



УКРАЇНА

(19) UA

(11) 121960

(13) C2

(51) МПК

C12N 15/82 (2006.01)

C12N 15/113 (2010.01)

A01H 5/10 (2018.01)

МІНІСТЕРСТВО РОЗВИТКУ
ЕКОНОМІКИ, ТОРГІВЛІ ТА
СІЛЬСЬКОГО ГОСПОДАРСТВА
УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВІНАХІД

(21) Номер заявки: **а 2014 13871**
(22) Дата подання заявки: **11.08.2010**
(24) Дата, з якої є чинними права на винахід: **25.08.2020**
(31) Номер попередньої заявки відповідно до Паризької конвенції: **61/238233, 09169017.2**
(32) Дата подання попередньої заявки відповідно до Паризької конвенції: **31.08.2009, 31.08.2009**
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: **US, EP**
(41) Публікація відомостей про заявку: **12.05.2015, Бюл.№ 9**
(46) Публікація відомостей про видачу патенту: **25.08.2020, Бюл.№ 16**
(62) Номер та дата подання попередньої заявки, з якої виділено заявку, позначену кодом (21): **а201203925, 11.08.2010**

(72) Винахідник(и):
**Кун Йозеф Мартін (DE),
Лоялл Лінда Патріція (DE),
Зіберт Мальте (DE),
Дувеніг Ельке (DE)**
(73) Власник(и):
**БАСФ ПЛАНТ САЕНС КОМПАНІ ГМБХ,
67056 Ludwigshafen, Germany (DE)**
(74) Представник:
Шамріна Олена Олексіївна, реєстр. №141
(56) Перелік документів, взятих до уваги експертизою:
WO 2008/104559 A1, 04.09.2008
WO 93/20216 A1, 14.10.1993
WO 2007/112326 A1, 04.10.2007
"Arabidopsis thaliana cDNA clone:RAFL22-53-N05, 5' end.", EMBL, 22.01.2005, Database accession no. BP820219, URL: EBI, XP002766827
Chen Z-L et al, "A DNA Sequence Element that Confers Seed-Specific Enhancement to a Constitutive Promoter", EMBO (European Molecular Biology Organization) Journal, 1988, vol. 7, no. 2, P. 297 - 302
Gidekel M et al, "The first intron of the Arabidopsis thaliana gene coding for elongation factor 1beta contains an enhancer-like element", Gene, Elsevier, Amsterdam, NL, 08.05.1996, vol. 170, no. 2, P. 201 - 206

(54) СПОСІБ ОДЕРЖАННЯ ВИСОКОЕКСПРЕСІЙНОГО НАСІННО-СПЕЦИФІЧНОГО ТА/АБО НАСІННО-СЕЛЕКТИВНОГО РОСЛИННОГО ПРОМОТОРУ**(57) Реферат:**

Винахід стосується способу одержання вискоекспресійного насінно-специфічного та/або насінно-селективного рослинного промотору, що включає (а) забезпечення насінно-специфічного та/або насінно-селективного рослинного промотору, (b) функціональне зв'язування з промотором молекули нуклеїнової кислоти, яка посилює експресію нуклеїнової кислоти (NEENA), гетерологічної зазначеному промотору. Винахід також стосується рекомбінантної експресійної послідовності, яка включає NEENA, трансгенної клітини або трансгенної рослини або її частини, яка включає рекомбінантний вектор експресії, який включає одну або кілька рекомбінантних експресійних послідовностей, трансгенної рослини або її частини та застосування NEENA.

UA 121960 C2

Опис винаходу

Даний винахід належить до галузі молекулярної біології рослин і забезпечує способи одержання вискоекспресійних насінно-специфічних та/або насінно-селективних промоторів та одержання рослин з посиленою насінно-специфічною та/або насінно-селективною експресією нуклеїнових кислот, причому нуклеїнові кислоти, які посилюють експресію нуклеїнових кислот (NEENA), є функціонально зв'язаними з зазначеними промоторами та/або включеними у рослини.

Експресія трансгенів у рослинах зазнає сильного впливу різних зовнішніх та внутрішніх чинників в результаті чого експресія трансгенів стає мінливою й непередбачуваною. Часто доводиться одержувати й аналізувати велику кількість трансформантів для розпізнавання ліній з потрібною силою експресії. Оскільки трансформація та відбір ліній з потрібною силою експресії є дорогою й трудомісткою операцією, існує потреба у високій експресії одного або кількох трансгенів у рослині. Ця проблема є особливо очевидною, коли кілька генів мають бути узгоджено експресовані у трансгенній рослині з метою досягнення специфічного ефекту, оскільки треба виявити рослину, в якій сильно експресується кожен з генів.

Наприклад, експресія трансгена може виявляти значні коливання, залежно від конструкції послідовності та ефектів позиції генів місця інсерції Т-ДНК в окремих випадках перетворення. Сильні промотори дозволяють частково долати ці проблеми. Однак наявність прийнятних промоторів, які виявляють сильну експресію з потрібною специфічністю, часто буває обмеженою. Для забезпечення наявності достатніх промоторів з потрібною специфічністю експресії, здійснюють ідентифікацію та характеризацію додаткових промоторів, що дозволяє заповнити цей пропуск. Однак природна наявність промоторів з відповідною специфічністю та силою і довготривала процедура характеризації можливих промоторів перешкоджає ідентифікації прийнятних нових промоторів.

У пошуках шляхів подолання цих проблем було виявлено, що різні генетичні елементи та/або мотиви позитивно впливають на експресію генів. Серед них було розпізнано деякі інтрони як генетичні елементи з сильним потенціалом поліпшення експресії генів. Хоча механізм ще здебільшого є нез'ясованим, було продемонстровано, що деякі інтрони позитивно впливають на кількість зрілої мРНК у стійкому стані, можливо, через посилену транскрипційну активність, поліпшене визрівання мРНК, посилений ядерний експорт мРНК та/або поліпшену ініціацію трансляції [наприклад, Huang and Gorman, 1990; Le Hir et al., 2003; Nott et al., 2004]. Оскільки було виявлено, що лише вибрані інтрони посилюють експресію, сплайсинг як такий навряд чи відповідає за ефект, що спостерігається.

Посилення експресії генів, яке спостерігається після функціонального зв'язування інтронів з промоторами, називається опосередкованим інтроном посиленням (IME) експресії генів і виявляється у різних однодольних рослинах [наприклад, Gallis et al., 1987; Vasil et al., 1989; Bruce et al., 1990; Lu et al., 2008] та дводольних рослинах [наприклад, Chung et al., 2006; Kim et al., 2006; Rose et al., 2008]. У цьому відношенні було продемонстровано, що позиція інтрона відносно сайту початку трансляції (ATG) є ключовою для опосередкованого інтроном посилення експресії генів (Rose et al., 2004).

Крім потенціалу посилення експресії генів, деякі інтрони також продемонстрували здатність впливати на тканинну специфічність у їх природному нуклеотидному середовищі у рослинах. Було виявлено, що експресія генів-репортерів залежить від присутності геномних ділянок, які містять до двох інтронів [Sieburth et al., 1997; Wang et al., 2004]. Також повідомлялося, що 5'UTR інтрони мають значення для належної функціональності промоторних елементів, можливо, через тканинно-специфічні елементи генного контролю, які містяться в інтронах [Fu et al., 1995a; Fu et al., 1995b; Vitale et al., 2003; Kim et al., 2006]. Однак ці дослідження також демонструють, що комбінація інтронів з гетерологічними промоторами може мати сильний негативний вплив на силу та/або специфічність експресії генів [Vitale et al., 2003; Kim et al., 2006, документи WO2006/003186, WO2007/098042]. Наприклад, сильно конститутивний промотор вірусу мозаїки цвітної капусти CaMV35S зазнає негативного впливу через комбінацію з 5'UTR інтроном кунжуту SeFAD2 (Kim et al., 2006). На відміну від цих спостережень, у деяких документах демонструється посилення експресії нуклеїнової кислоти під дією IME без впливу на тканинну специфічність відповідного промотора (Schunmann et al., 2004). Інтрони або NEENA, які посилюють насінноспецифічну та/або насінно-селективну експресію, будучи функціонально зв'язаними з гетерологічним промотором, у джерелах існуючого рівня техніки не описувалися.

У даній заявці описуються інші молекули нуклеїнових кислот, які посилюють експресію зазначених промоторів без зашкодження їх специфічності після функціонального зв'язування з насінно-специфічними та/або насінно-селективними промоторами. Ці молекули нуклеїнових кислот описуються у даній заявці як "нуклеїнові кислоти, які посилюють експресію нуклеїнових кислот"

(NEENA). Інтрони мають характерну особливість, яка полягає у сплайсингу з відповідної пре-мРНК. Натомість нуклеїнові кислоти, представлені у даній заявці, не обов'язково мають бути включені до мРНК або, у разі присутності в мРНК, не обов'язково мають бути сплайсовані з мРНК для посилення експресії, зумовленої промотором, з яким є функціонально зв'язаними NEENA.

Детальний опис винаходу

Перший варіант втілення винаходу включає спосіб одержання вискоекспресійного насінно-специфічного та/або насінно-селективного промотора, причому спосіб включає функціональне зв'язування з промотором однієї або кількох молекул нуклеїнової кислоти, яка посилює експресію нуклеїнових кислот (NEENA), що включає

i) молекулу нуклеїнової кислоти, яка має послідовність, визначену у будь-якій з послідовностей SEQ ID NO: з 1 по 15, або

ii) молекулу нуклеїнової кислоти, яка має послідовність з ідентичністю 80 % або більше будь-якій з послідовностей, визначених у SEQ ID NO: з 1 по 15, в оптимальному варіанті ідентичність складає 85 % або більше, у ще кращому варіанті ідентичність складає 90 % або більше, у ще кращому варіанті ідентичність складає 95 % або більше, 96 % або більше, 97 % або більше, 98 % або більше або 99 % або більше, у найкращому варіанті ідентичність складає 100 % будь-якій з послідовностей, визначених у SEQ ID NO: з 1 по 15 або

iii) фрагмент зі 100 або більшої кількості послідовних основ, в оптимальному варіанті - 150 або більшої кількості послідовних основ, у ще кращому варіанті - 200 послідовних основ або більше, у ще кращому варіанті - 250 або більшої кількості послідовних основ молекули нуклеїнової кислоти i) або ii), який має активність посилення експресії, наприклад, 65 % або більше, в оптимальному варіанті - 70 % або більше, у ще кращому варіанті - 75 % або більше, у ще кращому варіанті - 80 % або більше, 85 % або більше або 90 % або більше, у найкращому варіанті - має 95 % або більше активності посилення експресії відносно показника відповідної молекули нуклеїнової кислоти, яка має послідовність згідно з будь-якою з послідовностей, визначених у SEQ ID NO: з 1 по 15, або

iv) молекулу нуклеїнової кислоти, яка є комплементом або зворотним комплементом будь-якої з вищезгаданих молекул нуклеїнової кислоти з i) по iii), або

v) молекулу нуклеїнової кислоти, яку одержують шляхом ПЛР з застосуванням олігонуклеотидних праймерів, описаних у SEQ ID NO: з 20 по 29, з 34 по 41, з 44 по 51 із 54 по 57, як показано у Таблиці 2, або

vi) молекулу нуклеїнової кислоти з 100 нуклеотидів або більше, 150 нуклеотидів або більше, 200 нуклеотидів або більше або 250 нуклеотидів або більше, яка гібридизується в умовах, рівноцінних гібридизації у 7 % додецилсульфату натрію (SDS), 0,5 M NaPO₄, 1 mM EDTA при 50 °C з промиванням у 2 X SSC, 0,1 % SDS при 50 °C або 65 °C, в оптимальному варіанті - 65 °C, з молекулою нуклеїнової кислоти, яка включає принаймні 50, в оптимальному варіанті - принаймні 100, у ще кращому варіанті - принаймні 150, у ще кращому варіанті - принаймні 200, у найкращому варіанті - принаймні 250 послідовних нуклеотидів нуклеотидної послідовності, що посилює транскрипцію, описаної у SEQ ID NO: з 1 по 15, або її комплементу. В оптимальному варіанті зазначена молекула нуклеїнової кислоти гібридизується в умовах, рівноцінних гібридизації у 7 % додецилсульфату натрію (SDS), 0,5 M NaPO₄, 1 mM EDTA при 50 °C з промиванням у 1 X SSC, 0,1 % SDS при 50 °C або 65 °C, в оптимальному варіанті - 65 °C, з молекулою нуклеїнової кислоти, яка включає принаймні 50, в оптимальному варіанті - принаймні 100, у ще кращому варіанті принаймні 150, у ще кращому варіанті - принаймні 200, у найкращому варіанті - принаймні 250 послідовних нуклеотидів нуклеотидної послідовності, що посилює транскрипцію, описаної у SEQ ID NO: з 1 по 15, або її комплементом, у ще кращому варіанті зазначена молекула нуклеїнової кислоти гібридизується в умовах, рівноцінних гібридизації у 7 % додецилсульфату натрію (SDS), 0,5 M NaPO₄, 1 mM EDTA при 50 °C з промиванням у 0,1 X SSC, 0,1 % SDS при 50 °C або GSX, в оптимальному варіанті - 65 °C, з молекулою нуклеїнової кислоти, яка включає принаймні 50, в оптимальному варіанті - принаймні 100, у ще кращому варіанті - принаймні 150, у ще кращому варіанті - принаймні 200, у найкращому варіанті - принаймні 250 послідовних нуклеотидів нуклеотидної послідовності, що посилює транскрипцію, описаної у будь-якій з послідовностей, визначених у SEQ ID NO: з 1 по 15, або її комплементом.

В одному варіанті втілення одна або кілька NEENA є гетерологічними промоторами, з яким вони є функціонально зв'язаними.

Як описано вище під пунктом v), молекула нуклеїнової кислоти, яку одержують шляхом ПЛР з застосуванням олігонуклеотидів, визначених у SEQ ID з 20 по 29, з 34 по 41, з 44 по 51 і з 54

по 57, як показано у Таблиці 2 може бути одержана, наприклад, з геномної ДНК рослин *Arabidopsis*, таких, як *A. thaliana*, з застосуванням умов, описаних нижче у Прикладі 1.

Спеціалістам у даній галузі відомо про можливість змін у температурному профілі, кількості циклів та/або складу або концентрації буфера для одержання відповідної молекули NEENA.

5 Конкретна комбінація олігонуклеотидів, яка має застосовуватись у відповідній ПЛР-реакції для одержання відповідної молекули NEENA, описується у Таблиці 2.

Спеціалістам у даній галузі відомі способи перетворення односпрямованого промотора на двоспрямований та способи застосування комплемента або зворотного комплемента послідовності промотора для створення промотора, який має таку саму специфічність, що й у первісної послідовності. Такі способи описуються, наприклад, для конститутивних, а також індукцибельних промоторів у публікації Xie et al. (2001) "Bidirectionalization of polar promoters in plants", *Nature Biotechnology* 19, стор. 677-679. Автори описують, що достатньо додати мінімальний промотор на 5'-кінці будь-якого промотора для одержання промотора, який контролює експресію в обох напрямках з однаковою специфічністю. Таким чином,

10 високоекспресійний промотор, функціонально зв'язаний з NEENA, як описано вище, функціонує у комплементі або зворотному комплементі, а отже, NEENA також є функціональною у комплементі або зворотному комплементі.

В принципі, NEENA може бути функціонально зв'язана з будь-яким промотором, таким як тканинно-специфічні, індукцибельні, специфічні для розвитку або конститутивні промотори.

20 Відповідна NEENA веде до посиленої насінно-специфічної та/або насінно-селективної експресії гетерологічної нуклеїнової кислоти під контролем відповідного промотора, з яким одна або кілька NEENA є функціонально зв'язаними. Посилення експресії промоторів, відрізняючись від насінно-специфічних та/або насінно-селективних промоторів, наприклад, конститутивних промоторів або промоторів з відмінною тканинною специфічністю, надає цим промоторам специфічності. Експресія нуклеїнової кислоти під контролем відповідного промотора значною мірою підвищується у насінні, у якому транскрипт зазначеної нуклеїнової кислоти може не виявлятися або лише слабо виявлятися без NEENA, функціонально зв'язаної з її промотором.

25 Отже, тканинно-специфічний або специфічний для розвитку або будь-який інший промотор може бути перетворений на насінно-специфічний та/або насінно-селективний промотор через функціональне зв'язування однієї або кількох молекул NEENA, як було описано вище, з зазначеним промотором. Таким чином, ще один варіант втілення винаходу полягає у забезпеченні способу надання специфічності будь-якого даного промотора, що функціонує у рослині, насінно-специфічному та/або насінно-селективному промоторові шляхом зв'язування відповідного промотора з молекулою NEENA, яка включає послідовність, описану вище під

30 пунктами з i) по vi).

В оптимальному варіанті одна або кілька NEENA є функціонально зв'язаними з будь-яким насінно-специфічним та/або насінно-селективним промотором і посилюють експресію молекули нуклеїнової кислоти, яка перебуває під контролем зазначеного промотора. Насінно-специфічні та/або насінно-селективні промотори, які мають застосовуватися за будь-яким способом згідно з винаходом, можуть походити з рослин, наприклад, однодольних або дводольних рослин, з бактерій та/або вірусів, або можуть бути синтетичними промоторами. Насінно-специфічними та/або насінно-селективними промоторами, які мають застосовуватися, є, наприклад, SBP-промотор з *Viola faba*, промотор невідомого насінного білка (USP) з *Vicia faba*, промотор напину з *Brassica napus*, conlinin-промотор з *Linum usitatissimum*, промотор з гена *A. thaliana* At5g01670, який кодує подібний до пероксиредоксину білок, промотор подібного до пероксиредоксину білка з *Linum usitatissimum*, промотор глобуліноподібного білка з *Brassica napus*, промотор argelin5-1 з *Phaseolus vulgaris*, промотор зеїну з *Zea mays*, промотор глобуліну з *Zea mays*, промотор pKG86 з *Zea mays*, як описано нижче у Прикладі 6, і т. ін.

40

45

Високоекспресійні насінно-специфічні та/або насінно-селективні промотори згідно з винаходом, функціонально зв'язані з NEENA, можуть застосовуватись у будь-якій рослині, включаючи, наприклад, мох, папороть, голонасінні або покритонасінні, наприклад, однодольні або дводольні рослини. В оптимальному варіанті втілення зазначений промотор згідно з винаходом, функціонально зв'язаний з NEENA, може застосовуватись в однодольних або дводольних рослинах, в оптимальному варіанті - у культурних рослинах, таких, як кукурудза,

50 соя, канола, бавовна, картопля, цукровий буряк, рис, пшениця, сорго, ячмінь, бананові, цукрова тростина, міскантус і т. ін.. В оптимальному варіанті втілення винаходу зазначений промотор, який є функціонально зв'язаним з NEENA, може застосовуватися в однодольних культурних рослинах, таких, як кукурудза, рис, пшениця, сорго, ячмінь, бананові, міскантус або цукрова тростина. У варіанті втілення, якому віддають особливу перевагу, промотор, функціонально

55

зв'язаний з NEENA, може застосовуватись у дводольних культурних рослинах, таких, як соя, канولا, бавовна або картопля.

Високоекспресійний насінно-специфічний та/або насінно-селективний промотор, який застосовують згідно з цією заявкою, означає, наприклад, промотор, який є функціонально зв'язаним з NEENA, викликаючи посилену насінно-специфічну та/або насінно-селективну експресію промотора у насінні рослини або його частині, причому накопичення в насінні РНК або швидкість синтезу РНК, що походить з молекули нуклеїнової кислоти під контролем відповідного промотора, функціонально зв'язаного з NEENA, є вищими, в оптимальному варіанті - суттєво вищими за експресію у насінні, викликану таким самим промотором, у якому відсутня NEENA згідно з винаходом. В оптимальному варіанті кількість РНК відповідної нуклеїнової кислоти та/або швидкість синтезу РНК та/або стійкість РНК у рослині підвищується на 50 % або більше, наприклад, 100 % або більше, в оптимальному варіанті - 200 % або більше, у ще кращому варіанті - у 5 разів або більше, у ще кращому варіанті - у 10 разів або більше, у найкращому варіанті - у 20 разів або більше, наприклад, у 50 разів, порівняно з контрольною рослиною такого самого віку, вирощуваною за таких самих умов, які включають такий самий насінно-специфічний та/або насінно-селективний промотор, коли останній не є функціонально зв'язаним з NEENA згідно з винаходом.

У контексті даного опису значно вищий означає статистичну значущість, спосіб визначення якої є відомим спеціалістам у даній галузі, наприклад, з застосуванням статистичних критеріїв, таких, як критерій Ст'юдента, до відповідних наборів даних.

Способи виявлення експресії, яка забезпечується промотором, є відомими спеціалістам у даній галузі. Наприклад, промотор може бути функціонально зв'язаним з маркерним геном, таким як GUS, GFP або люцифераза, і у рослині або її частині може визначатись активність відповідного білка, кодованого відповідним маркерним геном. Як типовий приклад нижче детально описується спосіб виявлення люциферази. Іншими способами є, наприклад, вимірювання рівня у стійкому стані або вимірювання швидкості синтезу РНК молекули нуклеїнової кислоти під контролем промотора з застосуванням способів, відомих спеціалістам у даній галузі, наприклад, нозерн-блоттингу, кількісної ПЛР, кінетичних аналізів або інших способів, описаних у джерелах, які належать до даної галузі.

Спеціалістам у даній галузі відомі різні способи функціонального зв'язування двох або більшої кількості молекул нуклеїнових кислот. Такі способи можуть включати рестрикцію/лігування, лігазно-незалежне клонування, рекомбінування, рекомбінацію або синтез. Можуть застосовуватися й інші способи функціонального зв'язування двох або більшої кількості молекул нуклеїнових кислот.

Іншим варіантом втілення даного винаходу є спосіб одержання рослини або її частини з посиленою, порівняно з відповідною контрольною рослиною або її частиною, насінно-специфічною та/або насінно-селективною експресією однієї або кількох молекул нуклеїнової кислоти, який включає етапи включення у рослину або її частину однієї або кількох NEENA, які включають молекулу нуклеїнової кислоти, як визначено вище під пунктами з i) по vi), та функціонального зв'язування зазначених однієї або кількох NEENA з промотором, в оптимальному варіанті - насінно-специфічним та/або насінно-селективним промотором, і з молекулою нуклеїнової кислоти, що перебуває під контролем зазначеного промотора, в оптимальному варіанті - насінно-специфічного та/або насінно-селективного промотора, причому NEENA є гетерологічною зазначеній молекулі нуклеїнової кислоти.

NEENA може бути гетерологічною молекулою нуклеїнової кислоти, яка перебуває під контролем зазначеного промотора, з яким NEENA є функціонально зв'язаною, або вона може бути гетерологічною як промоторові, так і молекулі нуклеїнової кислоти під контролем зазначеного промотора.

Термін "гетерологічна" стосовно молекули нуклеїнової кислоти або ДНК означає молекулу нуклеїнової кислоти, яка є функціонально зв'язаною, або підданою маніпуляції для функціонального зв'язування, з молекулою нуклеїнової кислоти, з якою вона є функціонально зв'язаною у природі, або з якою вона у природних умовах є функціонально зв'язаною в іншому місці. Наприклад, NEENA згідно з винаходом у природному середовищі є функціонально зв'язаною з її природним промотором, тоді як згідно з винаходом, вона є зв'язаною з іншим промотором, який може походити з того самого організму, з іншого організму, або може бути синтетичним промотором. Це також може означати, що NEENA згідно з даним винаходом зв'язується з її природним промотором, але молекула нуклеїнової кислоти під контролем зазначеного промотора є гетерологічною промоторові, який включає природну NEENA. Також слід розуміти, що промотор та/або молекула нуклеїнової кислоти під контролем зазначеного промотора, що функціонально зв'язується з NEENA згідно з винаходом, є гетерологічними

зазначеній NEENA, оскільки їхня послідовність піддавалася маніпуляції, наприклад, мутації, такий, як інсерція, делеція і т. ін., таким чином, що природна послідовність промотора та/або молекули нуклеїнової кислоти під контролем зазначеного промотора є модифікованою, а отже, стала гетерологічною NEENA згідно з винаходом. Також слід розуміти, що NEENA є гетерологічною нуклеїновою кислотою, з якою вона є функціонально зв'язаною, коли NEENA є функціонально зв'язаною з природним промотором, причому позиція NEENA відносно зазначеного промотора є зміненою таким чином, що промотор після такої маніпуляції демонструє вищу експресію.

Рослина, яка демонструє посилену насінно-специфічну та/або насінно-селективну експресію молекули нуклеїнової кислоти, у контексті даного опису означає рослину, яка має вищу, в оптимальному варіанті - статистично значно вищу насінно-специфічну та/або насінно-селективну експресію молекули нуклеїнової кислоти порівняно з контрольною рослиною, вирощуваною за таких самих умов без відповідної NEENA, функціонально зв'язаної з відповідною молекулою нуклеїнової кислоти. Такою контрольною рослиною може бути рослина дикого типу або трансгенна рослина, яка включає такий самий промотор, який контролює такий самий ген, що й у рослині згідно з винаходом, у якій промотор не є зв'язаним з NEENA згідно з винаходом.

У контексті даного винаходу одержання рослини включає способи стійкої трансформації, такі як включення послідовності рекомбінантної ДНК у рослину або її частину за допомогою опосередкованої *Agrobacterium* трансформації, протопластної трансформації, бомбардування частинками або іншого подібного способу та, необов'язково, наступної регенерації трансгенної рослини. Воно також включає способи тимчасової трансформації рослини або її частини, наприклад, вірусну інфекцію або інфільтрацію *Agrobacterium*. Спеціалістам у даній галузі відомі інші способи стійкої та/або тимчасової трансформації рослини або її частини. Для одержання рослини згідно з винаходом також можуть застосовуватися такі підходи, як селекція або злиття протопластів, і охоплюються обсягом даного винаходу.

Спосіб згідно з винаходом може застосовуватися до будь-якої рослини, наприклад, голонасінних або покритонасінних, в оптимальному варіанті - покритонасінних, наприклад, дводольних або однодольних рослин, в оптимальному варіанті - дводольних рослин. Оптимальними однодольними рослинами є, наприклад, кукурудза, пшениця, рис, ячмінь, сорго, бананові, цукрова тростина, міскантус та коротконіжка, причому особливу перевагу віддають таким однодольним рослинам, як кукурудза, пшениця та рис. Оптимальними дводольними рослинами є, наприклад, соя, рапс, канولا, льон, бавовна, картопля, цукровий буряк, календула та *Arabidopsis*, причому особливу перевагу віддають таким дводольним рослинам, як соя, рапс, канولا та картопля.

В одному варіанті втілення винаходу визначені вище способи включають етапи

а) включення однієї або кількох NEENA, які містять молекулу нуклеїнової кислоти, як визначено вище у пунктах з i) по vi), у рослину або її частину та

б) об'єднання зазначених однієї або кількох NEENA у геном зазначеної рослини або її частини, таким чином, щоб зазначена одна або кілька NEENA були функціонально зв'язаними з ендогенною, в оптимальному варіанті - насінно-специфічною та/або насінно-селективною експресованою нуклеїновою кислотою, гетерологічною зазначеним однією або кількома NEENA, і, необов'язково,

с) регенерації рослини або її частини, що включає зазначені одну або кілька NEENA з зазначеної трансформованої клітини.

Одна або кілька молекул NEENA можуть бути включені у рослину або її частину за допомогою бомбардування частинками, електропорації протопластів, вірусної інфекції, опосередкованої *Agrobacterium* трансформації або будь-яким іншим способом, відомим спеціалістам у даній галузі. Молекула NEENA може бути включена, наприклад, у плазмідну або вірусну ДНК або вірусну РНК. Молекула NEENA також може перебувати на ВАС, YAC або штучній хромосомі перед включенням у рослину або частину рослини. Вона також може бути включена як лінійна молекула нуклеїнової кислоти, яка включає послідовність NEENA, причому додаткові послідовності можуть бути присутні суміжно з послідовністю NEENA на молекулі нуклеїнової кислоти. Ці послідовності, які межують з послідовністю NEENA, можуть включати від приблизно 20 п. о., наприклад, від 20 п. о. до кількох сотень пар основ, наприклад, 100 п. о. або більше, і можуть сприяти включенню у геном, наприклад, шляхом гомологічної рекомбінації. Може застосовуватися будь-який інший спосіб включення геному, чи то способи спрямованого включення, такі як гомологічна рекомбінація, чи способи випадкового включення, такі як незаконна рекомбінація.

Ендогенна, в оптимальному варіанті - насінно-специфічна та/або насінно-селективна експресована нуклеїнова кислота, з якою може функціонально зв'язуватися молекула NEENA, може бути будь-якою нуклеїновою кислотою, в оптимальному варіанті будь-якою насінно-специфічною та/або насінно-селективною експресованою молекулою нуклеїнової кислоти.

5 Молекула нуклеїнової кислоти може бути кодуною білок молекулою нуклеїнової кислоти або некодуною молекулою, такою, як анти смислова РНК, рРНК, тРНК, мікроРНК, та-міРНК, міРНК, длРНК, мяРНК, мноРНК або будь-яка некодуєча РНК, відома спеціалістам уданій галузі.

10 Спеціалістам у даній галузі відомі способи ідентифікації насінно-специфічних та/або насінно-селективних експресованих молекул нуклеїнових кислот, яких в оптимальному варіанті може стосуватися спосіб згідно з винаходом, наприклад, шляхом гібридизації мікрочіпів, кількісної ПЛР, нозерн-блоттингу, секвенування наступного покоління і т. ін..

Ще один шлях виконання способів згідно з винаходом може полягати у

15 а) забезпеченні експресійної послідовності, яка включає одну або кілька NEENA, які включають молекулу нуклеїнової кислоти, як визначено вище у пунктах з і) по vi), функціонально зв'язану з промотором, в оптимальному варіанті - насінно-специфічним та/або насінно-селективним промотором, як визначено вище, і з однією або кількома молекулами нуклеїнової кислоти, причому остання є гетерологічною зазначеним однією або кількома NEENA, перебуваючи під контролем зазначеного промотора, в оптимальному варіанті насінно-специфічного та/або насінно-селективного промотора, та

20 б) включенні зазначеної експресійної послідовності, яка включає зазначені одну або кілька NEENA, у геном зазначеної рослини або її частини та, необов'язково,

с) регенерації рослини або її частини, яка включає зазначені одну або кілька експресійних послідовностей, з зазначеної трансформованої рослини або її частини.

25 NEENA може бути гетерологічною молекулою нуклеїнової кислоти, яка перебуває під контролем зазначеного промотора, з яким NEENA є функціонально зв'язаною, або вона може бути гетерологічною як промоторові, так і молекулі нуклеїнової кислоти під контролем зазначеного промотора.

Експресійна послідовність може бути включена у геном відповідної рослини будь-яким способом, відомим спеціалістам у даній галузі. Включення може бути випадковим, з застосуванням таких способів, як бомбардування частинками або опосередкована Agrobacterium трансформація. В оптимальному варіанті втілення включення здійснюють через спрямоване включення наприклад, шляхом гомологічної рекомбінації. Останній спосіб дозволяє включати експресійну послідовність, яка включає вискоекспресійний промотор, функціонально зв'язаний з NEENA, у прийнятну геномну ділянку. Прийнятними геномними ділянками є, 35 наприклад, геномні ділянки, які включають гени, які сильно експресуються, наприклад, у насінні, а отже, можуть посилювати експресію, зумовлену зазначеною експресійною послідовністю, порівняно з геномною ділянкою, яка не демонструє транскрипційної активності.

