgtgcg gcaggataag aggacaagcg 6480

gcacagttca gacgggttgc cactttgatg cctgggcaag agcggggctc aatgtgaatg 6540

gggaccacta ctatcagatt gtggcgaccg agggctattt ctccagtggc tatgcgcgta 6600

taaccgtcgc tgatgttgga agcgagaagg acgagctgtg acctaggtcc ccgaatttcc 6660

ccgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg 6720

cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat 6780

gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 6840

acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat 6900

ctatgttact agatcgggaa ttggaattca tactaaagct tgcatgcctg caggtcgact 6960

ctagtaacgg ccgccagtgt gctggaatta attcggcttg tcgaccaccc aaccccatat 7020

cgacagagga tgtgaagaac aggtaaatca cgcagaagaa cccatctctg atagcagcta 7080

tcgattagaa caacgaatcc atattgggtc cgtgggaaat acttactgca caggaagggg 7140

gcgatctgac gaggccccgc caccggcctc gaccgaggcc cgaggccgac gaagcgcggg 7200

cgagtacggc gccgcggcgg cctctgcccg tgcctctgcc gcgtgggagg gagaggccgc 7260

ggtggtgggg gcgcgcgcgc gcgcgcgcgc agctggtgcg gcggcgcggg ggtcagccgc 7320

cgagccggcg gcgaccgagg agcagggcgg cgtggacgcg aacttccgat cggttggtca 7380

gagtgcgcga gttgggctta gccaattagg tctcaacaat ctattgggcc gtaaaattca 7440

tgggcctgg tttgtctagg cccaatatcc cgttcatttc agcccacaaa tatttcccca 7500

gaggattatt aaggcccaca cgcagcttat agcagatcaa gtacgatgtt tcctgatcgt 7560

tggatcggaa acgtacggtc ttgatcaggc atgccgactt cgtcaaagag aggcggcatg 7620
acctgacgcg gagttggttc cgggcaccgt ctggatggtc gtaccgggac cggacacgtg 7680
tcgcgcctcc aactacatgg acacgtgtgg tgctgccatt gggccgtacg cgtggcggtg 7740
accgcaccgg atgcrgcctc gcaccgcctt gcccacgctt tatatagaga ggttttctct 7800
ccattaatcg catagcgagt cgaatcgacc gaaggggagg gggagcgaag ctttgcgttc 7860
tctaatcgcc tcgtcaaggt aactaatcaa tcacctcgtc ctaatcctcg aatctctcgt 7920
ggtgcccgtc taatctcgcg atttgatgc tcgtggtgga aagcgtagga ggatcccgtg 7980
cgagttagtc tcaatctctc agggtttcgt gcgattttag ggtgatccac ctcttaatcg 8040
agttacggtt tcgtgcgatt ttagggtaat cctcttaatc tctcattgat ttagggtttc 8100
gtgagaatcg aggtagggat ctgtgttatt tatatcgatc taatagatgg attggttttg 8160
agattgttct gtcagatggg gattgtttcg atatattacc ctaatgatgt gtcagatggg 8220
gattgtttcg atatattacc ctaatgatgt gtcagatggg gattgtttcg atatattacc 8280
ctaatgatgg ataataagag tagttcacag ttatgttttg atcctgccac atagtttgag 8340
ttttgtgatc agatttagtt ttacttattt gtgcttagtt cggatgggat tgttctgata 8400
ttgttccaat agatgaatag ctcgttaggt taaaatcttt aggttgagtt aggcgacaca 8460
tagtttattt cctctggatt tggattggaa ttgtgttctt agtttttttc ccctggattt 8520
ggattggaat tgtgtggagc tgggttagag aattacatct gtatcgtgta cacctacttg 8580
aactgtagag cttgggttct aaggtcaatt taatctgtat tgtatctggc tctttgccta 8640
gttgaactgt agtgctgatg ttgtactgtg ttttttttacc cgttttattt gctttactcg 8700
tgcaaatcaa atctgtcaga tgctagaact aggtggcttt attctgtgtt cttacataga 8760
tctgttgtcc tgtagttact tatgtcagtt ttgttattat ctgaagatat ttttggttgt 8820
tgcttgttga tgtggtgtga gctgtgagca gcgctcttat gattaatgat gctgtccaat 8880
tgtagtgtag tatgatgtga ttgatatgtt catctatttt gagctgacag taccgatatc 8940
gtaggatctg gtgccaactt attctccagc tgcttttttt tacctatgtt aattccaatc 9000
ctttcttgcc tcttccagat ccagataatg tagaaactca ttaattcagt gcaaaactat 9060
gcctggggca gcaaaacggc gttgactgaa ctttatggta tggaaaatcc gtccagccag 9120
ccgatggccg agctgtggat gggcgcacat ccgaaaagca gttcacgagt gcagaatgcc 9180
gccggagata tcgtttcact gcgtgatgtg attgaggagtg ataaatcgac tctgctcgga 9240
gaggccgttg ccaaacgctt tggcgaactg cctttcctgt tcaaagtatt atgcgcagca 9300
cagccactct ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg ttttgccaaa 9360
gaaaatgccg caggtatccc gatggatgcc gccgagcgta actataaaga tcctaaccac 9420
aagccggagc tggtttttgc gctgacgcct ttccttgcga tgaacgcgtt tcgtgaattt 9480

tccgagattg tctccctact ccagccggtc gcaggtgcac atccggcgat tgctcacttt 9540
ttacaacagc ctgatgccga acgtttaagc gaactgttcg ccagcctgtt gaatatgcag 9600
ggtgaagaaa aatcccgcgc gctggcgatt ttaaaatcgg cctccgatag ccagcagggt 9660
gaaccgtggc aaacgattcg tttaatttct gaattttacc cggaagacag cggtctgttc 9720
tccccgctat tgctgaatgt ggtgaaattg aaccctggcg aagcgatgtt cctgttcgct 9780
gaaacaccgc acgcttacct gcaaggcgtg gcgctggaag tgatggcaaa ctccgataac 9840
gtgctgcgtg cgggtctgac gcctaaatac attgatattc cggaactggt tgccaatgtg 9900
aaattcgaag ccaaaccggc taaccagttg ttgaacccagc cggtgaaaca aggtgcagaa 9960
ctggacttcc cgattccagt ggatgatttt gccttctcgc tgcatgacct tagtgataaa 10020
gaaaccacca ttagccagca gagtgccgcc attttgttct gcgtcgaagg cgatgcaacg 10080
ttgtggaaag gttctcagca gttacagctt aaaccgggtg aatcagcgtt tattgccgcc 10140
aacgaatcac cggtgactgt caaaggccac ggccgtttag cgcgtgttta caacaagctg 10200
taagagctta ctgaaaaaat taacatctct tgctaagctg ggagctctag atccccgaat 10260
ttccccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt 10320
cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg 10380
taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt 10440
taatacgcga tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg 10500
tcatctatgt tactagatcg ggaattggcg agctcgaatt aattcagtac attaaaaacg 10560
tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc 10620
caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca ctcgatacag 10680
gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg cagactttgc 10740
tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg aaacacggat 10800
gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg cctgtgatca 10860
aatatcatct cctcgcaga gatccgaatt atcagccttc ttattcattt ctcgcttaac 10920
cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc tggataaagc 10980
cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat cgtcgaccgt 11040
acccccgatga attaattcgg acgtacgttc tgaacacagc tggatactta cttgggcgat 11100
tgtcatacat gacatcaaca atgtacccgt ttgtgtaacc gtcccttgga ggttcgtatg 11160
acactagtgg ttcccctcag cttgcgacta gatgttgagg cctaacattt tattagagag 11220
caggctagtt gcttagatac atgatcttca ggccgttatc tgtcagggca agcgaaaatt 11280
ggccatttat gacgaccaat gcccgcaga agctcccatc tttgccgcca tagacgccgc 11340
gcccccgttt tggggtgtag aacatccttt tgccagatgt ggaaaagaag ttcgttgtcc 11400

cattgttggc aatgacgtag tagccggcga aagtgcgaga cccatttgcg ctatatataa 11460
gcctacgatt tccgttgcga ctattgtcgt aattggatga actattatcg tagttgctct 11520
cagagttgtc gtaatttgat ggactattgt cgtaattgct tatggagttg tcgtagttgc 11580
ttggagaaat gtcgtagttg gatggggagt agtcataggg aagacgagct tcatccacta 11640
aaacaattgg caggtcagca agtgcctgcc ccgatgccat cgcaagtacg aggcttagaa 11700
ccaccttcaa cagatcgcgc atagtcttcc ccagctctct aacgcttgag ttaagccgcg 11760
ccgcgaagcg gcgtcggctt gaacgaattg ttagacatta tttgccgact accttggtga 11820
tctcgccttt cacgtagtga acaaattctt ccaactgatc tgcgcgcgag gccaagcgat 11880
cttcttgtcc aagataagcc tgcctagctt caagtatgac gggctgatac tgggccggca 11940
ggcgctccat tgcccagtcg gcagcgacat ccttcggcgc gattttgccg gttactgcgc 12000
tgtaccaaat gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg 12060
gcgagttcca tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg 12120
gatcaaagag ttcctccgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg 12180
tcagcaagat agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt 12240
cattgcgctg ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga 12300
tgtcgtcgtg cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg 12360
aagccgaagt ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta 12420
cggtcaccgt aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg 12480
agccgtacaa atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta 12540
cctctgatag ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactcct 12600
gaattaagcc gcgccgcgaa gcggtgtcgg cttgaatgaa ttgttaggcg tcatcctgtg 12660
ctcccgagaa ccagtaccag tacatcgctg tttcgttcga gacttgaggt ctagttttat 12720
acgtgaacag gtcaatggcg ccgagagtaa agccacattt tgcgtacaaa ttgcaggcag 12780
gtacattgtt cgtttgtgtc tctaatcgta tgccaaggag ctgtctgctt agtgcccact 12840
ttttcgcaaa ttcgatgaga ctgtgcgcga ctcctttgcc tcggtgcgtg tgcgacacaa 12900
caatgtgttc gatagaggct agatcgttcc atgttgagtt gagttcaatc ttcccgacaa 12960
gctcttggtc gatgaatgcg ccatagcaag cagagtcttc atcagagtca tcatccgaga 13020
tgtaatcctt ccggtagggg ctcacacttc tggtagatag ttcaaagcct tggtcggata 13080
ggtgcacatc gaacacttca cgaacaatga aatggttctc agcatccaat gtttccgcca 13140
cctgctcagg gatcaccgaa atcttcatat gacgcctaac gcctggcaca gcggatcgca 13200
aacctggcgc ggcttttggc acaaaaggcg tgacaggttt gcgaatccgt tgctgccact 13260
tgttaaccct tttgccagat ttggtaacta taatttatgt tagaggcgaa gtcttgggta 13320

aaaactggcc taaaattgct ggggatttca ggaaagtaaa catcaccttc cggctcgatg 13380
tctattgtag atatatgtag tgtatctact tgatcggggg atctgctgcc tcgcgcgttt 13440
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 13500
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 13560
tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat 13620
gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga 13680
tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg 13740
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 13800
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 13860
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 13920
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 13980
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 14040
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 14100
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 14160
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 14220
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 14280
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 14340
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 14400
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 14460
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 14520
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 14580
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 14640
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 14700
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 14760
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 14820
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 14880
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 14940
agtttgcgca acgttgttgc cattgctgca gggggggggg gggggggttt ccattgttca 15000
ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc tttcagcacc 15060
tgtcgtttcc tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa 15120
gaacggaaac gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg 15180
ccccgtaacc tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca 15240

tatcacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat 15300
cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg 15360
acactgaata cggggcaacc tcatgtcccc ccccccccc cctgcaggc atcgtggtgt 15420
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 15480
catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca 15540
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 15600
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 15660
gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg 15720
cgccacatag cagaactttа aaagtgctca tcattggaaa acgttcttcg gggcgaaaac 15780
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 15840
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 15900
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 15960
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 16020
gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 16080
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 16140
cctttcgtct tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt cgtggagttc 16200
ccgccacaga cccggattga aggcgagatc cagcaactcg cgccagatca tcctgtgacg 16260
gaactttggc gcgtgatgac tggccaggac gtcggccgaa agagcgacaa gcagatcacg 16320
cttttcgaca gcgtcggatt tgcgatcgag gatttttcgg cgctgcgcta cgtccgcgac 16380
cgcgttgagg gatcaagcca cagcagcccа ctcgaccttc tagccgaccc agacgagcca 16440
agggatcttt ttggaatgct gctccgtcgt caggctttcc gacgtttggg tggttgaaca 16500
gaagtcatta tcgcacggaa tgccaagcac tcccgagggg aacccctgtgg ttggcatgca 16560
catacaaatg gacgaacgga taaacctttt cacgcccttt taaatatccg attattctaa 16620
taaacgctct tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagttta 16680
aactgaaggc gggaaacgac aacctgatca tgagcggaga attaagggag tcacgttatg 16740
acccccgccg atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt 16800
tgaaggagcc actcagctta attaagtcta actcgagtta ctgytacgta ccaaatccat 16860
ggaatcaagg taccgtcgac tctagtaacg gccgccagtg tgctggaatt aattcggctt 16920
gtcgaccacc caacccata tcgacagagg atgtgaagaa caggtaaatc acgcagaaga 16980
acccatctct gatagcagct atcgattaga acaacgaatc catattgggt ccgtgggaaa 17040
tacttactgc acaggaaggg ggcgatctga cgaggccccg ccaccggcct cgaccgagg 17100
ccgaggccga cgaagcgccg gcgagtacgg cgccgcggcg gcctctgccc gtgccctctg 17160