В іншому оптимальному варіанті втілення зазначені одна або кілька NEENA є функціонально зв'язаними з промотором, в оптимальному варіанті - насінно-специфічним та/або насінно-селективним промотором, наближеним до сайта початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти.

Наближення до сайта початку транскрипції у контексті даного опису означає функціональне зв'язування однієї або кількох NEENA з промотором, в оптимальному варіанті насінно-специфічним та/або насінно-селективним промотором, 2500 п. о. або менше, в оптимальному 45 варіанті - 2000 п. о. або менше, ще краще - 1500 п. о. або менше, ще краще - 1000 п. о. або менше, найкраще - 500 п. о. або менше, від сайта початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти. Слід розуміти, що NEENA може бути включена на відповідній відстані від початку транскрипції відповідного промотора перед ним або після нього. Таким чином, одна або кілька NEENA не обов'язково мають бути включеними у транскрипт 50 відповідної гетерологічної нуклеїнової кислоти під контролем промотора, в оптимальному варіанті - насінно-специфічного та/або насінно-селективного промотора, з яким одна або кілька NEENA є функціонально зв'язаними. В оптимальному варіанті одна або кілька NEENA включаються після сайта початку транскрипції відповідного промотора, в оптимальному варіанті - насінно-специфічного та/або насінно-селективного промотора. Сайт включення після сайта 55 початку транскрипції може бути в 5'UTR, 3'UTR, екзоні або інтроні, або може замінювати інтрон або, частково або повністю, 5'UTR або 3'UTR гетерологічної нуклеїнової кислоти під контролем промотора, в оптимальному варіанті - насінно-специфічного та/або насінно-селективного промотора. В оптимальному варіанті одна або кілька NEENA включаються у 5'UTR або інтрон, або NEENA замінює інтрон або, частково або повністю, 5'UTR, у найкращому варіанті 60 включається до 5'UTR відповідної гетерологічної нуклеїнової кислоти.

Ще один варіант втілення винаходу стосується рекомбінантної експресійної послідовності, яка включає одну або кілька NEENA, які включають молекулу нуклеїнової кислоти, як визначено вище у пунктах з i) по vi).

5 Рекомбінантна експресійна послідовність також може включати один або кілька промоторів, в оптимальному варіанті - насінно-специфічний та/або насінно-селективний промотор, з яким одна або кілька NEENA є функціонально зв'язаними, та, необов'язково, одну або кілька експресованих молекул нуклеїнової кислоти, причому остання є гетерологічною зазначеним одній або кільком NEENA.

10 NEENA може бути гетерологічною молекулі нуклеїнової кислоти, яка перебуває під контролем зазначеного промотора, з яким NEENA є функціонально зв'язаною, або вона може бути гетерологічною як промоторів, так і молекулі нуклеїнової кислоти під контролем зазначеного промотора.

15 Експресійна послідовність може включати одну або більше, наприклад, дві або більше, наприклад, 5 або більше, наприклад, 10 або більше комбінацій промоторів, в оптимальному варіанті - насінно-специфічних та/або насінно-селективних промоторів, функціонально зв'язаних з NEENA, та молекулу нуклеїнової кислоти, яка має бути експресована, яка є гетерологічною відповідній NEENA. Експресійна послідовність також може включати інші промотори, які не включають NEENA, функціонально зв'язані з молекулами нуклеїнових кислот, які мають бути експресовані, гомологічні або гетерологічні відповідному промоторові.

20 Вектор рекомбінантної експресії, який включає одну або кілька рекомбінантних експресійних послідовностей, як визначено вище, є ще одним предметом втілення винаходу. Багато векторів експресії, які можуть застосовуватися згідно з даним винаходом, є відомими спеціалістам у даній галузі. Способи включення такого вектора, який включає таку експресійну послідовність, яка включає, наприклад, промотор, функціонально зв'язаний з NEENA, та, необов'язково, інші елементи, такі як термінатор, у геном рослини та відтворення трансгенних рослин з трансформованої клітини також є добре відомими спеціалістам у даній галузі. Залежно від способу, який застосовують для трансформації рослини або її частини, у геном зазначеної рослини або її частини може бути включений повний вектор, або ж у геном можуть бути включені певні компоненти вектора, такі, як, наприклад, Т-ДНК.

30 Трансгенна рослина або її частина, яка включає одну або кілька гетерологічних NEENA, як визначено вище у пунктах з i) по vi), також є предметом цього винаходу. NEENA вважається гетерологічною рослиною, якщо вона є синтетичною, походить з іншого організму або того ж самого організму, але відтворюється її природна геномна локалізація порівняно з контрольною рослиною, наприклад, рослиною дикого типу. Слід розуміти, що відтворена геномна локалізація означає, що NEENA розташовується на іншій хромосомі або на тій самій хромосомі, але є на 10 т. н. або більше, наприклад, 10 т. н., в оптимальному варіанті - 5 т. н. або більше, наприклад, 5 т. н., у ще кращому варіанті - 1000 п. о. або більше, наприклад, 1000 п. о., у ще кращому варіанті - 500 п. о. або більше, наприклад, 500 п. о., у найкращому варіанті - 100 п. о. або більше, наприклад, 100 п. о., у найкращому варіанті - 10 п. о. або більше, наприклад, 10 п. о. зсунутою відносно її природної геномної локалізації, наприклад, у рослині дикого типу.

40 Трансгенна клітина або трансгенна рослина або її частина, яка включає вектор рекомбінантної експресії, як визначено вище, або рекомбінантну експресійну послідовність, як визначено вище, є ще одним предметом втілення винаходу. Трансгенна клітина, трансгенна рослина або її частина може бути вибрана з групи, до якої належать бактерії, грибки, дріжджі, або клітини рослин, комах або ссавців або рослини. Оптимальними трансгенними клітинами є бактерії, грибки, дріжджі, рослинні клітини. Серед бактерій перевагу віддають кишковим бактеріям, таким як *E. coli*, та бактеріям роду *Agrobacteria*, наприклад, *Agrobacterium tumefaciens* та *Agrobacterium rhizogenes*. Серед рослин перевагу віддають однодольним або дводольним рослинам, наприклад, однодольним або дводольним культурним рослинам, таким як кукурудза, соя, канولا, бавовна, картопля, цукровий буряк, рис, пшениця, сорго, ячмінь, бананові, цукрова тростина, міскантус і т. ін. Серед культурних рослин перевагу віддають таким, як кукурудза, рис, пшениця, соя, канولا, бавовна або картопля. Серед дводольних культурних рослин особливу перевагу віддають таким, як соя, канولا, бавовна або картопля.

50 Серед однодольних культурних рослин особливу перевагу віддають таким, як кукурудза, пшениця та рис.

55 Трансгенна клітинна культура, трансгенне насіння, частини або матеріал для розмноження, які походять з трансгенної клітини або рослини або її частини, як визначено вище, які включають зазначену гетерологічну NEENA, як визначено вище у пунктах з i) по vi), або зазначену рекомбінантну експресійну послідовність або зазначений рекомбінантний вектор, як визначено вище, є іншими предметами втілення винаходу.

Трансгенні частини або матеріал для розмноження у контексті даного опису означають усі тканини та органи, наприклад, листя, стебла та плоди, а також матеріал, який застосовують для розмноження та/або регенерації рослин, такий як живці, пагони, відгілки, гілки або паростки, які включають відповідну NEENA, рекомбінантну експресійну послідовність або рекомбінантний вектор.

Інший варіант втілення винаходу стосується застосування NEENA, як визначено вище у пунктах з i) по vi), або рекомбінантної послідовності або рекомбінантного вектора, як визначено вище, для посилення експресії у рослинах або їх частинах.

Таким чином, дана заявка забезпечує молекули нуклеїнових кислот які посилюють насінно-специфічну та/або насінно-селективну експресію генів, які включають один або кілька промоторів, в оптимальному варіанті - насінно-специфічний та/або насінно-селективний промотор, функціонально зв'язаний з однією або кількома NEENA, Крім того, пропонується застосування таких молекул нуклеїнових кислот, які посилюють експресію генів, та експресійних послідовностей, векторів експресії, трансгенних рослин або їх частин та трансгенних клітин, які включають такі молекули нуклеїнових кислот, які посилюють експресію генів.

Застосування трансгенної клітинної культури, трансгенного насіння, частин або матеріалу для розмноження, які походять з трансгенної клітини або рослини або її частини, як визначено вище, для виробництва продуктів харчування, кормів для тварин, насіння, фармацевтичних продуктів або чистих хімікатів також охоплюється обсягом цього винаходу.

ВИЗНАЧЕННЯ

Абревіатури: NEENA - нуклеїнова кислота, яка посилює експресію нуклеїнових кислот,

GFP - білок зеленої флуоресценції,

GUS - бета-глюкуронідаза,

BAP - 6-бензиламінопурин; 2,4-D-2,4-дихлорофеноксіоцтової кислоти;

MS - середовище Мураклі-Скуга;

NAA-1-нафталіноцтова кислота;

MES, 2-(N-морфоліно-етансульфонова кислота, I

AA - індолицтова кислота;

Kan - сульфат канаміцину;

GA3 - гіберелова кислота;

Timentin™: тикарцилін динатрію/кпавуланат калію,

мкл - мікролітр.

Слід розуміти, що цей винахід не обмежується конкретною методологією або протоколами. Також слід розуміти, що вжита авторами термінологія має на меті лише опис конкретних варіантів втілення і не обмежує обсяг даного винаходу, який обмежується лише супровідною формулою винаходу. Слід зазначити, що вжиті в описі та супровідній формулі винаходу форми однини передбачають також форми множини, якщо контекстом не передбачається іншого. Таким чином, наприклад, посилання на "вектор" є посиланням на один або кілька векторів і включає еквіваленти, відомі спеціалістам у даній галузі, і т. д. Вжитий авторами термін "приблизно" означає "приблизно", "неточно", "близько" або "у межах". Якщо термін "приблизно" вживається у зв'язку з діапазоном числових значень, він змінює цей діапазон, розширюючи його у бік зменшення та збільшення за межі вказаних значень. Як правило, термін "приблизно" вживається авторами для зміни вказаних числових значень у бік зменшення та збільшення з 20-відсотковим коливанням, в оптимальному варіанті - з 10-відсотковим у бік зменшення або збільшення. Вжите авторами слово "або" означає будь-який предмет з певного переліку і також включає будь-яку комбінацію предметів з цього переліку. Слова "включати" "включаючи" та "включає", вжиті в цьому описі та супровідній формулі винаходу, вказують на наявність однієї або кількох зазначених особливостей, цілих чисел, компонентів або етапів, але вони не виключають наявності або додавання однієї або кількох інших особливостей, цілих чисел, компонентів, етапів, або їх груп. Для кращого розуміння деякі терміни, вжиті в описі, визначаються й вживаються таким чином:

Антипаралельний: "антипаралельні" у контексті даного опису означають дві нуклеотидні послідовності, з'єднані водневими зв'язками між комплементарними залишками основ з фосфодіестерними зв'язками, які проходять у 5'-3'-напрямку в одній нуклеотидній послідовності і у 3'-5'-напрямку в іншій нуклеотидній послідовності.

Антисмисловий: термін "антисмислова" стосується нуклеотидної послідовності, яка є інвертованою відносно її нормальної орієнтації для транскрипції або функції і, таким чином, експресує транскрипт РНК, який є комплементарним молекулі мРНК гена-мішені, що експресується у клітині-хазяїні (наприклад, може гібридизуватися з молекулою мРНК гена мішені або одноланцюговою геномною ДНК через створення пар основ Вотсона-Крика), або яка

є комплементарною молекулі ДНК-мішені, такий, як, наприклад, геномна ДНК, присутня у клітині-хазяїні.

Кодуюча ділянка: термін "кодуюча ділянка", вжитий авторами стосовно структурного гена, означає нуклеотидні послідовності, які кодують амінокислоти, які містяться у виникаючому поліпептиді в результаті трансляції молекули мРНК. В еукаріотах кодуюча ділянка зв'язується на 5'-стороні нуклеотидним триплетом "ATG", який кодує ініціатор метіонін, а на 3'-стороні - одним з трьох триплетів, які визначають стоп-кодони (тобто, TAA, TAG, TGA). Крім того, що вони містять інтрони, геномні форми гена також можуть включати послідовності, розташовані як на 5'-, так і на 3'-кінці послідовностей, які є присутніми на транскрипті РНК. Ці послідовності називаються "фланкуючими" послідовностями або ділянками (ці фланкуючі послідовності розташовуються у 5' або 3' напрямках відносно нетрансльованих послідовностей, присутніх на транскрипті мРНК). 5'-фланкуюча ділянка може містити регуляторні послідовності, такі як промотори та ехансери, які контролюють транскрипцію гена або впливають на неї. 3'-фланкуюча ділянка може містити послідовності, які спрямовують термінацію транскрипції, посттранскрипційне розщеплення та поліаденілування.

Комплементарний: "комплементарна" або "комплементарність" означає дві нуклеотидні послідовності, які включають антипаралельні нуклеотидні послідовності, здатні утворювати пари одна з одною (за правилами спарювання основ) після утворення водневих зв'язків між комплементарними залишками основ в антипаралельних нуклеотидних послідовностях. Наприклад, послідовність 5'-AGT-3' є комплементарною послідовності 5'-ACT-3'. Комплементарність може бути "частковою" або "повною". "Частковою" комплементарністю є та, в якій одна або кілька основ нуклеїнових кислот не є відповідними за правилами спарювання основ. "Повна" або "цілковита" комплементарність між молекулами нуклеїнових кислот є такою, в якій кожна основа нуклеїнової кислоти відповідає іншій основі за правилами спарювання основ. Ступінь комплементарності між ланцюгами молекул нуклеїнових кислот має суттєвий вплив на ефективність та силу гібридизації між ланцюгами молекул нуклеїнових кислот. "Комплемент" нуклеїновокислотної послідовності у контексті даного опису означає нуклеотидну послідовність, у якій молекули нуклеїнових кислот демонструють повну комплементарність молекулам нуклеїнових кислот нуклеїновокислотної послідовності.

Дволанцюгова РНК: молекула "дволанцюгової РНК" або молекула "длРНК" включає смисловий фрагмент РНК нуклеотидної послідовності та антисмисловий фрагмент РНК нуклеотидної послідовності, обидва з яких включають нуклеотидні послідовності, комплементарні одна одній, що дозволяє смисловим та антисмисловим фрагментам РНК з'єднуватись у пари й утворювати дволанцюгову молекулу РНК.

Ендогенний: "ендогенна" нуклеотидна послідовність означає нуклеотидну послідовність, яка є присутньою у геномі нетрансформованої рослинної клітини.

Посилена експресія: визначення "посиливати" або "підвищувати" стосовно експресії молекули нуклеїнової кислоти у рослинній клітині вживаються у рівноцінному сенсі й означають, що рівень експресії молекули нуклеїнової кислоти у рослині, частині рослини або рослинній клітині після застосування способу згідно з даним винаходом є вищим за її експресію у рослині, частині рослини або рослинній клітині до застосування способу або порівняно з контрольною рослиною, в якій відсутня рекомбінантна молекула нуклеїнової кислоти згідно з винаходом. Наприклад, контрольна рослина включає таку саму послідовність, але без відповідної NEENA. Терміни "посилена" або "підвищена" у контексті цього опису є синонімічними і означають вищу, а оптимальному варіанті - значно вищу експресію молекули нуклеїнової кислоти, яка має бути експресована. У контексті даного опису "посилення" або "підвищення" рівня агента, такого, як білок, мРНК або РНК, означає, що рівень підвищується відносно показника по суті ідентичної рослини, частини рослини або рослинної клітини, яку вирощували у по суті ідентичних умовах, але без рекомбінантної молекули нуклеїнової кислоти згідно з винаходом, наприклад, без молекули NEENA, рекомбінантної послідовності або рекомбінантного вектора згідно з винаходом. У контексті даного опису "посилення" або "підвищення" рівня агента, такого, як, наприклад, преРНК, мРНК, рРНК, тРНК, мноРНК, мяРНК, що експресується геном-мішенню, та/або білкового продукту, який ним кодується, означає, що рівень підвищується на 50 % або більше, наприклад, на 100 % або більше, в оптимальному варіанті - на 200 % або більше, у ще кращому варіанті - у 5 разів або більше, у ще кращому варіанті - у 10 разів або більше, у найкращому варіанті - у 20 разів або більше, наприклад, у 50 разів порівняно з клітиною або організмом без рекомбінантної молекули нуклеїнової кислоти згідно з винаходом. Посилення або підвищення може визначатися способами, відомими спеціалістам у даній галузі. Таким чином, посилення або підвищення кількості нуклеїнової кислоти або білка може визначатися, наприклад, шляхом імунологічного виявлення білка. Крім того, можуть застосовуватися такі

способи, як білковий спектр, флуоресценція, нозерн-гібридизація, захист від нуклеази, зворотна транскрипція (кількісна ЗТ-ПЛР), ELISA (ферментний імуносорбентний аналіз), вестерн-блотинг, радіоімунний аналіз (RIA) або інші імунологічні аналізи та сортування флуоресцентно-активованих клітин (FACS), для вимірювання конкретного білка або РНК у рослині або
 5 рослинній клітині. Залежно від типу індукованого білкового продукту, також може визначатись його активність або вплив на фенотип організму. Способи визначення кількості білка є відомими спеціалістам у даній галузі. Прикладами, які можуть бути наведені, є такі: мікробіуретовий спосіб (Goa J (1953) Scand J Clin Lab Invest 5:218-222), спосіб Фоліна-Чокальтеу (Lowry OH et al. (1951) J Biol Chem 193:265-275) або вимірювання поглинання СВВ G-250 (Bradford MM (1976)
 10 Analyt Blochem 72:248-254). Як один приклад кількісного визначення активності білка, нижче у Прикладах описується виявлення активності люциферази.

Експресія: "експресія" означає біосинтез генного продукту, в оптимальному варіанті - транскрипцію та/або трансляцію нуклеотидної послідовності, наприклад, ендегенного гена або гетерологічного гена, у клітині. Наприклад, у разі структурного гена, експресія включає
 15 транскрипцію структурного гена на мРНК та, необов'язково, наступну трансляцію мРНК на один або кілька поліпептидів. В інших випадках експресія може стосуватись лише транскрипції ДНК, яка включає молекулу РНК. Експресійна послідовність: "експресійна послідовність" у контексті даного опису означає послідовність ДНК, здатну спрямовувати експресію конкретної нуклеотидної послідовності у відповідних частинах рослини або рослинної клітини, яка включає
 20 промотор, який функціонує у зазначених частинах рослини або рослинної клітини, у яку її вводять, що функціонально зв'язується з потрібною нуклеотидною послідовністю, яка необов'язково функціонально зв'язується з сигналами термінації. Якщо вимагається трансляція, вона також зазвичай включає послідовності, необхідні для належної трансляції нуклеотидної послідовності. Кодуюча ділянка може кодувати потрібний білок, але також може кодувати потрібну функціональну РНК, наприклад, РНКа, міРНК, мноРНК, мяРНК, мікроРНК, та-міРНК
 25 або будь-яку іншу некодуючу регуляторну РНК, у смислового або антисмислового напрямку. Експресійна послідовність, яка включає потрібну нуклеотидну послідовність, може бути химерною, тобто, один або кілька її компонентів є гетерологічними одному або кільком іншим м компонентам. Експресійна послідовність також може бути такою, що трапляється у природі, але
 30 була одержана у рекомбінантній формі, яку застосовують для гетерологічної експресії. Однак зазвичай експресійна послідовність є гетерологічною хазяїнові, тобто, конкретна послідовність ДНК експресійної послідовності не трапляється у природі у клітині-хазяїні і мала бути включена у клітину-хазяїн або попередник клітини-хазяїна шляхом трансформації. Експресія нуклеотидної послідовності в експресійній послідовності може перебувати під контролем насінно-специфічного та/або насінно-селективного промотора або індукованого промотора, що
 35 започатковує транскрипцію лише тоді, коли клітина-хазяїн піддається дії певного зовнішнього подразника. У разі рослини промотор також може бути специфічним до конкретної тканини або органа або стадії розвитку.

Чужорідний: термін "чужорідний" стосується будь-якої молекули нуклеїнової кислоти
 40 (наприклад, послідовності гена), яку включають у геном клітини шляхом експериментальних маніпуляцій, і яка може включати послідовності, які містяться в цій клітині доти, доки включена послідовність містить певні модифікації (наприклад, точкову мутацію, наявність селектованого маркерного гена і т. ін.), і, таким чином, відрізняється від природної послідовності.

Функціональне зв'язування: термін "функціональне зв'язування" або "функціонально зв'язаний" слід розуміти, наприклад, як послідовне розташування регуляторного елемента
 45 (наприклад, промотора) з послідовністю нуклеїнової кислоти, яка має бути експресована, і, у відповідних випадках, іншими регуляторними елементами (такими, як, наприклад, термінатор або NEENA) таким чином, щоб кожен з регуляторних елементів міг виконувати свою призначену функцію для забезпечення можливості, модифікації, сприяння або іншого впливу на експресію
 50 зазначеної нуклеїновоокислотної послідовності. Результат експресії може залежати від розташування нуклеїновоокислотних послідовностей відносно смислових або антисмислових РНК. Для цього не обов'язково вимагається пряме зв'язування у хімічному сенсі. Послідовності генетичного контролю, такі як енхансерні послідовності, також можуть виконувати свою функцію стосовно послідовності-мішені з більш віддалених позицій, або й з інших молекул ДНК.
 55 Оптимальними варіантами розташування є ті, в яких нуклеїновоокислотна послідовність, яка підлягає рекомбінантній експресії, розташовується за послідовністю, що діє як промотор, таким чином, щоб дві послідовності ковалентно зв'язувалися одна з одною. Відстань між промоторною послідовністю та нуклеїновоокислотною послідовністю, яка підлягає рекомбінантній експресії, в оптимальному варіанті складає менше, ніж 200 пар основ, у ще кращому варіанті - менше, ніж
 60 100 пар основ, у найкращому варіанті - менше, ніж 50 пар основ. В оптимальному варіанті

втілення нуклеїновокислотна послідовність, яка підлягає транскрипції, розташовується за промотором таким чином, щоб початок транскрипції був ідентичним потрібному початкові химерної РНК згідно з винаходом. Функціональне зв'язування та експресійна послідовність можуть забезпечуватися за допомогою традиційних технологій рекомбінації та клонування, як описано [див., наприклад, Maniatis T, Fritsch EF and Sambrook J (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Silhavy et al. (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Ausubel et al. (1987) *Current Protocols in Molecular Biology*, Greene Publishing Assoc, and Wiley Interscience; Gelvin et al. (Eds) (1990) *Plant Molecular Biology Manual*; Kluwer Academic Publisher, Dordrecht, Нідерланди]. Однак інші послідовності, які діють, наприклад, як лінкери з сайтами специфічного розщеплення для рестрикційних ферментів або як сигнальні пептиди, також можуть розташовуватися між двома послідовностями. Інсерція послідовностей також може забезпечувати експресію злитих білків. В оптимальному варіанті експресійна послідовність, яка складається зі зв'язку регуляторної ділянки, наприклад, промотора, та нуклеїновокислотної послідовності, яка підлягає експресії, може існувати в інтегрованій у вектор формі і може бути вставлена у геном рослини, наприклад, шляхом трансформації.

Ген: термін "ген" стосується ділянки, функціонально з'єднаної з відповідними регуляторними послідовностями, здатними регулювати експресію генного продукту (наприклад, поліпептидом або функціональною РНК) у певний спосіб. Ген включає нетрансльовані регуляторні ділянки ДНК (наприклад, промотори, енхансери, репресори і т. ін.), розташовані до або після кодуєчої ділянки (відкрита рамка зчитування, ORF), а також, у відповідних випадках, вставні послідовності (тобто, інтрони) між окремими кодуєчими ділянками (тобто, екзонами). Термін "структурний ген" у контексті даного опису означає послідовність ДНК, яка транскрибується на мРНК, яка потім транслюється на послідовність амінокислот, характерну для певного поліпептиду.

Геном та геномна ДНК: терміни "геном" або "геномна ДНК" стосуються спадкової генетичної інформації організму-хазяїна. Зазначена геномна ДНК включає ДНК ядра (також відому як хромосомна ДНК), а також ДНК пластид (наприклад, хлоропластів) та інших клітинних органел (наприклад, мітохондрій). В оптимальному варіанті терміни "геном" або "геномна ДНК" стосуються хромосомної ДНК ядра.

Гетерологічний: термін "гетерологічна" стосовно молекули нуклеїнової кислоти або ДНК означає молекулу нуклеїнової кислоти, яка є функціонально зв'язаною, або підданою маніпуляції для функціонального зв'язування, з молекулою нуклеїнової кислоти, з якою вона є функціонально зв'язаною у природі, або з якою вона у природних умовах є функціонально зв'язаною в іншому місці. Гетерологічна експресійна послідовність, яка включає молекулу нуклеїнової кислоти та одну або кілька регуляторних молекул нуклеїнової кислоти (такі, як промотор або сигнал термінації транскрипції), які, наприклад, зв'язуються з нею, є послідовністю, одержаною через експериментальні маніпуляції, у яких а) зазначена молекула нуклеїнової кислоти або б) зазначена регуляторна молекула нуклеїнової кислоти, або с) обидві (тобто, (а) та (б)), знаходяться не у природному генетичному середовищі або були модифіковані через експериментальні маніпуляції, причому прикладом такої модифікації є заміщення, додавання, делеція, інверсія інсерція одного або кількох нуклеотидних залишків. Природне генетичне середовище означає природний хромосомний локус в організмі походження або присутність у геномній бібліотеці. У разі геномної бібліотеки природне генетичне середовище послідовності молекули нуклеїнової кислоти в оптимальному варіанті зберігається, принаймні частково. Середовище прилягає до нуклеїновокислотної послідовності принаймні з одного боку і має послідовність з принаймні 50 п. о., в оптимальному варіанті - принаймні 500 п. о., у ще кращому варіанті - принаймні 1000 п. о., у найкращому варіанті - принаймні 5000 п. о. Природна експресійна послідовність, наприклад, природна комбінація промотора з відповідним геном, стає трансгенною експресійною послідовністю, коли модифікується неприродними, синтетичними "штучними" способами, такими, як, наприклад, мутагенізація. Такі способи було описано (документи US 5,565,350; WO 00/15815). Наприклад, кодуєча білок молекула нуклеїнової кислоти, функціонально зв'язана з промотором, який не є природним промотором цієї молекули, вважається гетерологічною стосовно промотора. В оптимальному варіанті гетерологічна ДНК не є ендогенною або природно пов'язаною з клітиною, до якої вона включена, а була одержана з іншої клітини або синтезована. Гетерологічна ДНК також включає ендогенну послідовність ДНК, яка містить певну модифікацію, неприродні, повторювані копії ендогенної послідовності ДНК, або послідовність ДНК, яка не є природно пов'язаною з іншою послідовністю ДНК, яка з нею фізично з'єднується. Зазвичай, хоча й не обов'язково,

гетерологічна ДНК кодує РНК або білки, які зазвичай не продукуються клітиною, в яку вона експресується.

Високоекспресійний насінно-специфічний та/або насінно-селективний промотор: "високоекспресійний насінно-специфічний та/або насінно-селективний промотор" у контексті даного опису означає промотор, який викликає насінно-специфічну та/або насінно-селективну експресію у рослині або її частині, причому накопичення або швидкість синтезу РНК або стійкість РНК, що походить з молекули нуклеїнової кислоти, під контролем відповідного промотора є вищими, в оптимальному варіанті - суттєво вищими за експресію, викликану промотором без NEENA згідно з винаходом. В оптимальному варіанті кількість РНК та/або швидкість синтезу РНК та/або стійкість РНК підвищується на 50 % або більше, наприклад, 100 % або більше, в оптимальному варіанті - 200 % або більше, у ще кращому варіанті - у 5 разів або більше, у ще кращому варіанті - у 10 разів або більше, у найкращому варіанті - у 20 разів або більше, наприклад, у 50 разів, порівняно з насінно-специфічним та/або насінно-селективним промотором без NEENA згідно з винаходом.

Гібридизація: термін "гібридизація" у контексті даного опису включає "будь-який процес, за допомогою якого ланцюг молекули нуклеїнової кислоти з'єднується з комплементарним ланцюгом через створення пар основ" (J. Coombs (1994) Dictionary of Biotechnology, Stockton Press, New York). На гібридизацію та силу гібридизації (тобто, силу асоціації між молекулами нуклеїнових кислот) впливають такі чинники, як ступінь комплементарності між молекулами нуклеїнових кислот, жорсткість застосовуваних умов, т. п. утвореного гібриду та співвідношення G:C у молекулах нуклеїнових кислот. У контексті даного опису термін "т. п." означає "температуру плавлення". Температура плавлення є температурою, при якій популяція дволанцюгових молекул нуклеїнових кислот стає наполовину дисоційованою на окремі ланцюги. Рівняння для розрахунку т. п. молекул нуклеїнових кислот є добре відомим спеціалістам у даній галузі. Як вказується у стандартних джерелах, просте приблизне значення т. п. може бути розраховано за рівнянням: $T. п. = 81,5 + 0,41(\% G+C)$, коли молекула нуклеїнової кислоти перебуває у водному розчині при 1 M NaCl [див., наприклад, Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985)]. В інших джерелах представлено більш складні розрахунки, у яких для визначення т. п. враховуються структурні характеристики, а також характеристики послідовності. Жорсткі умови є відомими спеціалістам у даній галузі і описуються у публікації Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. "Ідентичність": термін "ідентичність", вжитий стосовно порівняння двох або більшої кількості молекул нуклеїнових кислот або амінокислот, означає, що послідовності зазначених молекул мають певний ступінь подібності послідовностей, і послідовності є частково ідентичними.

Для визначення відсотка ідентичності (у контексті даного опису взаємозамінно вживається термін "гомологія") двох амінокислотних послідовностей або двох молекул нуклеїнових кислот послідовності записують одну під одною для оптимального порівняння (наприклад, можуть бути вставлені пробіли у послідовність білка або нуклеїнової кислоти для забезпечення оптимального співставлення з іншим білком або іншою нуклеїною кислотою).

Потім порівнюють амінокислотні залишки або молекули нуклеїнових кислот у відповідних амінокислотних позиціях або нуклеотидних позиціях. Якщо позиція в одній послідовності займається тим самим амінокислотним залишком або тією самою молекулою нуклеїнової кислоти, що й відповідна позиція в іншій послідовності, молекули у цій позиції є гомологічними (тобто, "гомологія" амінокислоти або нуклеїнової кислоти у контексті цього опису відповідає "ідентичності" амінокислоти або нуклеїнової кислоти. Відсоток гомології між двома послідовностями залежить від кількості ідентичних позицій, які є спільними для послідовностей (тобто, $\% \text{ гомології} = \text{кількість ідентичних позицій} / \text{загальна кількість позицій} \times 100$). Таким чином, терміни "гомологія" та "ідентичність" слід розглядати як синонімічні.