cgcgtgggag ggagaggccg cggtggtggg ggcgcgcgcg cgcgcgcgcg cagctggtgc 17220

ggcggcgcgg gggtcagccg ccgagccggc ggcgacggag gagcagggcg gcgtggacgc 17280

gaacttccga tcggttggtc agagtgcgcg agttgggctt agccaattag gtctcaacaa 17340

tctattgggc cgtaaaattc atgggccctg gtttgtctag gcccaatatc ccgttcattt 17400

cagcccacaa atatttcccc agaggattat taaggcccac acgcagctta tagcagatca 17460

agtacgatgt ttcctgatcg ttggatcgga aacgtacggt cttgatcagg catgccgact 17520

tcgtcaaaga gaggcggcat gacctgacgc ggagttggtc ccgggcaccg tctggatggt 17580

cgtaccggga ccggacacgt gtcgcgcctc caactacatg gacacgtgtg gtgctgccat 17640

tgggccgtac gcgtggcggt gaccgcaccg gatgctgcct cgcaccgcct tgcccacgct 17700

ttatatagag aggttttctc tccattaatc gcatagcgag tcgaatcgac cgaaggggag 17760

ggggagcgaa gctttgcgtt ctcgaatcgc ctcgtcaagg taactaatca atcacctcgt 17820

cctaatcctc gaatctctcg tggtgcccgt ctaatctcgc gattttgatg ctcgtggtgg 17880

aaagcgtagg aggatcccgt gcgagttagt ctcaatctct cagggtttcg tgcgatttta 17940

gggtgatcca cctcttaatc gagttacggt ttcgtgcgat tttagggtaa tcctcttaat 18000

ctctcattga tttagggttt cgtgagaatc gaggtaggga tctgtgttat ttatatcgat 18060

ctaatagatg gattggtttt gagattgttc tgtcagatgg ggattgtttc gatatattac 18120

cctaatgatg tgtcagatgg ggattgtttc gatatattac cctaatgatg tgtcagatgg 18180

ggattgtttc gatatattac cctaatgatg gataatcaga gtagttcaca gttatgtttt 18240

gatcctgcca catagtttga gttttgtgat cagatttagt tttacttatt tgtgcttagt 18300

tcggatggga ttgttctgat attgttccaa tagatgaata gctcgttagg ttaaaatctt 18360

taggttgagt taggcgacac atagtttatt tcctcuggat ttggattgga attgtgttct 18420

tagttttttt cccctggatt tggattggaa ttgtgtggag ctgggttaga gaattacatc 18480

tgtatcgtgt acacctactt gaactgtaga gcttgggttc taaggtcaat ttaatctgta 18540

ttgtatctgg ctctttgcct agttgaactg tagtgctgat gttgtactgt gttttttac 18600

ccgttttatt tgctttactc gtgcaaatca aatctgtcag atgctagaac taggtggctt 18660

tattctgtgt tcttacatag atctgttgtc ctgtagttac ttatgtcagt tttgttatta 18720

tctgaagata tttttggttg ttgcttgttg atgtggtgtg agctgtgagc agcgctctta 18780

tgattaatga tgctgtccaa ttgtagtgta gtatgargtg attgatatgt tcatctattt 18840

tgagctgaca gtaccgatat cgtaggatct ggtgccaact tattctccag ctgcttttt 18900

ttacctatgt taattccaat cctttcttgc ctcttccaga tccagataat gctggaggac 18960

aagtctccca aactgcctga ttataagaac gaccttctgt acgaacgcac attcgacgag 19020

gggctctgct tccgtggca cacgtgcgaa gattcaggag ggaaatgcga ttttgccgtg 19080

gtcgacgttc caggcgagcc tgggaacaag gcgttcaggc tcactgttat cgataagggt 19140
cagaacaagt ggtcggtcca aatgagacac cggggtatca cgttggagca ggggcacaca 19200
tacaccgttc ggtttactat ctggagcgac aagagctgcc gcgtgtatgc caaaatcggc 19260
caaatgggtg aaccctacac ggagtactgg aacaataact ggaatccgtt caacctcact 19320
ccggggcaga aattgacggt ggaacagaac tttactatga attatccgac ggacgacacg 19380
tgtgagttta ccttccactt gggaggggaa ctggcagccg ggacccctta ctacgtgtac 19440
ctcgacgacg tttctctttа cgatccccgc tttgtcaagc cagtggaata cgtcctgcct 19500
caaccggatg tcagggttaa tcaagttgga tacctccctt ttgctaagaa atatgctact 19560
gtcgtgtcat cgagcacgtc cccattgaag tggcaacttc tgaatagtgc aaaccaagtt 19620
gtcttggagg gcaatacaat ccccaaggga ctggacaaag attcacaaga ctacgttcat 19680
tggatcgatt tctcgaactt taagaccgaa ggcaaggggt actatttcaa gttgcccact 19740
gtgaactccg atactaacta ctcccacccg tttgatattt ctgcagatat ctattcaaag 19800
atgaagttcg acgcgctcgc tttcttttac cataaaaggt cgggaatacc aatcgagatg 19860
ccctacgccg ggggagagca gtggacaagg cccgcagggc acattggtgt cgcgccgaac 19920
aagggcgaca cgaatgtgcc aacttggccc caggatgacg aatatgctgg acgcccccag 19980
aaatactata cgaaagacgt gaccggcggg tggtacgatg ccggtgacca cggcaagtac 20040
gtcgtgaacg gggtatcgc agtttggacc cttatgaata tgtacgagag agcaaagatt 20100
agaggaatcg ctaaccaggg tgcctacaaa gatggaggaa tgaatatccc ggaaaggaat 20160
aacggctatc ctgatattct ggacgaggcc agatgggaga tcgaattttt taagaagatg 20220
caagtcactg agaaagaaga tccgtcgatt gcaggtatgg tgcaccacaa gatccacgat 20280
ttcaggtgga cggcgctcgg aatgttgcct cacgaggacc cccagccacg ctaccttcgg 20340
cccgtcagca cagcggcaac cctgaatttc gcagcgaccc tcgctcagtc tgccagattg 20400
tggaaggatt acgacccgac ttttgcagcg gactgccttg agaaagctga aattgcctgg 20460
caagcagcac tcaaacaccc ggacatctac gctgagtaca cgccaggaag cggtgggccg 20520
ggtggaggtc cttataatga cgattatgtc ggggacgagt tctactgggc cgcttgtgaa 20580
ctctatgtga caaccggtaa ggatgagtac aagaattact tgatgaatag tccgcactat 20640
ctggaaatgc cagcgaagat gggcgagaac ggaggggcta acggcgagga caacggtctc 20700
tggggctgct ttacttgggg aacgacacag gggttgggta caattaccct tgccctcgtt 20760
gaaaacggcc tccttcggc ggatattcaa aaggcccgca acaatatcgc taaagccgca 20820
gataagtggc ttgagaatat tgaagaacaa ggttaccgcc tgcctatcaa acaagcggag 20880
gatgaacggg gcggataccc gtggggtagt aattctttca ttctcaacca gatgatcgtc 20940
atgggctacg cttacgactt cacgggaaac agcaagtatc ttgacgggat gcaggacggc 21000

atgtcctacc tgctcggtag aaacggactt gatcaatcgt acgttactgg gtacggggag 21060

aggccacttc agaacccccа cgaccgcttt tggacccctc aaacttcgaa gaaatttccg 21120

gccccaccgc ctggtattat cgcaggcggg ccgaatagcc ggtttgaaga tccaacgatc 21180

actgcagcgg ttaagaagga tacacccccg cagaagtgct atattgacca caccgattcc 21240

tggtctacta acgagatcac gattaattgg aacgcccccc tcgcgtgggt cacagcgtat 21300

ctggacgaaa ttgacttgat taccccaccc ggcggagtgg accctgaaga gccggaagtt 21360

atctacggtg attgtaacgg cgacggaaag gttaattcga ccgatgctgt ggccgttaaa 21420

aggtatatcc tccgcagcgg tatcttcgatc aacacggaca acgcggacgt taatgcagat 21480

ggtcgcgtga atagcactga cctcgctatt ttgaagcgct atattttgaa ggagatcgat 21540

gttcttcctc acaagtgacc taggtccccg aatttccccg atcgttcaaa catttggcaa 21600

taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg 21660

ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg 21720

gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag 21780

cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggaattg 21840

c 21841

<210> 83
<211> 20699
<212> DNA
<213> Штучна послідовність

<220>
<223> Синтетична послідовність, вектор, pAG2349

<400> 83
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tctctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg ccccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgcc ctctgcgcgt gggagggaga ggccgcggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag cggcggcgca cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggcccа 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgtggtgct gccattgggc cgtacgcgtg gcggtgaccg caccggatgc tgcctgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatcvgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt ttttccccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag cctgttgaat atgcagggtg aagaaaaatc ccgcgcgctg 2700

gcgattttaa aatcggcctt cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttaccccgga agacagcggt ctgttctccc cgcttattgct gaatgtggtg 2820

aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880

ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940

aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000

cagttgttga cccagccggt gaaacaaggt gcagaactgg acttcccgat tccagtggat 3060

gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120

gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180

cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240

ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300

atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3540

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600

ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660

taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720

accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780

tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840

gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900

attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960

cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020

actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080

tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140

acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200

acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260

cgactagatg ttgaggccta acattttatt agagagcagg ctagttgctt agatacatga 4320

tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380

cgcagaagct cccatctttg ccgccataga cgccgcgccc cctttttggg gtgtagaaca 4440

tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500

cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560

tgtcgtaatt ggatgaacta ttatcgtagt cgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaatgtcg tagttggatg 4680
gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgcccgga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc cagggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tccgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aacccttttg ccagatttgg 6360
tsactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480

tctacttgat cggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaaccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ttttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400

taacgtaaaa acaactttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccct gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gagggaaccc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
ccttttcacg ccctttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtactgc agtgcagcgt gacccggtcg tgcccctctc 9960
tagagataat gagcattgca tgtctaagtt ataaaaaatt accacatatt ttttttgtca 10020
cacttgtttg aagtgcagtt tatctatctt tatacatata tttaaacttt actctacgaa 10080
taatataatc tatagtacta caataatatc agtgttttag agaatcatat aaatgaacag 10140
ttagacatgg tctaaaggac aattgagtat tttgacaaca ggactctaca gttttatctt 10200
tttagtgtgc atgtgttctc cttttttttt gcaaatagct tcacctatat aatacttcat 10260
ccatttttatt agtacatcca tttagggttt agggttaatg gttttttatag actaatttttt 10320

ttagtacatc tattttattc tattttagcc tctaaattaa gaaaactaaa actctatttt 10380
agtttttta tttaataatt tagatataaa atagaataaa ataaagtgac taaaaattaa 10440
acaaatacca tttaagaaat taaaaaaact aaggaaacat ttttcttgtt tcgagtagat 10500
aatgccagcc tgttaaacgc cgtcgacgag tctaacggac accaaccagc gaaccagcag 10560
cgtcgcgtcg ggccaagcga agcagacggc acggcatctc tgtcgctgcc tctggacccc 10620
tctcgagagt tccgctccac cgttggactt gctccgctgt cggcatccag aaattgcgtg 10680
gcggagcggc agacgtgagc cggcacggca ggcggcctcc tcctcctctc acggcacggc 10740
agctacgggg gattcctttc ccaccgctcc ttcgctttcc cttcctcgcc cgccgtaata 10800
aatagacacc ccctccacac cctctttccc caacctcgtg ttgttcggag cgcacacaca 10860
cacaaccaga tctccccaaa atccacccgt cggcacctcc gcttcaaggt acgccgctcg 10920
tcctcccccc cccccctct ctaccttctc tagatcggcg ttccggtcca tggttagggc 10980
ccggtagttc tacttctgtt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct 11040
gctagcgttc gtacacggat gcgacctgta cgtcagacac gttctgattg ctaacttgcc 11100
agtgtttctc tttggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt 11160
catgatttt tttgtttcgt tgcatagggt ttggtttgcc cttttccttt atttcaatat 11220
atgccgtgca cttgtttgtc gggtcatctt ttcatgcttt tttttgtctt ggttgtgatg 11280
atgtggtctg gttgggcggt cgttctagat cggagtagaa ttctgtttca aactacctgg 11340
tggatttatt aattttggat ctgtatgtgt gtgccataca tattcatagt tacgaattga 11400
agatgatgga tggaaatatc gatctaggat aggtatacat gttgatgcgg gttttactga 11460
tgcatataca gagatgcttt ttgttcgctt ggttgtgatg atgtggtgtg gttgggcggt 11520
cgttcattcg ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa 11580
ttttggaact gtatgtgtgt gtcatacatc ttcatagtta cgagtttaag atggatggaa 11640
atatcgatct aggataggta tacatgttga tgtgggtttt actgatgcat atacatgatg 11700
gcatatgcag catctattca tatgctctaa ccttgagtac ctatctatta taataaacaa 11760
gtatgtttta taattatttt gatcttgata tacttggatg atggcatatg cagcagctat 11820
atgtggattt ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt 11880
gtcgatgctc accctgttgt ttggtgttac ttctgcagat ccagatcgga tcctaaacca 11940
tgcgcgtgct gctcgtggcc ctggccctgc tggctcttgc tgccagcgcc accctcagc 12000
agatcggcac ctacacgcgc gagacccacc caagcctgag ctggtccacc tgcaagagcg 12060
gcggttcctg cacgaccaac agcggcgcca tcaccttga tgcgaactgg cgctgggtgc 12120
acggcgtgaa caccagcacc aactgctaca cgggtaacac gtggaacacc gccatctgcg 12180
acaacggacgc ttcctgcgcc caggactgcg cgcttgatgg cgccgactac tccggcacct 12240

acggcatcac caectccggc aacagcctgc gcctgaactt cgtgaccggc agcaatgtgg 12300
gcagccgcac ctacctgatg gccgacaaca cccactacca gatcttcgac ctgctgaacc 12360
aggagttcac cttcaccgtc gacgtgtccc acctgccctg cggcctgaac ggcgccctct 12420
acttcgtgac gatggacgcc gacggcggcg tgtccaagta cccgaacaac aaggctggcg 12480
cccagtacgg tgtgggctac tgcgacagcc agtgccccgag ggacctgaag ttcatcgccg 12540
gccaggccaa cgtggagggc tggaccccga gcagcaacaa cgccaacacc ggcctgggca 12600
accacggcgc ctgctgcgcc gagctggaca tctgggaggc caacagcatc agcgaggccc 12660
tgaccccaca cccatgcgac acccccaggcc tgtctgtgtg caccaccgac gcctgcggcg 12720
gcacctactc cagcgaccgc tacgccggca cctgcgaccc agacggctgc gacttcaacc 12780
cgtaccgcct gggcgtgacc gacttctacg gcagcggcaa gaccgtggac accaccaagc 12840
cgatcaccgt ggtgaccccag ttcgtgaccg acgacggcac cagcaccggc accctgagcg 12900
agatccgccg ctactacgtc cagaacggcg tggtgatccc gcagccgagc agcaagatca 12960
gcggcgtgtc cggcaacgtg atcaacagcg acttctgcga cgccgagatc agcaccttcg 13020
gcgagaccgc cagcttcagc aagcacggcg gcctggccaa gatgggcgct ggcatggaag 13080
ccggcatggt gctggtgatg agcctgtggg acgactacte cgtgaacatg ctgtggctgg 13140
acagcaccta cccgaccaac gccaccggga cgccaggcgc tgccaggggc agctgcccaa 13200
ccacctcggg cgaccccaag accgtcgaga gccagagcgg cagcagctac gtgaccttca 13260
gcgacatccg cgtgggcccg ttcaactcca cgttcagcgg tggctctagc acgggcggct 13320
cctccaccac caccgccagc ggcaccacca ccaccaaggc ctccagcacg tctactagct 13380
ccacctctac cggcaccggc gttgctgccc attggggcca gtgcggtggc cagggctgga 13440
cgggtccaac gacttgcgcc tccggcacca cctgcaccgt ggtcaatccg tactactccc 13500
agtgcctgga cgagctgaag gccgaggcca agtgatcccc gaatttcccc gatcgttcaa 13560
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 13620
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 13680
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 13740
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 13800
atcgggaatt gggtaccgtc gactctagta acggccgcca gtgtgctgga attaattcgg 13860
cttgtcgacc acccaacccc atatcgacag aggatgtgaa gaacaggtaa atcacgcaga 13920
agaacccatc tctgatagca gctatcgatt agaacaacga atccatattg ggtccgtggg 13980
aaatacttac tgcacaggaa gggggcgatc tgacgaggcc ccgccaccgg cctcgacccg 14040
aggccgaggc cgacgaagcg ccggcgagta cggcgccgcg gcggcctctg cccgtgccct 14100
ctgcgcgtgg gagggagagg ccgcggtggt gggggcgcgc gcgcgcgcgc gcgcagctgg 14160

tgcggcggcg cggggctcag ccgccgagcc ggcggcgacg gaggagcagg gcggcgtgga 14220
cgcgaacttc cgatcggttg gtcagagtgc gcgagttggg cttagccaat taggtctcaa 14280
caatctattg ggccgtaaaa ttcatgggcc ctggtttgtc taggcccaat atcccgttca 14340
tttcagccca caaatatttc cccagaggat tattaaggcc cacacgcagc ttatagcaga 14400
tcaagtacga tgtttcctga tcgttggatc ggaaacgtac ggtcttgatc aggcatgccg 14460
acttcgtcaa agagaggcgg catgacctga cgcggagttg gttccgggca cogtctggat 14520
ggtcgtaccg ggaccggaca cgtgtcgcgc ctccaactac atggacacgt gtggtgctgc 14580
cattgggccg tacgcgtggc ggtgaccgca ccggatgctg cctcgcaccg ccttgcccac 14640
gctttatata gagaggtttt ctctccatta atcgcatagc gagtcgaatc gaccgaaggg 14700
gagggggagc gaagctttgc gttctctaat cgcctcgtca aggtaactaa tcaatcacct 14760
cgtcctaatc ctcgaatctc tcgtggtgcc cgtctaatct cgcgattttg atgctcgtgg 14820
tggaaagcgt aggaggatcc cgtgcgagtt agtctcaatc tctcagggtt tcgtgcgatt 14880
ttagggtgat ccacctctta atcgagttac ggtttcgtgc gattttaggg taatcctctt 14940
aatctctcat tgatttaggg tttcgtgaga atcgaggtag ggatctgtgt tatttatatc 15000
gatctaatag atggattggt tttgagattg ttctgtcaga tggggattgt ttcgatatat 15060
taccctaatg atgtgtcaga tggggattgt ttcgatatat taccctaatg atgtgtcaga 15120
tggggattgt ttcgatatat taccctaatg atggataata agagtagttc acagttatgt 15180
tttgatcctg ccacatagtt tgagttttgt gatcagattt agttttactt atttgtgctt 15240
agttcggatg ggattgttct gatattgttc caatagatga atagctcgtt aggttaaaat 15300
ctttaggttg agttaggcga cacatagttt atttcctctg gatttggatt ggaattgtgt 15360
tcttagtttt tttccccctgg atttggattg gaattgtgtg gagctgggtt agagaattac 15420
atctgtatcg tgtacaccta cttgaactgt agagcttggg ttctaaggtc aatttaatct 15480
gtattgtatc tggctctttg cctagttgaa ctgtagtgct gatgttgtac tgtgtttttt 15540
taccccgttct atttgcttta ctcgtgcaaa tcaaatctgt cagatgctag aactaggtgg 15600
ctttacttctg tgttcttaca tagatctgtt gtcctgtagt tacttatgtc agttttgtta 15660
ttatctgaag atatttttgg ttgttgcttg ttgatgtggt gtgagctgtg agcagcgctc 15720
ttatgattaa tgatgctgtc caattgtagt gtagtatgat gtgattgata tgttcatcta 15780
ttttgagctg acagtaccga tatcgtagga tctggtgcca acttattctc cagctgcttt 15840
cttttaccta tgttaattcc aatcctttct tgcctcttcc agatccagat aatggcccac 15900
gcccgcgtcc tcctcctggc gctcgccgtc ctggccaccg ccgccgtcgc cgtcgcctcc 15960
tcctcctcct tcgccgactc caaccccgatc cgcccggtga ccgaccgcgc cgcctccacc 16020
gcttacgact acaagcaggt gttcgcgggac tcgctactat tctatgaggc ccagagatcc 16080