Для визначення відсотка ідентичності двох або більшої кількості амінокислот або двох або більшої кількості нуклеотидних послідовностей було розроблено кілька комп'ютерних програм. Ідентичність двох або більшої кількості послідовностей розраховують, наприклад, за допомогою програми Pasta, яка у даний час застосовується у версії Pasta 3 (W. R. Pearson and D. J. Liptan, PNAS 85, 2444(1988); W. R. Pearson. Methods in Enzymology 183, 63 (1990); W. R. Pearson and D. J. Lipman, PNAS 85, 2444 (1988); W. R. Pearson, Enzymology 183, 63 (1990)). Іншою програмою, яка може застосовуватися для розрахунку ідентичності різних послідовностей, є стандартна програма Blast, яка є включеною до програми Pedant від Біотех (Вістах, Мюнхен, Федеративна Республіка Німеччина). На жаль, це іноді призводить до субоптимальних результатів, оскільки Blast не завжди включає повні послідовності наявних у базі даних та досліджуваних зразків. Незважаючи на це, завдяки високій ефективності цієї програми, вона може застосовуватися для

порівняння великої кількості послідовностей. Для порівняння послідовностей зазвичай застосовують такі настройки:

-p Назва програми [рядок]; -d База даних [рядок]; за замовчуванням = nr; -i Query File [File In]; за замовчуванням = stdin; -e очікуване значення (E) [Real]; за замовчуванням = 10,0; -m опції перегляду вирівнювання: 0 = попарно; 1=query-anchored з ідентичністю; 2 = query-anchored без ідентичності; 3 = flat query-anchored, показ ідентичності; 4=flat queryanchored, без ідентичності; 5 = query-anchored без ідентичності та тупих кінців; 6 = flat query-anchored, без ідентичності та тупих кінців; 7 = вивід XML Blast; 8 = таблиця; 9 таблиця з рядками коментарів [ціле число]; за замовчуванням = 0; -o файл виводу звіту BLAST [File Out] необов'язково; за замовчуванням = stdout; -F Filter query послідовність (DUST з blastn, SEG з іншими) [рядок]; за замовчуванням = T; -G плата за відкриття пробілу (нуль запускає поведінку за замовчуванням) [ціле число]; за замовчуванням = 0; -E плата за продовження пробілу (нуль запускає поведінку за замовчуванням) [ціле число]; за замовчуванням = 0; -X X зменшене значення за вирівнювання з пробілами (у бітах) (нуль запускає поведінку за замовчуванням); blastn 30, megablast 20, tblastx 0, усі інші 15 [ціле число]; за замовчуванням = 0; -l показ GI у deflines [T/F]; за замовчуванням = F; -q штраф за невідповідність нуклеотидів (лише blastn) [ціле число]; за замовчуванням = -3; -г винагорода за відповідність нуклеотидів (лише blastn) [ціле число]; за замовчуванням = 1; -v кількість послідовностей з бази даних для показу онлайн-описів для (V) [ціле число]; за замовчуванням = 500; -b номер послідовності з бази даних для показу вирівнювання для (B) [ціле число]; за замовчуванням = 250; -f поріг для продовження збігів, за замовчуванням, якщо нуль; blastp 11, blastn 0, blastx 12, tblastn 13; tblastx 13, megablast 0 [ціле число]; за замовчуванням = 0; -g виконання вирівнювання з пробілами (не передбачено з tblastx) [T/F]; за замовчуванням = T; -Q генетичний код досліджуваного зразка [ціле число]; за замовчуванням = 1; -D DB генетичний код (лише для tblast[nx]) [ціле число]; за замовчуванням = 1; -a кількість застосовуваних процесорів [ціле число]; за замовчуванням = 1; -O файл SeqAllgn [File Out] необов'язково; -J гадане значення query define [T/F], за замовчуванням = F; -M Matrix [рядок]; за замовчуванням = BLOSUM62; -W розмір слова, за замовчуванням, якщо нуль (blastn 11, megablast 28, усі інші 3) [ціле число], за замовчуванням = 0; -z ефективна довжина бази даних (використовуйте нуль для реального розміру) [Real], за замовчуванням = 0; -K кількість найкращих збігів у ділянці для дотримання (off за замовчуванням, у разі використання рекомендується значення 100) [ціле число]; за замовчуванням = 0; -P 0 для багатьох збігів, 1 для єдиного збігу [ціле число]; за замовчуванням = 0; -Y ефективна довжина області пошуку (використовуйте нуль для реального розміру) [Real]; за замовчуванням = 0; -S досліджувані ланцюги для пошуку у базі даних (для blast[nx] та tblastx); 3 - обидва, 1 - верхній, 2 - нижній [ціле число]; за замовчуванням = 3; вивід T Produce HTML [T/F]; за замовчуванням = F; -l обмеження пошуку у базі даних для списку GI [рядок] необов'язково; -U застосування фільтра нижнього реєстру послідовності FASTA [T/F] необов'язково; за замовчуванням = F; -у X значення dropoff для продовжень без пробілів у бітах (0,0 запускає поведінку за замовчуванням); blastn 20, megablast 10, усі інші 7 [Real]; за замовчуванням = 0,0; -Z X значення dropoff для остаточного вирівнювання з пробілами у бітах (0,0 запускає поведінку за замовчуванням); blastn/megablast 50, tblastx 0, усі інші 25 [ціле число]; за замовчуванням = 0; -R PSI-TBLASTN checkpoint file [File In] необов'язково; -p пошук MegaBlast [T/F]; за замовчуванням = F; -L розташування на досліджуваній послідовності [рядок] необов'язково; -A розмір вікна множинних збігів, за замовчуванням, якщо нуль (blastn/megablast 0, усі інші 40 [ціле число]; за замовчуванням = 0; -w штраф за зсув рамки (алгоритм OOF для blastx) [ціле число]; за замовчуванням = 0; -t довжина найбільшого інтрона, який допускається у tblastn для з'єднання HSP (0 унеможливорює з'єднання) [ціле число]; за замовчуванням = 0.

Результатів високої якості досягають при застосуванні алгоритму Нідлмана - Вунша або Сміта - Вотермана. Таким чином, перевагу віддають програмам на основі зазначених алгоритмів. Зручне порівняння послідовностей може бути здійснене за допомогою програми PileUp (J. Mol. Evolution., 25, 351 (1987), Higgins et al., CABIOS 5, 151 (1989)) або, в оптимальному варіанті, - за допомогою програм "Gap" та "Needle", обидві на основі алгоритмів Нідлмана - Вунша (J. Mol. Biol. 48; 443 (1970)), та "BestFit", на основі алгоритму Сміта - Вотермана (Adv. Appl. Math. 2; 482 (1981)). "Gap" та "BestFit" є частиною пакета програм GCG (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, США 53711 (1991); Altschul et al., (Nucleic Acids Res. 25, 3389 (1997)), "Needle" є частиною Європейського комплексу відкритих програм з молекулярної біології (EMBOSS) (Trends in Genetics 16 (6), 276 (2000)). Таким чином, в оптимальному варіанті розрахунки для визначення відсотка гомології послідовностей здійснюються за допомогою програм "Gap" або "Needle" по всіх послідовностях. Для "Needle" застосовували: матриця: EDNAFULL, штраф за пробіл: 10,0, штраф за

продовження: 0,5. Для "Gap" застосовували такі стандартні установки для порівняння нуклеїновокислотних послідовностей: значущість пробілу: 50, значущість довжини: 3, середня відповідність: 10,000, середня невідповідність: 0,000.

Наприклад, послідовність, яка має 80 % ідентичності з послідовністю SEQ ID NO: 1 на рівні нуклеїнової кислоти, розглядається як послідовність, яка, після порівняння з послідовністю, представленою у SEQ ID NO: 1, за допомогою зазначеної програми "Needle" з зазначеним набором параметрів, має 80 % ідентичності. В оптимальному варіанті гомологію розраховують по всій довжині досліджуваної послідовності, наприклад, SEQ ID NO:1.

Інтрон: означає відрізки ДНК (вставні послідовності) з геном, які не кодують частину білка, яку виробляє ген, і яка сплайсується з мРНК, що транскрибується з гена до перенесення з ядра клітини. Послідовність інтрона означає нуклеїновокислотну послідовність інтрона. Таким чином, інтрони є ділянками послідовностей ДНК, які транскрибуються разом з кодуючою ділянкою (екзонами), але видаляються під час утворення зрілої мРНК. Інтрони можуть розташовуватись у межах існуючої кодуючої ділянки або у 5' або 3' нетрансльованих лідерних послідовностях Пре-мРНК (несплайсованої мРНК). Інтрони у первинному транскрипті вирізаються, і кодуєчі послідовності одночасно і точно лігуються для утворення зрілої мРНК. Місця з'єднання інтронів та екзонів утворюють сайт сплайсингу. Послідовність інтрона починається з GU і закінчується AG. Крім того, у рослинах було описано два приклади AU-AC інтронів: чотирнадцятий інтрон RecA-подібного білкового гена та сьомий інтрон G5 гена з *Arabidopsis thaliana* є AT-AC інтронами, Пре-мРНК, які містять Інтрони, мають три короткі послідовності, які є - крім інших послідовностей - суттєвими для інтрона, який має бути точно сплайсований. Ці послідовності є 5'-сайтом сплайсингу, 3'-сайтом сплайсингу та точкою галуження. Сплайсинг мРНК є видаленням вставних послідовностей (інтронів), присутніх у первинних транскриптах мРНК, та з'єднання або лігування екзонних послідовностей. Це також називається цис-сплайсингом, який з'єднує два екзони на одній РНК з видаленням вставної послідовності (інтрона). Функціональні елементи інтрона включають послідовності, які розпізнаються й зв'язуються специфічними білковими компонентами сплайсингосоми (наприклад, сплайсинговими консенсусними послідовностями на кінцях інтронів). Взаємодія функціональних елементів зі сплайсингосомою в результаті веде до видалення послідовності інтрона з незрілої мРНК та повторного з'єднання послідовностей екзонів. Інтрони мають три короткі послідовності, які є суттєвими, хоча й недостатніми, для інтрона, який має бути точно сплайсований. Цими послідовностями є 5'-сайт сплайсингу, 3'-сайт сплайсингу та точка галуження. Послідовність точки галуження є важливою для сплайсингу та вибору сайтів сплайсингу у рослинах. Послідовність точки галуження зазвичай розташовується за 10-60 нуклеотидів перед 3'-сайтом сплайсингу.

Ізогенний: організми (наприклад, рослини), які є генетично ідентичними, за винятком відмінностей, які полягають у присутності або відсутності гетерологічної послідовності ДНК.

Виділений: термін "виділений" у контексті даного опису означає, що матеріал було видалено людиною, і він існує окремо від його первісного, природного середовища і, таким чином, не є природним продуктом. Виділений матеріал або молекула (наприклад, молекула ДНК або фермент) може існувати в очищеній формі або може існувати у неприродному середовищі, наприклад, у трансгенній клітині-хазяїні. Наприклад, природний полінуклеотид або поліпептид, присутній у живій рослині, не є виділеним, але той самий полінуклеотид або поліпептид, відокремлений від деяких або всіх співіснуючих матеріалів у природній системі, є виділеним. Такі полінуклеотиди можуть бути частиною вектора, та/або такі полінуклеотиди або поліпептиди можуть бути частиною композиції і вважатимуться виділеними у тому сенсі, що такі вектор або композиція не є частиною первісного середовища. В оптимальному варіанті термін "виділений", вжитий стосовно молекули нуклеїнової кислоти, як у "виділеній нуклеїновокислотній послідовності", означає нуклеїновокислотну послідовність, яка є ідентифікованою й відокремленою від принаймні однієї сторонньої молекули нуклеїнової кислоти, з якою вона є первісно зв'язаною у природному джерелі. Виділеною молекулою нуклеїнової кислоти є молекула нуклеїнової кислоти, присутня у формі або в оточенні, які відрізняються від тих, у яких вона існує у природі. Натомість невиділеними молекулами нуклеїнових кислот є молекули нуклеїнових кислот, такі як ДНК та РНК, які перебувають у стані, в якому вони існують у природі. Наприклад, дана послідовність ДНК (наприклад, ген) міститься на хромосомі клітини-хазяїна поблизу від сусідніх генів; послідовності РНК, такі як специфічна послідовність мРНК, яка кодує певний білок, існують у клітині як суміш з багатьма іншими мРНК, які кодують велику кількість білків. Однак виділена нуклеїновокислотна послідовність, яка включає, наприклад, SEQ ID NO: 1, включає, скажімо, такі нуклеїновокислотні послідовності у клітинах, які первісно містять SEQ ID NO:1, де нуклеїновокислотна послідовність перебуває у хромосомній або позахромосомній позиції, відмінній від позиції природних клітин, або іншим чином межує з нуклеїновокислотною

послідовністю, відмінною від тієї, що існує у природі. Виділена нуклеїновокислотна послідовність може бути присутньою в одоланцюговій або дволанцюговій формі. Якщо виділена нуклеїновокислотна послідовність має бути застосована для експресії білка, нуклеїновокислотна послідовність міститиме принаймні частину смислового або кодуєчого ланцюга (тобто, нуклеїновокислотна послідовність може бути одоланцюговою). В альтернативному варіанті вона може містити як смислові, так і антисмислові ланцюги (тобто, нуклеїновокислотна послідовність може бути дволанцюговою).

Мінімальний промотор: промоторні елементи, зокрема, ТАТА-елемент, які є неактивними або які мають сильно знижену промоторну активність за відсутності активації у 3'5'-напрямку. У присутності відповідного фактора транскрипції мінімальний промотор функціонує, забезпечуючи можливість транскрипції.

NEENA: див. "Нуклеїнова кислота, яка посилює експресію нуклеїнових кислот".

Некодуєчий: термін "некодуєчий" стосується послідовностей молекул нуклеїнових кислот, які не кодують частину експресованого білка або весь експресований білок. До некодуєчих послідовностей, крім інших, належать інтрони, енхансери, промоторні ділянки, 3'-нетрансльовані ділянки та 5'-нетрансльовані ділянки.

Нуклеїнова кислота, яка посилює експресію нуклеїнових кислот (NEENA): термін "нуклеїнова кислота, яка посилює експресію нуклеїнових кислот" стосується послідовності та/або молекули нуклеїнової кислоти з певною послідовністю, яка має природну властивість посилення експресії нуклеїнової кислоти під контролем промотора, з яким NEENA є функціонально зв'язаною. На відміну від промоторних послідовностей, NEENA як така, не здатна запускати експресію. Для виконання функції посилення експресії молекули нуклеїнової кислоти, функціонально зв'язаної з NEENA, NEENA сама має бути функціонально зв'язана з промотором. На відміну від енхансерних послідовностей, відомих спеціалістам у даній галузі, NEENA діє у цис-, а не транс-позиції, і має розташовуватись поблизу від сайту початку транскрипції нуклеїнової кислоти, яка підлягає експресії.

Нуклеїнові кислоти та нуклеотиди: терміни "нуклеїнові кислоти" та "нуклеотиди" стосуються природних або синтетичних або штучних нуклеїнових кислот або нуклеотидів. Терміни "нуклеїнові кислоти" та "нуклеотиди" охоплюють дезоксирибонуклеотиди або рибонуклеотиди або будь-які нуклеотидні аналоги та полімери або їх гібриди в одно- або дволанцюговій, смисловій або антисмисловій формі. Якщо не зазначено іншого, передбачається, що конкретна нуклеїновокислотна послідовність також охоплює її консервативно модифіковані варіанти (наприклад, заміщення вироджених кодонів) та комплементарні послідовності, а також прямо вказану послідовність. Термін "нуклеїнова кислота" вжито взаємозамінно з термінами "ген", "кДНК", "мРНК", "олігонуклеотид" та "полінуклеотид". До нуклеотидних аналогів належать нуклеотиди, які мають модифікації у хімічній структурі основи, цукру та/або фосфату, включаючи, крім інших, піримідинові модифікації у 5-позиції, пуринові модифікації у 8-позиції, модифікації в екзоциклічних амінах цитозину, заміщення 5-бромоурацилу і т. ін.; та модифікації цукрів у 2'-позиції, включаючи, крім інших, модифіковані цукрами рибонуклеотиди, у яких 2'-ОН є заміненним на групу, вибрану з H, OR, R, гало, SH, SR, NH₂, NHR, NR₂ або CN, Короткі шпилькові РНК (шРНК) також можуть включати неприродні елементи, такі як неприродні основи, наприклад, іозин та ксантин, неприродні цукри, наприклад, 2'-метоксирибозу, або неприродні фосфодіестерні зв'язки, наприклад, метилфосфонати, фосфоротіоати та пептиди.

Нуклеїновокислотна послідовність: фраза "нуклеїновокислотна послідовність" стосується одно- або дволанцюгового полімеру дезоксирибонуклеотидних або рибонуклеотидних основ, які читаються з 5'- до 3'-кінця. Вона включає хромосомну ДНК, самореplikовані плазмиди, інфекційні полімери ДНК або РНК та ДНК або РНК, які виконують здебільшого структурну роль. "Нуклеїновокислотна послідовність" також стосується впорядкованого списку скорочень, літер, знаків або слів, які представляють нуклеотиди. В одному варіанті втілення нуклеїнова кислота може бути "зондом", який є відносно короткою нуклеїновою кислотою, зазвичай меншою за 100 нуклеотидів завдовжки. Часто нуклеїновокислотний зонд має довжину від приблизно 50 нуклеотидів до приблизно 10 нуклеотидів. "Ділянка-мішень" нуклеїнової кислоти є частиною нуклеїнової кислоти, ідентифікованою як така, що являє інтерес. "Кодуюча ділянка" нуклеїнової кислоти є частиною нуклеїнової кислоти, яка транскрибується й трансльється у специфічний до послідовності спосіб для утворення конкретного поліпептиду або білка при розміщенні під контроль відповідних регуляторних послідовностей. Кодуюча ділянка кодує такий поліпептид або білок.

Олігонуклеотид: термін "олігонуклеотид" стосується олігомеру або полімеру рибонуклеїнової кислоти (РНК) або дезоксирибонуклеїнової кислоти (ДНК) або їх міметиків, а також олігонуклеотидів, які мають неприродні частини, які функціонують подібним чином. Таким

модифікованим або заміщеним олігонуклеотидам часто віддають перевагу перед природними формами завдяки потрібним властивостям, таким, як, наприклад, підвищене клітинне захоплення, підвищена афінність до нуклеїнової кислоти-мішені та підвищена стійкість у присутності нуклеаз. Олігонуклеотид в оптимальному варіанті включає два або більше нуклеомономерів, ковалентно зв'язаних один з одним зв'язками (наприклад, фос-фодіестерами) або замісними зв'язками.

Нависання: "нависання" є відносно короткою одноланцюговою нуклеотидною послідовністю на 5'- або 3'-гідроксильному кінці дволанцюгової олігонуклеотидної молекули (також називається "подовженням", "виступаючим кінцем" або "липким кінцем").

Рослина: під рослиною зазвичай розуміють будь-який еукаріотний одно- або багатоклітинний організм або його клітину, тканину, орган, частину або матеріал для розмноження (такий, як насіння або плоди), здатні до фотосинтезу. З точки зору даного винаходу, до них належать усі роди та види вищих та нижчих рослин з царства рослин. Перевагу віддають однорічним, багаторічним, однодольним та дводольним рослинам. Термін охоплює зрілі рослини, насіння, паростки та сіянці та одержані з них частини, матеріал для розмноження (такий, як насіння або мікроспори), органи рослин, тканини, протопласти, калюс та інші культури, наприклад, культури клітин та будь-який інший тип групи рослинних клітин для одержання функціональних або структурних одиниць. Зрілі рослини означають рослини на будь-якій потрібній стадії розвитку після стадії сіянців. Сіянець означає молодшу незрілу рослину на ранній стадії розвитку. Однорічні, дворічні, однодольні та дводольні рослини є оптимальними організмами-хазяями для створення трансгенних рослин. Крім того, експресія генів забезпечує переваги в усіх декоративних рослинах, корисних або декоративних деревах, квітах, квітах на зрізання, чагарниках або газонних травах. До рослин, які можна згадати для прикладу, але не для обмеження обсягу винаходу, належать покритонасінні, мохи, такі, як, наприклад, *Hepaticae* (печінкові мохи) та *Musci*; птеридофіти, такі як папороть, хвощі та плауни; голонасінні, такі як хвойні, цикадові, гінго та *Gnetatae*; водорості, такі як *Chlorophyceae*, *Phaeophyceae*, *Rhodophyceae*, *Mycophyceae*, *Xanthophyceae*, *Bacillariophyceae* (діатомові) та *Euglenophyceae*. Перевагу віддають рослинам, які застосовують для виробництва продовольства та кормів, наприклад, з родин *Leguminosae*, таким як горох, люцерна та соя; *Gramineae*, таким як рис, кукурудза, пшениця, ячмінь, сорго, просо, жито, тритикале або овес; родині *Umbelliferae*, зокрема, роду *Daucus*, особливо, таким видам, як *carota* (морква) та *Apium*, зокрема, таким видам, як *Graveolens dutce* (селера) та багато інших; родині *Solanaceae*, зокрема, роду *Lycopersicon*, зокрема, видові *esculentum* (томати) та роду *Solanum*, зокрема, видам *tuberosum* (картопля) та *melongena* (баклажан), а також багатьом іншим (таким, як тютюн); та роду *Capsicum*, зокрема, видові *annuum* (перець) та багатьом іншим; родині *Leguminosae*, особливо, роду *Glycine*, зокрема, видові *max* (соя), таким рослинам, як горох, люцерна, квасоля або арахіс та багато інших; та родині *Cruciferae* (*Brassicaceae*), особливо, роду *Brassica*, зокрема, видам *napus* (олійний рапс), *campestris* (буряк), *oleracea cv Tasty* (капуста), *oleracea cv Snowball Y* (цвітна капуста) та *oleracea cv Emperor* (броколі); та роду *Arabidopsis*, зокрема, видові *thaliana*, та багатьом іншим; родині *Compositae*, особливо, роду *Lactuca*, зокрема, видові *sativa* (lettuce) та багатьом іншим; родині *Asteraceae*, зокрема, таким рослинам, як соняшник, календула, латук, та багатьом іншим; родині *Cucurbitaceae*, зокрема, таким рослинам, як диня, гарбуз або цукіні, та льняному насінню. Також перевагу віддають таким рослинам, як бавовна, цукрова тростина, конопля, льон, чилі, а також різні види деревних рослин, горіхів та винограду.

Поліпептид: терміни "поліпептид", "пептид", "олігопептид", "поліпептид", "генний продукт", "продукт експресії" та "білок" у контексті даного опису вживаються взаємозамінно і стосуються полімеру або олігомеру з послідовних амінокислотних залишків.

Білок-попередник: білок, який зазвичай є спрямованим на клітинну органелу, наприклад, хлоропласт, і зберігає свій перехідний пептид.

Первинний транскрипт: термін "первинний транскрипт" у контексті даного опису означає незрілий РНК-транскрипт гена. "Первинний транскрипт", наприклад, зберігає інтрони та/або ще не містить поліА-хвоста або кеп-структури та/або не містить інших модифікацій, необхідних для його належної функції як транскрипта, наприклад, зачищення або коректування.

Промотор: терміни "промотор" або "промоторна послідовність" є рівноцінними і у контексті даного опису означають послідовність ДНК, яка, будучи лігваною з потрібною нуклеотидною послідовністю, здатна контролювати транскрипцію потрібної нуклеотидної послідовності в РНК. Такі промотори можна знайти, наприклад, у таких загальнодоступних базах даних: <http://www.grassius.org/grasspromdb.html>, <http://mendel.cs.rhul.ac.uk/mendel.php?topic=plantprom>, <http://ppdb.gene.nagoya-u.ac.jp/cgi-bin/index.cgi>. Перелічені у них промотори можуть застосовуватися згідно зі способами винаходу і, таким чином, є включеними шляхом посилання.

Промотор розташовується на 5'-кінці (тобто, у зворотному напрямку), поблизу від сайту початку транскрипції потрібної нуклеотидної послідовності, транскрипцію якої в мРНК він контролює, і забезпечує сайт специфічного зв'язування РНК-полімеразою та інші та інші фактори транскрипції для започаткування транскрипції. Зазначений промотор включає, наприклад, 5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 принаймні 10 т. н., наприклад, 5 т. н. або 2 т. н, поблизу від сайту початку транскрипції. Він також може включати принаймні 1500 п. о, поблизу від сайту початку транскрипції, в оптимальному варіанті принаймні 1000 п. о., у ще кращому варіанті - принаймні 500 п. о., у ще кращому варіанті - принаймні 400 п. о., принаймні 300 п. о., принаймні 200 п. о. або принаймні 100 п. о. У ще одному оптимальному варіанті втілення промотор включає принаймні 50 п. о. поблизу від сайту початку транскрипції, наприклад, принаймні 25 п. о. Промотор не включає екзонних та/або інтронних ділянок або 5'-нетрансльованих ділянок. Промотор може бути, наприклад, гетерологічним або гомологічним відповідній рослині. Полінуклеотидна послідовність є "гетерологічною" організмові або другій полінуклеотидній послідовності, якщо вона походить з іншого виду, або, якщо походить з того самого виду, є модифікованою з первісної форми. Наприклад, промотор, функціонально зв'язаний з гетерологічною кодуючою послідовністю, означає кодуючу послідовність з виду, відмінного від того, з якого походить промотор, або, якщо походить з того самого виду, кодуючу послідовність, яка у природі не є зв'язаною з промотором (наприклад, піддану генній інженерії кодуючу послідовність або алель з іншого екотипу або сорту). Прийнятні промотори можуть походити з генів клітин-хазяїв, у яких має відбуватися експресія, або з патогенів для цих клітин-хазяїв (наприклад, рослин або рослинних патогенів, таких, як рослинні віруси). Рослинно-специфічний промотор є промотором, придатним для регулювання експресії у рослині. Він може походити з рослини, але також і з рослинних патогенів, або ж може бути синтетичним промотором, побудованим людиною. Якщо промотор є індукованим промотором, швидкість транскрипції збільшується у відповідь на дію індукуючого агента. Крім того, промотор може регулюватись у тканинно-специфічний спосіб, таким чином, що він є виключно або переважно активним у транскрипції асоційованої кодуючої ділянки у конкретному(их) типі(ях) тканин, таких, як листя, коріння або меристема. Термін "тканинно-специфічний", якщо стосується промотора, означає промотор, здатний спрямовувати вибірково експресію потрібної нуклеотидної послідовності на певний тип тканини (наприклад, пелюстки) за відносної відсутності експресії тієї ж самої потрібної нуклеотидної послідовності в іншому типі тканини (наприклад, корінні). Тканинна специфічність промотора може визначатися, наприклад, шляхом функціонального зв'язування гена-репортера з промоторною послідовністю для утворення репортерної послідовності, включення репортерної послідовності у геном рослини, таким чином, щоб репортерної послідовності були включені у кожену тканину одержаної в результаті трансгенної рослини, та виявлення експресії гена-репортера (наприклад, виявлення мРНК, білка або активності білка, який кодується геном-репортером) в інших тканинах трансгенної рослини. Виявлення вищого рівня експресії гена-репортера в одній або кількох тканинах відносно рівня експресії гена-репортера в інших тканинах свідчить, що промотор є специфічним для тканин, у яких виявляється вищий рівень експресії. Термін "специфічний для типу клітин" стосовно промотора означає промотор, здатний спрямовувати вибірково експресію потрібної нуклеотидної послідовності у певний тип клітини за відносної відсутності експресії тієї ж самої потрібної нуклеотидної послідовності в іншому типі клітин у межах однієї тканини. Термін "специфічний для типу клітин" стосовно промотора також означає промотор, здатний сприяти вибірковій експресії потрібної нуклеотидної послідовності у ділянці у межах однієї тканини. Специфічність промотора для типу клітин визначають способами, добре відомими спеціалістам у даній галузі, наприклад, за допомогою забарвлення GUS на виявлення активності, зеленого флуоресцентного білка або імуногістохімічного забарвлення. Термін "конститутивний" стосовно промотора або експресії, зумовленої промотором, означає, що промотор здатен спрямовувати транскрипцію функціонально зв'язаної молекули нуклеїнової кислоти за відсутності подразника (наприклад, термічного шоку, хімікатів, світла і т. ін.) у більшості рослинних тканин та клітин протягом практично всієї тривалості життя рослини або частини рослини. Як правило, конститутивні промотори здатні спрямовувати експресію трансгена практично у будь-якій клітині та будь-якій тканині.

Промотор специфічність: Термін "специфічність" стосовно промотора означає характер експресії, зумовлений відповідним промотором. Специфічність характеризує тканини та/або стан розвитку рослини або її частини, у якій промотор забезпечує експресію молекули нуклеїнової кислоти під контролем відповідного промотора. Специфічність промотора також може включати умови середовища, за яких промотор може активуватися або пригнічуватися, наприклад, викликання або пригнічення через біологічні стреси або вплив навколишнього середовища, такі як холод, посуха, пошкодження або інфікування.

Очищений: у контексті даного опису термін "очищена" стосується молекул, нуклеїновокислотних або амінокислотних послідовностей, видалених з їх природного середовища, виділених або відокремлених. "Суттєво очищені" молекули є принаймні на 60 % вільними, в оптимальному варіанті - принаймні на 75 % вільними, у ще кращому варіанті

5 принаймні на 90 % вільними від інших компонентів, з якими вони є зв'язаними у природі. Очищена нуклеїновокислотна послідовність може бути виділеною нуклеїновокислотою послідовністю.

Рекомбінантний: термін "рекомбінантна" стосовно молекул нуклеїнових кислот означає молекули нуклеїнових кислот, одержані за допомогою технологій рекомбінантних ДНК.

10 Рекомбінантні молекули нуклеїнових кислот також можуть включати молекули, що як такі не існують у природі, але є модифікованими, зміненими, мутованими або іншим чином підданими маніпуляціям з боку людини. В оптимальному варіанті "рекомбінантна молекула нуклеїнової кислоти" є неприродною молекулою нуклеїнової кислоти, яка відрізняється за послідовністю від природної молекули нуклеїнової кислоти принаймні на одну нуклеїнову кислоту. "Рекомбінантна

15 молекула нуклеїнової кислоти" також може включати "рекомбінантну послідовність", яка включає, в оптимальному варіанті - є функціонально зв'язаною, з послідовностями молекул нуклеїнових кислот, які не є природними у цьому порядку. Оптимальні способи одержання зазначеної рекомбінантної молекули нуклеїнової кислоти можуть включати клонування, спрямований або неспрямований мутагенез, синтез або способи рекомбінації.

20 "Насінно-специфічний промотор" у контексті цього винаходу означає промотор, який регулює транскрипцію молекули нуклеїнової кислоти під контролем відповідного промотора у насінні, причому транскрипція у будь-якій тканині або клітині насіння складає більш, ніж 90 %, в оптимальному варіанті - більш, ніж 95 %, у ще кращому варіанті - більш, ніж 99 % загальної кількості РНК, транскрибованої з зазначеної нуклеїновокислотної послідовності в усій рослині

25 протягом будь-якої стадії її розвитку. Терміни "насінно-специфічна експресія" та "насінно-специфічна NEENA" слід розуміти відповідним чином. Отже, "насінно-специфічна NEENA" посилює транскрипцію насінно-специфічного або насінно-селективного промотора таким чином, що транскрипція у насінні, зумовлена зазначеним промотором, функціонально зв'язаним з відповідною NEENA, складає більш, ніж 90 %, в оптимальному варіанті - більш, ніж 95 %, у ще

30 кращому варіанті - більш, ніж 99 % загальної кількості РНК, транскрибованої з відповідного промотора, функціонально зв'язаного з NEENA, в усій рослині протягом будь-якої стадії її розвитку.

"Насінно-селективний промотор" у контексті цього винаходу означає промотор, який регулює транскрипцію молекули нуклеїнової кислоти під контролем відповідного промотора у насінні,

35 причому транскрипція у будь-якій тканині або клітині насіння складає більш, ніж 50 %, в оптимальному варіанті - більш, ніж 70 %, у ще кращому варіанті - більш, ніж 80 % загальної кількості РНК, транскрибованої з зазначеної нуклеїновокислотної послідовності в усій рослині протягом будь-якої стадії її розвитку. Терміни "насінно-селективна експресія" та "насінно-селективна NEENA" слід розуміти відповідним чином. Отже, "насінно-селективна NEENA"

40 посилює транскрипцію насінно-специфічного або насінно-селективного промотора таким чином, що транскрипція у насінні, зумовлена зазначеним промотором, функціонально зв'язаним з відповідною NEENA, складає більш, ніж 50 %, в оптимальному варіанті - більш, ніж 70 %, у ще кращому варіанті - більш, ніж 80 % загальної кількості РНК, транскрибованої з відповідного промотора, функціонально зв'язаного з NEENA, в усій рослині протягом будь-якої стадії її

45 розвитку.

Смислова: термін "смислова" означає молекулу нуклеїнової кислоти, яка має послідовність, комплементарну або ідентичну послідовності-мішені, наприклад, послідовність, яка зв'язується з фактором транскрипції білка і бере участь в експресії даного гена. Згідно з оптимальним

50 варіантом втілення, молекула нуклеїнової кислоти включає потрібний ген та елементи, які забезпечують можливість експресії зазначеного потрібного гена.

Значне підвищення або зниження: підвищення або зниження, наприклад, ферментної активності або експресії генів, яке перевищує межу похибки, характерну для даного способу вимірювання, в оптимальному варіанті - приблизно дворазове або більше підвищення або

55 зниження активності контрольного ферменту або експресії у контрольній клітині, у ще кращому варіанті - підвищення або зниження приблизно у 5 разів або більше, у найкращому варіанті - підвищення або зниження приблизно у 10 разів або більше.

Малі молекули нуклеїнових кислот: "малі молекули нуклеїнових кислот" слід розуміти як молекули, які складаються з нуклеїнових кислот або їх похідних, таких, як РНК або ДНК. Вони можуть бути дволанцюговими або одноланцюговими і включають від приблизно 15 до

60 приблизно 30 п. о., наприклад, від 15 до 30 п. о., у ще кращому варіанті - від приблизно 19 до

приблизно 26 п. о., наприклад, від 19 до 26 п. о., у ще кращому варіанті - від приблизно 20 до приблизно 25 п. о., наприклад, від 20 до 25 п. о. У варіанті втілення, якому віддають найбільшу перевагу, олігонуклеотиди складають від приблизно 21 до приблизно 24 п. о., наприклад, від 21 до 24 п. о. У найкращому варіанті малі молекули нуклеїнових кислот складають приблизно 21 п. о., та приблизно 24 п. о., наприклад, 21 п. о. та 24 п. о.

Суттєво комплементарна: у найширшому сенсі термін "суттєво комплементарна", вжитий у контексті даного опису стосовно нуклеотидної послідовності для контрольної або заданої нуклеотидної послідовності, означає нуклеотидну послідовність, яка має відсоток ідентичності між суттєво комплементарною нуклеотидною послідовністю та точно комплементарною послідовністю зазначеної контрольної або заданої нуклеотидної послідовності принаймні 60 %, більш бажано - принаймні 70 %, більш бажано - принаймні 80 % або 85 %, в оптимальному варіанті - принаймні 90 %, у ще кращому варіанті - принаймні 93 %, у ще кращому варіанті - принаймні 95 % або 96 %, у ще кращому варіанті - принаймні 97 % або 98 %, у ще кращому варіанті - принаймні 99 %, або у найкращому варіанті - 100 % (остання є рівноцінною термінові "ідентична" у цьому контексті). В оптимальному варіанті ідентичність визначають на відрізок принаймні 19 нуклеотидів завдовжки, в оптимальному варіанті - принаймні 50 нуклеотидів, у ще кращому варіанті - по всій довжині нуклеїновокислотної послідовності, з зазначеною контрольною послідовністю (якщо нижче не зазначено іншого). Порівняння послідовностей здійснюють, застосовуючи стандартний GAPаналіз з GCG Університету шт. Вісконсін, платформа SEQWEB для GAP, на основі алгоритму Нідлмана - Вунша (Needleman and Wunsch (1970) J Mol. Biol. 48: 443-453; як визначено вище). Нуклеотидна послідовність, "суттєво комплементарна" контрольній нуклеотидній послідовності, гібридується з контрольною нуклеотидною послідовністю в умовах низької жорсткості, в оптимальному варіанті - і в умовах середньої жорсткості, у найкращому варіанті - в умовах високої жорсткості (як визначено вище).

Трансген: термін "трансген" у контексті даного опису стосується будь-якої нуклеїновокислотної послідовності, яку вводять у геном клітини шляхом експериментальних маніпуляцій. Трансген може бути "ендогенною послідовністю ДНК" або "гетерологічною послідовністю ДНК" (тобто, "чужорідною ДНК"). Термін "ендогенна послідовність ДНК" стосується нуклеотидної послідовності, яка у природі міститься у клітині, до якої вона включена, якщо вона не містить певних модифікацій (наприклад, точкової мутації, наявності селектованого маркерного гена і т. ін.) відносно природної послідовності.

Трансгенний: термін "трансгенний" стосовно організму означає трансформований, в оптимальному варіанті - стійко трансформований, молекулою рекомбінантної ДНК, яка в оптимальному варіанті включає відповідний промотор, функціонально з'єднаний з потрібною послідовністю ДНК.

Вектор: у контексті даного опису термін "вектор" стосується молекули нуклеїнової кислоти, здатної переносити іншу молекулу нуклеїнової кислоти, з якою він зв'язується. Один тип вектора являє собою геномний інтегрований вектор або "інтегрований вектор", який може бути інтегрований у хромосомну ДНК клітини-хазяїна. Іншим типом вектора є епісомний вектор, тобто, молекула нуклеїнової кислоти, здатна до позахромосомної реплікації. Вектори, здатні спрямовувати експресію генів, з якими вони є функціонально зв'язаними, у даному описі називаються "векторами експресії". У даному описі терміни "плазміда" та "вектор" вживаються взаємозаміно, якщо контекст прямо не передбачає іншого. Вектори експресії, побудовані для утворення описаних авторами РНК *in vitro* або *in vivo*, можуть містити послідовності, які розпізнаються будь-якою РНК-полімеразою, включаючи мітохондріальну РНК-полімеразу, РНК pol I, РНК pol II та РНК pol III. Ці вектори можуть застосовуватися для транскрипції потрібної молекули РНК у клітині згідно з винаходом. Вектор трансформації рослин слід розуміти як вектор, прийнятний для процесу трансформації рослин.

Дикий тип; терміни "дикий тип", "природний" або "природне походження" стосовно організму, поліпептиду або нуклеїновокислотної послідовності означає, що зазначений організм є природним або наявним у принаймні одному природному організмі, який не був змінений, мutowаний або підданий іншим маніпуляціям з боку людини.

ПРИКЛАДИ

Хімічні речовини та загальні способи

Якщо не зазначено іншого, процедури клонування, які здійснюють згідно з даним винаходом, включаючи рестрикційне картування, електрофорез в агарозному гелі, очищення нуклеїнових кислот, лігування нуклеїнових кислот, трансформація, відбір та культивування бактеріальних клітин, здійснювали, як описано (Sambrook et al. 1989). Аналізи послідовностей рекомбінантних ДНК здійснювали за допомогою пристрою для лазерного флуоресцентного секвенування ДНК (Applied Biosystems, Foster City, CA, США) з застосуванням технології Сенджера (Sanger et al.,

1977). Якщо не описується іншого, хімічні речовини та реагенти одержували від Sigma Aldrich (Sigma Aldrich. St. Louis, США), від Promega (Madison, WI, США), Duchefa (Haarlem, Нідерланди) або Invitrogen (Carlsbad, CA, США). Рестрикційні ендонуклеази одержували від New England Biolabs (Ipswich, MA, США) або Roche Diagnostics GmbH (Penzberg, Німеччина).

5 Олігонуклеотиди були синтезовані компанією Eurofins MWG Operon (Ebersberg, Німеччина).

Приклад 1: Ідентифікація нуклеїнових кислот-кандидатів, які посилюють експресію нуклеїнових кислот (NEENA), з генів з насінно-специфічною або насінно-селективною експресією

1.1 Ідентифікація молекул NEENA з генів *A. thaliana*

10 Застосовуючи загально доступні геномні послідовності ДНК (наприклад, <http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>) та дані експресії транскриптів (наприклад, <http://www.weigelworld.org/resources/microarray/AtGenExpress/>), набір з 19 NEENA-кандидатів, що походять з транскриптів *Arabidopsis thaliana* з насінно-специфічною або насінно-селективною експресією було відібрано для детального аналізу. Кандидати зазначені нижче:

15

Таблиця 1: насінно-специфічні NEENA-кандидати (NEENAss).

Назва NEENA	Локус	Анотація	SEQ ID NO
NEENAss1	At1g62290	білок групи аспартил-протеаз	1
NEENAss2	At1g65090	експресований білок	2
NEENAss15	At2g27040	білок, який містить PAZ-домен	3

NEENAss18	At1g01170	реагуючий на озон стрес-білок, гаданий	4
NEENAss14	At5g63190	білок, який містить MA3-домен	5
NEENAss4	At5g07830	білок, який містить N-кінцевий домен, подібний до бета-глюкуронідази AtGUS2, з родини 79 глікозил-гідролаз	6
NEENAss13	At2g04520	фактор ініціації еукаріотної трансляції 1A, гаданий / eIF-1A	7
NEENAss3	At5g60760	пов'язаний з 2-фосфогліцераткіназою	8
NEENAss5	At1g11170	експресований білок містить Pfam-профіль PF05212	9
NEENAss11	At4g37050	PLA V/PLP4 (пататиноподібний білок 4)	10
NEENAss8	At1g56170	HAP5B (гем-активуючий білок (дріжджі) гомолог 5B)	11
NEENAss16	At1g54100	альдегіддегідрогеназа, гаданий / антиквітін	12
NEENAss9	At3g12670	СТР-синтаза, гаданий / UTP-аміак-лігаза, гаданий	13
NEENAss20	At4g04460	білок групи аспартил-протеаз	14
NEENAss10	At1g04120	ATMRP5 (білок множинної лікарської резистентності 5 <i>Arabidopsis thaliana</i>)	15
NEENAss6	At2g41070	основний фактор транскрипції лейцинової застіжки (BZIP12)	16
NEENAss12	At1g05450	пов'язаний з інгібітором протеази / запасанням у насінні/білком перенесення ліпідів (LTP)	17
NEENAss7	At4g03050	2-оксоглутарат-залежна діоксигеназа, гаданий (AOP3)	18
NEENAss17	At3g12490	інгібітор цистеїнпротеази, гаданий / цистатин	19

1.2 Виділення NEENA-кандидатів

- Геномну ДНК видобували з зеленої тканини *A. thaliana*, застосовуючи комплект Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Німеччина). Фрагменти геномної ДНК, які містять гадані молекули NEENA виділяли за допомогою традиційної полімеразної ланцюгової реакції (ПЛР). Праймери будували на основі геномної послідовності *A. thaliana* з багатьма NEENA-кандидатами. Реакція включала 19 наборів праймерів (Таблиця 2), і її здійснювали за протоколом Phusion High Fidelity DNA Polymerase (Cat No F-540L, New England Biolabs, Ipswich, MA, США). Виділену ДНК використовували як матричну ДНК у ПЛР-ампліфікації з застосуванням таких праймерів:

Таблиця 2: Праймерні послідовності

Назва праймера	Послідовність	SEQ ID NO	ПЛР, що забезпечує SEQ ID NO
NEENAss1_for	aataatggtacctggtgcttaaacactctggtgagt	20	1
NEENAss1_rev	aataatccatggttgacctacaaaatcaaagcagtca	21	
NEENAss2_for	tttttggtaccagttctttgcttcgaagtgc	22	2
NEENAss2_rev	tttttccatggtactacgtactgtttcaattct	23	
NEENAss3_for	aaaaaagggtaccattccacacgcttctatcatttc	24	8
NEENAss3_rev	aaaaaaccatggttatctctctctaaaaataaaaaacgaatc	25	
NEENAss4_for	aataaagggtaccgtccagaattttctccattga	26	6
NEENAss4_rev	aataaaccatggtcttcactatccaaagctctca	27	
NEENAss5_for	tttttggtaccgtctactttcattacagtactctg	28	9
NEENAss5_rev	tttttccatggttatattttacgtgcaacacaattcaa	29	
NEENAss6_for	ttttatggtaccactcgaatactgcatgcaa	30	16
NEENAss6_rev	ttttatccatggttatgtagcctttacacagaaaacaa	31	
NEENAss7_for	tatataggtaccaacaactatggcctgagggt	32	18
NEENAss7_rev	tatataccatggttatcttactgttttaacaaaaataaaaat	33	
NEENAss8_for	tttttaggtaccatcttagggtttcgagatctca	34	11
NEENAss8_rev	tttttccatggtgctaagctatctctgtaataataaaattg	35	
NEENAss9_for	tttttggtaccattttgttggtgaaaggtaga	36	13
NEENAss9_rev	ttttaccatggttacgtttttgtctctgctctct	37	
NEENAss10_for	tatatgggtacctctgggaatatcgattttgatct	38	15
NEENAss10_rev	tatataccatggtctcaccacatcccaaagctc	39	
NEENAss11_for	ttttatggtaccgcacaaatcttagcttaccttgaa	40	10
NEENAss11_rev	ttttatccatggttatattaaccacaagccttgctc	41	
NEENAss12_for	tttttaggtacctgtcggagaagtgggcg	42	17
NEENAsskor-rev	ttttaccatggagaagtgggoggacg	43	
NEENAss13_for	ttttatggtacctagcttaatctcagattcgaatcgt	44	7
NEENAss13_rev	ttttatccatggttagtatctacataccaatcatacaaatg	45	
NEENAss14_for	tttttggtaccttcacgatttgaattga	46	5
NEENAss14_rev	tttttccatggtctacaacattaaaaacgaccatta	47	
NEENAss15_for	tatataggtaccagggtttcgttttgttca	48	3
NEENAss15_rev	tatataccatggttatctctgctcaaaagaaacca	49	
NEENAss16_for	tttataggtaccagaagctcatttctcgatac	50	12
NEENAss16_rev	tttataccatggtctctgogcaaaaattcacc	51	
NEENAss17_for	tatatgggtacctctaaaaatacagggcacc	52	19
NEENAss17_rev	tatatccatggttactcttcgttcagaagccta	53	
NEENAss18_for	tatataggtaccactgtttaagcttcactgtct	54	4
NEENAss18_rev	tatataccatggtttctttaaagctgaaagt	55	
NEENAss20_for	tatataggtaccttaagcttttaagaatctctactaca	56	14
NEENAss20(2)_rev	atatatccatggttaaattttacctgtcatcaaaaacaaca	57	

Ампліфікацію під час ПЛР здійснювали з такою композицією (50 мкл);

3,00 мкл геномної ДНК *A. thaliana* (50 нг/мкл)

10,00 мкл 5x Phusion HP буфера

4,00 мкл дНТФ (2,5 мМ)

2,50 мкл прямого праймера (10 мкМ)

2,50 мкл зворотного праймера (10 мкМ)

0,50 мкл ДНК-полімерази Phusion HP (2U/мкл)

Для ПЛР застосовували метод touch-down з такими параметрами: 98,0 °C протягом 30 сек (1 цикл), 98,0 °C протягом 30 сек, 56,0 °C протягом 30 сек та 72,0 °C протягом 60 сек (4 цикли), 4 додаткові цикли кожен для температури випалу 54,0 °C, 51,0 °C та 49,0 °C, з наступними 20 циклами при 98,0 °C протягом 30 сек, 46,0 °C протягом 30 сек та 72,0 °C протягом 60 сек (4 цикли) та 72,0 °C протягом 5 хв. Продукти ампліфікації подавали на 2 % (маса/об'єм) агарозний гель і відокремлювали при 80 В. Продукти ПЛР видаляли з гелю й очищали за допомогою комплекту Qiagen Gel Extraction Kit (Qiagen, Hilden, Німеччина). Після рестрикції ДНК за допомогою рестрикційних ендонуклеаз NcoI (10 О/мкл) та KpnI (10 О/мкл) гідролізовані продукти знову очищали за допомогою комплекту Qiagen Gel Extraction Kit (Qiagen, Hilden, Німеччина).

1.3 Побудова вектора

Застосовуючи систему Multisite Gateway (Invitrogen, Carlsbad, CA, США), касети промотор: NEENA: ген-репортер збирали у бінарні послідовності для трансформації рослин. Насінно-специфічний промотор *A. thaliana* p-AtPXR (At1g48130, GenBank AC023673.3; документ WO2006089950; з префіксом p-, який означає промотор) застосовували у послідовності гена-репортера, і люциферазу світляка (Promega, Madison, WI, США) застосовували як білок-репортер для кількісного визначення ефекту посилення експресії, який забезпечується молекулами NEENA, що підлягають аналізу.

Вектор pENTR/A, який містить промотор p-AtPXR, клонували через сайт-специфічну рекомбінацію (BP-реакцію) між вектором pDONR/A та продуктами ампліфікації p-AtPXR з праймерами p-AtPXR-for та p-AtPXR-rev (Таблиця 3) на геномній ДНК (див, вище) з сайтами сайт-специфічної рекомбінації на будь-якому з кінців згідно з інструкціями виробника (Invitrogen, Carlsbad, CA, США). Позитивні клони pENTR/A піддавали аналізу послідовності для перевірки правильності промотора p-AtPXR.

Таблиця 3: Праймерні послідовності (p-AtPXR)

Назва прай- мера	Послідовність	SEQ ID NO.
p-AtPXR-for	ggggacaacttggatagaaaagttggccacatcatgtttagacttacc	58
p-AtPXR-rev	ggggactgctttttgtacaaactgtttaccttttatatttatatag	59

Генерували вектор ENTR/B, який містить кодуючу послідовність люциферази світляка (Promega, Madison, WI, США) з наступним термінатором транскрипції нопалінсинтази t-nos (Genbank V00087). ПЛР-фрагменти NEENA-кандидата (див. вище) клонували окремо перед кодуючою послідовністю люциферази світляка, застосовуючи рестрикційні ферменти KpnI та NcoI. Одержані в результаті вектори pENTR/B показано у Таблиці 4, з промоторними молекулами, які мають префікс p-, кодуючими послідовностями, які мають префікс C-, та термінуючими молекулами, які мають префікс t-.

Таблиця 4: усі вектори рENTR/B плюс та мінус NEENA-кандидати

Вектор рENTR/B	Склад касети часткової експресії SEQ ID NO::ген-репортер::термінатор
LJK01	MCS::c-LUC::t-nos
LJK19	SEQ ID NO1::c-LUC::t-nos
LJK20	SEQ ID NO2::c-LUC::t-nos
LJK21	SEQ ID NO8::c-LUC::t-nos
LJK22	SEQ ID NO6::c-LUC::t-nos
LJK23	SEQ ID NO9::c-LUC::t-nos
LJK24	SEQ ID NO16::c-LUC::t-nos
LJK25	SEQ ID NO18::c-LUC::t-nos
LJK26	SEQ ID NO11::c-LUC::t-nos
LJK27	SEQ ID NO13::c-LUC::t-nos
LJK28	SEQ ID NO15::c-LUC::t-nos
LJK29	SEQ ID NO10::c-LUC::t-nos
LJK30	SEQ ID NO17::c-LUC::t-nos
LJK31	SEQ ID NO7::c-LUC::t-nos
LJK32	SEQ ID NO5::c-LUC::t-nos
LJK33	SEQ ID NO3::c-LUC::t-nos
LJK34	SEQ ID NO12::c-LUC::t-nos
LJK35	SEQ ID NO19::c-LUC::t-nos
LJK36	SEQ ID NO4::c-LUC::t-nos
LJK38	SEQ ID NO14::c-LUC::t-nos

Вектор рENTR/C будували шляхом включення сайту множинного клонування (SEQ ID NO60) через сайти рестрикції KpnI та HindIII. Шляхом виконання сайт-специфічної рекомбінації (LR-реакції) створені рENTR/A, рENTR/B та рENTR/C комбінували з вектором призначення рSUN (похідною рSUN) згідно з інструкціями Multisite Gateway від виробника (Invitrogen, Carlsbad, CA, США). Реакції забезпечували 1 бінарний вектор з промотором р-AtPXR, кодууючою послідовністю люцифери світляка c-LUC та термінатором t-nos і 19 векторів, які включали SEQ ID NO1, NO2, NO3, NO4, NO5, NO6, NO7, NO8, NO9, NO10, NO11, NO12, NO13, NO14, NO15, NO16, NO17, NO18 та NO19 безпосередньо перед кодууючою послідовністю люцифери світляка (Таблиця 5), прикладом якої є комбінація з SEQ ID NO1 (SEQ ID NO61). За винятком зміни SEQ ID NO2 на NO19, нуклеотидна послідовність є ідентичною в усіх векторах (Таблиця 5). Одержані в результаті вектори трансформації рослин показано у Таблиці 5:

Таблиця 5: Рослинні вектори експресії для трансформації *A. thaliana*

Рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	SEQ ID NO
LJK134	p-AtPXR::c-LUC::t-nos	
LJK71	p-AtPXR::SEQ ID NO1::c-LUC::t-nos	61
LJK72	p-AtPXR::SEQ ID NO2::c-LUC::t-nos	
LJK73	p-AtPXR::SEQ ID NO8::c-LUC::t-nos	
LJK74	p-AtPXR::SEQ ID NO6::c-LUC::t-nos	
LJK75	p-AtPXR::SEQ ID NO9::c-LUC::t-nos	
LJK76	p-AtPXR::SEQ ID NO16::c-LUC::t-nos	
LJK77	p-AtPXR::SEQ ID NO18::c-LUC::t-nos	
LJK78	p-AtPXR::SEQ ID NO11::c-LUC::t-nos	
LJK79	p-AtPXR::SEQ ID NO13::c-LUC::t-nos	
LJK80	p-AtPXR::SEQ ID NO15::c-LUC::t-nos	
LJK81	p-AtPXR::SEQ ID NO10::c-LUC::t-nos	
LJK82	p-AtPXR::SEQ ID NO17::c-LUC::t-nos	
LJK83	p-AtPXR::SEQ ID NO7::c-LUC::t-nos	
LJK84	p-AtPXR::SEQ ID NO5::c-LUC::t-nos	
LJK85	p-AtPXR::SEQ ID NO3::c-LUC::t-nos	
LJK86	p-AtPXR::SEQ ID NO12::c-LUC::t-nos	
LJK87	p-AtPXR::SEQ ID NO19::c-LUC::t-nos	
LJK88	p-AtPXR::SEQ ID NO4::c-LUC::t-nos	
LJK90	p-AtPXR::SEQ ID NO14::c-LUC::t-nos	

Одержані в результаті вектори потім застосовували для одержання трансгенних рослин *A. thaliana*.

Приклад 2; Відбір молекул NEENA, які посилюють експресію генів у трансгенних рослинах *A. thaliana*.

Цей приклад показує, що лише вибрані молекули NEENA-кандидатів здатні посилювати експресію генів.

Усі бінарні послідовності, які містили вибрані молекули NEENA-кандидатів, описаних у Прикладі 1, стійко трансформували у рослини *Arabidopsis thaliana* разом з контрольною послідовністю без NEENA. Для одержання трансгенних рослин *A. thaliana* *Agrobacterium tumefaciens* (штам C58C1 pGV2260) трансформували різними описаними вище послідовностями векторів. Для трансформації *A. thaliana* застосовували спосіб Floral Dip (Clough and Bent, 1998, Plant Journal 16: 735-743). Трансгенні рослини T1 відбирали шляхом пророщування та вирощування сіянців на канаміцині. Через 12 днів сім'ядолі трансформантів та контрольних рослин дикого типу відбирали й розподіляли по 96-лункових планшетах, у які попередньо поміщали 50 мкл 0,5х середовища Мурацлре-Скура і піддавали аналізу гена-репортера люциферази (змінений протокол Weigel and Glazebrook, 2002, *Arabidopsis*, лабораторний довідник. Cold Spring Harbor Laboratory Press, Chapter 7, ISBN 0-87969-572-2). Люмінесцентність сім'ядолі визначали у розчині, який містив 0,1 мМ D-люциферину (Кат. №: L-8220, BioSynth, Staad, Швейцарія) та 0,01 % Tween20 (Sigma Aldrich, St Louis, США) у Microlumat Plus LB96V (Berthold Technologies, Bad Wildbad, Німеччина), і записували через 60 хв після додавання D-люциферину. Інструментальні показання усереднювали для кожної послідовності і на основі цих середніх значень експресії розраховували значення кратних змін для оцінки впливу присутності гаданої NEENA порівняно з послідовностями гена-репортера без відповідної гаданої NEENA. Порівняно з послідовностями гена-репортера, які містили лише насінно-специфічний промотор p-AtPXR без NEENA, 19 випробуваних послідовностей, які містили NEENA-кандидати, демонстрували негативний, а також позитивний вплив, від 0,8-разової до 22,2-разової індукції

активності люциферази (Фіг. 1). Загалом 15 гаданих молекул NEENA, які включали послідовності з SEQ ID NO1, NO2, NO3, NO4, NO5, NO6, NO7, NO8, NO9, NO10, NO11, NO12, NO13, NO14 та NO15, забезпечували більш, ніж 2,5-разове підвищення експресії генів на основі активності гена-репортера люциферази порівняно з послідовністю гена-репортера, що не містить NEENA і містить лише промотор (Фіг. 1), а отже, є функціональними молекулами NEENA. Оскільки багато випробуваних молекул NEENA-кандидатів мають незначний або навіть негативний вплив на посилення експресії генів, не всі гадані молекули NEENA опосередковують загальний стимулюючий вплив, а лише вибрані послідовності NEENA забезпечують значне посилення експресії генів (SEQ ID NO з 1 по 15).

Приклад 3: Випробування молекул NEENA на посилення насінно-специфічної експресії генів у рослинах олійного рапсу.

Цей приклад показує, що молекули NEENA можуть застосовуватись у різних видах для посилення експресії генів тканинно-специфічного промотора порівняно з відсутністю NEENA і присутністю лише промотора.

Молекули NEENA, які опосередковують найбільше посилення експресії генів у попередньому відборі (пор. Приклад 2, SEQ ID NO1, NO2, NO3, NO4, NO5, NO6 та NO7), вибирали для визначення посилення рівня експресії генів у трансгенних рослинах олійного рапсу.

3.1 Побудова вектора для трансформації рослин *B. napus*

Для трансформації рослин олійного рапсу касети експресії генів-репортерів без молекул контролю експресії генів та з ними (SEQ ID NO1-NO7) комбінували з касетою експресії генів, яка містила селектований маркерний ген, для виявлення ліній трансгенних рослин у межах вектора рENTR/C. Шляхом виконання сайт-специфічної рекомбінації (LRреакції), як було описано вище (див. вище, 1,3), рENTR/A, рENTR/B та рENTR/C, які містили касету селектованого маркера, комбінували з вектором призначення рSUN згідно з інструкціями Multisite Gateway від виробника (Invitrogen, Carlsbad, CA, США). Реакції забезпечували один бінарний вектор з промотором р-AtPXR, кодуючою послідовністю люциферази світляка c-LUC, термінатором t-nos та касетою селектованого маркера, а також 7 векторів, які включали SEQ ID NO1, NO2, NO3, NO4, NO5, NO6 та NO7 безпосередньо перед кодуючою послідовністю люциферази світляка (Таблиця 6), прикладом якої є комбінація з SEQ ID NO1 (SEQ ID NO62). За винятком зміни SEQ ID NO2 на NO7, нуклеотидна послідовність є ідентичною в усіх векторах (Таблиця 6). Одержані в результаті вектори трансформації рослин показано у Таблиці 6:

Таблиця 6: Рослинні вектори експресії для трансформації *B. napus*

Рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	SEQ ID NO
LJK148	p-AtPXR::c-LUC::t-nos	
LJK156	p-AtPXR::SEQ ID NO1::c-LUC::t-nos	62
LJK157	p-AtPXR::SEQ ID NO2::c-LUC::t-nos	
LJK158	p-AtPXR::SEQ ID NO7::c-LUC::t-nos	
LJK159	p-AtPXR::SEQ ID NO5::c-LUC::t-nos	
LJK160	p-AtPXR::SEQ ID NO4::c-LUC::t-nos	
LJK161	p-AtPXR::SEQ ID NO6::c-LUC::t-nos	
LJK162	p-AtPXR::SEQ ID NO3::c-LUC::t-nos	

3.2 Створення трансгенних рослин олійного рапсу (змінений протокол згідно з Moloney et al., 1992, Plant Cell Reports, 8: 238-242).

При підготуванні до створення трансгенних рослин олійного рапсу бінарні вектори трансформували у *Agrobacterium tumefaciens* C58C1:pGV2260 (Deblaere et al., 1985, Nucl. Acids, Res. 13: 4777-4788). Розведену у пропорції 1:50 витриману протягом доби культуру *Agrobacteria*, яка містила відповідну послідовність, вирощували у середовищі МурасігеСкуга (Murashige and Skoog, 1962, Physiol. Plant 15, 473), доповненому 3 % сахарозою (ЗМ5-середовище). Для трансформації рослин олійного рапсу черешки або гіпокотиль стерильних рослин інкубували з розчином 1:50 *Agrobacterium* протягом 5-10 хвилин з наступною триденною спільною інкубацією у темряві при 25 °C на ЗМС. Середовище доповнювали 0,8 % бактоагаром. Через три дні експлантати переносили до MS-середовища, яке містило 500 мг/л клафорану (Cefotaxime-

Sodium), 100 нМ імазетапіру, 20 мкМ бензиламінопурину (BAP) та 1,6 г/л глюкози у світловому режимі 16 год. світла/8 год. темряви, який повторювали з тижневою періодичністю. Паростки, що розвивалися, переносили до MS-середовища, яке містило 2 % сахарози, 250 мг/л клафрану та 0,8 % бактоагару. Через 3 тижні до середовища додавали гормон росту 2-індолбутилову кислоту для сприяння формуванню коріння. Паростки переносили у ґрунт після початку розвитку коріння, вирощували протягом двох тижнів у камері штучного клімату і вирощували до дозрівання в оранжерейних умовах.

3.3 Аналіз рослин

Зразки тканин збирали з листя, квіток та насіння одержаних трансгенних рослин на різних стадіях розвитку, які зберігали у морозильній камері при -80 °C, і піддавали аналізу гена-репортера люциферази (змінений протокол Ow et al., 1986). Після подрібнення заморожені зразки тканин ресуспендували у 800 мкл буфера I (0,1 М фосфатного буфера, pH 7,8, 1 мМ DTT (Sigma Aldrich, St Louis, MO, США), 0,05 % Tween 20 (Sigma Aldrich, St. Louis, MO, США) з наступним центрифугуванням при 10000 g протягом 10 хв 75 мкл водного супернатанту переносили до 96-лункових планшетів. Після додавання 25 мкл буфера II (80 мМ гліцин-гліцилу (Carl Roth, Karlsruhe, Німеччина), 40 мМ MgSO₄ (Duchefa, Haarlem, Нідерланди), 60 мМ ATP (Sigma Aldrich, St. Louis, MO, США), pH 7,8) та D-люциферину до кінцевої концентрації 0,5 мМ (Кат. №: L-8220, BioSynth, St. Gallen, Швейцарія) записували люмінесценцію у MicroLumat Plus LB96V (Berthold Technologies, Bad Wildbad, Німеччина), одержуючи показник у відносних світлових одиницях RLU за хвилину (RLU/XB).

Для нормалізації активності люциферази між зразками визначали концентрацію білка у водному супернатанті паралельно активності люциферази (адаптований спосіб згідно з Bradford, 1976, Anal. Biochem. 72, 248). 5 мкл водного екстракту клітин у буфері I змішували з 250 мкл реагента Бредфорда (Sigma Aldrich, St. Louis, MO, США), інкубували протягом 10 хв при кімнатній температурі. Поглинання визначали при 595 нм у планшетридері (Thermo Electron Corporation, Multiskan Ascent 354). Загальну кількість білка у зразках розраховували за попередньо побудованою стандартною кривою концентрації. Значення за результатами співвідношення RLU/XB та мг білка/мл зразка усереднювали для трансгенних рослин, які містили ідентичні послідовності, і розраховували значення кратних змін для оцінки впливу присутності молекули NEENA порівняно з послідовностями гена-репортера без NEENA.