ggccggctcc cagccgacca gaaggtcacg tggaggaagg atagcgcgct gaatgaccag 16140
ggtgaccagg gacaagactt gaccggcggc tactttgacg ctggggactt cgtcaagttc 16200
gggttcccca tggcttatac cgcaacgtg ctggcatggg gcctcataga ttttgaggcc 16260
ggctacagca gtgccgggge cttggatgat ggacggaagg ctgtcaaatg ggccaccgac 16320
tatttcataa aggcccacac aagtcaaaat gagttctatg gtcaggtcgg ccagggtgac 16380
gccgatcacg ctttctgggg aagaccagag gatatgacga tggcgcgccc ggcgtacaag 16440
atagacacct caaggcctgg ctctgatctg gcaggcgaga cagcggctgc tcttgccgct 16500
gcttcaatcg tgttccggaa cgtcgatggc acttactcaa ataacctgtt aacacacgct 16560
cgccagctat tcgacttcgc gaacaactac cggggaaagt atagtgactc tattactgac 16620
gcaagaaatt tctacgcaag cgcagactac agagacgagt tggtttgggc tgctgcgtgg 16680
ttatacagag cgaccaacga caacacctac ctcaacactg ctgagtcact gtacgatgag 16740
tttgggctac agaactgggg ggggggcctg aactgggata gcaaggtgtc tggcgtgcag 16800
gtgttgttgg ccaagcttac caataagcag gcctacaagg acacggtgca gtcttacgtc 16860
aattacctaa ttaataacca gcagaagact cccaagggcc tcctctacat cgacatgtgg 16920
ggcaccctte gccacgctgc caacgccgca ttcatcatgc tcgaagccgc cgagctgggc 16980
ttgtccgcct cctcttatag acagttcgcg caaacgcaaa tcgactacgc cctgggcgat 17040
ggtggccgct cctttgtgtg cgggttcggg agtaatcctc ctacgagacc gcaccacaga 17100
tcctcgtcgt gcccgccagc tcccgctact tgcgactgga atacattcaa ctcacctgac 17160
ccaaactacc acgtcctctc tggggccctg gtgggcggac ctgatcagaa tgacaactac 17220
gtcgatgacc gttcagacta tgttcacaac gaagtcgcca ctgattacaa cgcgggtttc 17280
cagtccgcgt tagctgcttt ggtggccctt ggttacagcg agaaggacga gctgtgacct 17340
aggtccccga atttccccga tcgttcaaac atttggcaat aaagtttctt aagattgaat 17400
cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta 17460
ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg 17520
caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 17580
tcgcgcgcgg tgtcatctat gttactagat cgggaattgg gtacgtcga ctctagtaac 17640
ggccgccagt gtgctggaat taattcggct tgtcgaccac ccaacccoat atcgacagag 17700
gatgtgaaga acaggtaaat cacgcagaag aacccatctc tgatagcagc tatcgattag 17760
aacaacgaat ccatattggg tccgtgggaa atacttactg cacaggaagg gggcgatctg 17820
acgaggcccc gccaccggcc tcgacccgag gccgaggccg acgaagcgcc ggcgagtacg 17880
gcgccgcggc ggcctctgcc cgtgccctct gcgcgtggga gggagaggcc gcggtggtgg 17940
gggcgcgcgc gcgcgcgcgc gcagctggtg cggcggcgcg ggggtcagcc gccgagccgg 18000

cggcgacgga ggagcagggc ggcgtggacg cgaacttccg atcggttggt cagagtgcgc 18060
gagttgggct tagccaatta ggtctcaaca atctattggg ccgtaaaatt catgggccct 18120
ggtttgtcta ggcccaatat cccgttcatt tcagcccaca aatatttccc cagaggatta 18180
ttaaggccca cacgcagctt atagcagatc aagtacgatg tttcctgatc gttggatcgg 18240
aaacgtacgg tcttgatcag gcatgccgac ttcgtcaaag agaggcggca tgacctgacg 18300
cggagttggt tccgggcacc gtctggatgg tcgtaccggg accggacacg tgtcgcgcct 18360
ccaactacat ggacacgtgt ggtgctgcca ttgggccgta cgcgtggcgg tgaccgcacc 18420
ggatgctgcc tcgcaccgcc ttgcccacgc tttatataga gaggttttct ctccattaat 18480
cgcatagcga gtcgaatcga ccgaaggggа gggggagcga agctttgcgt tctctaatcg 18540
cctcgtcaag gtaactaatc aatcacctcg tcctaatcct cgaatctctc gtggtgcccg 18600
tctaatctcg cgattttgat gctcgtggtg gaaagcgtag gaggatcccg tgcgagttag 18660
tctcaatctc tcagggtttc gtgcgatttt agggtgatcc acctcttaat cgagttacgg 18720
tttcgtgcga ttttagggta atcctcttaa tctctcattg atttagggtt tcgtgagaat 18780
cgaggtaggg atctgtgtta tttatatcga tctaatagat ggattggttt tgagattgtt 18840
ctgtcagatg gggattgttt cgatatatta ccctaatgat gtgtcagatg gggattgttt 18900
cgatatatta ccctaatgat gtgtcagatg gggattgttt cgatatatta ccctaatgat 18960
ggataataag agtagttcac agttatgttt tgatcctgcc acatagtttg agttttgtga 19020
tcagatttag ttttacttat ttgtgcttag ttcggatggg attgttctga tattgttcca 19080
atagatgaat agctcgttag gttaaaatct ttaggttgag ttaggcgaca catagtttat 19140
ttcctctgga tttggattgg aattgtgttc ttagtttttt tcccctggat ttggattgga 19200
attgtgtgga gctgggttag agaattacat ctgtatcgtg tacacctact tgaactgtag 19260
agcttgggtt ctaaggtcaa tttaatctgt attgtatctg gctctttgcc tagttgaact 19320
gtagtgctga tgttgtactg tgttttttta cccgttttat ttgctttact cgtgcaaatc 19380
aaatctgtca gatgctagaa ctaggtggct ttattctgtg ttcttacata gatctgttgt 19440
cctgtagtta cttatgtcag ttttgttatt atctgaagat atttttggtt gttgcttgtt 19500
gatgtggtgt gagctgtgag cagcgctctt atgattaatg atgctgtcca attgtagtgt 19560
agtatgatgt gattgatatg ttcatctatt ttgagctgac agtaccgata tcgtaggatc 19620
tggtgccaac ttattctcca gctgcttttt tttacctatg ttaattccaa tcctttcttg 19680
cctcttccag atccagataa tggcgaacaa acatttgtcc ctctccctct tcctcgtcct 19740
ccttggcctg tcggccagct tggcctccgg gcaagtcttc ccagctggaa acgcaacgga 19800
attggagaaa agacaaacca cccctaactc tgagggctgg catgacggat actactactc 19860
ttggtggagc gatggtggtg cacaggccac ctatacaaac ctcgaaggcg gcacttatga 19920

gatttcatgg ggtgacggtg gcaaccttgt cggcggaaag gggtggaacc ccggacttaa 19980

cgccagggca atccacttcg aaggggtgta ccagcccaat ggcaactcat acctggccgt 20040

ctacgggtgg acgcgcaatc cgctggttga gtactatatc gtggagaatt tcggaacttа 20100

tgaccctagc tccggtgcca cggacctcgg gacagtcgag tgtgacggaa gcatctacag 20160

gctgggtaaa actacccgcg ttaatgctcc atcgatcgac ggcacgcaaa catttgatca 20220

atactggtcc gtgcggcagg ataagaggac aagcggcaca gttcagacgg gttgccactt 30280

tgatgcctgg gcaagagcgg ggctcaatgt gaatggggac cactactatc agattgtggc 20340

gaccgagggc tatttctcca gtggctatgc gcgtataacc gtcgctgatg ttggaagcga 20400

gaaggacgag ctgtgaccta ggtccccgaa tttccccgat cgttcaaaca tttggcaata 20460

aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat aatttctgtt 20520

gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt 20580

ttttatgatt agagtcccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg 20640

cgcaaactag gataaattat cgcgcgcggt gtcatctatg ttactagatc gggaattgg 20699

<210> 84
<211> 14002
<212> ДНК
<213> Штучна послідовність

<220>
<223> Синтетична конструкція, pAG2042

<400> 84
aattcatact aaagcttgca tgcctgcagg tcgactctag taacggccgc cagtgtgctg 60

gaattaattc ggcttgtcga ccacccaacc ccatatcgac agaggatgtg aagaacaggt 120

aaatcacgca gaagaaccca tcctgatag cagctatcga ttagaacaac gaatccatat 180

tgggtccgtg ggaaatactt actgcacagg aaggggcga tctgacgagg cccgccacc 240

ggcctcgacc cgaggccgag gccgacgaag cgccggcgag tacggcgccg cggcggcctc 300

tgcccgtgct ctctgcgcgt gggagggaga ggccgcgggtg gtgggggcgc gcgcgcgcgc 360

gcgcgcagct ggtgcggcgg cgcgggggtc agccgccgag ccggcggcga cggaggagca 420

gggcggcgtg gacgcgaact tccgatcggt tggtcagagt gcgcgagttg ggcttagcca 480

attaggtctc aacaatctat tgggccgtaa aattcatggg ccctggtttg tctaggcccа 540

atatcccgtt catttcagcc cacaaatatt tccccagagg attattaagg cccacacgca 600

gcttatagca gatcaagtac gatgtttcct gatcgttgga tcggaaacgt acggtcttga 660

tcaggcatgc cgacttcgtc aaagagaggc ggcatgacct gacgcggagt tggttccggg 720

caccgtctgg atggtcgtac cgggaccgga cacgtgtcgc gcctccaact acatggacac 780

gtgcggtgct gccattgggc cgtacgcgtg gcggtgaccg cacccgatgc tgcctcgcac 840

cgccttgccc acgctttata tagagaggtt ttctctccat taatcgcata gcgagtcgaa 900

tcgaccgaag gggaggggga gcgaagcttt gcgttctcta atcgcctcgt caaggtaact 960

aatcaatcac ctcgtcctaa tcctcgaatc tctcgtggtg cccgtctaat ctcgcgattt 1020

tgatgctcgt ggtggaaagc gtaggaggat cccgtgcgag ttagtctcaa tctctcaggg 1080

tttcgtgcga ttttagggtg atccacctct taatcgagtt acggtttcgt gcgattttag 1140

ggtaatcctc ttaatctctc attgatttag ggtttcgtga gaatcgaggt agggatctgt 1200

gttatttata tcgatctaat agatggattg gttttgagat tgttctgtca gatggggatt 1260

gtttcgatat attaccctaa tgatgtgtca gatggggatt gtttcgatat attaccctaa 1320

tgatgtgtca gatggggatt gtttcgatat attaccctaa tgatggataa taagagtagt 1380

tcacagttat gttttgatcc tgccacatag tttgagtttt gtgatcagat ttagttttac 1440

ttatttgtgc ttagttcgga tgggattgtt ctgatattgt tccaatagat gaatagctcg 1500

ttaggttaaa atctttaggt tgagttaggc gacacatagt ttatttcctc tggatttgga 1560

ttggaattgt gttcttagtt tttttcccct ggatttggat tggaattgtg tggagctggg 1620

ttagagaatt acatctgtat cgtgtacacc tacttgaact gtagagcttg ggttctaagg 1680

tcaatttaat ctgtattgta tctggctctt tgcctagttg aactgtagtg ctgatgttgt 1740

actgtgtttt tttacccgtt ttatttgctt tactcgtgca aatcaaatct gtcagatgct 1800

agaactaggt ggctttattc tgtgttctta catagatctg ttgtcctgta gttacttatg 1860

tcagttttgt tattatctga agatattttt ggttgttgct tgttgatgtg gtgtgagctg 1920

tgagcagcgc tcttatgatt aatgatgctg tccaattgta gtgtagtatg atgtgattga 1980

tatgttcatc tattttgagc tgacagtacc gatatcgtag gatctggtgc caacttattc 2040

tccagctgct ttttttttacc tatgttaatt ccaatccttt cttgcctctt ccagatccag 2100

ataatgcaga aactcattaa ctcagtgcaa aactatgcct ggggcagcaa aacggcgttg 2160

actgaacttt atggtatgga aaatccgtcc agccagccga tggccgagct gtggatgggc 2220

gcacatccga aaagcagttc acgagtgcag aatgccgccg gagatatcgt ttcactgcgt 2280

gatgtgattg agagtgataa atcgactctg ctcggagagg ccgttgccaa acgctttggc 2340

gaactgcctt tcctgttcaa agtattatgc gcagcacagc cactctccat tcaggttcat 2400

ccaaacaaac acaattctga aatcggtttt gccaaagaaa atgccgcagg tatcccgatg 2460

gatgccgccg agcgtaacta taaagatcct aaccacaagc cggagctggt ttttgcgctg 2520

acgcctttcc ttgcgatgaa cgcgtttcgt gaattttccg agattgtctc cctactccag 2580

ccggtcgcag gtgcacatcc ggcgattgct cacttttttac aacagcctga tgccgaacgt 2640

ttaagcgaac tgttcgccag ccttgttgaat atgcagggtg aagaaaaatc cgcgcgctg 2700

gcgatttttaa aatcggccct cgatagccag cagggtgaac cgtggcaaac gattcgttta 2760

atttctgaat tttaccggga agacagcggt ctgttctccc cgctattgct gaatgtggtg 2820
aaattgaacc ctggcgaagc gatgttcctg ttcgctgaaa caccgcacgc ttacctgcaa 2880
ggcgtggcgc tggaagtgat ggcaaactcc gataacgtgc tgcgtgcggg tctgacgcct 2940
aaatacattg atattccgga actggttgcc aatgtgaaat tcgaagccaa accggctaac 3000
cagttgttga cccagccggt gaaacaaggt gcagaactgg actttccgat tccagtggat 3060
gattttgcct tctcgctgca tgaccttagt gataaagaaa ccaccattag ccagcagagt 3120
gccgccattt tgttctgcgt cgaaggcgat gcaacgttgt ggaaaggttc tcagcagtta 3180
cagcttaaac cgggtgaatc agcgtttatt gccgccaacg aatcaccggt gactgtcaaa 3240
ggccacggcc gtttagcgcg tgtttacaac aagctgtaag agcttactga aaaaattaac 3300
atctcttgct aagctgggag ctctagatcc ccgaatttcc ccgatcgttc aaacatttgg 3360
caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3420
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3480
tgggtttta tgattagagt cccgcaatta tacatttaat acgcgatagg aaacaaaata 3540
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa 3600
ttggcgagct cgaattaatt cagtacatta aaaacgtccg caatgtgtta ttaagttgtc 3660
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 3720
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcagtc cgggacggcg 3780
tcagcgggag agccgttgta aggcggcaga ctttgctcat gttaccgatg ctattcggaa 3840
gaacggcaac taagctgccg ggtttgaaac acggatgatc tcgcggaggg tagcatgttg 3900
attgtaacga tgacagagcg ttgctgcctg tgatcaaata tcatctccct cgcagagatc 3960
cgaattatca gccttcttat tcatttctcg cttaaccgtg acaggctgtc gatcttgaga 4020
actatgccga cataatagga aatcgctgga taaagccgct gaggaagctg agtggcgcta 4080
tttctttaga agtgaacgtt gacgatcgtc gaccgtaccc cgatgaatta attcggacgt 4140
acgttctgaa cacagctgga tacttacttg ggcgattgtc atacatgaca tcaacaatgt 4200
acccgtttgt gtaaccgtct cttggaggtt cgtatgacac tagtggttcc cctcagcttg 4260
cgactagatg ttgaggccta acatttttatt agagagcagg ctagttgctt agatacatga 4320
tcttcaggcc gttatctgtc agggcaagcg aaaattggcc atttatgacg accaatgccc 4380
cgcagaagct cccatctttg ccgccataga cgccgcgccc ccctttgggg gtgtagaaca 4440
tcctttgcc agatgtggaa aagaagttcg ttgtcccatt gttggcaatg acgtagtagc 4500
cggcgaaagt gcgagaccca tttgcgctat atataagcct acgatttccg ttgcgactat 4560
tgtcgtaatt ggatgaacta ttatcgtagt tgctctcaga gttgtcgtaa tttgatggac 4620
tattgtcgta attgcttatg gagttgtcgt agttgcttgg agaaacgtcg tagttggatg 4680

gggagtagtc atagggaaga cgagcttcat ccactaaaac aattggcagg tcagcaagtg 4740
cctgccccga tgccatcgca agtacgaggc ttagaaccac cttcaacaga tcgcgcatag 4800
tcttccccag ctctctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac 4860
gaattgttag acattatttg ccgactacct tggtgatctc gcctttcacg tagtgaacaa 4920
attcttccaa ctgatctgcg cgcgaggcca agcgatcttc ttgtccaaga taagcctgcc 4980
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag 5040
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa 5100
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt 5160
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg 5220
gacctaccaa ggcaacgcta tgtttcttcg cttttgtcag caagatagcc agatcaatgt 5280
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt 5340
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga 5400
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt 5460
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa 5520
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca 5580
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt 5640
cggcgatcac cgcttccctc atgatgttta actcctgaat taagccgcgc cgcgaagcgg 5700
tgtcggcttg aatgaattgt taggcgtcat cctgtgctcc cgagaaccag taccagtaca 5760
tcgctgtttc gttcgagact tgaggtctag ttttatacgt gaacaggtca atgccgccga 5820
gagtaaagcc acattttgcg tacaaattgc aggcaggtac attgttcgtt tgtgtctcta 5880
atcgtatgcc aaggagctgt ctgcttagtg cccacttttt cgcaaattcg atgagactgt 5940
gcgcgactcc tttgcctcgg tgcgtgtgcg acacaacaat gtgttcgata gaggctagat 6000
cgttccatgt tgagttgagt tcaatcttcc cgacaagctc ttggtcgatg aatgcgccat 6060
agcaagcaga gtcttcatca gagtcatcat ccgagatgta atccttccgg taggggctca 6120
cacttctggt agatagttca aagccttggt cggataggtg cacatcgaac acttcacgaa 6180
caatgaaatg gttctcagca tccaatgttt ccgccacctg ctcagggatc accgaaatct 6240
tcatatgacg cctaacgcct ggcacagcgg atcgcaaacc tggcgcggct tttggcacaa 6300
aaggcgtgac aggtttgcga atccgttgct gccacttgtt aaccctttg ccagatttgg 6360
taactataat ttatgttaga ggcgaagtct tgggtaaaaa ctggcctaaa attgctgggg 6420
atttcaggaa agtaaacatc accttccggc tcgatgtcta ttgtagatat atgtagtgta 6480
tctacttgat cgggggatct gctgcctcgc gcgtttcggt gatgacggtg aaaacctctg 6540
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca 6600

agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc 6660
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg 6720
agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6780
aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 6840
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 6900
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 6960
ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt 7020
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 7080
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 7140
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 7200
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgccttа 7260
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 7320
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 7380
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 7440
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 7500
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 7560
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 7620
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 7680
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 7740
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 7800
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 7860
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 7920
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 7980
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 8040
gctgcagggg ggggggggg ggggttccat tgttcattcc acggacaaaa acagagaaag 8100
gaaacgacag aggccaaaaa gctcgctttc agcacctgtc gtttcctttc ctttcagagg 8160
gtattttaaa taaaaacatt aagttatgac gaagaagaac ggaaacgcct taaaccggaa 8220
aattttcata aatagcgaaa acccgcgagg tcgccgcccc gtaacctgtc ggatcaccgg 8280
aaaggacccg taaagtgata atgattatca tctacatatc acaacgtgcg tggaggccat 8340
caaaccacgt caaataatca attatgacgc aggtatcgta ttaattgatc tgcatcaact 8400
taacgtaaaa acaacttcag acaatacaaa tcagcgacac tgaatacggg gcaacctcat 8460
gtcccccccc cccccccctt gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8520

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8580
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8640
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8700
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8760
gttgctcttg cccggcgtca acacgggata taccgcgcc acatagcaga actttaaaag 8820
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8880
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8940
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9000
cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 9060
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9120
gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca 9180
tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaattggtcg 9240
acgatcttgc tgcgttcgga tattttcgtg gagttcccgc cacagacccg gattgaaggc 9300
gagatccagc aactcgcgcc agatcatcct gtgacggaac tttggcgcgt gatgactggc 9360
caggacgtcg gccgaaagag cgacaagcag atcacgcttt tcgacagcgt cggatttgcg 9420
atcgaggatt tttcggcgct gcgctacgtc cgcgaccgcg ttgagggatc aagccacagc 9480
agcccactcg accttctagc cgacccagac gagccaaggg atcttttttgg aatgctgctc 9540
cgtcgtcagg ctttccgacg tttgggtggt tgaacagaag tcattatcgc acggaatgcc 9600
aagcactccc gaggggaacc ctgtggttgg catgcacata caaatggacg aacggataaa 9660
cctttcacg cccttttaaa tatccgatta tctaataaa cgctcttttc tcttaggttt 9720
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaacc 9780
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa 9840
gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc agcttaatta 9900
agtctaactc gagttactgg tacgtaccaa atccatggaa tcaaggtacc gtcgactcta 9960
gtaacggccg ccagtgtgct ggaattaatt cggcttgtcg accacccaac cccatatcga 10020
cagaggatgt gaagaacagg taaatcacgc agaagaaccc atctctgata gcagctatcg 10080
attagaacaa cgaatccata ttgggtccgt gggaaatact tactgcacag gaaggggcg 10140
atctgacgag gccccgccac cggcctcgac ccgaggccga ggccgacgaa gcgccggcga 10200
gtacggcgcc gcggcggcct ctgcccgtgc cctctgcgcg tgggagggag aggccgcggt 10260
ggtgggggcg cgcgcgcgcg cgcgcgcagc tggtgcggcg gcgcgggggt cagccgccga 10320
gccggcggcg acggaggagc agggcggcgt ggacgcgaac ttccgatcgg ttggtcagag 10380
tgcgcgagtt gggcttagcc aattaggtct caacaatcta ttgggccgta aaattcatgg 10440

gccctggttt gtctaggccc aatatcccgt tcatttcagc ccacaaatat ttccccagag 10500

gattattaag gcccacacgc agcttatagc agatcaagta cgatgtttcc tgatcgttgg 10560

atcggaaacg tacggtcttg atcaggcatg ccgacttcgt caaagagagg cggcatgacc 10620

tgacgcggag ttggttccgg gcaccgtctg gatggtcgta ccgggaccgg acacgtgtcg 10680

cgcctccaac tacatggaca cgtgtggtgc tgccattggg ccgtacgcgt ggcggtgacc 10740

gcaccggatg ctgcctcgca ccgccttgcc cacgctttat atagagaggt tttctctcca 10800

ttaatcgcat agcgagtcga atcgaccgaa ggggagcggg agcgaagctt tgcgttctct 10860

aatcgcctcg tcaaggtaac taatcaatca cctcgtccta atcctcgaat ctctcgtggt 10920

gcccgtctaa tctcgcgatt ttgatgctcg tggtggaaag cgtaggagga tcccgtgcga 10980

gttagtctca atctctcagg gtttcgtgcg attttagggt gatccacctc ttaatcgagt 11040

tacggtttcg tgcgatttta gggtaatcct cttaatctct cattgattta gggtttcgtg 11100

agaatcgagg tagggatctg tgttatttat atcgatctaa tagatggatt ggttttgaga 11160

ttgttctgtc agatggggat tgtttcgata tattacccta atgatgtgtc agatggggat 11220

tgtttcgata tattacccta atgatgtgtc agatggggat tgtttcgata tattacccta 11280

atgatggata ataagagtag ttcacagtta tgttttgatc ctgccacata gtttgagttt 11340

tgtgatcaga tttagtttta cttatttgtg cttagttcgg atgggattgt tctgatattg 11400

ttccaataga tgaatagctc gttaggttaa aatctttagg ttgagttagg cgacacatag 11460

tttatttcct ctggatttgg attggaattg tgttcttagt ttttttcccc tggatttgga 11520

ttggaattgt gtggagctgg gttagagaat tacatctgta tcgtgtacac ctacttgaac 11580

tgtagagctt gggttctaag gtcaatttaa tctgtattgt atctggctct ttgcctagtt 11640

gaactgtagt gctgatgttg tactgtgttt ttttaccccgt tttatttgct ttactcgtgc 11700

aaatcaaatc tgtcagatgc tagaactagg tggctttatt ctgtgttctt acatagatct 11760

gttgtcctgt agttacttat gtcagttttg ttattatctg aagatatttt tggttgttgc 11820

ttgttgatgt ggtgtgagct gtgagcagcg ctcttatgat taatgatgct gtccaattgt 11880

agtgtagtat gatgtgattg atatgttcat ctattttgag ctgacagtac cgatatcgta 11940

ggatctggtg ccaacttatt ctccagctgc ttttttttac ctatgttaat tccaatcctt 12000

tcttgcctct tccagatcca gataatggcc cacgcccgcg tcctcctcct ggcgctcgcc 12060

gtcctggcca ccgccgccgt cgccgtcgcc tcctcctcct ccttcgccga ctccaacccg 12120

atccgccccgg tgaccgaccg cgccgcctcc accgctggag gaggatactg gcacacttcc 12180

ggcagggaga tcctcgacgc aaataacgtt ccagtcagaa tcgccgggat taattggttt 12240

ggcttcgaaa cgtgtaacta cgtggttcac ggcctgtggt ctcgggatta cagatcaatg 12300

ctcgaccaga tcaaatcctt ggggtataat acaattaggc tgccctacag cgatgacatt 12360

cttaagcctg gaaccatgcc gaactcgatt aatttctacc aaatgaacca ggatctgcag 12420

ggattgactt ctctgcaggt tatggacaag atcgtggcgt acgccggcca aatcgggctc 12480

agaattattt tggatcggca caggccagac tgctcaggtc agtcggcgct gtggtacaca 12540

agctccgtgt cagaggcaac atggatttca gatcttcaag ccctcgcaca acgctataaa 12600

ggcaacccca cggttgtggg attcgacctt cacaacgaac ctcacgatcc ggcctgttgg 12660

ggctgcgggg accttcgat cgactggaga ctggcagcgg agagggctgg taacgccgtt 12720

ctcagcgtca atcccaactt gctgatcttt gtggagggag ttcagtccta caacggcgat 12780

tcttactggt ggggcggaaa tctccaaggc gcagggcagt atcctgtcgt gcttaacgtt 12840

ccgaatcgcc tggtctactc agcacacgac tacgcgacta gcgtgtaccc acagacgtgg 12900

ttctccgatc ccacatttcc taacaatatg ccgggaatct ggaacaagaa ttggggttac 12960

ttgtttaacc aaaacattgc tccagtttgg ttgggtgaat ttggcaccac tcttcagtcg 13020

acgacagacc aaacctggct gaaaaccctc gtccagtatt tgcggccaac tgctcagtac 13080

ggagcagatt cttttcaatg gacgttctgg tcttggaatc ctgactccgg ggatacaggc 13140

ggtatcctga aagacgattg gcagaccgtg gacactgtta aggacgggta cttggcgccg 13200

attaaaagct cgatctttga cccagtcggc gctagcgctt ccccatcttc acaaccttcg 13260

ccgagcgtca gccccagccc aagcccaagc ccgtctgcca gcagaacccc cactcccaca 13320

cctaccccca cggcctcacc aactccgacg ctcactccta cggcgacgcc aacaccaact 13380

gcttcaccca ctcctagccc caccgcagcg agcggggcta ggtgcacccc ttcttaccag 13440

gtcaactctg actggggtaa tggcttcacc gtgactgtgg cggtcactaa ctcaggaagc 13500

gtcgcgacga aaacctggac tgtgtcctgg acgttcgggg gcaaccaaac aatcaccaac 13560

agctggaacg ctgcagttac gcagaatggg caaagcgtca cggcgcgcaa tatgagctac 13620

aacaacgtga ttcaaccagg ccagaatacc acattcggtt ttcaagcaag ctataccggg 13680

tcaaacgctg ccccaactgt cgcttgtgct gcctcatgac ctaggtcccc gaatttcccc 13740

gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg 13800

atgattatca tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc 13860

atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 13920

gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct 13980

atgttactag atcgggaatt gg 14002

<210> 85
<211> 9983
<212> ДНК
<213> Штучна послідовність

<220>

<223> Синтетична конструкція, pAG2004

<400> 85

ccgatggccg agctgtggat gggcgcacat ccgaaaagca gttcacgagt gcagaatgcc 60

gccggagata tcgtttcact gcgtgatgtg attgagagtg ataaatcgac tctgctcgga 120

gaggccgttg ccaaacgctt tggcgaactg cctttcctgt tcaaagtatt atgcgcagca 180

cagccactct ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg ttttgccaaa 240

gaaaatgccg caggtatccc gatggatgcc gccgagcgta actataaaga tcctaaccac 300

aagccggagc tggtttttgc gctgacgcct ttccttgcga tgaacgcgtt tcgtgaattt 360

tccgagattg tctccctact ccagccggtc gcaggtgcac atccggcgat tgctcacttt 420

ttacaacagc ctgatgccga acgtttaagc gaactgttcg ccagcctgtt gaatatgcag 480

ggtgaagaaa aatcccgcgc gctggcgatt ttaaaatcgg ccctcgatag ccagcagggt 540

gaaccgtggc aaacgattcg tttaatttct gaattttacc cggaagacag cggtctgttc 600

tcccgctat tgctgaatgt ggtgaaattg aaccctggcg aagcgatgtt cctgttcgct 660

gaaacaccgc acgcttacct gcaaggcgtg gcgctggaag tgatggcaaa ctccgataac 720

gtgctgcgtg cgggtctgac gcctaaatac attgatattc cggaactggt tgccaatgtg 780

aaattcgaag ccaaaccggc taaccagttg ttgacccagc cggtgaaaca aggtgcagaa 840

ctggacttcc cgattccagt ggatgatttt gccttctcgc tgcatgacct tagtgataaa 900

gaaaccacca ttagccagca gagtgccgcc attttgttct gcgtcgaagg cgatgcaacg 960

ttgtggaaag gttctcagca gttacagctt aaaccgggtg aatcagcgtt tattgccgcc 1020

aacgaatcac cggtgactgt caaaggccac ggccgtttag cgcgtgttta caacaagctg 1080

taagagctta ctgaaaaaat taacatctct tgctaagctg ggagctctag atccccgaat 1140

ttcccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt 1200

cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg 1260

taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt 1320

taatacgcga tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg 1380

tcatctatgt tactagatcg ggaattggcg agctcgaatt aattcagtac attaaaaacg 1440

tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc 1500

caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca ctcgatacag 1560

gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg cagactttgc 1620

tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg aaacacggat 1680

gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg cctgtgatca 1740

aatatcatct cctcgcaga gatccgaatt atcagccttc ttattcattt ctcgcttaac 1800

cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc tggataaagc 1860

cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat cgtcgaccgt 1920

acccccgatga attaattcgg acgtacgttc tgaacacagc tggatactta cttgggcgat 1980

tgtcatacat gacatcaaca atgtacccgt ttgtgtaacc gtctcttgga ggttcgtatg 2040

acactagtgg ttcccctcag cttgcgacta gatgttgagg cctaacattt tattagagag 2100

caggctagtt gcttagatac atgatcttca ggccgttatc tgtcagggca agcgaaaatt 2160

ggccatttat gacgaccaat gcccccgcaga agctcccatc tttgccgcca tagacgccgc 2220

gccccccttt tggggtgtag aacatccttt tgccagatgt ggaaaagaag ttcgttgtcc 2280

cattgttggc aatgacgtag tagccggcga aagtgcgaga cccatttgcg ctatatataa 2340

gcctacgatt tccgttgcga ctattgtcgt aattggatga actattatcg tagttgctct 2400

cagagttgtc gtaatttgat ggactattgt cgtaattgct tatggagttg tcgtagttgc 2460

ttggagaaat gtcgtagttg gatggggagt agtcataggg aagacgagct tcatccacta 2520

aaacaattgg caggtcagca agtgcctgcc ccgatgccat cgcaagtacg aggcttagaa 2580

ccaccttcaa cagatcgcgc atagtcttcc ccagctctct aacgcttgag ttaagccgcg 2640

ccgcgaagcg gcgtcggctt gaacgaattg ttagacatta tttgccgact accttggtga 2700

tctcgccttt cacgtagtga acaaattctt ccaactgatc tgcgcgcgag gccaagcgat 2760

cttcttgtcc aagataagcc tgcctagctt caagtatgac gggctgatac tgggccggca 2820

ggcgctccat tgcccagtcg gcagcgacat ccttcggcgc gattttgccg gttactgcgc 2880

tgtaccaaat gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg 2940

gcgagttcca tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg 3000

gatcaaagag ttcctccgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg 3060

tcagcaagat agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt 3120

cattgcgctg ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga 3180

tgtcgtcgtg cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg 3240

aagccgaagt ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta 3300

cggtcaccgt aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg 3360

agccgtacaa atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta 3420

cctctgatag ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactcct 3480

gaattaagcc gcgccgcgaa gcggtgtcgg cttgaatgaa ttgttaggcg tcatcctgtg 3540

ctcccgagaa ccagtaccag tacatcgctg tttcgttcga gacttgaggt ctagttttat 3600

acgtgaacag gtcaatgccg ccgagagtaa agccacattt tgcgtacaaa ttgcaggcag 3660

gtacattgtt cgtttgtgtc tctaatcgta tgccaaggag ctgtctgctt agtgccccact 3720

ttttcgcaaa ttcgatgaga ctgtgcgcga ctcctttgcc tcggtgcgtg tgcgacacaa 3780

caatgtgttc gatagaggct agatcgttcc atgttgagtt gagttcaatc ttcccgacaa 3840
gctcttggtc gatgaatgcg ccatagcaag cagagtcttc atcagagtca tcatccgaga 3900
tgtaatcctt ccggtagggg ctcacacttc tggtagatag ttcaaagcct tggtcggata 3960
ggtgcacatc gaacacttca cgaacaatga aatggttctc agcatccaat gtttccgcca 4020
cctgctcagg gatcaccgaa atcttcatat gacgcctaac gcctggcaca gcggatcgca 4080
aacctggcgc ggcttttggc acaaaaggcg tgacaggttt gcgaatccgt tgctgccact 4140
tgttaaccct tttgccagat ttggtaacta taatttatgt tagaggcgaa gtcttgggta 4200
aaaactggcc taaaattgct ggggatttca ggaaagtaaa catcaccttc cggctcgatg 4260
tctattgtag atatatgtag tgtatctact tgatcggggg atctgctgcc tcgcgcgttt 4320
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 4380
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 4440
tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat 4500
gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga 4560
tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg 4620
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 4680
tccacagaat cagggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 4740
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 4800
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 4860
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 4920
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 4980
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 5040
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 5100
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 5160
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 5220
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 5280
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 5340
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 5400
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 5460
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 5520
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 5580
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 5640
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagatttа 5700