3.4 Послідовності NEENA опосередковують значне посилення експресії генів у насінні олійного рапсу

Для оцінки потенціалу вибраних молекул NEENA у посиленні експресії генів (SEQ ID NO1, NO2, NO3, NO4, NO5, NO6, NO7) у насінні олійного рапсу насіння на ідентичних стадіях розвитку збирали з окремих ліній трансгенних рослин олійного рапсу, які включали послідовність гена-репортера лише з промотором або послідовності гена-репортера люциферази з вмістом NEENA (SEQ ID NO1, NO2, NO3, NO4, NO5, NO6 або NO7). З кожної трансгенної події збирали 10 зерен, обробляли й аналізували на активність люциферази, як описано вище (Приклад 3.3).

Порівняно з послідовностями гена-репортера, які містили лише насінноспецифічний промотор р-AtPXR без NEENA, усі 7 випробуваних молекул NEENA опосередковували значне посилення експресії генів, від 54-разової до 380-разової індукції активності люциферази у насінні каноли (Фіг. 2a). Порівнянне посилення експресії виявляли у насінні олійного рапсу на пізніших стадіях визрівання (дані не показано).

3.5 Молекули NEENA підвищують тканинно-специфічну експресію генів у насінні олійного рапсу

Для оцінки посилення тканинно-специфічної експресії генів, опосередкованої молекулами NEENA (SEQ ID NO1, NO2, NO3, NO4, NO5, NO6 або NO7) визначали активність люциферази у повністю розвинутому листі та розкритих квітках трансгенних рослин олійного рапсу, які включали вищезазначені послідовності гена-репортера. Три зразки листя ідентичного розміру, а також цілі квітки збирали з кожної рослини окремо і піддавали аналізу гена-репортера люциферази, як описано вище (Приклад 3.3). 5 (SEQ ID NO1, NO2, NO3, NO4, NO5) із 7 випробуваних молекул NEENA демонстрували рівень експресії люциферази, який був порівняним з рівнем послідовності промотора р-AtPXR, що не містить NEENA, у листі та квітках, і, таким чином, не змінює тканинну специфічність насінноспецифічного промотора р-AtPXR (Фіг. 2, b та c). 2 молекули NEENA (SEQ ID NO6, NO7) дещо підвищували активність люциферази у листі та квітках підданих аналізу рослин олійного рапсу (Фіг. 2, b та c) порівняно з рослинами, які включали послідовність, що не містить NEENA. Отже, ці NEENA, SEQ ID NO 6 та 7, є насінно-селективними NEENA, а інші NEENA, SEQ ID NO з 1 по 5, є насінно-специфічними NEENA.

Приклад 4: Аналіз NEENA для насінно-специфічного підвищення сильних насінно-специфічних промоторів.

Цей приклад показує, що здатність молекул NEENA до посилення експресії може використовуватись у комбінації з різними промоторними молекулами для підвищення рівня

5 тканинно-специфічної експресії порівняно з застосуванням лише промоторів.

4.1 Побудова вектора для трансформації рослин *B. napus*

Вибрані молекули NEENA з групи, яку випробували у Прикладі 3 (SEQ ID NO1, NO2, NO3, NO5 та NO6) випробували на їх вплив на посилення тканинно-специфічної експресії генів сильних насінно-специфічних промоторів p-LuPXR (документ WO2006089950, послідовність 9) та p-VfUSP (X56240, Baeumlein et al., 1991). Побудову вектора здійснювали, як описано вище (пор. Приклад 1.3 та 3.1), з праймерними послідовностями, вказаними у Таблиці 7, та вектором LJB765 (документ WO2009016202) як матричною ДНК. Позитивні клони pENTR/A піддавали аналізу послідовності для перевірки правильності промоторів p-LuPXR та p-VfUSP.

Таблиця 7: Праймерні послідовності для p-LuPXR та p-VfUSP

Назва праймера	Послідовність	SEQ ID NO.
p-LuPXR-for	ggggacaactttgtatagaaaagttcacgggcaggacataggactactac	63
p-LuPXR-rev	ggggactgctttttgtacaaacttggatttatgataaaaatgtcggttc	64
p-VfUSP-for	ggggacaactttgtatagaaaagttctgcagcaaatattcacattgccac	65
p-VfUSP-rev	ggggactgctttttgtacaaacttgcactggctatgaagaaattataatc	66

Шляхом виконання сайт-специфічної рекомбінації (LR-реакції), як було описано вище (див. вище, 1.3), вектори pENTR/A, pENTR/B та вектор pENTR/C, який містив касету селектованого маркера, комбінували з вектором призначення pSUN згідно з інструкціями Multisite Gateway від виробника (Invitrogen, Carlsbad, CA, США). Реакції забезпечували 1 бінарний вектор з промотором p-LuPXR, кодуючою послідовністю люциферази світляка та термінатором t-nos, а також касетою селектованого маркера та 4 вектори, які включали SEQ ID NO1, NO2, NO3 та NO6 безпосередньо перед кодуючою послідовністю люциферази світляка (Таблиця 8), прикладом якої є комбінація з SEQ ID NO1 (SEQ ID NO67). За винятком зміни SEQ ID NO2, NO3 та NO6, нуклеотидна послідовність є ідентичною в усіх векторах (Таблиця 8). Подібним чином промотор p-VfUSP застосовували для одержання послідовності LJK219 лише з промотором, а також послідовностей LJK220, LJK221, LJK224 та LJK225, які містили SEQ ID NO1, NO2, NO3 та NO5 (Таблиця 8). Одержані в результаті вектори трансформації рослин показано у Таблиці 8:

Таблиця 8: Рослинні вектори експресії для трансформації *B. napus*

рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	SEQ ID NO
LJK212	p-LuPXR::c-LUC::t-nos	
LJK213	p-LuPXR::SEQ ID NO1::c-LUC::t-nos	67
LJK214	p-LuPXR::SEQ ID NO2::c-LUC::t-nos	
LJK215	p-LuPXR::SEQ ID NO6::c-LUC::t-nos	
LJK218	p-LuPXR::SEQ ID NO3::c-LUC::t-nos	
LJK219	p-VfUSP::c-LUC::t-nos	
LJK220	p-VfUSP::SEQ ID NO1::c-LUC::t-nos	
LJK221	p-VfUSP::SEQ ID NO2::c-LUC::t-nos	
LJK224	p-VfUSP::SEQ ID NO5::c-LUC::t-nos	
LJK225	p-VfUSP::SEQ ID NO3::c-LUC::t-nos	

4.2 Послідовності NEENA опосередковують посилення тканинно-специфічної експресії генів сильних насінно-специфічних промоторів у насінні олійного рапсу.

Створення трансгенних рослин олійного рапсу та аналізу рослин здійснювали, як описано вище (приклад 3.2 та 3.3).

Для випробування впливу вибраних молекул NEENA у комбінації з насінноспецифічними промоторами (SEQ ID NO1, NO2, NO3, NO5 та NO6) у насінні олійного рапсу насіння на ідентичних стадіях розвитку збирали з окремих ліній трансгенних рослин олійного рапсу, які включали послідовність гена-репортера лише з промотором (LJK212 та LJK219) або послідовності гена-репортера люциферази з вмістом NEENA (SEQ ID NO1, NO2, NO3, NO5 та NO6) (Таблиця 9). З кожної трансгенної події збирали 10 зерен, обробляли й аналізували на активність люциферази, як описано вище (Приклад 3.3).

Порівняно з послідовностями гена-репортера, які містили лише насінно-специфічні промотори p-LuPXR та p-VfUSP без NEENA, усі випробувані молекули NEENA опосередковували значне посилення експресії генів у насінні олійного рапсу середньої зрілості у комбінації з обома промоторами, p-LuPXR та p-VfUSP (Таблиця 9). Подібне посилення експресії виявляли у насінні олійного рапсу на пізніших стадіях визрівання.

Таблиця 9: Експресія LUC у насінні стійко трансформованих рослин олійного рапсу.

Рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	Експресія LUC у насінні олійного рапсу*	
LJK212	p-LuPXR::c-LUC::t-nos	+	20%**
LJK213	p-LuPXR::SEQ ID NO1::c-LUC::t-nos	++++	80%
LJK214	p-LuPXR::SEQ ID NO2::c-LUC::t-nos	++++	80%
LJK215	p-LuPXR::SEQ ID NO6::c-LUC::t-nos	++++	80%
LJK218	p-LuPXR::SEQ ID NO3::c-LUC::t-nos	++++	80%
LJK219	p-VfUSP::c-LUC::t-nos	++	40%
LJK220	p-VfUSP::SEQ ID NO1::c-LUC::t-nos	+++++	100%
LJK221	p-VfUSP::SEQ ID NO2::c-LUC::t-nos	+++++	100%
LJK224	p-VfUSP::SEQ ID NO5::c-LUC::t-nos	++++	80%
LJK225	p-VfUSP::SEQ ID NO3::c-LUC::t-nos	+++++	100%

* Експресія LUC, представлена як показник активності люциферази світляка (- від відсутності експресії до +++++ дуже високої експресії), відносна експресія LUC порівняно з експресією промотора p-LuPXR льону у відповідній тканині.

** Відносна експресія люциферази порівняно з експресією, контрольованою промотором пероксиредоксину льону p-LuPXR.

Для оцінки посилення тканинно-специфічної експресії генів, опосередкованої молекулами NEENA (SEQ ID NO1, NO2, NO3, NO5 та NO6), активність люциферази визначали у повністю розвинутому листі трансгенних рослин олійного рапсу, які включали вищезазначені послідовності гена-репортера. З зразки листа ідентичного розміру збирали з кожної рослини окремо і піддавали аналізам гена-репортера люциферази, як описано вище (Приклад 3.2). Тканинна специфічність випробуваних молекул NEENA (SEQ ID NO1, NO2, NO3, NO5, NO6) у комбінації з промотором p-LuPXR та промотором p-VfUSP є подібною до показників, випробуваних з промотором p-AtPXR, який аналізували раніше (Приклад 3.5). Як і у разі промотора p-AtPXR (Приклад 3.5), молекули NEENA (SEQ ID NO1, NO2, NO3 та NO5) не демонстрували зміни тканинної специфічності промотора p-LuPXR або p-VfUSP (Таблиця 10). Подібно до активності з промотором p-AtPXR (Приклад 3.5), NEENA (SEQ ID NO6) забезпечувала посилення активності люциферази у насінні, але й опосередковувала експресію люциферази у листі підданих аналізу рослин олійного рапсу (Таблиця 10).

Таблиця 10: Експресія LUC у листі стійко трансформованих рослин олійного рапсу.

Рослинний вектор	Склад експресійної касети Промотор::SEQ ID NO::ген-	Експресія LUC у листі олійного
------------------	--	--------------------------------

експресії	репортер::термінатор	рапсу*	
LJK212	p-LuPXR::c-LUC::t-nos	-	0%**
LJK213	p-LuPXR::SEQ ID NO1::c-LUC::t-nos	-	0%
LJK214	p-LuPXR::SEQ ID NO2::c-LUC::t-nos	-	0%
LJK215	p-LuPXR::SEQ ID NO6::c-LUC::t-nos	+	100%
LJK218	p-LuPXR::SEQ ID NO3::c-LUC::t-nos	-	0%
LJK219	p-VfUSP::c-LUC::t-nos	-	0%
LJK220	p-VfUSP::SEQ ID NO1::c-LUC::t-nos	-	0%
LJK221	p-VfUSP::SEQ ID NO2::c-LUC::t-nos	-	0%
LJK224	p-VfUSP::SEQ ID NO5::c-LUC::t-nos	-	0%
LJK225	p-VfUSP::SEQ ID NO3::c-LUC::t-nos	-	0%

* Експресія LUC, представлена як показник активності люциферази світляка (- від відсутності експресії до +++++ дуже високої експресії), відносна експресія LUC порівняно з експресією промотора p-LuPXR льону у відповідній тканині.

** Відносна експресія люциферази порівняно з експресією, контрольованою промотором пероксиредоксину льону p-LuPXR.

Приклад 5: Аналіз посилення тканинно-специфічної експресії генів у рослинах сої.

Цей приклад показує, що заявлені молекули NEENA можуть застосовуватися у багатьох видах рослин і у різних видах з різних родин рослин для посилення експресії генів тканинно-специфічно порівняно з відсутністю NEENA і присутністю лише промотора.

Послідовності молекул NEENA, які опосередковують найбільше посилення експресії генів у попередньому відборі (пор. Приклад 2, SEQ ID NO1, NO2, NO4, NO5, NO6 та NO7), вибирали для визначення посилення рівня експресії генів у трансгенних рослинах сої. Рослинні вектори експресії LJK148, LJK156, LJK157, LJK158, LJK159, LJK160 та LJK161 (пор. приклад 3.1) застосовували для стійкої трансформації сої.

5.1 Одержання трансгенних рослин сої (змінений протокол згідно з документом WO2005/121345; Olhoft et al., 2007).

Пророщування, розмноження насіння сої, приготування *A. rhizogenes* та додаткових експлантатів меристем, та інокуляцію здійснювали, як було описано вище (документ WO2005/121345; Olhoft et al., 2007), за винятком того, що послідовності LJK148, LJK156, LJK157, LJK158, LJK159, LJK160 та LJK161 (пор. Приклад 3.1) містила мутований ген AHAS під дією убіквітинового промотора петрушки PcUbi4-2, що опосередковує толерантність до імідазолінових гербіцидів для відбору.

5.2 Послідовності NEENA опосередковують значне посилення експресії генів у рослинах сої зі збереженням тканинної специфічності промотора.

Зразки тканин збирали з листя, квіток та насіння одержаних трансгенних рослин. Зразки тканин обробляли й аналізували, як описано вище (пор. Приклад 3.3).

Порівняно з насінно-специфічною послідовністю гена-репортера LJK148, що містить лише промотор p-AtPXR без NEENA, усі сім випробуваних молекул NEENA опосередковували значне посилення експресії генів у насінні сої на основі активності люциферази (Фіг. 3а). Натомість у листі та квітках сої не виявлялося значних змін активності люциферази, опосередкованої молекулами NEENA (SEQ ID NO1, NO2, NO4, NO5, NO6 та NO7) (Фіг. 3, b та c).

Приклад 6: Аналіз активності NEENA в однодольних рослинах

У цьому прикладі описується аналіз послідовностей NEENA з SEQ ID NO 1, 2, 3, 4, 5, 6 та 7 в однодольних рослинах.

6.1 Побудова вектора

Для аналізу послідовностей NEENA з SEQ ID NO 1, 2, 3, 4, 5, 6 та 7 в однодольних рослинах застосовують вектор експресії на основі pUC, який містив експресійну касету, яка включала насінно-специфічний промотор p-kg86of однодольних, що не містить NEENA, з *Z. mais*, комбінують з кодуючою послідовністю гена бѳаттлюкуронідази (GUS) з наступним термінатором транскрипції нопалінсинтази (NOS). Геномну ДНК видобувають з зеленої тканини *A. thaliana*, застосовуючи комплект Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Німеччина). Фрагменти геномної ДНК, які містять молекули NEENA, виділяють шляхом традиційної полімеразної ланцюгової реакції (ПЛР). Праймери будують на основі геномної послідовності *A. thaliana* з

багатьма NEENA-кандидатами. Реакція включає 7 наборів праймерів (Таблиця 11) і здійснюється за протоколом, визначеним Phusion High Fidelity DNA Polymerase (Кат. № F-540L, New England Biolabs, Ipswich, MA, США) з застосуванням таких праймерів:

Таблиця 11: Праймерні послідовності

Назва праймера	Послідовність	SEQ ID NO	ПЛР, що забезпечує SEQ ID NO
NEENAss1_forII	aataatggcgcgccctgggtctaaacactctggtagt	68	1
NEENAss1_revII	aataatggcgcgcccttgacctacaaaatcaaagcagtca	69	
NEENAss2_forII	tttttggcgcgccagttctttgcttcgaagtgc	70	2
NEENAss2_revII	tttttggcgcgccctactactgttttcaattct	71	
NEENAss4_forII	aataaaggcgcgccgtccagaattttctccattga	72	6
NEENAss4_revII	aataaaggcgcgcccttctactatccaaagctctca	73	
NEENAss13_forII	ttttatggcgcgccctagcttaatctcagattcgaatcgt	74	7
NEENAss13_revII	ttttatggcgcgccctagtatctacataccaatcatacaaatg	75	
NEENAss14_forII	tttttggcgcgcccttcacgaatttgaatttga	76	5
NEENAss14_revII	tttttggcgcgccctctacaacattaaaacgaccatta	77	
NEENAss15_forII	tatataggcgcgccagggtttcgtttttgtttca	78	3
NEENAss15_revII	tatataggcgcgccctatctcctgctcaaagaaacca	79	
NEENAss18_forII	tatataggcgcgccactgtttaagcttctactgtct	80	4
NEENAss18_revII	tatataggcgcgcccttcttctaagctgaaagt	81	

Ампліфікацію під час ПЛР та очищення продуктів ампліфікації здійснюють, як детально описано вище (приклад 1.2). Після рестрикції ДНК за допомогою рестрикційної ендонуклеази AclI (10 О/мкл) гідролізовані продукти очищують за допомогою комплексу Qiagen Gel Extraction Kit (Qiagen, Hilden, Німеччина).

ПЛР-фрагменти NEENA (див. вище) клонують окремо перед кодуючою послідовністю бета-глюкуронідази з застосуванням сайтів рестрикції AclI. Реакція забезпечує один бінарний вектор з промотором р-KG86, кодуючою послідовністю бета-глюкуронідази с-GUS та термінатором t-pos та 7 векторами, які включали SEQ ID NO1, NO2, NO3, NO4, NO5, NO6 та NO7, безпосередньо перед кодуючою послідовністю бета-глюкуронідази (Таблиця 12), прикладом якої є комбінація з SEQ ID NO1 (SEQ ID NO82). За винятком зміни SEQ ID NO2-NO7, нуклеотидна послідовність є ідентичною в усіх векторах (Таблиця 12). Одержані в результаті вектори показано у Таблиці 12, з промоторними молекулами, які мають префікс р-, кодуючими послідовностями, які мають префікс с-, та термінуючими молекулами, які мають префікс t-.

Таблиця 12: Рослинні вектори експресії

Рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	SEQ ID NO
RTP2933	p-KG86::c-GUS::t-nos	
LJK351	p-KG86::SEQ ID NO1::c-GUS::t-nos	82
LJK352	p-KG86::SEQ ID NO2::c-GUS::t-nos	
LJK353	p-KG86::SEQ ID NO3::c-GUS::t-nos	
LJK354	p-KG86::SEQ ID NO4::c-GUS::t-nos	
LJK355	p-KG86::SEQ ID NO5::c-GUS::t-nos	
LJK356	p-KG86::SEQ ID NO6::c-GUS::t-nos	
LJK357	p-KG86::SEQ ID NO7::c-GUS::t-nos	

Одержані в результаті вектори використовують для аналізу молекул NEENA в описаних нижче експериментах (приклад 6.2).

6.2 Аналіз молекул NEENA, які посилюють експресію генів у тканинах однодольних рослин.

Ці експерименти здійснюють шляхом бомбардування тканин або культивованих клітин однодольної рослини (Приклад 6.2.1) або шляхом опосередкованої *Agrobacterium* трансформації (Приклад 6.2.2). Тканина-мішень для цих експериментів може являти собою рослинні тканини (наприклад, тканина листя), культивовані рослинні клітини (наприклад, кукурудзи Black Mexican Sweetcorn (BMS), або рослинні ембріони для протоколів з застосуванням *Agrobacterium*.

6.2.1 Перехідний аналіз з застосуванням бомбардування мікрочастинками

Плазмідні послідовності виділяють, застосовуючи набір плазмід Qiagen (cat# 12143). ДНК осаджують на частинки золота розміром 0,6 мкМ (Bio-Rad cat# 165-2262) згідно з протоколом, описаним Sanford et al. (1993) (Optimizing the biolistic process for different biological applications. Methods In Enzymology, 217: 483-509), і спрямовують з прискоренням на тканини-мішені (наприклад, двотижневе листя кукурудзи, культивовані клітини BMS і т. ін.) з застосуванням системного пристрою PDS-1000/He (Bio-Rad). Усі етапи осадження ДНК та бомбардування здійснюють у стерильних умовах при кімнатній температурі. Суспензію культивованих клітин кукурудзи Black Mexican Sweet (BMS) поширюють у рідкому середовищі для культури клітин BMS [солі Мурасіге-Скура (MS) (4,3 г/л), 3 % (маса/об'єм) цукроза, міоїнозит (100 мг/л), 3 мг/л 2,4-дихлорофеноксіоцтової кислоти (2,4-D), гідролізат казеїну (1 г/л), тіамін (10 мг/л) та L-пролін (1,15 г/л), pH 5,8]. Щотижня 10 мл стаціонарної клітинної культури переносять у 40 мл свіжого середовища і культивують на ротаційному шейкері, який працює зі швидкістю 110 об/хв при 27 °C, у 250 мл колбі.

60 мг частинок золота у силіконізованій пробірці Еппендорфа ресуспендують у 100 % етанолі з наступним центрифугуванням протягом 30 секунд. Осад промивають один раз у 100 % етанолі і двічі стерильною водою з центрифугуванням після кожного промивання. Остаточний осад ресуспендують в 1 мл стерильного 50 % гліцерину. Потім суспензію частинок золота розділяють на аліквоти по 50 мкл і зберігають при 4 °C. До однієї аліквоти додають такі реагенти: 5 мкл 1 мкг/мкл повної ДНК, 50 мкл 2,5 М CaCl₂, 20 мкл 0,1 М спермідину, у формі вільної основи. Розчин ДНК піддають вихруванню протягом 1 хвилини і тримають при -80 °C протягом 3 хв з наступним центрифугуванням протягом 10 секунд. Супернатант видаляють. Осад обережно ресуспендують в 1 мл 100 % етанолу шляхом постукування по пробірці з наступним центрифугуванням протягом 10 секунд. Супернатант видаляють і осад обережно ресуспендують у 50 мкл 100 % етанолу й тримають при 80 °C до використання (від 30 хв до 4 год. до бомбардування). Якщо у розчині спостерігаються агрегати частинок золота, пробірки обробляють ультразвуком протягом однієї секунди у сонікаторі з водяною ванною безпосередньо перед застосуванням.

Для бомбардування двотижневе кукурудзяне листя нарізають на шматки приблизно по 1 см завдовжки і поміщають адаксіальною стороною догори на середовище, що викликає осмос, M-N6-702 [солі N6 (3,96 г/л), 3 % (маса/об'єм) цукроза, 1,5 мг/л 2,4 дихлорофеноксіоцтової кислоти (2,4-D), гідролізат казеїну (100 мг/л), та L-пролін (2,9 г/л), MS вітамінний основний розчин (1 мг/л), 0,2 М маніту, 0,2 М сорбіту, pH 5,8]. Ці шматки інкубують протягом 1-2 годин.

У разі культивованих клітин BMS однотижневу суспензію клітин центрифугують при 1000 g у центрифугу Beckman/Coulter Avanti J25 і надосадову рідину зливають. Клітини поміщають на круглі беззольні фільтри Whatman № 42 шаром 1/16 дюйма завтовшки за допомогою лопатки. Фільтрувальний папір з рослинним матеріалом поміщають на середовище, що викликає осмос, при 27 °C у темряві на 1-2 години перед бомбардуванням. Безпосередньо перед бомбардуванням фільтри виймають із середовища і поміщають на кілька шарів стерильного фільтрувального паперу для часткового висушування поверхні.

Кожну пластинку двічі обстрілюють 6 мкл розчину золота-ДНК при 1800 psi для листяних матеріалів і при 1100 psi для культивованих клітин BMS. Для збереження позиції рослинних матеріалів на зразок згори укладають стерилізоване дротяне сито. Після бомбардування фільтри зі зразками переносять на середовище M-N6-702 без маніту та сорбіту й інкубують протягом 2 днів у темряві при 27 °C перед перехідними аналізами.

Тимчасову трансформацію інших однодольних рослин через бомбардування мікрочастинками здійснюють, застосовуючи, наприклад, спосіб, описаний у Wang et al., 1988 (Transient expression of foreign genes in rice, wheat and soybean cells following particle bombardment, *Plant Molecular Biology*, 11 (4), 433-439), Christou, 1997 (Rice transformation: bombardment, *Plant Mol Biol.* 35 (1-2)).

Рівень експресії експресованих генів в описаних вище послідовностях (приклад 6.1) визначають шляхом GUS-забарвлення, кількісного визначення люмінесценції/флуоресценції, ЗТ-ПЛР, protein abundance (виявлення специфічних антитіл) або метаболічних продуктів, одержаних через описані вище експресійні касети з застосуванням відомих спеціалістам протоколів. GUS-забарвлення здійснюють шляхом інкубації рослинного матеріалу у розчині GUS [100 mM NaHPO₄, 10 mM EDTA, 0,05 % Triton X100, 0,025 % розчину X-Gluc (5-бромо-4-хлоро-3-індоліл-бета-D-глюкуронова кислота, розчинена у DMSO), 10 % метанол, pH 7,0] при 37 °C протягом 16-24 годин. Рослинні тканини двічі піддають вакуум-інфільтрації протягом 15 хвилин для сприяння рівномірному забарвленню. Аналізи активності люциферази здійснюють, як описано вище (див. Приклади 2 та 3.3).

Порівняно з послідовностями гена-репортера, які містять лише насінноспецифічний промотор p-ZmKG86 без NEENA, усі молекули NEENA опосередковують значне посилення експресії генів у цих аналізах.

6.2.2 Трансформація та регенерація однодольних культурних рослин

Опосередкована *Agrobacterium* трансформація рослин з застосуванням стандартних способів трансформації та регенерації також може здійснюватися з метою трансфермації культурних рослин (Gelvin and Schilperoort, 1995, *Plant Molecular Biology Manual*, 2nd Edition, Dordrecht: Kluwer Academic Publ. ISBN 0-7923-2731-4; Click and Thompson (1993) *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton: CRC Press, ISBN 0-84935164-2). Трансформація кукурудзи або інших однодольних рослин може здійснюватися з застосуванням, наприклад, способу, описаного в документі US 5,591,616. Трансформація рослин з застосуванням бомбардування частинками, опосередкованого поліетиленгліколем поглинання ДНК або через застосування волокон карбонату кремнію описується, наприклад, у публікації Freeling & Walbot (1993) "The maize handbook" ISBN 3-540-97826-7, Springer Verlag New York).

Рівень експресії експресованих генів визначають шляхом GUS-забарвлення, кількісного визначення люмінесценції або флуоресценції, ЗТ-ПЛР або protein abundance (виявлення специфічних антитіл) з застосуванням відомих спеціалістам протоколів. GUS-забарвлення здійснюють шляхом інкубації рослинного матеріалу у розчині GUS [100 mM NaHPO₄, 10 mM EDTA, 0,05 % Triton X100, 0,025 % розчину X-Gluc (5-бромо-4-хлоро-3-індоліл-бета-D-глюкуронова кислота, розчинена у DMSO), 10 % метанол, pH 7,0] при 37 °C протягом 16-24 годин. Рослинні тканини двічі піддають вакуум-інфільтрації протягом 15 хвилин для сприяння рівномірному забарвленню. Аналізи активності люциферази здійснюють, як описано вище (Приклади 2 та 3.3).

Порівняно з послідовностями гена-репортера, які містять лише насінно-специфічний промотор p-ZmKG86 без NEENA, молекули NEENA опосередковують сильне тканинно-специфічне підвищення експресії генів у рослинах.

Приклад 7: Кількісний аналіз активності NEENA в однодольних рослинах У цьому прикладі описується аналіз послідовностей NEENA з SEQ ID NO 1 та 2 у рослинах кукурудзи.

7.1 Побудова вектора

Для кількісного аналізу послідовностей NEENA з SEQ ID NO 1 та 2 в однодольних рослинах, вектор експресії на основі pUC, який містив експресійну касету, яка включала насінно-специфічний промотор p-KG86 однодольних, що не містить NEENA, з Z, mais, комбінували з кодувальною послідовністю гена люциферази світляка (LUC) (Promega, Madison, WI, США) з

наступним термінатором транскрипції нопалінсинтази (NOS). Геномну ДНК видобували з зеленої тканини *A. thaliana*, застосовуючи комплект Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Німеччина). Фрагменти геномної ДНК, які містять молекули NEENA, виділяли шляхом традиційної полімеразної ланцюгової реакції (ПЛР). Праймери будували на основі геномної послідовності *A. thaliana* з багатьма NEENA-кандидатами. Реакція включала 2 наборів праймерів (Таблиця 11), і її здійснювали за протоколом Phusion High Fidelity DNA Polymerase (Кат. № F-540L, New England Biolabs, Ipswich, MA, США) з застосуванням таких праймерів:

Таблиця 11: Праймерні послідовності

Назва праймера	Послідовність	SEQ ID No	ПЛР, що забезпечує SEQ ID NO
NEENAss1_forIII	atatacgcgtggtgcttaaacactctggtgagt	83	1
NEENAss1_revIII	atatggcgcgcctttgacctacaaaatcaaagcagtc	84	
NEENAss2_forIII	atatacgcgtagttctttgcttgaagtgc	85	2
NEENAss2_revIII	atatggcgcgcctactacgtactgttttcaattct	86	

Ампліфікацію під час ПЛР та очищення продуктів ампліфікації здійснювали, як детально описано вище (приклад 1.2). Після рестрикції ДНК за допомогою рестрикційних ендонуклеаз ДНК (10 О/мкл) та *Ascl* (10 О/мкл) гідролізовані продукти очищали за допомогою комплексу Qiagen Gel Extraction Kit (Qiagen, Hilden, Німеччина).

ПЛР фрагменти NEENA (див. вище) клонували окремо перед кодуючою послідовністю люциферази світляка з застосуванням сайтів рестрикції *Ascl*. Реакція забезпечувала один бінарний вектор з промотором p-KG86, кодуючою послідовністю люциферази світляка C-LUC та термінатором t-nos і два вектори, які включали SEQ ID NO1 та NO2, безпосередньо перед кодуючою послідовністю люциферази світляка (Таблиця 12), прикладом якої є комбінація з SEQ ID NO1 (SEQ ID NO87). За винятком зміни SEQ ID NO2, нуклеотидна послідовність у векторах є ідентичною (Таблиця 12). Одержані в результаті вектори показано у Таблиці 12, з промоторними молекулами, які мають префікс p-, кодуючими послідовностями, які мають префікс c-, та термінуючими молекулами, які мають префікс t-.

Таблиця 12: Рослинні вектори експресії

Рослинний вектор експресії	Склад експресійної касети	SEQ ID NO
	Промотор::SEQ ID NO::ген-репортер::термінатор	
RTP5679	p-KG86::c-LUC::t-nos	
RTP5683	p-KG86::SEQ ID NO1::c-LUC::t-nos	87
RTP5684	p-KG86::SEQ ID NO2::c-LUC::t-nos	

Одержані в результаті вектори застосовували для аналізу молекул NEENA в описаних нижче експериментах (Приклад 7.2).

7.2 Одержання трансгенних рослин кукурудзи

Пророщування, розмноження кукурудзи, приготування *A. tumefaciens* та інокуляцію здійснювали, як було описано вище (документи WO2006136596, US20090249514), за винятком того, що послідовності RTP5679, RTP5683 та RTP5684 (пор. приклад 7.1) містила мутований ген *ANAS* під дією убіквітинового промотора p-Ubi кукурудзи, що опосередковує толерантність до імідазолінових гербіцидів для відбору.