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 5760
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 5820
agtttgcgca acgttgttgc cattgctgca gggggggggg gggggggttt ccattgttca 5880
ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc tttcagcacc 5940
tgtcgtttcc tttcttttca gagggtattt taaataaaaa cattaagtta tgacgaagaa 6000
gaacggaaac gccttaaacc ggaaaatttt cataaatagc gaaaacccgc gaggtcgccg 6060
ccccgtaacc tgtcggatca ccggaaagga cccgtaaagt gataatgatt atcatctaca 6120
tatcacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg acgcaggtat 6180
cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata caaatcagcg 6240
acactgaata cggggcaacc tcatgtcccc cccccccccc ccctgcaggc atcgtggtgt 6300
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 6360
catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca 6420
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 6480
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 6540
gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg 6600
cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac 6660
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 6720
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 6780
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 6840
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 6900
gtatttagaa aaataaacaa atagggggtc cgcgcacatt tccccgaaaa gtgccacctg 6960
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 7020
cctttcgtct tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt cgtggagttc 7080
ccgccacaga cccggattga aggcgagatc cagcaactcg cgccagatca tcctgtgacg 7140
gaactttggc gcgtgatgac tggccaggac gtcggccgaa agagcgacaa gcagatcacg 7200
cttttcgaca gcgtcggatt tgcgatcgag gatttttcgg cgctgcgcta cgtccgcgac 7260
cgcgttgagg gatcaagcca cagcagccca ctcgaccttc tagccgaccc agacgagcca 7320
agggatcttt ttggaatgct gctccgtcgt caggctttcc gacgtttggg tggttgaaca 7380
gaagtcatta tcgcacggaa tgccaagcac tcccgagggg aaccctgtgg ttggcatgca 7440
catacaaatg gacgaacgga taaacctttt cacgccattt taaatatccg attattctaa 7500
taaacgctct tttctcttag gtttacccgc caatatatcc tgtcaaacac tgatagttta 7560
aactgaaggc gggaaacgac aacctgatca tgagcggaga attaagggag tcacgttatg 7620

accccgccg atgacgcggg acaagccgtt ttacgtttgg aactgacaga accgcaacgt 7680

tgaaggagcc actcagctta attaagtcta actcgagtta ctggtacgta ccaaatccat 7740

ggaatcaagg taccatcaat cccgggtatt catcctaggt atccaagaat tcatactaaa 7800

gcttgcatgc ctgcaggtcg actctagtaa cggccgccag tgtgctggaa ttaattcggc 7860

ttgtcgacca cccaacccca tatcgacaga ggatgtgaag aacaggtaaa tcacgcagaa 7920

gaaccatct ctgatagcag ctatcgatta gaacaacgaa tccatattgg gtccgtggga 7980

aatacttact gcacaggaag ggggcgatct gacgaggccc cgccaccggc ctcgacccga 8040

ggccgaggcc gacgaagcgc cggcgagtac ggcgccgcgg cggcctctgc ccgtgccctc 8100

tgcgcgtggg agggagaggc cgcggtggtg ggggcgcgcg cgcgcgcgcg cgcagctggt 8160

gcggcggcgc gggggtcagc cgccgagccg gcggcgacgg aggagcaggg cggcgtggac 8220

gcgaacttcc gatcggttgg tcagagtgcg cgagttgggc ttagccaatt aggtctcaac 8280

aatctattgg gccgtaaaat tcatgggccc tggtttgtct aggccccaata tcccgttcat 8340

ttcagcccac aaatatttcc ccagaggatt attaaggccc acacgcagct tatagcagat 8400

caagtacgat gtttcctgat cgttggatcg gaaacgtacg gtcttgatca ggcatgccga 8460

cttcgtcaaa gagaggcggc atgacctgac gcggagttgg ttccgggcac cgtctggatg 8520

gtcgtaccgg gaccggacac gtgtcgcgcc tccaactaca tggacacgtg tggtgctgcc 8580

attgggccgt acgcgtggcg gtgaccgcac cggatgctgc ctcgcaccgc cttgcccacg 8640

ctttatatag agaggttttc tctccattaa tcgcatagcg agtcgaatcg accgaagggg 8700

aggggagcg aagctttgcg ttctctaatc gcctcgtcaa ggtaactaat caatcacctc 8760

gtcctaatcc tcgaatctct cgtggtgccc gtctaatctc gcgattttga tgctcgtggt 8820

ggaaagcgta ggaggatccc gtgcgagtta gtctcaatct ctcagggttt cgtgcgattt 8880

tagggtgatc cacctcttaa tcgagttacg gtttcgtgcg attttagggt aatcctctta 8940

atctctcatt gatttagggt ttcgtgagaa tcgaggtagg gatctgtgtt atttatatcg 9000

atctaataga tggattggtt ttgagattgt tctgtcagat ggggattgtt tcgatatatt 9060

accctaatga tgtgtcagat ggggattgtt tcgatatatt accctaatga tgtgtcagat 9120

ggggattgtt tcgatatatt accctaatga tggataataa gagtagttca cagttatgtt 9180

ttgatcctgc cacatagttt gagttttgtg atcagattta gttttactta tttgtgctta 9240

gttcggatgg gattgttctg atattgttcc aatagatgaa tagctcgtta ggttaaaatc 9300

tttaggttga gttaggcgac acatagttta tttcctctgg atttggattg gaattgtgtt 9360

cttagttttt ttccccttgga tttggattgg aattgtgtgg agctgggtta gagaattaca 9420

tctgtatcgt gtacacctac ttgaactgta gagcttgggt tctaaggtca atttaatctg 9480

tattgtatct ggctctttgc ctagttgaac tgtagtgctg atgttgtact gcgttttttt 9540

```
acccgtttta tttgctttac tcgtgcaaat caaatctgtc agatgctaga actaggtggc     9600
tttattctgt gttcttacat agatctgttg tcctgtagtt acttatgtca gttttgttat     9660
tatctgaaga tatttttggt tgttgcttgt tgatgtggtg tgagctgtga gcagcgctct     9720
tatgattaat gatgctgtcc aattgtagtg tagtatgatg tgattgatat gttcatctat     9780
tttgagctga cagtaccgat atcgtaggat ctggtgccaa cttattctcc agctgctttt     9840
ttttacctat gttaattcca atcctttctt gcctctttcc gatccagata atgcagaaac     9900
tcattaactc agtgcaaaac tatgcctggg gcagcaaaac ggcgttgact gaactttatg     9960
gtatggaaaa tccgtccagc cag                                             9983
```

ФОРМУЛА ВИНАХОДУ

1. Спосіб одержання розчинних цукрів з одержаного способами генної інженерії рослинного матеріалу, який включає:
попередню обробку шляхом змішування одержаного способами генної інженерії рослинного матеріалу з перетворюючим у пульпу складом, який включає бісульфіт амонію і карбонат амонію, з утворенням суміші з pH в діапазоні від 5,0 до 10,0, де одержаний способами генної інженерії рослинний матеріал включає першу полінуклеотидну послідовність, яка кодує перший білок, і другу полінуклеотидну послідовність, що кодує другий білок, і де перший білок містить амінокислотну послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю, вибраною з групи, яка складається з: SEQ ID NO: 1, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 і SEQ ID NO: 8, і другий білок містить амінокислотну послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю, вибраною з групи, яка складається з: SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 і SEQ ID NO: 12; і
забезпечення умов гідролізу.

2. Спосіб за п. 1, який додатково включає додавання до суміші іншого рослинного матеріалу.

3. Спосіб за п. 1, який додатково включає третю полінуклеотидну послідовність, яка кодує третій білок, вибраний із групи, що складається з: ксиланази, ендоглюканази, екзоглюканази, ферулоїлестерази, модифікованої інтеїном ксиланази, модифікованої інтеїном ендоглюканази, модифікованої інтеїном екзоглюканази і модифікованої інтеїном ферулоїлестерази; де білок, вибраний як третій білок, відрізняється від білка, вибраного як перший білок, і відрізняється від білка, вибраного як другий білок.

4. Спосіб за п. 3, де третій білок включає амінокислотну послідовність зі щонайменше 90 % ідентичністю з третьою референсною послідовністю, вибраною з групи, що складається з SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10 і SEQ ID NO: 11.

5. Спосіб за п. 3, де третій білок включає амінокислотну послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю SEQ ID NO: 12.

6. Спосіб за будь-яким з пп. 1 або 3, де щонайменше одна з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності та третьої полінуклеотидної послідовності додатково містить першу націлювальну полінуклеотидну послідовність, яка кодує відповідний націлювальний пептид, незалежно вибраний для кожної з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності з групи, що складається з: націлювального на амілопласт сигналу, націлювального на клітинну стінку пептиду, націлювального на мітохондрії пептиду, сигналу локалізації в цитозолі, націлювального на хлоропласти сигналу, націлювального в ядро пептиду і націлювального на вакуоль пептиду, де відповідний націлювальний пептид злитий з першим білком, другим білком або третім білком.

7. Спосіб за п. 6, де відповідна перша націлювальна полінуклеотидна послідовність розташована вище першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності, і націлювальний пептид незалежно вибраний із групи, що складається з: SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 і SEQ ID NO: 17.

8. Спосіб за п. 6, де перша націлювальна полінуклеотидна послідовність у комбінації зі щонайменше однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом кодують амінокислотну

послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю, вибраною із групи, що складається з: SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20 і SEQ ID NO: 21.

9. Спосіб за п. 6, де рослина додатково містить другу націлювальну полінуклеотидну послідовність, яка кодує карбоксинацілювальний пептид, злитий щонайменше з одним з першого білка, другого білка або третього білка, де карбоксинацілювальний пептид незалежно вибраний із групи, що складається з: SEQ ID NO: 22, SEQ ID NO: 23 і SEQ ID NO: 24.

10. Спосіб за п. 6, де рослина додатково містить другу націлювальну полінуклеотидну послідовність, яка кодує карбоксинацілювальний пептид, і де перша націлювальна полінуклеотидна послідовність і друга націлювальна полінуклеотидна послідовність у комбінації з однією з першої полінуклеотидної послідовності, другої полінуклеотидної послідовності або третьої полінуклеотидної послідовності разом кодують амінокислотну послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30 і SEQ ID NO: 31.

11. Спосіб за п. 1, де бісульфіт амонію знаходиться в концентрації від 0,02 М до 0,35 М і карбонат амонію знаходиться в концентрації від 0,025 М до 0,25 М.

12. Спосіб за п. 11, де суміш має співвідношення рідини і твердої речовини, вибране з величини, що менша або дорівнює одному з 10, 9, 8, 7, 6, 5, 4, 3, 2 або 1.

13. Спосіб за п. 12, де попередня обробка включає інкубацію суміші протягом періоду, що менший або дорівнює одному з 16 годин, 15 годин, 14 годин, 13 годин, 12 годин, 11 годин, 10 годин, 9 годин, 8 годин, 7 годин, 6 годин, 5 годин, 4 годин, 3 годин, 2 годин або 1 години.

14. Спосіб за п. 13, де попередня обробка включає забезпечення температури суміші від 40 °С до 95 °С.

15. Спосіб за п. 14, який додатково включає очищення суміші механічним подрібнюванням.

16. Спосіб за п. 1, де умови гідролізу включають від 2 % до 25 % твердих речовин.

17. Спосіб за п. 16, де умови гідролізу включають інкубацію суміші протягом періоду аж до 144 годин.

18. Спосіб за п. 17, де умови гідролізу включають температуру суміші, що становить 100 °С або менше.

19. Спосіб за п. 18, де температура суміші становить 65 °С або менше.

20. Спосіб за п. 19, де температура суміші становить від 48 °С до 50 °С.

21. Спосіб за п. 18, де умови гідролізу включають pH суміші від 4,8 до 5,0.

22. Спосіб за п. 21, який додатково включає контактування суміші з ферментуючим організмом для одержання біохімічного продукту.

23. Спосіб за п. 22, де ферментуючим організмом є дріжджі, вибрані з групи, що складається з: *Saccharomyces, Kluyveromyces, Pichia, Yorrowia, Spothaspora* і *Scheffersomyces.*

24. Спосіб за п. 1, який додатково включає додавання одного або декількох екзогенних ферментів.

25. Спосіб за п. 24, де один або декілька екзогенних ферментів включають ендоглюканазу, целобіогідролазу, глікозидазу і ксиланазу.

26. Спосіб за п. 25, де один або декілька екзогенних ферментів додатково включають β-ксилозидазу.

27. Спосіб за п. 24, де один або декілька екзогенних ферментів включають целюлазу, виділену з *Trichoderma reesii.*

28. Одержана способами генної інженерії рослина, що включає першу полінуклеотидну послідовність, яка кодує амінокислотну послідовність зі щонайменше 90 % ідєнтичністю з першою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 і SEQ ID NO: 8, та яка додатково включає другу полінуклеотидну послідовність, яка кодує амінокислотну послідовність зі щонайменше 90 % ідентичністю з другою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 і SEQ ID NO: 12.

29. Одержана способами генної інженерії рослина за п. 28, яка додатково включає третю полінуклеотидну послідовність, яка кодує амінокислотну послідовність зі щонайменше 90 % ідентичністю з третьою референсною послідовністю, вибраною з групи, що складається з SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 і SEQ ID NO: 12, де SEQ ID NO, вибрана як третя референсна послідовність, відрізняється від SEQ ID NO, вибраної як друга референсна послідовність.

30. Одержана способами генної інженерії рослина за п. 28 або 29, де перша полінуклеотидна послідовність, друга полінуклеотидна послідовність або третя полінуклеотидна послідовність

додатково включає націлювальну полінуклеотидну послідовність, що кодує націлювальний пептид, вибраний із групи, що складається з: націлювального на амілопласт сигналу, націлювального на клітинну стінку пептиду, націлювального на мітохондрії пептиду, сигналу локалізації в цитозолі, націлювального на хлоропласти сигналу, націлювального в ядро пептиду і націлювального на вакуоль пептиду.

31. Одержана способами генної інженерії рослина за п. 30, де націлювальний пептид вибраний із групи, що складається з: SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 22, SEQ ID NO: 23 і SEQ ID NO: 24.

32. Одержана способами генної інженерії рослина за п. 28 або 29, де модифікована рослина вибрана з групи, що складається з: кукурудзи, цукрової тростини, цукрового буряку, сорго, проса, міскантуса, евкаліпта, верби і тополі.

33. Одержана способами генної інженерії рослина, яка включає щонайменше одну полінуклеотидну послідовність, яка кодує амінокислотну послідовність зі щонайменше 90 % ідентичністю з послідовністю, вибраною з групи, що складається з: SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30 і SEQ ID NO: 31.

34. Касета експресії, яка містить: першу полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з першої референсної послідовності, вибраної з групи, що складається з: SEQ ID NO: 32, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38 і SEQ ID NO: 39; і другу полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з другої референсної послідовності, вибраної з групи, що складається з: SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42 і SEQ ID NO: 43.

35. Касета експресії за п. 34, яка додатково містить третю полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з третьою референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42 і SEQ ID NO: 43, де SEQ ID NO, вибрана як третя референсна послідовність, відрізняється від SEQ ID NO, вибраної як друга референсна послідовність.

36. Касета експресії за будь-яким з пп. 34 або 35, де перша полінуклеотидна послідовність, друга полінуклеотидна послідовність або третя полінуклеотидна послідовність додатково включає націлювальну полінуклеотидну послідовність, що кодує націлювальний пептид, вибраний із групи, що складається з: націлювального на амілопласт сигналу, націлювального на клітинну стінку пептиду, націлювального на мітохондрії пептиду, сигналу локалізації в цитозолі, націлювального на хлоропласти сигналу, націлювального в ядро пептиду і націлювального на вакуоль пептиду.

37. Касета експресії за п. 36, де націлювальна полінуклеотидна послідовність вибрана з групи, що складається з: SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50 і SEQ ID NO: 51.

38. Касета експресії за п. 37, яка містить полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з референсної послідовності, вибраної з групи послідовностей, що складається з: SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SRQ ID NO: 60, SEQ ID NO: 61 і SEQ ID NO: 62.

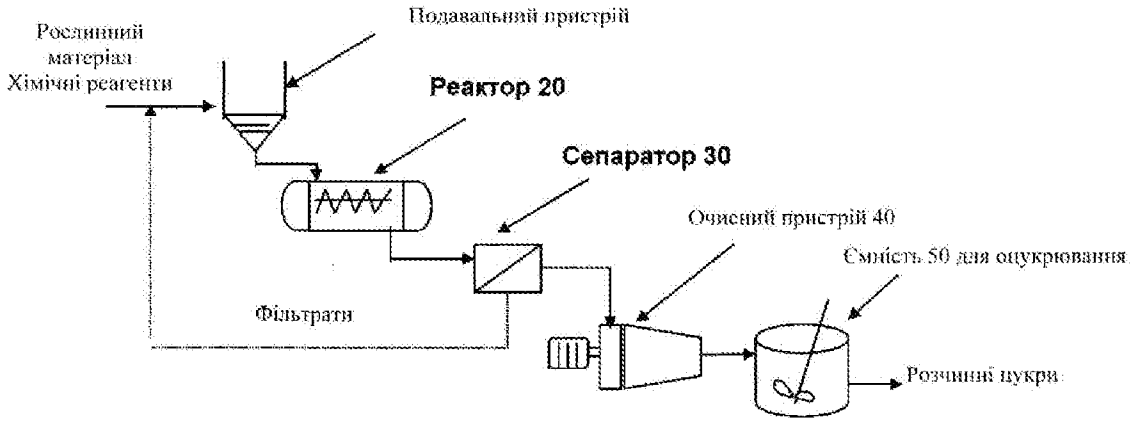
39. Вектор експресії, який містить полінуклеотидну послідовність, здатну гібридизуватися в умовах помірної жорсткості з нуклеїновою кислотою, що складається з послідовності, вибраної з групи, що складається з: SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82 і SEQ ID NO: 83.

40. Спосіб отримання розчинних цукрів з отриманого способами генної інженерії рослинного матеріалу, що включає:

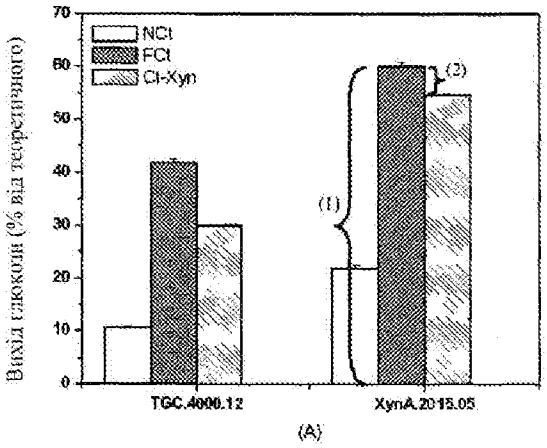
попередню обробку шляхом змішування отриманого способами генної інженерії рослинного матеріалу з перетворюючим в пульпу складом, який включає бісульфіт амонію і карбонат амонію, з утворенням суміші з pH в діапазоні від 5,0 до 9,0,

де бісульфіт амонію знаходиться в концентрації від 0,02 М до 0,18 М, і карбонат амонію знаходиться в концентрації від 0,025 М до 0,20 М, при цьому отриманий способами генної інженерії рослинний матеріал включає першу полінуклеотидну послідовність, яка кодує перший білок, і другу полінуклеотидну послідовність, яка кодує другий білок, і де перший білок містить амінокислотну послідовність зі щонайменше 90 % ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 1, SEQ ID NO: 5, SEQ ID NO: 6 , SEQ ID NO: 7 і SEQ ID NO: 8, і другий білок містить амінокислотну послідовність зі щонайменше 90 %

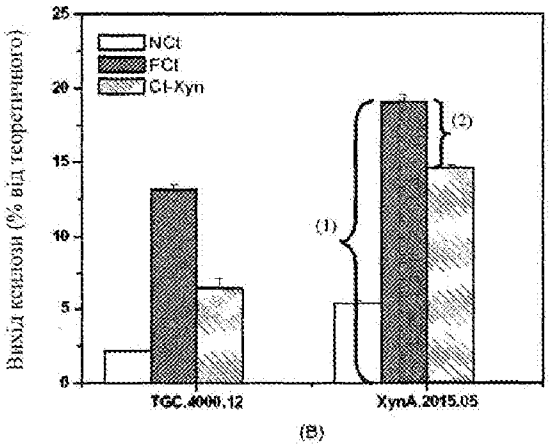
ідентичністю з референсною послідовністю, вибраною з групи, що складається з: SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 і SEQ ID NO: 12; і забезпечення умов гідролізу.



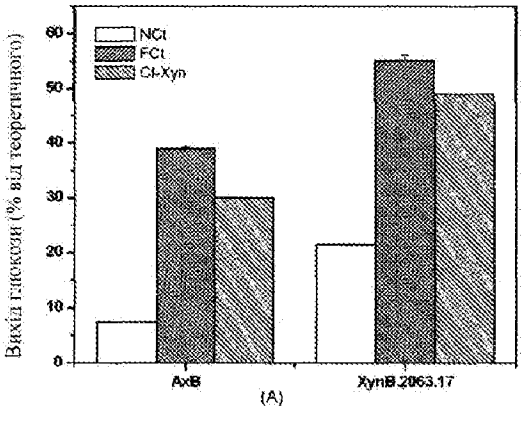
Фіг. 1



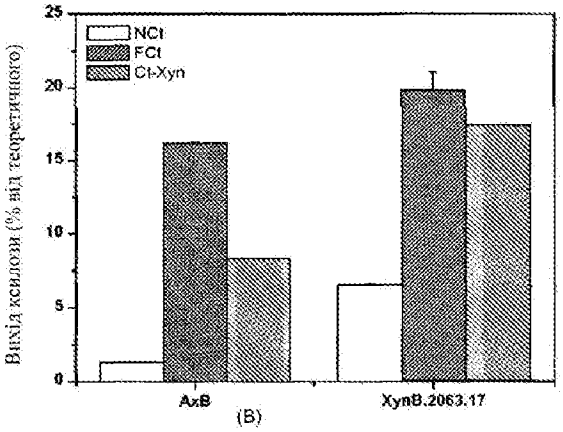
Фіг. 2A



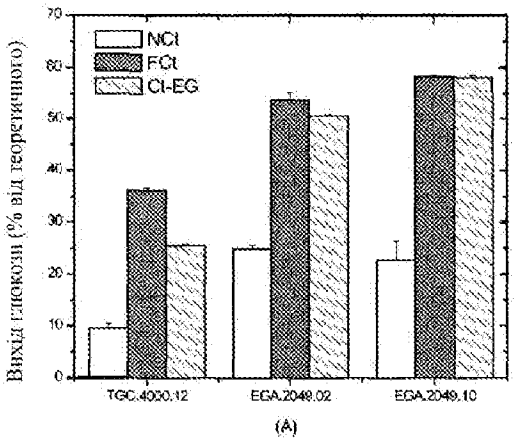
Фіг. 2B



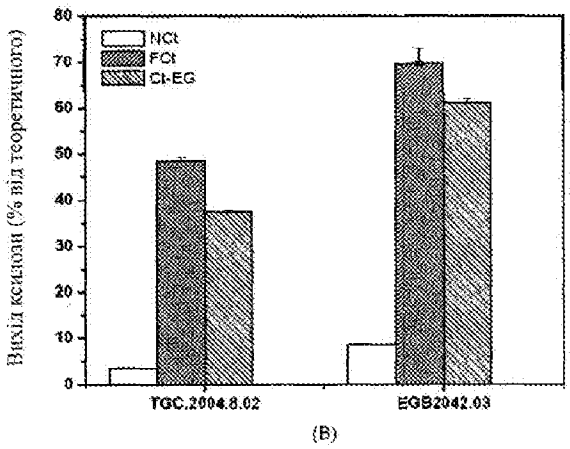
Фіг. 3A



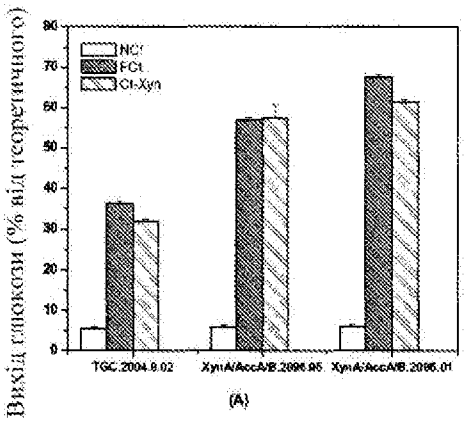
Фіг. 3B



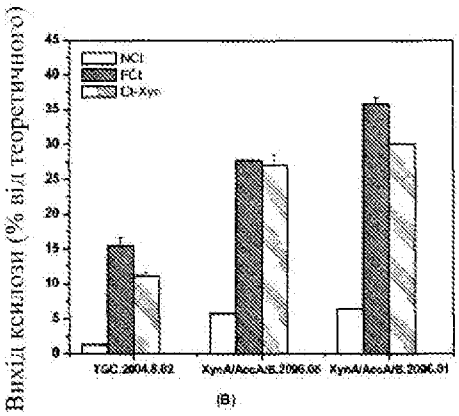
Фіг. 4A



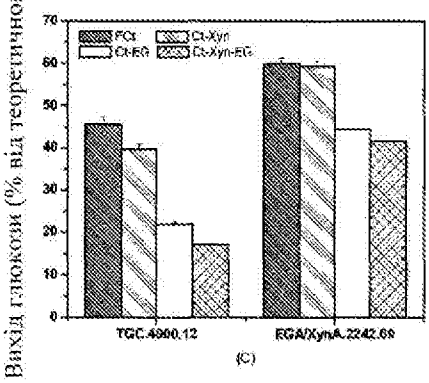
Фіг. 4B



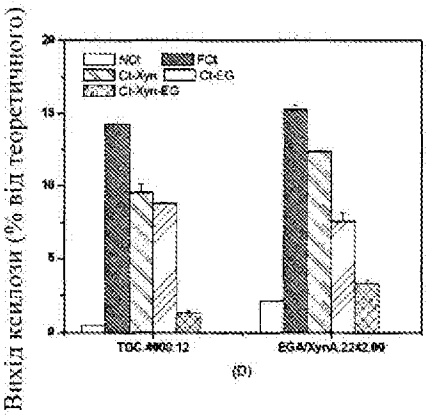
Фіг. 5A



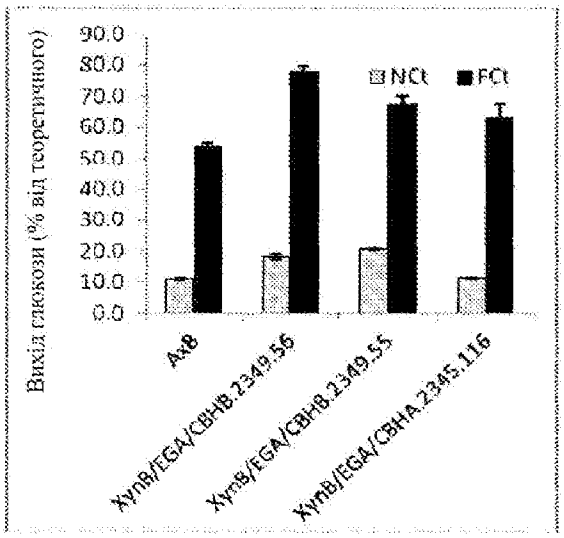
Фіг. 5B



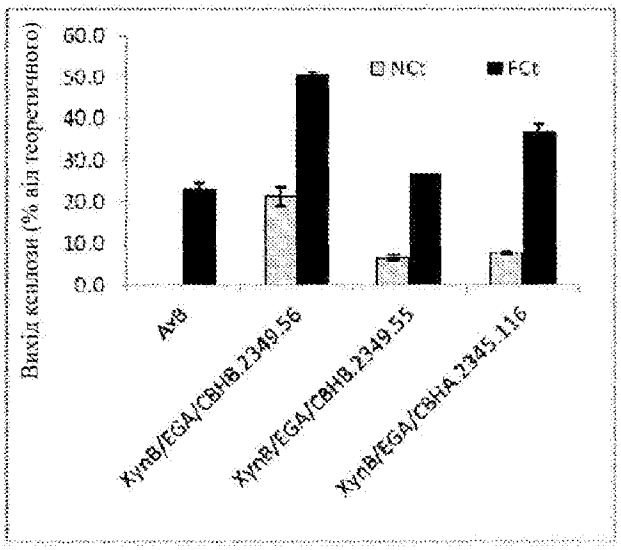
Фіг. 5C



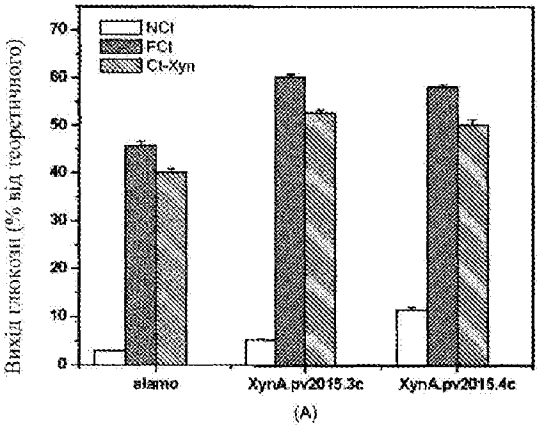
Фіг. 5D



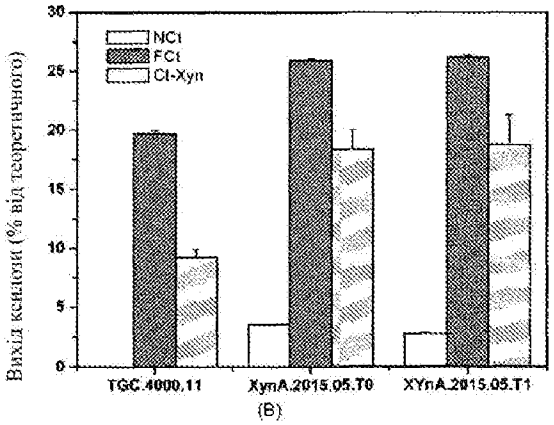
Фіг. 6А



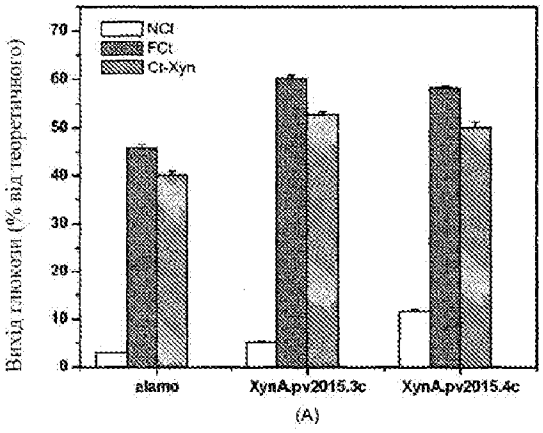
Фіг. 6В



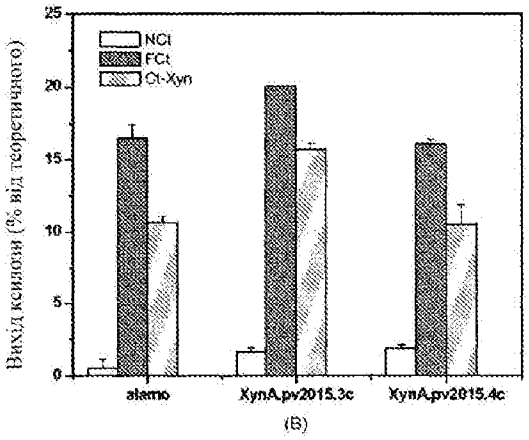
Фіг. 7А



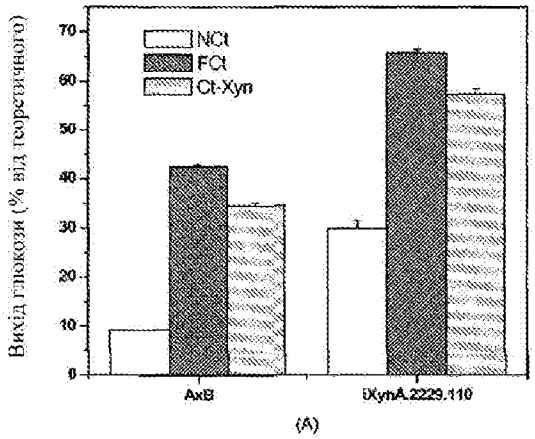
Фіг. 7В



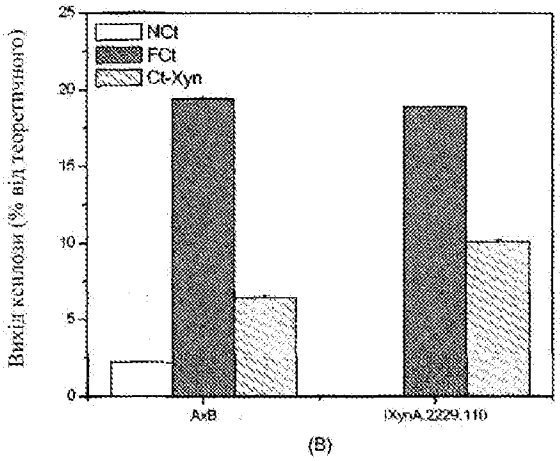