7.3 Послідовності NEENA опосередковують сильне тканинно-специфічне підвищення експресії генів у рослинах кукурудзи

Зразки тканин збирали з листя та зерен одержаних трансгенних рослин. Зразки тканин обробляли й аналізували, як описано вище (пор. приклад 3.3)

Порівняно з послідовністю гена-репортера, яка включає лише насінно-специфічний промотор P-KG86 без NEENA, дві випробувані молекули NEENA (SEQ ID NO1 та NO2) опосередковували значне посилення експресії генів у зернах (Фіг. 4а). Натомість у листі

кукурудзи не виявляли значних змін активності люциферази, опосередкованої молекулами NEENA (SEQ ID NO1 та NO2) (Фіг. 4b).

Приклад 8: Кількісний аналіз активності NEENA у рослинах рису

У цьому прикладі описується аналіз послідовностей NEENA з SEQ ID NO 1 та 2 у рослинах рису.

8.1 Побудова вектора

Для кількісного аналізу послідовностей NEENA з SEQ ID NO 1 та 2 у рослинах рису вектори рENTR/B LJK1, UK19 та LJK20 (пор. з прикладом 1.3) комбінували з вектором призначення, який включав насінно-селективний промотор рису PRO0090 перед сайтом рекомбінації, застосовуючи сайт-специфічну рекомбінацію (LR-реакцію) згідно з інструкціями виробника (Invitrogen, Carlsbad, CA, США). Реакції забезпечували один бінарний вектор з промотором PRO0090, кодуючою послідовністю люциферази світляка c-LUC та термінатором t-nos, а також 2 вектори, які включали SEQ ID NO1 та NO2 безпосередньо перед кодуючою послідовністю люциферази світляка (Таблиця 13). За винятком зміни SEQ ID NO2, нуклеотидна послідовність у векторах є ідентичною (Таблиця 13). Одержані в результаті вектори показано у Таблиці 13, з промоторними молекулами, які мають префікс р, кодуючими послідовностями, які мають префікс с-, та термінуючими молекулами, які мають префікс t-.

Таблиця 13: Рослинні вектори експресії

рослинний вектор експресії	Склад експресійної касети Промотор::SEQ ID NO::ген-репортер::термінатор	SEQ ID NO
CD30977	p- PRO0090::c-LUC::t-nos	
CD30971	p- PRO0090::SEQ ID NO1::c-LUC::t-nos	-
CD30972	p- PRO0090::SEQ ID NO2::c-LUC::t-nos	-

Одержані в результаті вектори застосовували для аналізу молекул NEENA в описаних нижче експериментах (Приклад 8.2).

8.2 Одержання трансгенних рослин рису

Agrobacterium, що містить відповідний вектор експресії, застосовували для трансформації рослин *Oryza sativa*. Зрілі сухі зерна рису японського сорту Nipponbare вилущували. Стерилізацію здійснювали шляхом інкубації протягом однієї хвилини у 70 % етанолу, а потім 30 хвилин у 0,2 % HgCl_2 , з наступним шестиразовим промиванням по 15 хвилин стерильною дистильованою водою. Стерильне насіння потім пророщували на середовищі, яке містило 2,4-D (середовище для калюсогенезу). Після інкубації у темряві протягом чотирьох тижнів ембріогенні калюси зі щитка вирізали й розмножували на тому самому середовищі. Через два тижні калюси розмножували або поширювали шляхом субкультивування на тому самому середовищі ще протягом 2 тижнів. Фрагменти ембріогенного калюсу субкультивували на свіжому середовищі за 3 дні до спільного культивування (для стимулювання ділення клітин).

Штам *Agrobacterium* LBA4404, який містив відповідний вектор експресії, застосовували для спільного культивування. *Agrobacterium* інокулювали на АВ-середовищі з відповідними антибіотиками і культивували протягом 3 днів при 28 °C. Потім бактерії збирали й суспендували у рідкому середовищі для спільного культивування до густини (OD_{600}) приблизно 1. Потім суспензію переносили у чашку Петрі і калюси занурювали у суспензію на 15 хвилин. Після цього тканини калюсу наносили для видалення рідини на фільтрувальний папір переносили на затвердле середовище для спільного культивування й інкубували протягом 3 днів у темряві при 25 °C. Спільно культивовані калюси вирощували на середовищі, яке містило 2,4-D, протягом 4 тижнів у темряві при 28 °C у присутності агента для відбору. Протягом цього періоду розвивалися швидко зростаючі острівці резистентного калюсу. Після перенесення цього матеріалу на середовище регенерації та інкубації при світлі виявлявся ембріогенний потенціал, і паростки розвивалися протягом наступних чотирьох-п'яти тижнів. Паростки вирізали з калюсу й інкубували протягом 2-3 тижнів на середовищі, яке містило ауксин, з якого їх переносили у ґрунт. Загартовані паростки вирощували в умовах високої вологості у режимі короткого дня в оранжереї.

Для однієї послідовності було одержано приблизно 35 незалежних трансформантів рису T0. Первинні трансформанти переносили з камери для культивування тканини до оранжереї. Після кількісного ПЛР-аналізу для перевірки кількості копій вставки Т-ДНК лише одиночні трансгенні рослини, які виявляли толерантність до агента для відбору, зберігали для збирання насіння T1. Потім насіння збирали через три-п'ять місяців після розсаджування. Цей спосіб забезпечував

показник трансформантів з єдиним локусом понад 50 % (Aldemita and Hodges 1996, Chan et al. 1993, Hiei et al. 1994).

8.3 Послідовності NEENA опосередковують сильне тканинно-специфічне підвищення експресії генів у рослинах рису

5 Зразки тканин збирали з листя та насіння одержаних трансгенних рослин. Зразки тканин обробляли й аналізували, як описано вище (пор. приклад 3.3)

Порівняно з послідовністю гена-репортера, яка включає лише насінно-специфічний промотор PRO0090 без NEENA, випробувані молекули NEENA (SEQ ID NO1 та NO2) опосередковували значне посилення експресії генів у насінні (Фіг. 5a). Натомість у листі рису не виявляли значних змін активності люциферази, опосередкованої молекулами NEENA (SEQ ID NO1 та NO2) (Фіг. 5b).

Пояснення до фігур:

15 Фіг. 1: Аналіз експресії генів-репортерів люциферази у сім'ядолях стійко трансформованих рослин *A. thaliana* послідовностей, що не містять NEENA (LJK134) і містять NEENA (LJK71-LJK90), які представляють гадані молекули NEENA, які походять з насінно-селективних експресованих генів під контролем промотора p-AtPXR. Значення експресії показано відносно контрольної послідовності, що не містить NEENA (LJK134 = 1).

20 Фіг. 2: Стовпчасті діаграми активності гена-репортера люциферази, показаної як відносні світлові одиниці (RLU) незалежних ліній трансгенних рослин олійного рапсу, які включають послідовності гена-репортера, що не містять NEENA (LJK148) або містять NEENA, які представляють молекули NEENA з насінно-селективних експресованих генів (LJK156-LJK162) під контролем промотора p-AtPXR і після нормалізації стосовно вмісту білка у кожному зразку. Значення експресії рослин, які включають послідовності, що містять NEENA, показано стосовно рослин, які експресують контрольну послідовність, що не містить NEENA (LJK148) (усереднені значення, аналізовані тканини 20 незалежних трансгенних рослин). А) насіння, В) тканина

25 листя, С) квітки.
Фіг. 3: Стовпчасті діаграми активності гена-репортера люциферази, показаної як відносні світлові одиниці (RLU) незалежних ліній трансгенних рослин сої які включають послідовності гена-репортера, що не містять NEENA (LJK148) або містять NEENA які представляють молекули NEENA з насінно-селективних експресованих генів (LJK156 - LJK161) під контролем промотора p-AtPXR і після нормалізації стосовно вмісту білка у кожному зразку. Значення експресії рослин, які включають послідовності, що містять NEENA, показано стосовно рослин, які експресують контрольну послідовність, що не містить NEENA (LJK148) (усереднені значення, аналізовані тканини 10 незалежних трансгенних рослин). А) насіння В) тканина листя, С) квітки.

35 Фіг. 4: Стовпчаста діаграма активності гена-репортера люциферази, показаної як відносні світлові одиниці (RLU) незалежних ліній трансгенних рослин кукурудзи, які включають послідовності гена-репортера, що не містять NEENA (RTP5679) або містять NEENA, які представляють молекули NEENA з насінно-селективних експресованих генів (RTP5683-RTP5684) під контролем промотора p-KG86 і після нормалізації стосовно вмісту білка у кожному зразку (усереднені значення, аналізовані тканини 15 незалежних трансгенних рослин). А) зерна. В) тканина листя.

40 Фіг. 5: Стовпчаста діаграма активності гена-репортера люциферази, показаної як відносні світлові одиниці (RLU) незалежних ліній трансгенних рослин рису, які включають послідовність гена-репортера, що не містить NEENA (CD30977) або містить NEENA, які представляють молекулу NEENA з насінно-селективних експресованих генів (CD30971 - CD3G972) під контролем промотора PRO0090 рису і після нормалізації стосовно вмісту білка у кожному зразку (усереднені значення, аналізовані тканини 15 незалежних трансгенних рослин). А) насіння, В) тканина листя.

ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<110> Басф Планта Саенс Компані ГМБХ

<120> Регуляторні молекули нуклеїнових кислот для посилення насінно-специфічної та/або насінно-селективної експресії генів у рослинах

<130> PF 62516

<160> 87

<170> PatentIn, версія 3.5

<210> 1

<211> 847

<212> ДНК

<213> *Arabidopsis thaliana*

<400> 1

tggtgcttaa acactctggt gagttctagt acttctgcta tgatcgatct cattaccatt	60
tcttaaattt ctctccctaa atattccgag ttcttgattt ttgataactt cagggtttct	120
ctttttgata aatctggctt ttccattttt ttttttttgt ggtaattta gtttcctatg	180
ttcttcgatt gtattatgca tgatctgtgt ttggattctg ttagattatg tattggtgaa	240
tatgtatgtg tttttgcatg tctggttttg gtcttaaaaa tgttcaaact tgatgatttg	300
attgaagctt ttttagtggt ggtttgattc ttctcaaaac tactgttaat ttactatcat	360
gttttccaac ttgattcat gatgacactt ttgttctgct ttgttataaa attttggttg	420
gtttgatttt gtaattatag tgtaattttg ttaggaatga acatgtttta atactctgtt	480
ttcgatttgt cacacattcg aattattaat cgataattta actgaaaatt catggttcta	540
gatcttggtg tcatcagatt atttgtttcg ataattcacc aaatatgtag tccttttgc	600
gatttgcgac tgtttcattt ttctcaaaa ttgttttttg ttaagtttat ctaacagtta	660
tcgttgctca aagtctcttt cattttgcaa aatcttcttt ttttttttgt ttgtaacttt	720
gttttttaag ctacacattt agtctgtaaa atagcatcga ggaacagttg tcttagtaga	780
cttgcatgtt cttgtaactt ctatttggtt cagtttggtg atgactgctt tgattttgta	840
ggtcaaa	847

<210> 2

<211> 455

<212> ДНК

<213> *Arabidopsis thaliana*

<400> 2

agttctttgc tttcgaagtt gccgcaacct aaacaggttt ttctttcttc tttcttctta	60
ttaactacga ccttgtcctt tgcctatgta aaattactag gttttcatca gttacactga	120
ttaagttcgt tatagtggaa gataaaatgc cctcaaagca ttttgcagga tatctttgat	180
ttttcaaaga tatggaactg tagagtttga tagtgttctt gaatgtggtt gcatgaagtt	240

```

tttttggctc gcatgttatt ttttctcga aatatgtttt gagtccaaca agtgattcac      300
ttgggattca gaaagttgtt ttctcaatat gtaacagttt ttttctatgg agaaaaatca      360
tagggaccgt tgggttttggc ttctttaatt ttgagctcag attaaacca ttttaccggg      420
tgttcttggc agaattgaaa acagtacgta gtacc                                455

```

```

<210> 3
<211> 758
<212> ДНК
<213> Arabidopsis thaliana

```

```

<400> 3
aggggttcgt ttttgtttca tcgataaact caaagggtgat gatttttaggg tcttgtgagt      60
gtgctttttt gtttgattct actgtagggt ttatgttctt tagctcatag gttttgtgta      120
ttctcttagaa atgtggcttc tttaatctct ggggtttgtga ctttttgtgt ggtttctgtg      180
tttttcatat caaaaacctt ttttttcga gttttttttt acaaattctt actctcaagc      240
ttgaatactt cacatgcagt gttcttttgt agattttaga gttaatgtgt taaaaagttt      300
ggatttttct tgettataga gcttcttcac ttgattttg tgggtttttt tgttttaag      360
gtgagatttt tgatgagggt ttgcttcaa agatgtcacc tttctgggtt tgtcttttga      420
ataaagctat gaactgtcac atggctgacg caattttgtt actatgtcat gaaagctgac      480
gtttttccgt gttatacatg ttgcttaca cttgcacgcg tcaaaaaaat tggggctttt      540
tagttttagt caaagatttt acttctcttt tgggatttat gaaggaaagt tgcaaacttt      600
ctcaaatttt accatttttg ctttgatgtt tgttttagatt gcgacagAAC aaactcatat      660
atgttgaaat ttttcttgg ttttgtatag gattgtgtct tttgcttata aatgttgaaa      720
tctgaacttt tttttgttt ggtttctttg agcaggag                                758

```

```

<210> 4
<211> 252
<212> ДНК
<213> Arabidopsis thaliana

```

```

<400> 4
actgtttaag cttcactgtc tctgaatcgg caaaggtaaa cgtatcaatt attctacaaa      60
cccttttatt tttcttttga attaccgtct tcattgggtt tatgataact tgataagtaa      120
agcttcaata attgaatttg atctgtgttt ttttggcctt aataactaat ccttacataa      180
gctttgttgc ttctctctct gtgagttgag tggttaagttg taataatggt tcactttcag      240
ctttagaaga aa                                                                252

```

```

<210> 5
<211> 377
<212> ДНК
<213> Arabidopsis thaliana

```

<400> 5
 ttccacgatt tgggaatttga ttccctcgat cacaggatat acaggttaga ttttgttttg 60
 tatagtttga tacatacttc tttgtgatgt tttgtttact taatcgaatt tttggagtgt 120
 ttttaaggctct ctcgttttaga aatcgtggaa aatatcactg tgtgtgtgtt cttatgattc 180
 acagtgttta tgggtttcat gttctttgtt ttatcattga atgggaagaa atttcgttgg 240
 gatacaaatt tctcatgttc ttactgatcg ttattaggag tttggggaaa aaggaagagt 300
 ttttttggtt ggttcgagt attatgaggt tttttctgta tttgatttat gagttaatgg 360
 tcgttttaaat gttgtag 377

<210> 6
 <211> 718
 <212> ДНК
 <213> *Arabidopsis thaliana*

<400> 6
 gtccagaatt ttctccattg aagctggatt ctaaggtcag ttcttaectc ttatctctaa 60
 tctgatgatt ccatatcgaa agtcttactt ttccacttca atttcaatct gatgattcta 120
 agatctttga ttcgaggctg atctctgata gttactacat gtttctgggt ttattttattt 180
 ttaatccata tagtaattaa aaactcttat gaggtttaat tatggttact tgagaatttg 240
 caatcgteat ctttctttga ctctatcca ttttttggtt ttctcttctg ttaattctg 300
 ttccataaatt gtaattgtaa attaaccaaa acaaatgat cagaaacctt ttctctatgg 360
 aatatttctc acacgcaagc ctgtgagttg tgactctgta atcacttcct tgttctggta 420
 atttcagtgg ttaaggctct ctttttttct gatgttgta gcaaaagtta gtttttcttc 480
 ttctttaatg ggttaattac acctaaatct ctggttatta aacaatccag aaagaaaaaa 540
 agtttattcc ttctctatg tatatagttt cacatgcaag catcacttgt ttgtctgac 600
 aaattgcaga gttttgagtt ctgttttttt ttttttctaa tgttttgtct ttaagaaagt 660
 tctgtttttt ttctctcagg aaagtatca aaagtttga gagctttgga tagtgaag 718

<210> 7
 <211> 495
 <212> ДНК
 <213> *Arabidopsis thaliana*

<400> 7
 ctagcttaat ctccagattcg aatcgttcca tagtggtag ctctgtgttc ttctttctgc 60
 tcttactcct gattctcgat tttaggggtt tcagtaattg cgtcggcggc gaaagtcttt 120
 atcgccgac gatcttctt atctagaaat tattgatcag aaactgttgg gttttgtttg 180
 attcttgtca agttttgatt ttctatgca aattgtctaa tcccaattca aagttacgat 240
 ttttattgaa aacctagat tggtttcttc aagtttgtca ctttgattca atctaatagc 300
 ttagcttaat cgttaagtct ctttttttgt tttaggttct atttgcgatt taaaggttct 360

tggttttgga ttgtttttgc ttggttcctt taagtttgag aggccttatgt agattataag 420
agagaagagt attgctttgc atgttttaaag gaagaacttt taactgaaca ttgtatgat 480
tggtatgtag atact 495

<210> 8
<211> 139
<212> ДНК
<213> Arabidopsis thaliana

<400> 8
atttccacac gctttctatc atttccaccc aaaaggtaac gcgcttttta ttctctttcc 60
tgcattcata aatttgcttc ctgcatgttg aaaaaaaaaa atttcatcg agattcgttt 120
ttatttttta gagagagat 139

<210> 9
<211> 889
<212> ДНК
<213> Arabidopsis thaliana

<400> 9
gtctactttc attacagtga ctctgcatgc ttcaggcttc gtctaattct tgaattctct 60
tcttttctgt tccgtaattt actttctagg gtctctagat ttgtgtctcc tctaacaaaa 120
gactctatct ttgcacaaat ttaatttcat cattgacett tgtcgattcc attctctctc 180
tatctctctg tttcttcgaa aacctagagg ttttgaattt aatgattcct ttttatgtca 240
ataaatttgc aatcaatggg agctttttta aatcatcgtt atatctataa acaaaaaaac 300
agtaattact cttcttagat ctaaaaaat taataaatct ttcccttttt tctcatcata 360
attttttctg atttaactct tgtaaaaatt tgcttagccg ttctgctttc tcaggcccca 420
ggtgattcgt gtcttctagg tcagcttggt aaacctgaga gaagccatct ttgttttgcg 480
gttacaaact ttgccgcttc aatatttcat tgctgttttc tgggaaaacc tttttctagt 540
tttttcggct tattatgcct ttttaacttt ttgtcattta acatttattg ttagtgcttt 600
gcttagtgta aagtagtagt tctctttgta atattaccat aagggtcaga agtaaatatt 660
tctaaaattg ttttcttggt ggaaattcag actgatttca gcaacatgca tgggcttaaa 720
atcagcttct aagactgaga tttagtgage agtgtggtgg tgtcttgctc tctgttcttg 780
ggagaacaca aaggcagtgt ggagtctggt gaggtttctg attcttgaaa agatttataa 840
attttcttgc aaaattagtc tttatgttga attgtgttgc aggtaaaaat 889

<210> 10
<211> 433
<212> ДНК
<213> Arabidopsis thaliana
<400> 10

```

gcacaatctt agcttacctt gaatcacaa ctcagggtata tgtaactgat tctaaattga      60
agatttgtgtg caaatcttat atccattttt tattattaaa tttattgaaa aagctagcgg      120
tgtaaattaa tgtcacaaaa tcagtatat tttagttttt gttttttttg aagtttttatg      180
caaatcttca aaaagtatat tcagtgtgtg aattgacaaa tagagactct agttcttttt      240
ttttttttct tttttttaac atctgactct tataagagact ctagtctcatg tacacttttt      300
ttaatggaaa aacaaatttg aaactgaata tcttatttcc acgtagattg tatattagtt      360
taatttgatt gttatatattg taaatgtcta ctaaacagga attggatggg gaggaggcaa      420
ggcttgttga tta                                                                433

```

```

<210> 11
<211> 354
<212> ДНК
<213> Arabidopsis thaliana

```

```

<400> 11
atcttagggg ttgcgagat ctcaactctca ctgggtatgtc tgtgttttctt ctccatttt      60
ctgtttctat tggaaacttc tctctccaat ttctgtttctt tcacttcttt gatccttttag      120
ctttgacaaa accgtagtaa aggatcaaaa gttatcatct ttgggtccatg ttgtgaatcg      180
tgctctgctt gggtcgtgac tcccaaatcc ggatttgaaa ccagcatatc tgagcttaat      240
tcgagcatgc atgcgcttct tttttctga ttttttttag actttgggtc taaatccctt      300
aactttggat taactgtcaa tctacaattt tatattaaca gagatagctt agca          354

```

```

<210> 12
<211> 143
<212> ДНК
<213> Arabidopsis thaliana

```

```

<400> 12
cagaagctca tttcttcgat acgatcaacc attagggtgat tttttctctt gatcttcgag      60
ttctgataat tgctcttttt tctctggctt tggtatcgat aatttctctg gattttcttt      120
ctggggtgaa tttttgcgca gag                                                                143

```

```

<210> 13
<211> 182
<212> ДНК
<213> Arabidopsis thaliana

```

```

<400> 13
atttttgttg gtgaaaggta gaattcgtaa atttctctctg ctcaactttat tgtttcgact      60
catacccgat aatctcttct atgtttggta gagatatctt ctcaaagtct tatctttcct      120
taccgtgttc tgtgtttttt gatgatttag gtgaagaaga agaagcagag acaaaaaacga      180
tt                                                                182

```

```

<210> 14
<211> 665
<212> ДНК
<213> Arabidopsis thaliana

<400> 14
ttaagctttt aagaatctct actcacattt tctctgtgag tgttctttta tacttctttg      60
ttatttccaa tttttcttct tttctctctaa aaatttttagg aactattgaa tcatttaatt      120
tctgttttgt gataaaattt ccatcaactg ttctcggcct accgatgcat tttttgtaaa      180
accgtctttt ttgtgtgaat aaaattttta attcatacaa aaaaaaaca tatttgatac      240
tatttttagct ccattgtatc tgaatcttca ttgtttaatt tttttgttct ctctgttctc      300
acttgaattt tggaatat ttctctaggt ttacctata ttcttcactt taagaactat      360
atgaagattt gattggaagt aataatattc ggtgatagaa tctgagtttg ttgtattctg      420
gtgtggggct tatactaac tttttcttct gtaccaatac attttcaatt ttacattttt      480
gattagctta aaatgtgaag gataccttgt aaataactat tacactattg cttgtcttag      540
tctaatagtc ttcactaata ttttgtgcag tagaagtaaa tattataaag agttgttgtt      600
tgattataga gagttgttgt ctattcttta acttgatgtg atgttgtttt tgatgacagg      660
taaaa                                             665

<210> 15
<211> 252
<212> ДНК
<213> Arabidopsis thaliana

<400> 15
tctgggaat atcgattttg atctattaag agctgggtgag agccaaagtt tcctttttgt      60
ttgtttgttt gtttgtttgt tgtttgtatt ttgtatctc tgtgatcgtc tctacgtgtt      120
gggtcatgca gagaaactca ttttgttttg atttgcaatg tgtcaattcc actttgaaat      180
ataagattca tcgcctctct ctcttttgtt ttttttcttc ttctgcagct acgagctttg      240
ggatgtggtg ag                                             252

<210> 16
<211> 186
<212> ДНК
<213> Arabidopsis thaliana

<400> 16
tattcacaat ctctgccac ctctcatttc tctagttgag ttgttatctg cgttttttaag      60
cactcgaata ctgcatgcaa attccctgat tgtttgttag taccttagag attctcgatt      120
ttttagttgt ttagattgaa ccaggattac taaattgtta ttgttttctg tgtaaaggct      180
acatat                                             186

```


<210> 17
 <211> 345
 <212> ДНК
 <213> *Arabidopsis thaliana*

<400> 17
 ctttgcagct tctgcagcac ctctccctac tccagggtact tatgtttttg ataattttat 60
 tgatagactc tttacaatta tacttaagct tggtactttt tattggtacc aacaaaagct 120
 aatgtatagt tcataactca caggtcctgc gtctttcggg cggaccactt ctctacaga 180
 ttgcgaaact tctgatcctg aagggtactcg cgaacttttt actgcaactt ctagtcttaa 240
 ctccaaaaca ttttgttcag aatttgtttc taaaagattt tggggtttgt tgacgtcaca 300
 taactgcgag ggtctgcttc tttcgtccg cccacttctc cgaca 345

<210> 18
 <211> 285
 <212> ДНК
 <213> *Arabidopsis thaliana*

<400> 18
 aacaactatg gcttgagggt aacaagagta tcagggtatat gtgaaaactc tacttttgaa 60
 gtttaccaaa aaaaatactc tacttttgga aagacattgc tcctaaaatc ttattagtgt 120
 tatataattt actaaaacac atagtctctg aattcttggt aatgagcatg ttaccttgga 180
 caagtgacct tttttctaca ttttgttttt ctatcacacg tcatgctgtt tgattgtttc 240
 cttacgagtt ttaattttat ttttggtta aaaacagtaa gataa 285

<210> 19
 <211> 137
 <212> ДНК
 <213> *Arabidopsis thaliana*

<400> 19
 tctaaaaata cagggcaccg aaccaataa aggtgagaat gatgagaagc cgtttcttac 60
 tcttcattgt tttctctct ctatccctct tcatttctc tctgatcgcc agtgatttag 120
 gcttctgcaa cgaagag 137

<210> 20
 <211> 36
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 20
 aataatggtc cctggtgctt aaacactctg gtgagt 36

<210> 21
 <211> 38
 <212> ДНК

<213> штучна

<220>
<223> синтетична послідовність

<400> 21
aataatccat ggtttgacct acaaaatcaa agcagtc 38

<210> 22
<211> 34
<212> ДНК

<213> штучна

<220>
<223> синтетична послідовність

<400> 22
ttttttggta ccagttctttt gctttcgaag ttgc 34

<210> 23
<211> 35
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 23
ttttttccat ggtactacgt actgttttca attct 35

<210> 24
<211> 37
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 24
aaaaaaggta ccatttccac acgcttttcta tcatttc 37

<210> 25
<211> 42
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 25
aaaaaacccat ggttatctct ctctaaaaa taanaacgaa tc 42

<210> 26
<211> 33
<212> ДНК
<213> штучна

<220>
 <223> синтетична послідовність

<400> 26
 aataaaggta ccgccagaa tttctccat tga 33

<210> 27
 <211> 34
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 27
 aataaaccat ggtcttcaat atccaaagct ctca 34

<210> 28
 <211> 37
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 28
 ttttttgga cgtctactt tcattacagt gactctg 37

<210> 29
 <211> 39
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 29
 tttttccat gggtatattt tacctgcaac acaattcaa 39

<210> 30
 <211> 32
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 30
 ttttatgga cccactcga tactgcatgc aa 32

<210> 31
 <211> 38
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 31

ttttatccat gggtatgtag cctttacaca gaaaacaa 38

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

<220>
<223> синтетична послідовність

<400> 32
tatataggta ccaacaacta tggcctgagg gt 32

<210> 33
<211> 44
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 33
tatataccat gggtatctta ctgtttttaa ccaaaaaata aaat 44

<210> 34
<211> 36
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 34
tttttaggta ccatcttagg gtttcgag atctca 36

<210> 35
<211> 42
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 35
tttttccat ggtgctaagc tatctctgtt aatataaaat tg 42

<210> 36
<211> 34
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 36
ttttttggta ccatttttgt tggtgaaagg taga 34

<210> 37

```

<211> 37
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 37
tttttaccat ggttacgttt ttgtctctgc ttcttct

37

<210> 38
<211> 36
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 38
tatattggta cctctgggaa atatcgattt tgatct

36

<210> 39
<211> 33
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 39
tatataccat ggtctcacca catcccaaaag etc

33

<210> 40
<211> 35
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 40
ttttatggta ccgcacaatc ttagcttacc ttgaa

35

<210> 41
<211> 37
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 41
ttttatccat ggttatcttaa tccacaagcc ttgcctc

37

<210> 42
<211> 29
<212> ДНК
<213> штучна

```

<220>
 <223> синтетична послідовність

 <400> 42
 tttttaccat ggtgtcggag aagtgggag 29

 <210> 43
 <211> 27
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 43
 tttttaccat ggagaagtgg gcggacg 27

 <210> 44
 <211> 37
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 44
 ttttatggta cctagcttaa tctcagattc gaatcgt 37

 <210> 45
 <211> 40
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 45
 tttttaccat ggtagtatct acataccaat catacaaatg 40

 <210> 46
 <211> 32
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 46
 ttttttggtt cctttcacga ttggaaattt ga 32

 <210> 47
 <211> 36
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 47

ttttttccat ggtctacaac attaaaacga ccatta	36
<p><210> 48 <211> 32 <212> ДНК <213> штучна</p>	
<p><220> <223> синтетична послідовність</p>	
<400> 48 tatataggta ccagggtttc gtttttgttt ca	32
<p><210> 49 <211> 35 <212> ДНК <213> штучна</p>	
<p><220> <223> синтетична послідовність</p>	
<400> 49 tatataccat ggttatctcc tgctcaaaga aacca	35
<p><210> 50 <211> 33 <212> ДНК <213> штучна</p>	
<p><220> <223> синтетична послідовність</p>	
<400> 50 tttataggta ccagaagctc atttcttcga tac	33
<p><210> 51 <211> 32 <212> ДНК <213> штучна</p>	
<p><220> <223> синтетична послідовність</p>	
<400> 51 tttataccat ggtctctgcg caaaaattca cc	32
<p><210> 52 <211> 31 <212> ДНК <213> штучна</p>	
<p><220> <223> синтетична послідовність</p>	
<400> 52 tatattggta cctctaaaaa tacagggcac c	31
<210> 53	

```

<211> 35
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 53
tatattccat gggtactctt cgttgcagaa gccta 35

<210> 54
<211> 33
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 54
tatataggta ccactgttta agcttcactg tct 33

<210> 55
<211> 32
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 55
tatataccat gggttcttct aaagctgaaa gt 32

<210> 56
<211> 39
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 56
tatataggta ccttaagctt ttaagaatct ctactcaca 39

<210> 57
<211> 41
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 57
atatatccat gggttaaattt tacctgtcat caaaaacaac a 41

<210> 58
<211> 49
<212> ДНК
<213> штучна

```



```

<220>
<223> синтетична послідовність

<400> 58
ggggacaact ttgtatagaa aagttggcca catcatgttt agacttatac
49

<210> 59
<211> 50
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 59
ggggactgct tttttgtaca aacttggtta ctttttatat ttatatatag
50

<210> 60
<211> 73
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 60
ggtaccgggg gatcctctag catatgctcg aggcgggcgc agatatcaga tctggtcgac
60
ggcatgcaag ctt
73

<210> 61
<211> 12197
<212> ДНК
<213> штучна

<220>
<223> векторна послідовність

<400> 61
gacatacaaaa tggacgaacg gataaacctt ttcacgccct tttaaataac cgattattct
60
aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt
120
taaaactgaag gcgggaaaacg acaatcagat ctagttagaa acagctatga ccatgattac
180
gccaaactat cgattacgcc aagctatcaa ctttgtatag aaaagttggc cacatcatgt
240
ttagacttat ctccataaag aaaaccactc atcaaaagcta atacaaaagc tctagtgtga
300
cacttcaacg tcacttcate aggatcagca ggtaaatcc gaaaattctc acgcagccac
360
gctaaggaca catgagaacc atgaagatcc ttgggacctg gcctatgacc aagcaaatcc
420
tcacataaat cagcccagtt gtattttgta ctaccagtta ctgcaggtec atcgacacgt
480
agacccaaca aaatattcac atcttgtaaa gtcacagtga tctctccagc aggaagatga
540
aaagtatgcg ttctgggtct ccatctctcc accaaagctg ttatcagagc ataatcaagt
600
tgtataaagg caaccttgta aactccatat agaccaaaact ctatcaactt ttgacacacg
660