Фіг. 8А



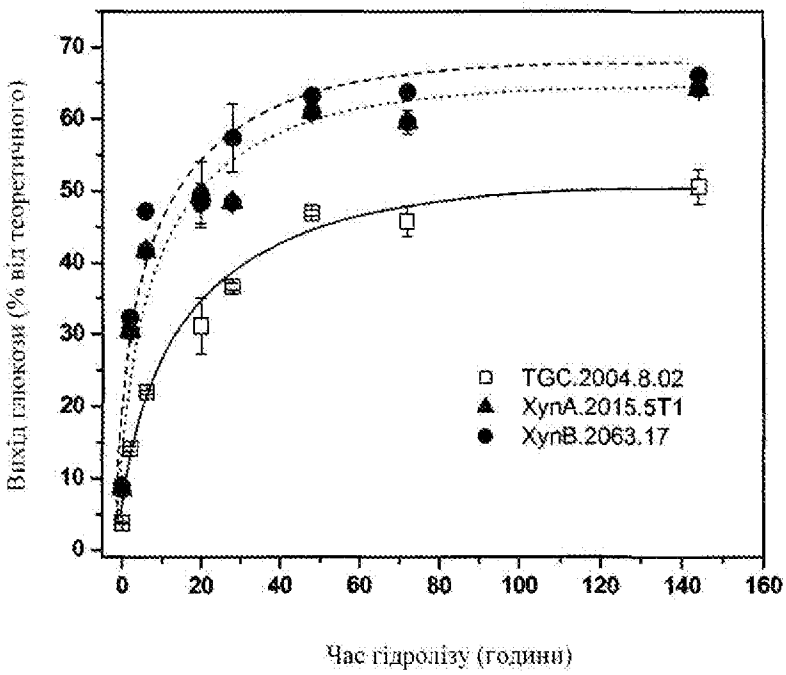
Фіг. 8В



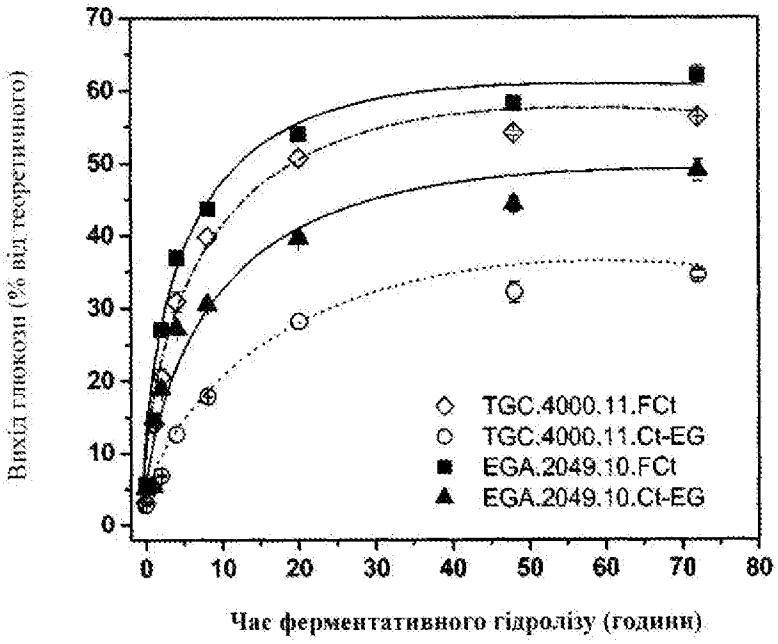
Фіг. 9A



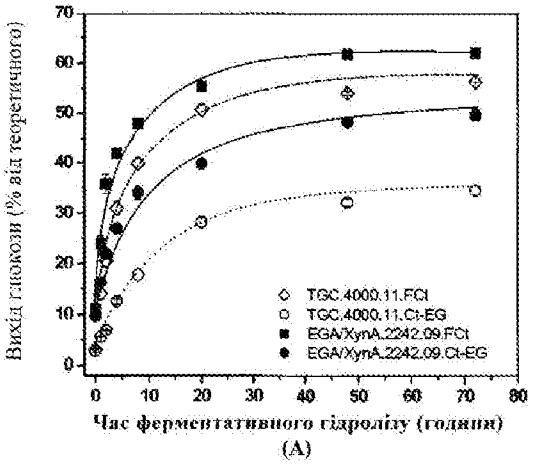
Фіг. 9B



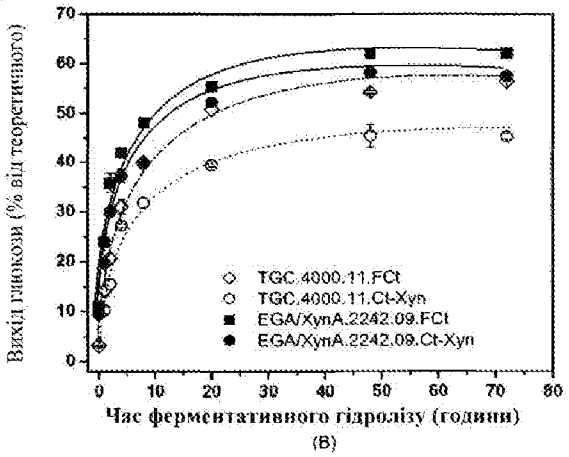
Фіг. 10



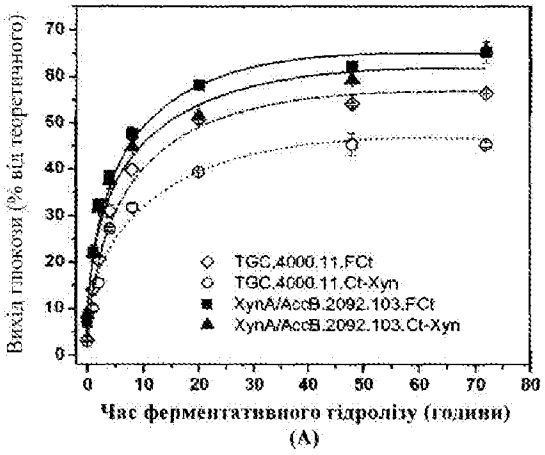
Фіг. 11



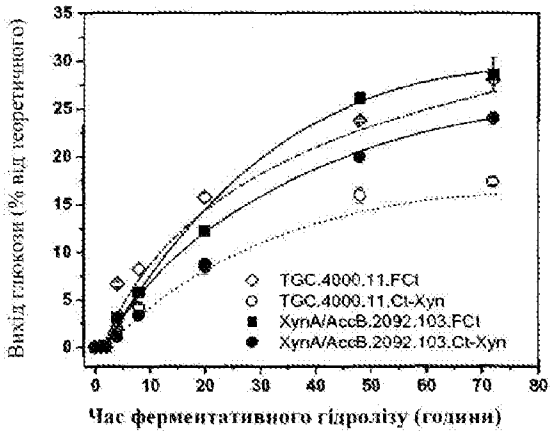
Фіг. 12A



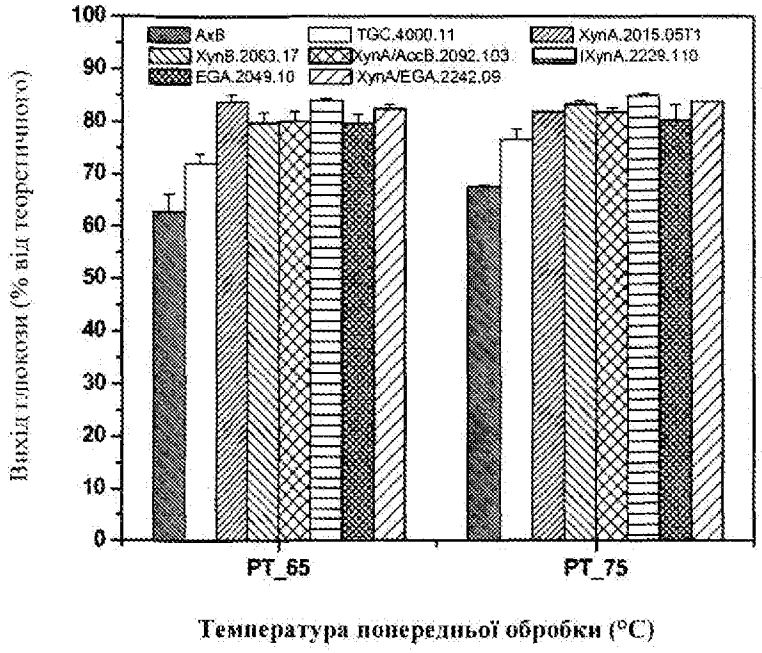
Фіг. 12B



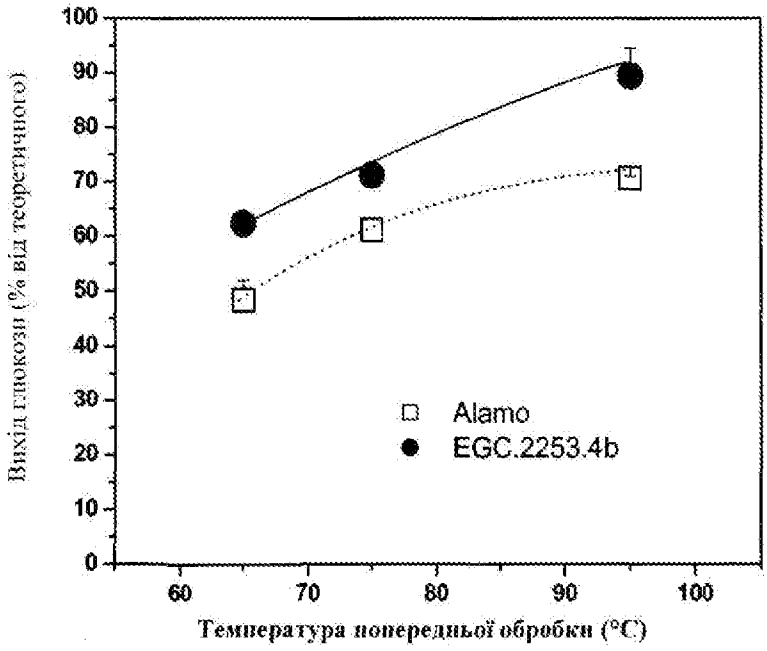
Фіг. 13A



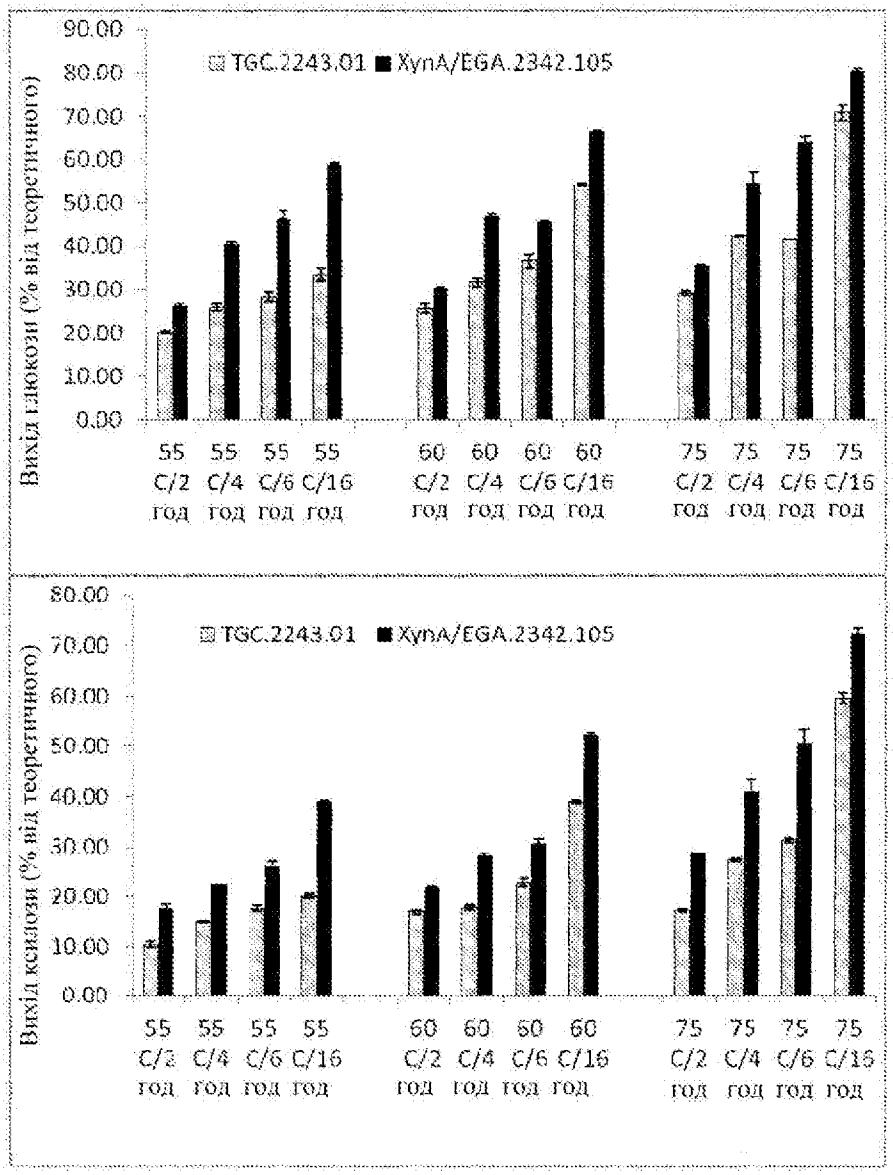
Фіг. 13B


AxB
TGC.4000.11
XynA.2015.0513
XynB.2063.17
XynA/AccB.2092.103
IXynA.2229.118
EGA.2049.10
XynA/EGA.2242.09
Вихід глюкози (% від теоретичного)
PT_65
PT_75
Температура попередньої обробки (°C)

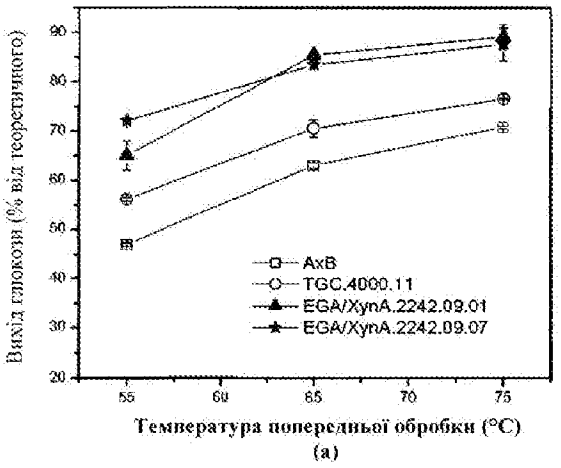
Фіг. 14


Вихід глюкози (% від теоретичного)
Alamo
EGC.2253.4b
Температура попередньої обробки (°C)

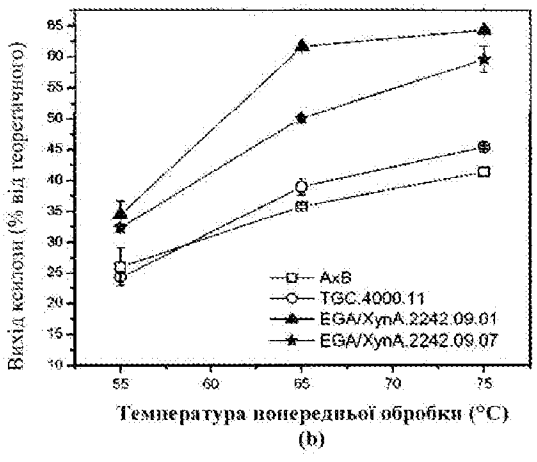
Фіг. 15



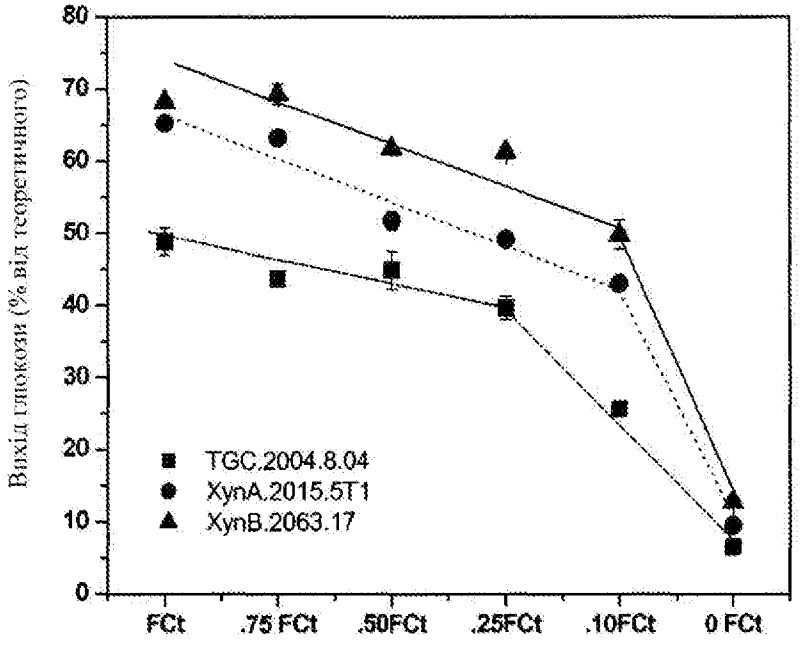
Фіг. 16



Фіг. 17A

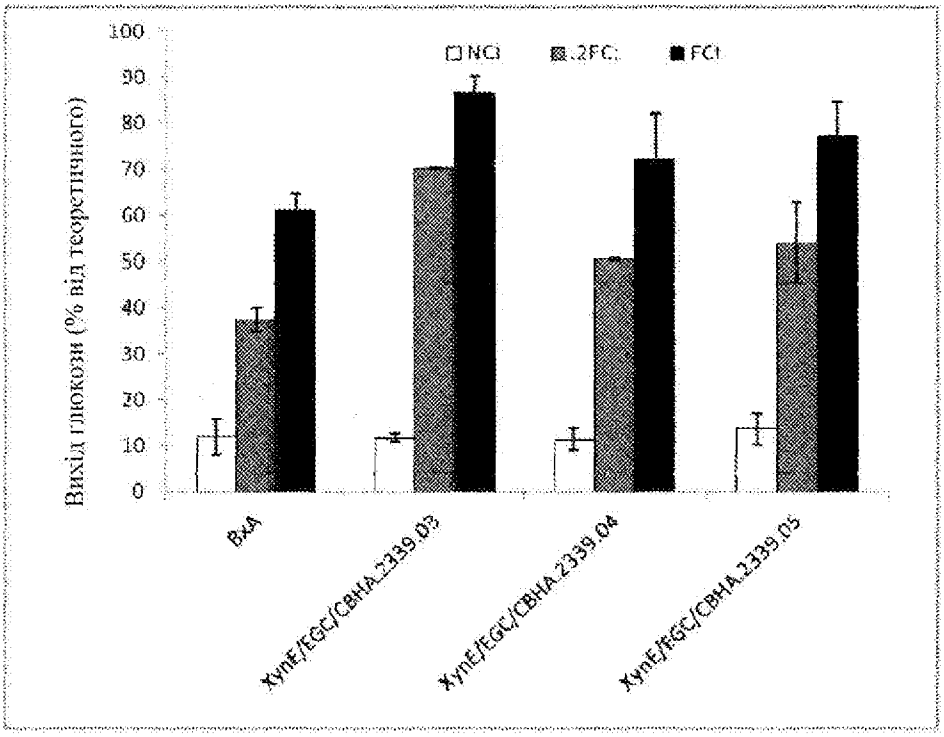


Фіг. 17B

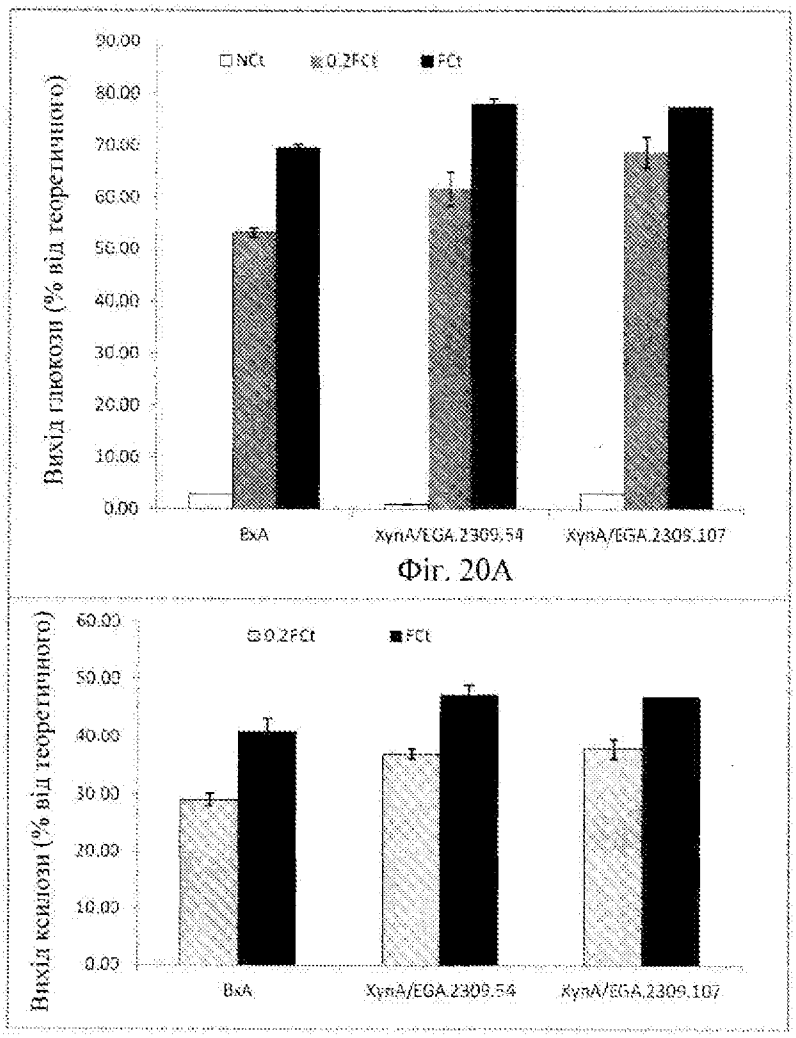


Власний коктейль #1

Фіг. 18

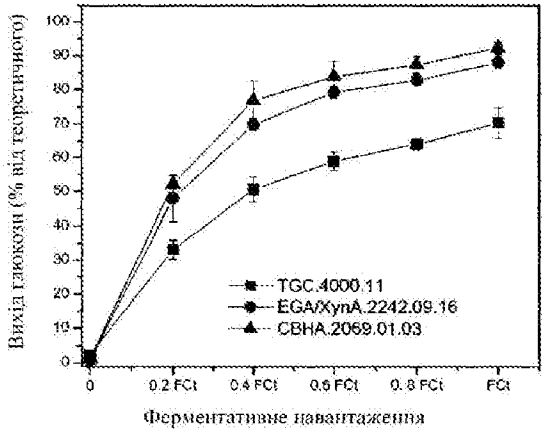


Фіг. 19

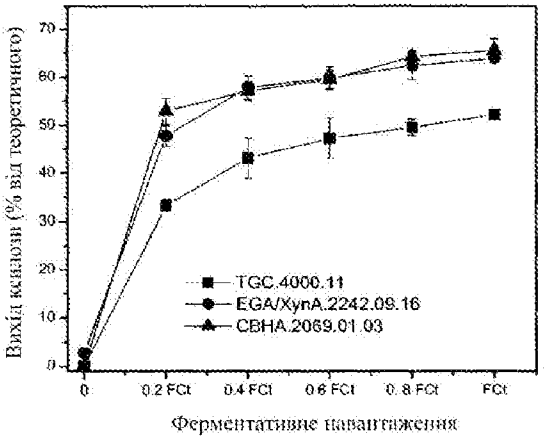


Фіг. 20A

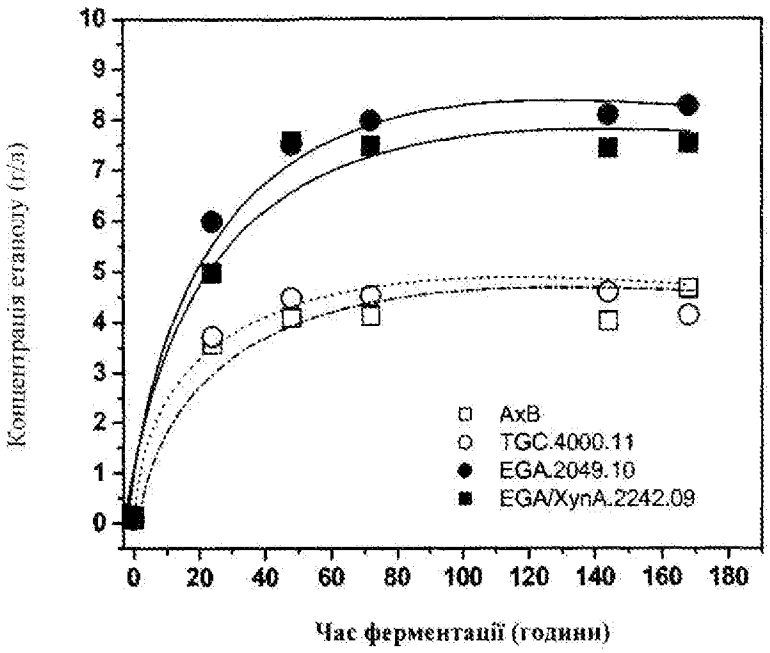
Фіг. 20B



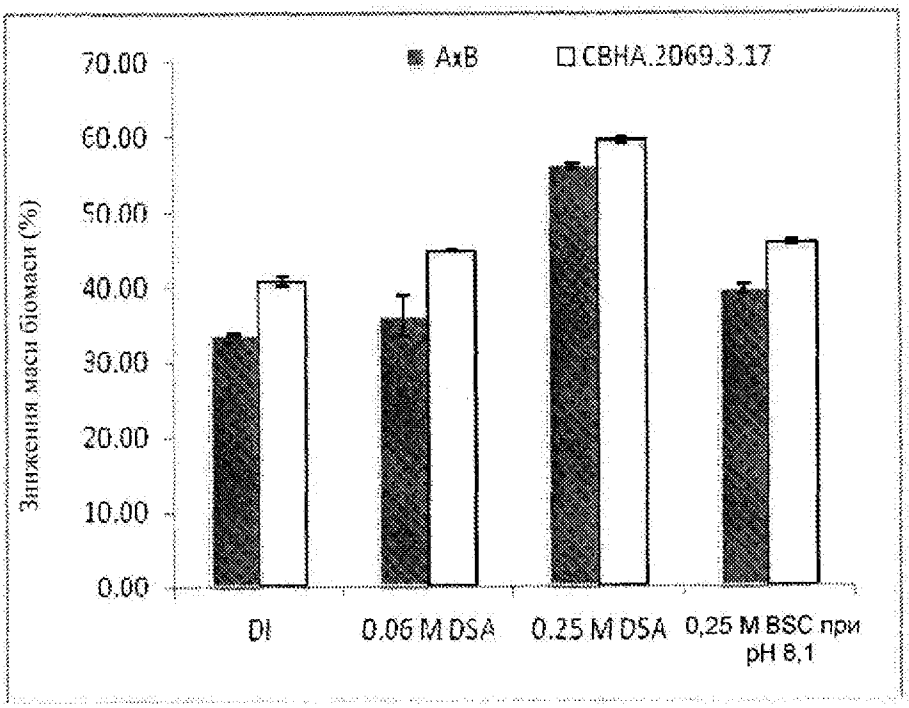
Фіг. 21A



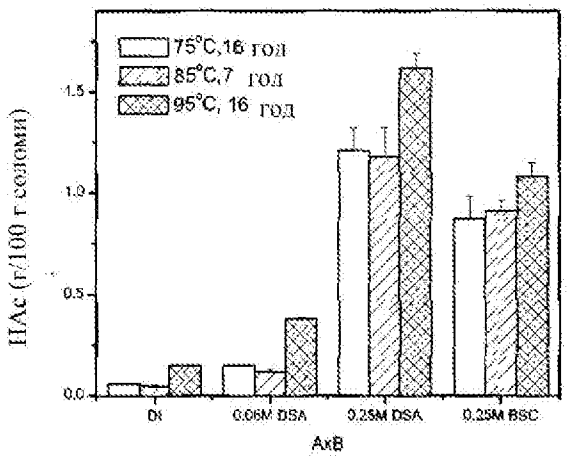
Фіг. 21B


Концентрація етанолу (г/л)
Час ферментації (години)
AxB
TGC.4000.11
EGA.2049.10
EGA/XynA.2242.09

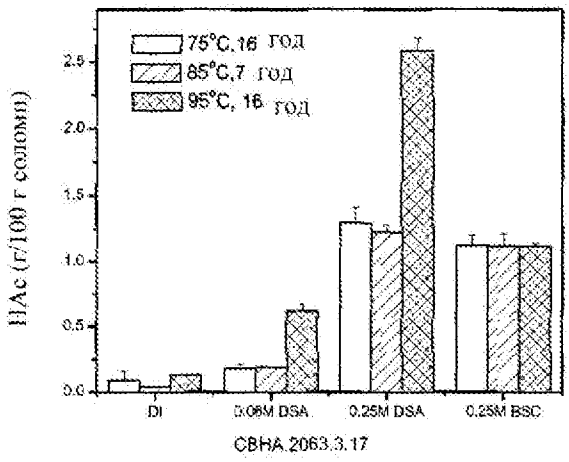
Фіг. 22


AxB
CBHA.2069.3.17
Зниження маси біомаси (%)
DI
0.06 M DSA
0.25 M DSA
0,25 M BSC при pH 8,1

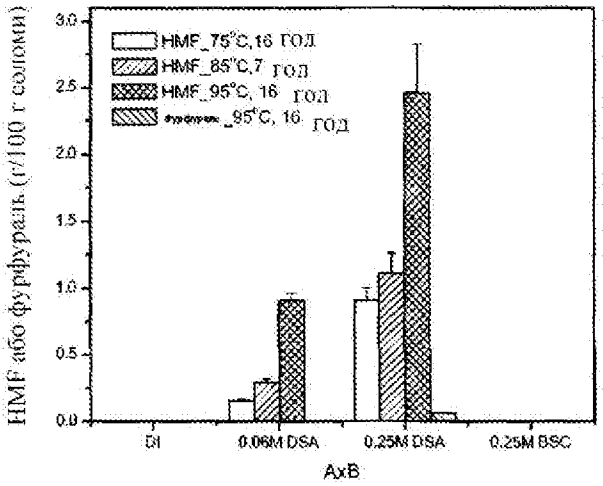
Фіг. 23



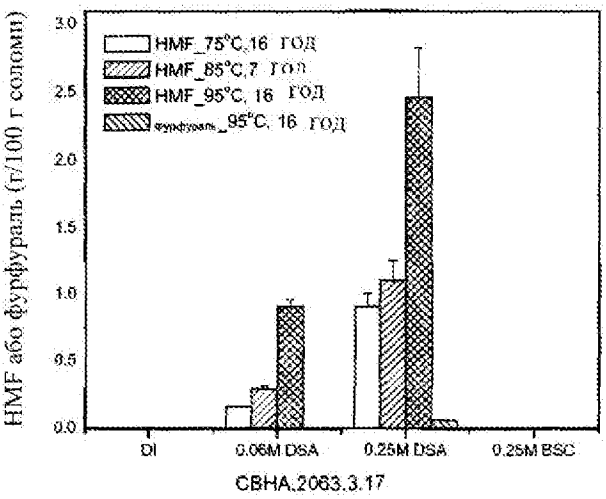
Фіг. 24А



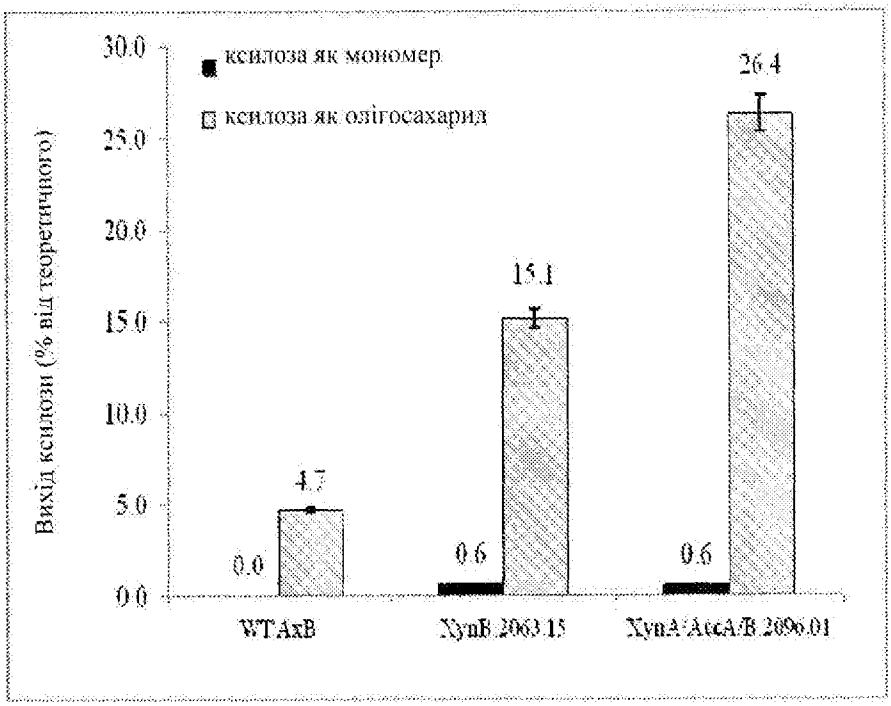
Фіг. 24В



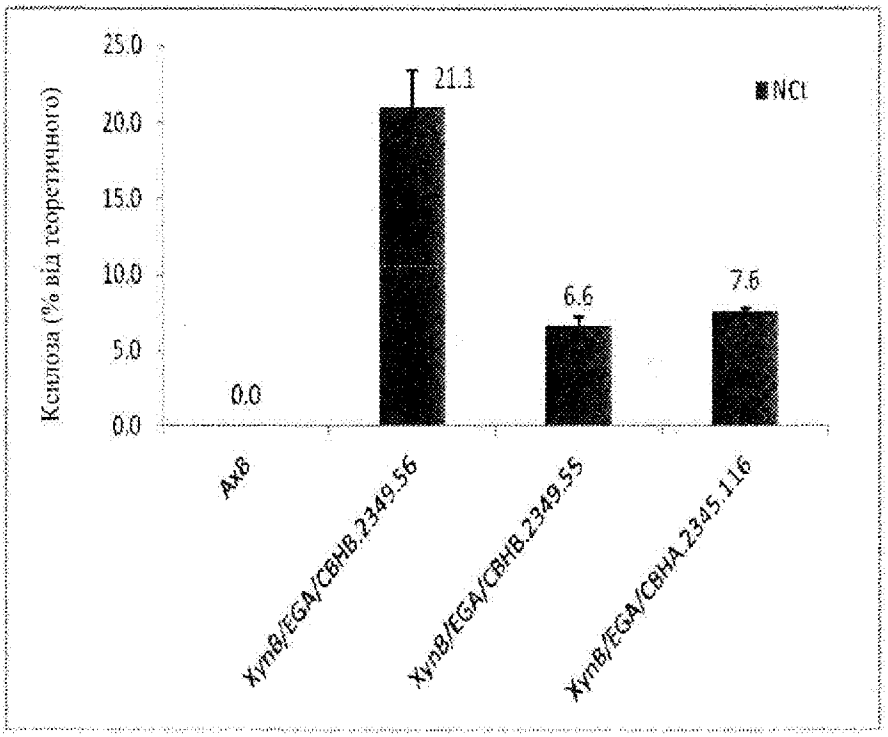
Фіг. 25A



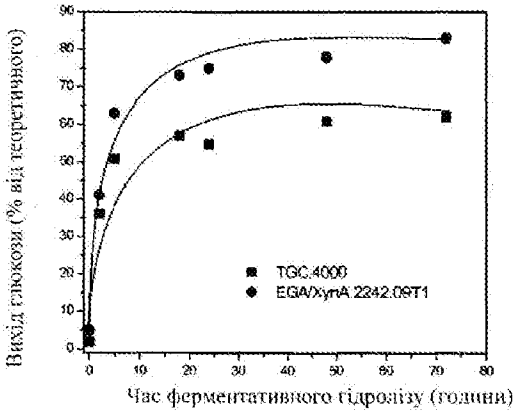
Фіг. 25B


ксилоза як мономер
ксилоза як олігосахарид
Вихід ксилози (% від теоретичного)
30.0
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.7
0.6
15.1
0.6
26.4
WT AxB
XynB 2063.15
XynA/AccA/B 2096.01

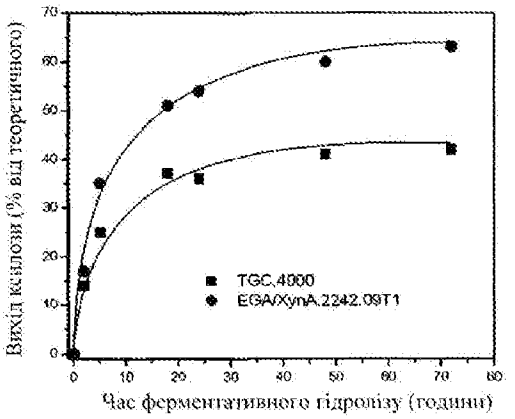
Фіг. 26


NCl
Ксилоза (% від теоретичного)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
21.1
6.6
7.6
AxB
XynB/EGA/CBHB.2349.56
XynB/EGA/CBHB.2349.55
XynB/EGA/CBHA.2345.116

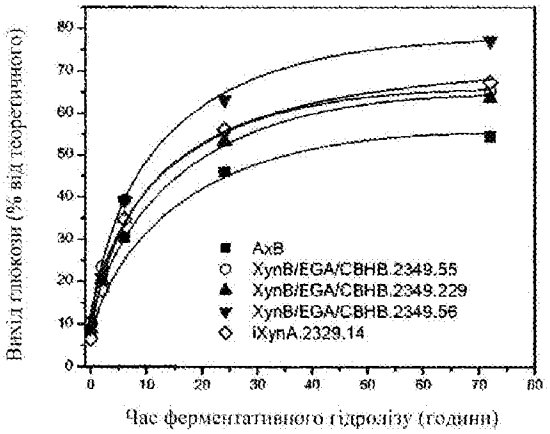
Фіг. 27



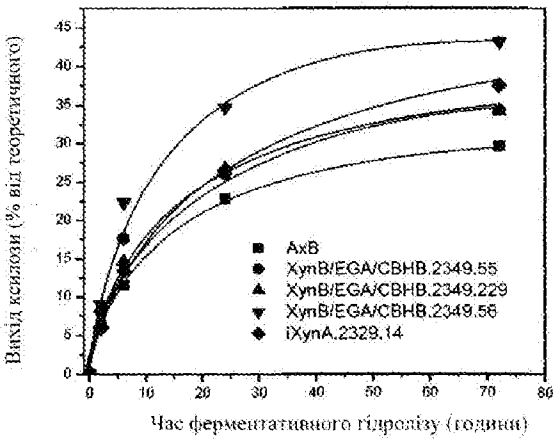
Фіг. 28A



Фіг. 28B



Фіг. 28C



Фіг. 28D

Комп'ютерна верстка М. Мацело

Державна служба інтелектуальної власності України, вул. Василя Липківського, 45, м. Київ, МСП, 03680, Україна

ДП "Український інститут інтелектуальної власності", вул. Глазунова, 1, м. Київ – 42, 01601