```

agaggatcca gaggccaatc tcgcatcccc aataacttgt gccgacatgt cagttcacga	720
ggaggaaacct gaatgtgaag tataacggta aaaaggaaat aattaaaaca acggaagcaa	780
aacaagaaac aagatgaaat gagaactag taacacacct catcttccca tatagcagct	840
gatctatgct catgttgcca caccaatata gattgatcaa ctggaccagg atccaaatca	900
aagtttaata gactttgcac ctccatctat ataatatatc acaggacaat aaacacaatg	960
atcagtgatt atacaacatc aaagaaaact tgtaattctg ggaatataac tgagaaatga	1020
gaattaaaga ttcataattt gaacacaaga aatcctaacc tggtagcaaa gaaaaattgc	1080
tcaacaaaaa aatcaagcta attactcgta tacaagaca cgaagaacta atacaagaaa	1140
caagaaacaa caaaccacaa agagattgaa ccaatccaaa ttcgacaaca taaaccaagt	1200
gtgtgggtga ttggaatcag aggacgtacc aaagaaaagc gtccgtgac aacaaaaacc	1260
aaaaagagac gtttgaaata aaccagagga agacgaagaa taattaagca aagagaagcg	1320
ttaagcggga gcgagaaagg aaacgagaga aagagagagc ttccagatcc gacagaagtt	1380
ttcggcttct tctttttcgt ttaagaactt ctgactctcc taggtctgtc cgaagaacta	1440
atctttttga ggtaacgacg ccgtttttct caaacatcgg gccattaac catagtctcg	1500
gccccaacga aacttaatac gacaatgttt ggggtgtaac gcaaagattt tgtcgattat	1560
cacaagtaaa aaaataaata caaacacttg agtctctcta gacatcgtgc atcgcttag	1620
ctttaagttt tttctcgaaa caaaagagtt attttatttg aactttgaag attatacgaa	1680
gacacgtggc gtgaacccaa ttcataacaa cgccacgcta tactcttttg catgcacctc	1740
aatttgaaca tcatcaagtc tctctctctt tttctgactt tgatccacga acctaacag	1800
cttgcgatct ctatttaato ggtcctcgac gcaacttcaa cttctactac atccattcac	1860
atcaaatcaa tacagaaagt tttttctata tataaatata aaaggtaaac aagtttgtag	1920
aaaaaagcag gctggtacct ggtgcttaaa cactctggtg agttctagta cttctgctat	1980
gatcgatctc attaccattt cttaaatttc tctccctaaa tattccgagt tcttgatttt	2040
tgataacttc aggtttttct tttttgataa atctggctct tccatttttt tttttttgtg	2100
gttaatttag tttcctatgt tcttcgattg tattatgcat gatctgtgtt tggattctgt	2160
tagattatgt attggtgaat atgtatgtgt ttttgcattg ctgggttttg tcttaaaaaat	2220
gttcaaatct gatgatttga ttgaagcttt tttagtgttg gtttgattct tctcaaaact	2280
actgttaatt tactatcatg ttttccaaact ttgattcatg atgacacttt tgttctgctt	2340
tggtataaaa ttttgggttg tttgattttg taattatagt gtaattttgt taggaatgaa	2400
catgttttaa tactctgttt tcgatttgtc acacattcga attattaatc gataatttaa	2460
ctgaaaattc atggttctag atcttgttgt catcagatta tttgtttcga taattcatca	2520
aatatgtagt ccttttgcgt atttgcgact gtttcatttt tctcaaaaat tgttttttgt	2580

taagtttato taacagttat cgttgtcaaa agtctctttc attttgcaaa atcttctttt	2640
tttttttggt tgtaactttg ttttttaagc tacacattta gtctgtaaaa tagcatcgag	2700
gaacagttgt cttagtagac ttgcatgttc ttgtaacttc tatttggttc agtttggtga	2760
tgactgtctt gattttgtag gtcaaaccat ggaagacgcc aaaaacataa agaaaggccc	2820
ggcgccattc tatccgctgg aagatggaac cgttgagag caactgcata aggctatgaa	2880
gagatacgcc ctggttcctg gaacaattgc ttttacagat gcacatatcg aggtggacat	2940
cacttacgct gagtacttcg aaatgtccgt tcggttgga gaagctatga aacgatatgg	3000
gctgaataca aatcacagaa tcgtcgtatg cagtgaatac tctcttcaat tctttatgcc	3060
ggtgttggtc gcgttatcta tcggagtgc agttgcgcc gcgaacgaca ttataaatga	3120
acgtgaattg ctcaacagta tgggcatttc gcagcctacc gtggtgttcg tttccaaaa	3180
ggggttgcaa aaaattttga acgtgcaaaa aaagctccca atcatccaaa aaattattat	3240
catggattct aaaacggatt accagggatt tcagtcgatg tacacgttcg tcacatctca	3300
tctacctccc gggtttaatg aatacgattt tgtgccagag tccttcgata gggacaagac	3360
aattgcactg atcatgaact cctctggatc tactggtctg cctaaagggt tcgctctgcc	3420
tcatagaact gcctgcgtga gattctcgca tgccagagat cctatttttg gcaatcaaat	3480
cattccgat actgcgattt taagtgtgtg tccattccat cacggttttg gaatgtttac	3540
tacactcgga tatttgatat gtggatttcg agtcgtctta atgtatagat ttgaagaaga	3600
gctgtttctg aggagccttc aggattacaa gattcaaat gcgctgctgg tgccaaccct	3660
attctccttc ttcgccaaa gcactctgat tgacaaatac gatttatcta atttacagca	3720
aattgcttct ggtggcgctc cctctctaa ggaagtcggg gaagcgggtg ccaagagggt	3780
ccatctgcca ggtatcaggc aaggatatgg gtcactgag actacatcag ctattctgat	3840
tacaccgag ggggatgata aaccgggcgc gtcggtaaa gttgttccat ttttgaaagc	3900
gaagggtgtg gatctggata ccgggaaaac gctgggcgtt aatcaaagag gcgaactgtg	3960
tgtgagaggt cetatgatta tgtccggtta tgtaaacat ccggaagcga ccaacgcctt	4020
gattgacaag gatggatggc tacattctgg agacatagct tactgggacg aagacgaaca	4080
cttcttcttc gttgaccgcc tgaagtctct gattaagtac aaaggctatc aggtggctcc	4140
cgctgaattg gaatccatct tgctccaaca ccccaacatc ttcgacgcag gtgtcgacgg	4200
tcttcccagc gatgacgcc gtgaacttcc cgcgcgctt gttgttttgg agcacggaaa	4260
gacgatgacg gaaaaagaga tcgtggatta cgtcgccagt caagtaacaa ccgcgaaaaa	4320
gttgcgcgga ggagtttgtt ttgtggacga agtaccgaaa ggtcttaccg gaaaactcga	4380
cgcaagaaaa atcagagaga tcctcataaa ggccaagaag ggcggaagaa tcgccgtgta	4440

actcgagcat atgggctega atttccccga tcgttcaaac atttggcaat aaagtttctt	4500
aagattgaat cctgttgccg gtcttgcat gattatcata taatttctgt tgaattacgt	4560
taagcatgta ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat	4620
tagagtcccg caattataca tttaatcgc gatagaaaac aaaatatagc gcgcaaaacta	4680
ggataaatta tcgcgcgcgg tgcatctat gttactagat cgggaattca agcttggcgt	4740
aatcatggac ccagctttct tgtacaaagt ggggtaccgg gggatcctct agcatatgct	4800
cgaggcgcc gcagatatca gatctggctg acggcatgca agcttggcgt aatcatggca	4860
actttattat acatagttga taattcactg gccggataat tcaactggcg tcgttttaca	4920
acgactcagg atcctgtcaa acaactgtag tttaaactga aggcgggaaa cgacaatctg	4980
atcatgagcg gagaattaag ggagtcacgt tatgaccccc gccgatgacg cgggacaagc	5040
cgttttacgt ttggaactga cagaaccgca acgttgaagg agccactcag ccgcgggttt	5100
ctggagtta atgagctaag cacatactg agaaaccatt attgcgcgtt caaaagtcgc	5160
ctaaggtcac tatcagctag caaatatttc ttgtcaaaa tgctccactg acgttccata	5220
aattcccctc ggtatccaat tagagtctca tttcactct caatccaaat aatctgcacc	5280
ggatctggat cgtttcgcat gattgaacaa gatggattgc acgcaggttc tccggccgct	5340
tgggtggaga ggctattcgg ctatgactgg gcacaacaga caatcgctg ctctgatgcc	5400
gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt ttgtcaagac cgacctgtcc	5460
ggtgccctga atgaactgca ggacgaggca gcgcggctat cgtggctggc cagcaggggc	5520
gttccctgcg cagctgtgct cgaactgtc actgaagcgg gaagggaactg gctgtattg	5580
ggcgaagtgc cggggcagga tctcctgtca tctcacttg ctctgcccga gaaagtatcc	5640
atcatggctg atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac	5700
caccaagcga aacatcgcat cgagcgagca cgtactcgga tgggaagccgg tcttgtcgat	5760
caggatgac tggaacgaaga gcacagggg ctccgcgcag ccgaactggt cgcaggctc	5820
aaggcgcgca tgcccagcgg cgaggatctc gtctgaccc atggcgatgc ctgcttgccg	5880
aatatcatgg tggaaaatgg ccgcttttct ggattcatcg actgtggcgg gctgggtgtg	5940
gcggaccgct atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc	6000
gaatgggctg accgcttctt cgtgctttac ggtatcgccg ctcccgatcc gcagcgcatc	6060
gccttctatc gccttcttga cgagtcttct tgagcgggac ccaagctcta gatcttgctg	6120
cgttcggata ttttcgtgga gttcccgcca cagaccgga tgatccccga tcgttcaaac	6180
atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcat gattatcata	6240
taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt	6300
atgagatggg tttttatgat tagagtcccg caattataca tttaatcgc gatagaaaac	6360

aaaatatagc gcgcaaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat	6420
cgggcctcct gtcaagctct gcttggtaat aattgtcatt agattgtttt tatgcataga	6480
tgcactcgaa atcagccaat tttagacaag tatcaaacgg atgttaattc agtacattaa	6540
agacgtccgc aatgtgttat taagttgtct aagcgtcaat ttgtttacac cacaatatat	6600
cctgccacca gccagccaac agctccccga ccggcagctc ggcacaaaat caccacgcgt	6660
taccaccacg ccggccggcc gcatgggtgt gaccgtgttc gccggcattg ccgagttcga	6720
gcgttcccta atcatcgacc gcacccggag cgggcgcgag gccgccaaagg cccgaggcgt	6780
gaagtttggc ccccgcccta cctcaccac gccacagatc gcgcacgccc gcgagctgat	6840
cgaccaggaa ggccgcaccg tgaaagaggc ggctgcactg cttggcgtgc atcgctcgac	6900
cctgtaccgc gcacttgagc gcagcgagga agtgacgccc accgaggcca ggccggcgcg	6960
tgcttccgt gaggacgcat tgaccgagge cgacgccctg gcggccgccg agaatgaacg	7020
ccaaggaggaa caagcatgaa accgcaccag gacggccagg acgaaccgtt ttctattacc	7080
gaagagatcg aggcggagat gatcgcggcc gggtacgtgt tcgagccgcc cgcgcacgtc	7140
tcaaccgtgc ggctgcatga aatcctggcc ggtttgtctg atgccaagct ggccggcctgg	7200
ccggccagct tggccgctga agaaaccgag cgccgccgtc taaaagggtg atgtgtattt	7260
gagtaaaaca gcttgcgtca tgcggtcgtc gcgtatatga tgcgatgagt aaataaaca	7320
atacgcaagg ggaacgcatg aaggttatcg ctgtacttaa ccagaaaggc gggtcaggca	7380
agacgaccat cgcaacccat ctagcccgcc cctgcaact cgccggggcc gatgttctgt	7440
tagtcgattc cgatccccag ggcagtgcgc gcgattgggc ggcctgtcgc gaagatcaac	7500
cgctaaccgt tgtcggcatc gaccgccga cgattgaccg cgacgtgaag gccatcggcc	7560
ggcgcgactt cgtagtgate gacggagcgc ccaggcgcc ggacttggct gtgtcccgca	7620
tcaaggcagc cgaattcgtg ctgattccgg tgcagccaag cccttacgac atatgggcca	7680
ccgccgacct ggtggagctg gtttaagcagc gcattgaggt caccgatgga aggctacaag	7740
cggcctttgt cgtgtcgcgg gcgatcaaaag gcacgcgcac cggcggtgag gttgccgagg	7800
cgctggccgg gtacgagctg cccattcttg agtcccgat caccgagcgc gtgagctacc	7860
caggcactgc cgccgccggc acaaccgttc ttgaatcaga acccgagggc gacgctgccc	7920
gcgaggteca ggcgtggcc gctgaaatta aatcaaaaet catttgagtt aatgaggtaa	7980
agagaaaatg agcaaaagca caaacacgct aagtgcgggc cgtccgagcg caccgagcag	8040
caaggctgca acgttgcca gcctggcaga caccgccgc atgaagcggg tcaactttca	8100
gttgccggcg gaggatcaca ccaagctgaa gatgtacgg gtacgccaag gcaagaccat	8160
taccgagctg ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa	8220
tgagtagatg aattttagcg gctaaaggag gcggcatgga aaatcaagaa caaccaggca	8280

ccgacgccgt ggaatgcccc atgtgtggag gaacggggcg ttggccaggc gtaagcggct	8340
gggttgtctg cgggcctgc aatggcactg gaacccccaa gcccgaggaa tcggcgtgag	8400
cggtcgcaaa ccacccggcc cggtaaaaat cggcgccggc ctgggtgatg acctggtgga	8460
gaagtgaag gccgcgcagg ccgccagcg gcaacgcacg gaggcagaag cacgccccgg	8520
tgaatcgtgg caagcggcgg ctgatcgaat ccgcaaaaga tcccggcaac cggcggcagc	8580
cgggtgcgcg tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc	8640
gatgctctat gacgtgggca ccgcgatag tcgcagcacc atggacgtgg ccgttttccg	8700
tctgtcgaag cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca	8760
cgtagagggt tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact	8820
gatggcgggt tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa	8880
gcccggccgc gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga	8940
tggcggaag cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt	9000
tgccatgcag cgtacgaaga aggccaaaga cggccgcctg gtgacgggat ccgaggggtga	9060
agccttgatt agccgctaca agatcgtaaa gagcgaacc gggcgcccg agtacatcga	9120
gatcgagcta gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct	9180
gacggttcac cccgattact ttttgatcga tcccggcacc gccggttttc tctaccgct	9240
ggcacgccgc gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg	9300
cagtggcagc gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc	9360
aaatgacctg ccggagtacg atttgaaggga ggaggcgggg caggctggcc cgatcctagt	9420
catgcgtac cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca	9480
gatgctaggg caaatgccc tagcagggga aaaaggtcga aaaggctctt ttcctgtgga	9540
tagcacgtac attgggaacc caaagccgta cattgggaac cggaaccctg acattgggaa	9600
cccaaagccg tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa	9660
aggcgatttt tccgcctaaa actcttttaa acttattaaa actcttaaaa cccgcctggc	9720
ctgtgcataa ctgtctggcc agcgacacgc cgaagagctg caaaaagcgc ctacccttcg	9780
gtcgtgcgc tccctacgcc ccgcgccttc gcgtcggcct atcgcgcccg ctggccgctc	9840
aaaaatggct ggcctacggc caggcaatct accagggcgc ggacaagccg cgcgctcgcc	9900
actcgaccgc cggcgcccac atcaaggcac cctgcctcgc gcgtttcggg gatgacgggtg	9960
aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg	10020
ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca	10080
tgaccacgtc acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca	10140
gattgtactg agagtgcacc atatgagggtg tgaataaccg cacagatgcg taaggagaaa	10200

ataccgcac	aggcgctctt	ccgcttcttc	gtcactgac	tccgtgcgct	cggtcgttcg	10260
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggttaata	cggttatcca	cagaatcagg	10320
ggataacgca	ggaagaaca	tgtgagcaa	aggccagcaa	aaggccagga	accgtaaaaa	10380
ggccgcgttg	ctggcgtttt	tccataggct	ccgccccctt	gacgagcatc	acaaaaatcg	10440
acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttccccc	10500
tggaagctcc	ctcgtgcgct	ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	10560
ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	cgctgtaggt	atctcagttc	10620
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	10680
ctgcgcctta	tccggttaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	10740
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	10800
gttcttgaag	tgttgcccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	10860
tctgtgaag	ccagttacct	tcggaaaaag	agttggtage	tcttgatccg	gcaaaaaaac	10920
caccgctggg	agcggtgggt	tttttggttg	caagcagcag	attacgcgca	gaaaaaaagg	10980
atctcaagaa	gaccccttga	tcttttctac	gggtctctgac	gctcagtggc	acgaaaactc	11040
acgttaaggg	attttggtca	tgcctgatct	atctcccaat	ttgtgtaggg	cttattatgc	11100
acgcttaaaa	ataataaaaag	cagacttgac	ctgatagttt	ggctgtgagc	aattatgtgc	11160
ttagtgcac	taacgcttga	gttaagccgc	gccgcgaagc	ggcgctcggt	tgaacgaatt	11220
tctagctaga	cattatttgc	cgaactacct	ggtgatctcg	cctttcacgt	agtggacaaa	11280
ttcttccaac	tgatctgcgc	gcgaggccaa	gcgatcttct	tcttgtccaa	gataagcctg	11340
tctagcttca	agtatgacgg	gctgatactg	ggcgggcagg	cgctccattg	cccagtcggc	11400
agcgacatcc	ttcggcgcgga	ttttgccggt	tactgcgctg	taccaaattg	gggacaacgt	11460
aagcactaca	tttcgctcat	cgcagccca	gtcgggcggc	gagttccata	gcgttaagggt	11520
ttcattttag	gcctcaaata	gacccgtgtc	aggaaccgga	tcaaagagtt	cctccgcgcg	11580
tggacctacc	aaggcaacgc	tatgttctct	tgcctttgtc	agcaagatag	ccagatcaat	11640
gtcgatcgtg	gctggctcga	agataacctg	aagaatgtca	ttgcgctgcc	attctccaaa	11700
ttgcagttcg	cgtttagctg	gataacgccca	cggaaatgatg	tcgtcgtgca	caacaatggt	11760
gacttctaca	gcgcggagaa	tctcgtcttc	tccaggggaa	gccgaagttt	ccaaaaggtc	11820
gttgatcaaa	gctcgcgcg	ttgtttcatc	aagccttacg	gtcaccgtaa	ccagcaaadc	11880
aatatcactg	tgtggcttca	ggccgccatc	cactgcggag	ccgtacaaat	gtacggccag	11940
caacgtcggg	tcgagatggc	gctcgatgac	gccaaactac	tctgatagtt	gagtcgatac	12000
ttcggcgatc	accgcttccc	ccatgatggt	taactttggt	ttagggcgac	tgccctgctg	12060
cgtaacatcg	ttgctgctcc	ataacatcaa	acatcgaccc	acggcgtaac	gcgcttgctg	12120

```

cttggatgcc cgaggcatag actgtacccc aaaaaaacag tcataacaag ccatgaaaaac 12180
cgccactgcg ttccatg 12197

<210> 62
<211> 14901
<212> ДНК
<213> штучна

<220>
<223> векторна послідовність

<400> 62
gacatacaaaa tggacgaacg gataaacctt ttcacgccct tttaaataac cgattattct 60
aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt 120
taaaactgaag gcgggaaacg acaatcagat ctagtaggaa acagctatga ccatgattac 180
gccaaactat cgattacgcc aagctatcaa ctttgtatag aaaagttggc cacatcatgt 240
ttagacttat ctccataaag aaaaccactc atcaaaagta atacaaaagc tctagtgtga 300
cacttcaacg tcaattcacc aggatcagca ggtaaaattcc gaaaattctc acgcagccac 360
gctaaggaca catgagaacc atgaagatcc ttgggacctg gcctatgacc aagcaaatcc 420
tcacataaat cagcccagtt gtattttgta ctaccagtta ctgcaggtcc atcgacacgt 480
agacccaaca aaatattcac atcttgtaaa gtcacagtga tctctccagc aggaagatga 540
aaagtatgcg ttccgggtct ccatctctcc accaaagctg ttatcagagc ataatcaagt 600
tgtataaagg caaccttgta aactccatat agaccaaact ctatcaactt ttgacacacg 660
agaggatcca gaggccaatc tcgcatcccc aataacttgt gccgacatgt cagttcacga 720
ggaggaacct gaatgtgaag tataacggtg aaaaggaaat aattaaaaa acggaagcaa 780
aacaagaaac aagatgaaat gagaactag taacacacct catcttccca tatagcagct 840
gatctatgct catgttgcca caccaatata gattgatcaa ctggaccagg atccaaatca 900
aagtttaata gactttgcac ctccatctat ataatatatc acaggacaat aaacacaatg 960
atcagtgatt atacaacatc aaagaaaact tgtaattctg ggaatataac tgagaaatga 1020
gaattaaaga ttcataattt gaacacaaga aatcctaaac tggtagcaaa gaaaaattgc 1080
tcaacaaaaa aatcaagcta attactcgta tacaagaca cgaagaacta atacaagaaa 1140
caagaacaaa caaacacaaa agagattgaa ccaatccaaa ttcgacaaca taaaccaagt 1200
gtgtgggtga ttggaatcag aggacgtacc aaagaaaagc gtccgtgac aacaaaaacc 1260
aaaaagagac gtttgaaata aaccagagga agacgaagaa taattaagca aagagaagcg 1320
ttaagcggga gcgagaaagg aaacgagaga aagagagagc ttccagatcc gacagaagtt 1380
ttcggcttct tctttttctg ttaagaactt ctgactctcc taggtctgtc cgaagaacta 1440
atctttttga ggtaacgacg ccgtttttct caaaacatgg gccattaac catagtctcg 1500

```

gccccaaacga aacttaatac gacaatgttt ggggtgtaaac gcaaagattt tgtcgattat	1560
cacaagtataa aaaataaata caaacacttg agtctctcta gacatcgtag atcgccctag	1620
ctttaagttt tttctcgaaa caaaagagtt attttatttg aactttgaag attatacgaa	1680
gacacgtggc gtgaacccaa ttcataacaa cgccacgcta tactcttttg catgcacctc	1740
aatttgaaca tcataagtc tctctctctt tttctgactt tgatccacga acctaacgag	1800
cttgcgatct ctatttaate ggtcctcgac gcaacttcaa cttctactac atccattcac	1860
atcaaatcaa tacagaaagt tttttctata tataaatata aaaggtaaac aagtttgtac	1920
aaaaaagcag gctggtagct ggtgcttaaa cactctgggt agttctagta cttctgctat	1980
gatcgatctc attaccattt cttaaatttc tctccctaaa tattccgagt tcttgatttt	2040
tgataacttc aggtttttctc tttttgataa atctgggtctt tccatttttt tttttttgtg	2100
gttaatttag tttcctatgt tcttcgattg tattatgcat gatctgtgtt tggattctgt	2160
tagattatgt attggtagat atgtatgtgt ttttgcattg ctggtttttg tcttaaaaaa	2220
gttcaaatct gatgatttga ttgaagcttt tttagtgttg gtttgattct tctcaaaact	2280
actgttaatt tactatcatg ttttccaaact ttgattcatg atgacacttt tgttctgctt	2340
tgttataaaa ttttgggttg tttgattttg taattatagt gtaattttgt taggaatgaa	2400
catgttttaa tactctgttt tctgatttgc acacattcga attattaatc gataatttaa	2460
ctgaaaatc atgggtctag atcttgtgt catcagatta tttgtttcga taattcatca	2520
aatatgtagt ccttttgctg atttgcgact gtttcatttt tcttcaaaat tgttttttgt	2580
taagtttatc taacagttat cgttgtcaaa agtctctttc attttgcaa atcttctttt	2640
tttttttgtt tgtaaacttg ttttttaagc tacacattta gtctgtaaaa tagcatcgag	2700
gaacagttgt cttagtagac ttgcatgttc ttgtaacttc tatttgtttc agtttgttga	2760
tgactgcttt gatttttagt gtcaaaccat ggaagacgac aaaaacataa agaagggccc	2820
ggcgccattc tatccgctgg aagatggaac cgctggagag caactgcata aggctatgaa	2880
gagatacgcc ctggttcctg gaacaattgc ttttacagat gcacatatcg aggtggacat	2940
cacttacgct gagtacttcg aaatgtccgt tcggttgcca gaagctatga aacgatatgg	3000
gctgaatata aatcacagaa tcgtcgtatg cagtgaatac tctcttcaat tctttatgcc	3060
gggtgtgggc gcgttattta tcggagtgc agttgcgccc gcgaacgaca tttataatga	3120
acgtgaattg ctcaacagta tgggcatttc gcagcctacc gtggtgttcg tttccaaaaa	3180
gggggtgcaa aaaattttga acgtgcaaaa aaagctccca atcatccaaa aaattattat	3240
catggattct aaaaaggatt accagggatt tcagtcgatg tacacgttcg tcacatctca	3300
tctacctccc ggttttaatg aatacgattt tgtgccagag tcttctgata gggacaagac	3360
aattgcactg atcatgaact cctctggatc tactgggtctg cctaaagggt tcgctctgcc	3420

tcatagaact	gcctgcgtga	gattctcgca	tgccagagat	cctatttttg	gcaatcaaat	3480
cattccggat	actgcgattt	taagtgttgt	tccattccat	cacggttttg	gaatgtttac	3540
tacactcgga	tatttgatat	gtggatttcg	agtcgtctta	atgtatagat	ttgaagaaga	3600
gctgtttctg	aggagccttc	aggattacaa	gattcaaagt	gcgctgctgg	tgccaaccct	3660
attctccttc	ttcgccaaaa	gcactctgat	tgacaaatac	gattttatct	atttacacga	3720
aattgcttct	ggtggcgctc	ccctctctaa	ggaagtcggg	gaagcggttg	ccaagagggt	3780
ccatctgcca	ggtatcaggc	aaggatatgg	gctcactgag	actacatcag	ctattctgat	3840
tacacccgag	ggggatgata	aaccgggcgc	ggtcggtaaa	gttgttccat	tttttgaagc	3900
gaagggtgtg	gatctggata	ccgggaaaac	gctgggcgtt	aatcaaagag	gcgaactgtg	3960
tgtgagaggt	cctatgatta	tgtccggtta	tgtaaacaaat	ccggaagcga	ccaacgcctt	4020
gattgacaag	gatggatggc	tacattctgg	agacatagct	tactgggacg	aagacgaaca	4080
cttcttcctc	gttgaccgcc	tgaagtctct	gattaagtac	aaaggctatc	aggtggctcc	4140
cgtgaattg	gaatccatct	tgctccaaca	ccccaacatc	ttcgacgcag	gtgtcgcagg	4200
tcttcccgac	gatgacgcgc	gtgaacttcc	cgccgcgctt	gttgtttttg	agcacggaaa	4260
gacgatgacg	gaaaaagaga	tcgtggatta	cgtcgccagt	caagtaacaa	ccgcgaaaaa	4320
gttgccggga	ggagtgtgtg	ttgtggacga	agtaccgaaa	ggtcttaccg	gaaaactcga	4380
cgcaagaaaa	atcagagaga	tcctcataaa	ggccaagaag	ggcggaaaga	tcgccgtgta	4440
actcgagcat	atgggctcga	atttccccga	tcgttcaaac	atttggaat	aaagtttctt	4500
aagattgaat	cctgttgccg	gtcttgcgat	gattatcata	taatttctgt	tgaattacgt	4560
taagcatgta	ataattaaca	tgtaatgcac	gacgttattt	atgagatggg	tttttatgat	4620
tagagtcccc	caattataca	tttaatacgc	gatagaaaac	aaaatatagc	gcgcaacta	4680
ggataaatta	tcgcgcgcgg	tgctcatctat	gttactagat	cgggaattca	agcttggcgt	4740
aatcatggac	ccagctttct	tgtacaaaagt	ggggtaccaa	ttcgaatcca	aaaattacgg	4800
atatgaatat	agggatatcc	gtatccgaat	tatccgtttg	acagctagca	acgattgtac	4860
aattgcttct	ttaaaaaagg	aagaaagaaa	gaaagaaaag	aatcaacatc	agcgtaacaa	4920
aacggccccc	ttacggccca	aacggtcata	tagagtaacg	gcgttaagcg	ttgaaagact	4980
cctatcgaaa	tacgtaaccg	caaacgtgtc	atagtcagat	ccccctcttc	ttcaccgeet	5040
caaacacaaa	aataatcttc	tacagcctat	atatacaacc	cccccttcta	tctctccttt	5100
ctcacaattc	atcatcttct	ttctctctac	cccaatttta	agaaatcctc	tctctctctc	5160
ttcattttca	aggtaaatct	ctctctctct	ctctctctct	gttattcctt	gttttaatta	5220
ggatatgtatt	attgctagtt	tgtaaatctg	cttatcttat	gtatgcotta	tgtgaatata	5280
tttatcttgt	tcctctcctc	cgtttagaag	ctataaaatt	gttgatttga	ctgtgtatct	5340

acacgtgggt atgtttatat ctaatcagat atgaatttct tcatattgtt gcgtttgtgt	5400
gtaccaatcc gaaatcgttg atttttttca tttaatcgtg tagctaattg tacgtataca	5460
tatggatcta cgtatcaatt gttcatctgt ttgtgtttgt atgtatacag atctgaaaac	5520
atcacttctc tcatctgatt gtgtgtttac atacatagat atagatctgt tatatcattt	5580
tttttattaa ttgtgtatat atatatgtgc atagatctgg attacatgat tgtgattatt	5640
tacatgattt tgttatttac gtatgtatat atgtagatct ggactttttg gagtgtgtga	5700
cttgattgta ttgtgtgtg tatatgtgtg ttctgatctt gatattgtat gtatgtgcag	5760
ctgaaccatg gcggcggcaa caacaacaac aacaacatct tcttcgatct ccttctccac	5820
caaacatctc cttctctctt ccaaatcacc attaccaatc tcagattctt cctctccatt	5880
ctccctaaac cccaacaaat catcctctct ctcgcgcgc gcgggtatca aatccagctc	5940
tcctctctcc atctccgcg tgctcaacac aaccaccaat gtcacaacca ctccctctcc	6000
aaccaaacct accaaaccgg aaacattcat ctcccgattc gctccagatc aaccccgcaa	6060
aggcgctgat atcctcgtcg aagctttaga acgtcaaggc gtagaaaccg tattcgctta	6120
cctggagggt acatcaatgg agattcacca agccttaacc cgctcttctt caatccgtaa	6180
cgctcttctt cgtcacgaac aaggagggtg attcgcagca gaaggatag ctcgatcctc	6240
aggtaaacca ggtatctgta tagccacttc aggtcccga gctacaaatc tcgttagcgg	6300
attagccgat gcgttgtag atagtgttc tcttgtagca atcacaggac aagtcctctg	6360
tcgtatgatt ggtacagatg cgtttcaaga gactccgatt gttgaggtaa cgcgttcgat	6420
tacgaagcat aactatcttg tgatggatgt tgaagatata cctaggatta ttgaggaagc	6480
tttcttttta gctactctg ttagacctgg acctgttttg gttgatgttc ctaaagatat	6540
tcaacaacag cttgcgattc ctaattggga acaggctatg agattacctg gttatatgtc	6600
taggatgcct aaacctccgg aagattctca tttggagcag attgttaggt tgatttctga	6660
gtctaagaag cctgtgttgt atgttgggtg tggttgtttg aattctagcg atgaattggg	6720
taggtttgtt gagcttacgg ggatccctgt tgcgagtacg ttgatggggc tgggatctta	6780
tccttgatga gatgagttgt cgttacatat gcttggaatg catgggactg tgtatgcaaa	6840
ttacgctgtg gagcatagtg atttgttgtt ggcgtttggg gtaagggttg atgatcgtgt	6900
cacgggtaag cttgaggctt ttgctagtag ggctaagatt gttcatattg atattgactc	6960
ggctgagatt gggagaata agactctca tgtgtctgtg tgtggtgatg ttaagctggc	7020
tttgcaaggg atgaataagg ttcttgagaa ccgagcggag gagcttaagc ttgattttgg	7080
agtttgagg aatgagttga acgtacagaa acagaagttt ccgttgagct ttaagacgtt	7140
tggggaagct attcctccac agtatcgat taaggctcct gatgagttga ctgatggaaa	7200
agccataata agtactggtg tcgggcaaca tcaaatgtgg gcggcgcagt tctacaatta	7260

caagaaacca aggcagtggc tatcatcagg aggccttga gctatggat ttggacttcc	7320
tgctgcgatt ggagcgtctg ttgctaacc tgatgcgata gttgtggata ttgacggaga	7380
tggaagcttt ataatgaatg tgcaagagct agccactatt cgtgtagaga atcttccagt	7440
gaaggtactt ttattaaaca accagcatct tggcatggt atgcaatggg aagatcgggt	7500
ctacaaagct aaccgagctc acacattttc cggggatccg gctcaggagg acgagatatt	7560
cccgaacatg ttgctgtttg cagcagcttg cgggattcca gcggcgaggg tgacaaagaa	7620
agcagatctc cgagaagcta ttcagacaat gctggatata ccaggacctt acctgttggg	7680
tgtgatttgt ccgcaccaag aacatgtgtt gccgatgatc ccgaatggtg gcactttcaa	7740
cgatgtcata acggaaggag atggccggat taaatactga gagatgaaac cgggtgattat	7800
cagaaccttt tatggtcttt gtatgcatac ggtaaaaaa cttagtgttc aatttcctgt	7860
ttgttttggg aatttgagtt tcttttagtt gttgatctgc ctgctttttg gtttacgtca	7920
gactactact gctgtgttg tttggtttcc tttctttcat tttataaata aataatccgg	7980
ttcggtttac tcttgtgac tggctcagtt tggttattgc gaaatcgaa tggtaaattg	8040
agtaattgaa attcgttatt agggttctaa gctgttttaa cagtcaactg gttaatatct	8100
ctcgaatctt gcattgaaaa tgcctctacc attgggtttt aattgaaatg tgctcatatg	8160
ggcgtggtt tccaaattaa ataaaactac gatgtcatcg agaagtaaaa tcaactgtgt	8220
ccacattatc agttttgtgt atacgatgaa atagggtaat tcaaaatcta gcttgatatg	8280
ccttttgggt cattttaacc ttctgtaaac attttttcag attttgaaca agtaaatcca	8340
aaaaaaaaa aaaaaaatct caactcaaca ctaaattatt ttaattgtata aaagatgctt	8400
aaaacatttg gcttaaaaga aagaagctaa aaacatagag aactcttgta aattgaagta	8460
tgaaaatata ctgaattggg tattatatga atttttctga tttaggattc acatgatcca	8520
aaaaggaaat ccagaagcac taatcagaca ttggaagtag gaatatttca aaaagttttt	8580
tttttttaag taagtacaa aagcttttaa aaaatagaaa agaaactagt attaaagttg	8640
taaatttaat aaacaaaaga aattttttat attttttcat ttctttttcc agcatgaggt	8700
tatgatggca ggatgtggat ttcatttttt tccttttgat agccttttaa ttgatctatt	8760
ataattgacg aaaaaatatt agttaattat agatatattt taggtagtat tagcaattta	8820
cacttccaaa agactatgta agttgtaaat atgatgcgtt gatctcttca tcattcaatg	8880
gtagtcaaaa aaaataaaag cttaactagt aaactaaagt agtcaaaaat tgtactttag	8940
tttaaaatat tacatgaata atccaaaacg acatttatgt gaaacaaaaa caatatctag	9000
agtcgactta attaaactag tggcgcgcca attgactagt aggcctatcg attaattaag	9060
gccgcctcga gcatatgggc atgcaagctt ggcgtaatca tggcaacttt attatacata	9120
gttgataatt cactggccgg atctgcttgg taataattgt cattagattg tttttatgca	9180

tagatgcact cgaaatcagc caattttaga caagtatcaa acggatgtta attcagtaca	9240
ttaaagacgt ccgcaatgtg ttattaagtt gtctaagcgt caatttgttt acaccacaat	9300
atatacctgcc accagccagc caacagctcc ccgaccggca gctcggcaca aaatcaccac	9360
gcgttaccac caccgcgcc gcccgcatgg tgttgaccgt gttcggcgcc attgcccagt	9420
tgcagcgctc cctaatacgc gaccgcaccc ggagcggggc cgaggccgcc aaggcccag	9480
gcgtgaagtt tggccccgc cctaccctca cccgggcaca gatcgcgcac gcccgcgagc	9540
tgatcgacca ggaaggccgc accgtgaaag aggcggctgc actgcttgcc gtgcatcgct	9600
cgaccctgta ccgcgcactt gagcgcagcg aggaagtgc gccaccgag gccaggcgcc	9660
gcggtgcctt ccgtgaggac gcattgaccg aggcgcagcc cctggcgcc cccgagaatg	9720
aacgccaaga ggaacaagca tgaaccgcga ccaggacggc caggacgaac cgtttttcat	9780
taccgaagag atcgaggcgg agatgatcgc ggccgggtac gtgttcgagc cgcgcgcga	9840
cgtctcaacc gtgcggctgc atgaatacct ggccgggttg tctgatgcc agetggcgcc	9900
ctggccggcc agcttgccgc ctgaagaaac cgagcgcgc cgtctaaaaa ggtgatgtgt	9960
atttgagtaa aacagcttgc gtcatgcggt cgtgcgtat atgatcgat gagtaataa	10020
acaaatacgc aaggggaacg catgaaggtt atcgctgtac ttaaccagaa aggcgggtca	10080
ggcaagacga ccatacgcaac ccatactagc cgcgccctgc aactcgcggc ggccgatgtt	10140
ctgttagtcg attccgatcc ccagggcagt gcccgcgatt ggccggccgt gcgggaagat	10200
caaccgctaa ccgttgctcg catcgaccgc ccgacgattg accgcgacgt gaaggccatc	10260
ggccggcgcc acttcgtagt gatcgacgga ggcggccagg cggcggaact ggtgtgtcc	10320
gcgatcaagg cagccgactt cgtgtgatt ccggtgcagc caagccctta cgacatatgg	10380
gccaccgccc accctggtgga gctggttaag cagcgcatg aggtcacgga tggaaggcta	10440
caagcggcct ttgtcgtgc gccggcgatc aaaggcacgc gcacggcgcg tagggttgc	10500
gaggcgctgg ccgggtacga gctgccatt cttgagtccc gtatcacgca gcgcgtgagc	10560
taccagga caagccgcgc cggcacaacc gttcttgaat cagaaccgga gggcgacgct	10620
gcccgcgagg tccaggcgct ggccgctgaa attaaatcaa aactcatttg agttaatgag	10680
gtaaagagaa aatgagcaaa agcacaacaa cgctaagtgc cggccgtccg agcgacgca	10740
gcagcaaggc tgcaacgttg gccagcctgg cagacacgcc agccatgaag cgggtcaact	10800
ttcagttgcc ggccggaggat cacaccaagc tgaagatgta cgcggtacgc caaggcaaga	10860
ccattaccga gctgctatct gaatacatcg cgcagctacc agagtaaatg agcaaatgaa	10920
taaatgagta gatgaatttt agcggctaaa ggaggcggca tggaaaatca agaacaacca	10980
ggcaccgacg ccgtggaatg ccccatgtgt ggagggaacgg gcggttgcc aggcgtaagc	11040
ggctgggttg tctgcgggcc ctgcaatggc actggaaccc ccaagccgga ggaatcgcg	11100

tgagcgggtcg caaaccatcc ggcccgggtac aaatcggcgc ggcgctgggt gatgacctgg	11160
tggagaagtt gaaggccgcg caggccgccc agcggcaacg catcgaggca gaagcacgcc	11220
ccggtgaatc gtggcaagcg gccgctgate gaatccgcaa agaatcccgg caaccgcccgg	11280
cagccgggtgc gccgtcgatt aggaagccgc ccaagggcga cgagcaacca gattttttcg	11340
ttccgatget ctatgacgtg ggcacccgcg atagtgcag catcatggac gtggccgttt	11400
tccgtctgtc gaagcgtgac cgaagagctg gcgaggtgat ccgctacgag ctccagacg	11460
ggcacgtaga ggtttccgca gggccggccg gcatggccag tgtgtgggat tacgacctgg	11520
tactgatggc ggtttcccat ctaaccgaat ccatgaaccg ataccggga a ggggaaggag	11580
acaagcccgg ccgctgttgc cgtccacacg ttgcggacgt actcaagttc tgccggcgag	11640
ccgatggcgg aaagcagaaa gacgacctgg tagaaacctg cattcggtta aacaccacgc	11700
acgttgccat gcagcgtacg aagaaggcca agaaccggccg cctggtgacg gtatccgagg	11760
gtgaagcctt gattagccgc tacaagatcg taaagagcga aaccggggcg ccggagtaca	11820
tcgagatcga gctagctgat tggatgtacc gcgagatcac agaaggcaag aaccgggacg	11880
tgttgacggt tcccccgat tactttttga tegtcccg ccgctggcgt tttctctacc	11940
gcctggcacg ccgcgcgcga ggcaaggcag aagccagatg gttgttcaag acgatctacg	12000
aacgcagtgg cagcgcggga gagttcaaga agttctgttt caccgtgcgc aagctgatcg	12060
ggtcaaatga cctgccggag tacgatttga aggaggaggc ggggcaggct ggcccgatcc	12120
tagtcatgcg ctaccgcaac ctgatcgagg gcgaagcacc cgcgggttcc taatgtacgg	12180
agcagatgct agggcaaat gcccctagcag gggaaaaagg tcgaaaaagg ctctttctcg	12240
tggatagcac gtacattggg aaccctaaagc cgtacattgg gaaccggaac ccgtacattg	12300
ggaacccaaa gccgtacatt gggaaccggc cacacatgta agtgactgat ataaaagaga	12360
aaaaaggcga tttttccgcc taaaactctt taaaacttat taaaactctt aaaacccgcc	12420
tggcctgtgc ataactgtct ggccagcgca cagccgaaga gctgcaaaaa gcgcctaccc	12480
ttcggctcgt gcgctcccta cgcgccgcgc ctctcgctcg gcctatcgcg gccgctggcc	12540
gctcaaaaat ggctggccta cggccaggca atctaccagg gcgcgggaca gccgcgcgt	12600
cgcactcga ccgcggcgcc ccacatcaag gcaacctgcc tcgcgcgttt cgggtgatgac	12660
ggtgaaaacc tctgacacat gcagctcccg gagacgggtc cagcttctct gtaagcggat	12720
gccgggagca gacaagcccg tcaggcgcg gcagcgggtg ttggcgggtg tcggggcgca	12780
gccatgaccc agtcacgtag cgatagcgga gtgtatactg gcttaactat gcggcatcag	12840
agcagattgt actgagagt caccatatgc ggtgtgaaat accgcacaga tgcgtaagga	12900
gaaaataccg catcaggcgc tcttcgctt cctcgctcac tgactcgtg cgctcggtcg	12960
ttcggctcgc gcgagcggta tcagctcact caaaggcggc aatacgggta tccacagaat	13020

caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	13080
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa	13140
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	13200
cccctggaag ctccctcgtg cgctctctg ttcggacct gccgcttacc ggatacctgt	13260
ccgcctttct cccttcggga agcgtggcg tttctcatag ctacacgctgt aggtatctca	13320
gttcgggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg	13380
accgctgcgc cttatccggt aactatcgtc ttgagtcctaa cccggtaaga cagcacttat	13440
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	13500
cagagtctct gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct	13560
gcgctctgct gaagccagtt accctcgaa aaagagttgg tagctcttga tccggcaaac	13620
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa	13680
aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaaacgaa	13740
actcacgtta agggattttg gtcatgcatg atatatctcc caatttgtgt agggcttatt	13800
atgcacgctt aaaaataata aaagcagact tgacctgata gtttggtgt gagcaattat	13860
gtgcttagtg catctaacgc ttgagttaag ccgcgcgcgc aagcggcgct ggcttgaacg	13920
aattttctagc tagacattat ttgccgacta ccttggtgat ctgcctttc acgtagtggg	13980
caaattcttc caactgatct gcgcgcgagg ccaagcgatc ttcttcttgt ccaagataag	14040
cctgtctagc ttcaagtatg acgggctgat actgggcggc caggcgctcc attgccagt	14100
cggcagcgac atccttcggc gcgattttgc cggttactgc gctgtaccaa atcggggaca	14160
acgtaagcac tacatttcgc tcatcgccag cccagtcggg cggcgagtgc catagcgta	14220
aggtttcatt tagcgctca aatagatcct gttcaggaa cggatcaaag agttcctccg	14280
ccgctggacc taccaaggca acgctatggt ctcttgcttt tgtcagcaag atagccagat	14340
caatgtcgat cgtggctggc tcgaagatac ctgcaagaat gtcattgcgc tgcattctc	14400
caaattgcag ttcgcgctta gctggataac gccacggaat gatgtcgtcg tgcacaacaa	14460
tggtgacttc tacagcgcg agaatctcgc tctctccagg ggaagccgaa gtttccaaaa	14520
ggtcgttgat caaagctcgc cgcgttggtt catcaagcct tacggtcacc gtaaccagca	14580
aatcaatata actgtgtggc ttcaggccgc catccactgc ggagccgtac aaatgtacgg	14640
ccagcaacgt cggttcgaga tggcgctcga tgacgccaac tacctctgat agttgagtcg	14700
atacttcggc gatcaccgct tccccatga tgtttaactt tgttttaggg cgactgccct	14760
gctgcgtaac atcgttgctg ctccataaca tcaaacatcg acccagggcg taacgcgctt	14820
gctgcttga tgcggaggc atagactgta ccccaaaaa acagtcataa caagccatga	14880
aaaccgccac tgcgttccat g	14901

```

<210> 63
<211> 51
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 63
ggggacaact ttgtatagaa aagttcacgg gcaggacata gggactacta c 51

<210> 64
<211> 51
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 64
ggggactgct tttttgtaca aacttggatt tatgataaaa atgtcgggtt c 51

<210> 65
<211> 50
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 65
ggggacaact ttgtatagaa aagttctgca gcaaatttac acattgccac 50

<210> 66
<211> 50
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 66
ggggactgct tttttgtaca aacttgactg gctatgaaga aattataatc 50

<210> 67
<211> 15029
<212> ДНК
<213> штучна

<220>
<223> векторна послідовність

<400> 67
gacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaataac cgattattct 60
aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt 120
taaaactgaag gcgggaaacg acaatcagat ctagtaggaa acagctatga ccatgattac 180

```

gccaaagctat cgattacgcc aagctatcaa ctttgtatag aaaagttgcc atgattacgc	240
caagcttgca tgcccatatg ctcgaggcgg ccgcagatat cagatctggt cgaccacggg	300
caggacatag ggactactac aagcatagta tgcttcagac aaagagctag gaaagaactc	360
ttgatggagg ttaagagaaa aaagtgctag aggggcatag taatcaaact tgtcaaaacc	420
gtcatcatga tgagggatga cataatataa aaagttgact aaggtcttgg tagtactctt	480
tgattagtat tatatattgg tgagaacatg agtcaagagg agacaagaaa ccgaggaacc	540
atagtttagc aacaagatgg aagttgcaa gttgagctag ccgctcgatt agttacatct	600
cctaagcagt actacaagga atggtctcta tactttcatg tttagcacat ggtagtgcgg	660
attgacaagt tagaaacagt gcttaggaga caaagagtc gtaaaggat tgaaagagt	720
aagttgatgc tcgacaggtc aggagaagtc cctccgccag atggtgacta ccaaggggtt	780
ggtatcagct gagacccaaa taagattctt cgggtgaacc agtggttcga ccgagactct	840
tagggtggga tttcactgta agatttgtgc attttgttga atataaattg acaatttttt	900
ttatttaatt atagattatt tagaatgaat tacatattta gtttctaaca aggatagcaa	960
tggtatggga tgggtacagg ttaaaccatat ctattaccca cccatctagt cgtcgggttt	1020
tacacgtacc caccggtta cataaaccag accggaattt taaaccgtac ccgtccgtta	1080
gcgggtttca gatttaccgg tttaatcggg taaaacctga ttactaaata tatatttttt	1140
atttgataaa caaaacaaaa atgttaatat tttcatattg gatgcaattt taagaaacac	1200
atattcataa atttccatat ttgtaggaaa ataaaaagaa aaatatattc aagaacacaa	1260
atttcaccga catgactttt attacagagt tggaattaga tctaacaatt gaaaaattaa	1320
aattaagata gaatatgttg aggaacatga catagtataa tgctgggtta cccgtcgggt	1380
aggatcagag gcggatacta ctaaatccat cccactcgct atccgataat cactgggttc	1440
gggtataccc attcccgta acaggccttt ttaaccggat aatttcaact tatagtgaat	1500
gaattttgaa taaatagtta gaataccaaa atcctggatt gcatttgcaa tcaaattttg	1560
tgaaccgtta aattttgcat gtacttggga tagatataat agaaccgaat tttcattagt	1620
ttaatttata acttactttg ttcaaagaaa aaaaatatct atccaattta cttataataa	1680
aaaataatct atccaagtta cttattataa tcaacttgta aaaaggtaag aatacaaatg	1740
tggtagcgta cgtgtgatta tatgtgacga aatgttatat ctaacaaaag tccaaattcc	1800
catggtaaaa aaatcaaaa tgcatggcag gctgtttgta acctgggaat aagatgttgg	1860
ccaattctgg agccgccacg tacgcaagac tcagggccac gttctcttca tgcaaggata	1920
gtagaacacc actccaccca cctcctatat tagaccttgg cccaacctc cccaactttc	1980
ccatcccatc cacaagaaa ccgacatttt tatcataaat cactagtccc gggtaaccaa	2040
gtttgtacaa aaaagcaggc tggtaacctg tgcttaacaa ctctgggtgag ttctagtact	2100

tctgctatga tgcattctcat taccattttct taaattttctc tccctaaata ttccgagttc	2160
ttgattttttg ataaacttcag gttttctctt tttgataaat ctggtctttc catttttttt	2220
tttttggtgt taatttagtt tcttatgttc ttcgattgta ttatgcata tctgtgtttg	2280
gattctgtta gattatgtat tggatgaatat gtatgtgttt ttgcattgtc ggttttggtc	2340
ttaaaaatgt tcaaatctga tgatttgatt gaagcttttt tagtggttgg ttgattcttc	2400
tcaaaactac tgttaattta ctatcatgtt ttccaacttt gattcatgat gacacttttg	2460
ttctgctttg ttataaaatt ttggttggtt tgattttgta attatagtgt aattttgtta	2520
ggaatgaaca tgttttaata ctctgttttc gatttgtcac acattcgaat tattaatcga	2580
taatttaact gaaaattcat ggttctagat cttgttgta tcagattatt tgtttcgata	2640
attcatcaaa tatgtagtcc ttttgcgtat ttgcgactgt ttcatttttt ctcaaaattg	2700
ttttttgtta agtttatcta acagttatcg ttgtcaaaag tctctttcat ttgcaaaat	2760
cttctttttt tttttgttg taactttgtt ttttaagcta cacatttagt ctgtaaaata	2820
gcacgcagga acagttgtct tagtagactt gcattgttctt gtaacttcta tttgtttcag	2880
ttgtttgatg actgccttga ttttgtagggt caaacatgg aagacgcca aaacataaag	2940
aaaggcccg cgccattcta tccgctggaa gatggaaaccg ctggagagca actgcataag	3000
gctatgaaga gatacgccct ggttctcgga acaattgctt ttacagatgc acatatcgag	3060
gtggacatca cttacgctga gtacttcgaa atgtccgttc ggttggcaga agctatgaaa	3120
cgatatgggc tgaatacaaa tcacagaatc gtcgtatgca gtgaaaactc tcttcaattc	3180
tttatgccgg tgttgggcgc gttatttate ggagtgcag ttgcgccgc gaacgacatt	3240
tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt ggtgttcgtt	3300
tccaaaaagg ggttgcaaaa aattttgaac gtgcaaaaa agctcccaat catccaaaa	3360
attattatca tggattctaa aacggattac cagggatttc agtcgatgta cacttctgtc	3420
acatctcacc tacctcccg ttttaatgaa tacgattttg tgccagagtc cttcgatagg	3480
gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc taaagggtgc	3540
gctctgcctc atagaactgc ctgcgtgaga ttctgcgatg ccagagatcc tatttttggc	3600
aatcaaatca ttccggatac tgcgatttta agtggtgttc cattccatca cggttttgga	3660
atgtttacta cactcggata tttgatatgt ggatttcgag tctgtttaat gtatagattt	3720
gaagaagagc tgtttctgag gagccttcag gattacaaga ttcaaagtgc gctgctggtg	3780
ccaaccctat tctccttctt cgccaaaagc actctgattg acaaatcga tttatctaat	3840
ttacacgaaa ttgcttctgg tggcgtccc ctctctaagg aagtcgggga agcgggtgcc	3900
aagaggttcc atctgccag taccaggcaa ggatatgggc tcaactgagac tacatcagct	3960
attctgatta cacccgagg ggatgataaa cggggcgagg tcggtaaagt tgttccattt	4020

tttgaagcga aggttgtgga tctggatacc gggaaaacgc tgggcgttaa tcaaagaggc	4080
gaactgtgtg tgagaggtec tatgattatg tccggttatg taaacaatcc ggaagcgacc	4140
aacgccttga ttgacaagga tggatggcta cattctggag acatagctta ctgggacgaa	4200
gacgaacact tcttcacgtg tgaccgcctg aagtctctga ttaagtacaa aggetatcag	4260
gtggctcccc ctgaattgga atccatcttg ctccaacacc ccaacatctt cgacgcaggt	4320
gtcgcaggtc tccccgacga tgacgcccgt gaacttcccg ccgcccgtgt tgttttgag	4380
cacggaaga cgatgacgga aaaagagatc gtggattacg tcgccagtca agtaacaacc	4440
gcgaaaaagt tgcgcggagg agttgtgtt gtggacgaag taccgaaagg tcttaccgga	4500
aaactcgacg caagaaaaat cagagagatc ctcataaagg ccaagaaggg cggaaagatc	4560
gccgtgtaac tcgagcatat gggctcgaat tccccgac gtccaacat ttggcaataa	4620
agtttcttaa gattgaatcc tgttgccgtt ctgcgatga ttatcatata atttctgtt	4680
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt	4740
tttatgatta ggtcccgca attatacatt taatacgcga tagaaaaaa aatatagcgc	4800
gcaaaactagg ataaattatc gcgcgcgggt tcatctatgt tactagatcg ggaattcaag	4860
cttggcgtaa tcatggaccc agctttcttg tacaagtgg ggtaccaatt cgaatccaaa	4920
aattacggat atgaatatag gcataatcgt atccgaatta tccgtttgac agtagcaac	4980
gattgtacaa ttgcttcttt aaaaaaggaa gaaagaaaga aagaaaaaa tcaacatcag	5040
cgtaacaaa cggccccgtt acggcccaaa cggtcataa gagtaacggc gttaagcgtt	5100
gaaagactcc tatcgaaata cgtaaccgca aacgtgtcat agtcagatcc cctcttctt	5160
caccgcctca aacacaaaaa taatcttcta cagcctatat atacaacccc ccttctatc	5220
tctcttttct cacaattcat catctttctt tctctacccc caattttaag aaatcctct	5280
ttctctctct cattttcaag gtaaatctct ctctctctct ctctctctgt tatctctgt	5340
tttaattagg tatgtattat tgctagtgtg ttaatctgct tatcttatgt atgccttat	5400
tgaatatctt tatcttggtc atctcatccg tttagaagct ataaatttgt tgatttgact	5460
gtgtatctac acgtgggtat gtttatatct aatcagatat gaatttcttc atattgttg	5520
gtttgtgtgt accaatccga aatcgttgat ttttttctt taatcgtgta gctaattgta	5580
cgtatacata tggatctacg tatcaattgt tcatctgttt gtgtttgtat gtatacagat	5640
ctgaaaacat cacttctctc atctgattgt gttgttacat acatagatat agatctgtta	5700
tatcattttt ttatttaatt gtgtatatat atatgtgcat agatctggat tacatgattg	5760
tgattattta catgattttg ttatttacgt atgtatatat gtagatctgg acttttttga	5820
gttgttgact tgattgtatt tgtgtgtgta tatgtgtgt ctgactctga tatgttatgt	5880

atgtgcagct gaacctggc ggcggaaca acaacaaca caacatcttc ttogatctcc	5940
ttctccacca aacctctcc ttctctctcc aaatcaccat taccaatctc cagattctcc	6000
ctcccattct ccttaaaccc caacaaatca tcctctctct cccgcgcgcg cgggatcaaa	6060
tccagctctc cctctccat ctccgcgtg ctcaacacaa ccaccaatgt cacaaccact	6120
ccctctccaa ccaaacctac caaacccgaa acattcatct cccgattcgc tccagatcaa	6180
ccccgcaaag gcgctgatat cctcgtcgaa gctttagaac gtcaaggcgt agaaaccgta	6240
ttcgcttacc ctggaggtag atcaatggag attcaccag ccttaaccgc ctctctctca	6300
atccgtaacg tccttctctg tcacgaacaa ggagggtgat tcgcagcaga aggatacgct	6360
cgatcctcag gtaaaccagg tatctgtata gccacttcag gtcccggagc tacaaatctc	6420
gttagcggat tagccgatgc gttgttagat agtgttcttc ttgtagcaat cacaggacaa	6480
gtccctcgtc gtatgattgg tacagatgcg ttcaagaga ctccgattgt tgaggtaacg	6540
cgttcgatta cgaagcataa ctatcttgtg atggatgttg aagatatccc taggattatt	6600
gaggaagctt tctttttagc tacttctggt agacctggac ctgttttggg tgatgttctc	6660
aaagatattc aacaacagct tgcgattcct aattgggaac aggctatgag attacctggt	6720
tatatgtcta ggatgcctaa acctccgaa gattctcatt tggagcagat tgttaggttg	6780
atttctgagt ctaagaagcc tgtgttgtat gttggtggtg gttgtttgaa ttctagcgat	6840
gaattgggta ggtttgttga gcttacgggg atccctgttg cgagtaacgt gatggggctg	6900
ggatcttata cttgtgatga tgagttgtcg ttacatatgc ttggaatgca tgggactgtg	6960
tatgcaaat acgctgtgga gcatagtgat ttgtgttgg cgtttggggg aaggtttgat	7020
gatcgtgtca cgggtaagct tgaggctttt gctagtaggg ctaagattgt tcatattgat	7080
attgactcgg ctgagattgg gaagaataag actcctcatg tgtctgtgtg tggatgatt	7140
aagctggctt tgcaagggat gaataagggt cttgagaacc gagcggagga gcttaagctt	7200
gattttggag ttggaggaa tgagttgaac gtacagaaac agaagtttcc gttgagcttt	7260
aagacgtttg gggaagctat tcctccacag tatgcgatta aggtccttga tgagttgact	7320
gatggaaaag ccataataag tactggtgtc gggcaacatc aaatgtgggc ggcgcagttc	7380
tacaattaca agaaaccaag gcagtggcta tcatcaggag gccttgagc tatgggattt	7440
ggacttctcg ctgcgattgg agcgtctgtt gctaaccctg atgcgatagt tgtggatatt	7500
gacggagatg gaagctttat aatgaatgtg caagagctag ccactattcg tgtagagaat	7560
cttcagtgta aggtactttt attaaacaac cagcatcttg gcatggttat gcaatgggaa	7620
gatcggttct acaaagctaa ccgagctcac acatttctcg gggatccgc tcaggaggac	7680
gagatattcc cgaacatgtt gctgtttgca gcagcttgcg ggattccagc ggcgagggtg	7740
acaaagaaag cagatctccg agaagctatt cagacaatgc tggatacacc aggaccttac	7800

ctgttgatg tgatttgtcc gcaccaagaa catgtgttgc cgatgatccc gaatggtggc	7860
actttcaacg atgtcataac ggaaggagat ggcgggatta aatactgaga gatgaaaccg	7920
gtgattatca gaacctttta tgggtcttgt atgcatatgg taaaaaaact tagtttgcaa	7980
tttctgtttt gttttggtaa tttgagtttc ttttagttgt tgatctgcct gctttttggt	8040
ttacgtcaga ctactactgc tgttgttgtt tggtttccct tctttcattt tataaataaa	8100
taatccggtt cggtttactc cttgtgactg gctcagtttg gttattgcga aatgcgaatg	8160
gtaaattgag taattgaaat tcgttattag ggttctaagc tgttttaaca gtcactgggt	8220
taatattctc cgaatcttgc atggaaaatg ctcttaccat tgggtttttaa ttgaaatgtg	8280
ctcatatggg ccgtggtttc caaattaaat aaaactacga tgtcatcgag aagtaaaatc	8340
aactgtgtcc acattatcag ttttgtgtat acgatgaaat agggtaattc aaaatctagc	8400
ttgatatgcc ttttggttca ttttaacctt ctgtaaacat tttttcagat tttgaacaag	8460
taaatccaaa aaaaaaaaaa aaaatctca actcaacact aaattatttt aatgtataaa	8520
agatgcttaa aacatttggc ttaaaagaaa gaagctaaaa acatagagaa ctcttgtaaa	8580
ttgaagtatg aaaatatact gaattgggta ttatatgaat ttttctgatt taggattcac	8640
atgatccaaa aaggaaatcc agaagcacta atcagacatt ggaagtagga atatttcaa	8700
aagttttttt tttttaagta agtgacaaaa gcttttaaaa aatagaaaag aaactagtat	8760
taaagttgta aatttaataa acaaaagaaa ttttttatat tttttcattt ctttttccag	8820
catgaggtta tgatggcagg atgtggattt catttttttc cttttgatag ccttttaatt	8880
gatctattat aattgacgaa aaaatattag ttaattatag atatatatta ggtagtatta	8940
gcaatttaca cttccaaaag actatgtaag ttgtaaatat gatgcgttga tctcttcac	9000
attcaatggt tagtcaaaaa aataaaagct taactagtaa actaaagtag tcaaaaattg	9060
tacttttagt taaaatatta catgaataat ccaaaacgac atttatgtga acaaaaaaca	9120
atatctagag tcgacttaat taaactagtg gcgcgccaat tgactagtag gcctatcgat	9180
taattaaggc cgcctcgagc atatgggcat gcaagcttgg cgtaatcatg gcaactttat	9240
tatacatagt tgataattca ctggccggat ctgcttggtg ataattgtca ttagattggt	9300
tttatgcata gatgcactcg aaatcagcca attttagaca agtatcaaac ggatgttaat	9360
tcagtacatt aaagacgtcc gcaatgtgtt attaagttgt ctaagcgtca atttgtttac	9420
accacaatat atcctgccac cagccagcca acagctcccc gaccggcagc tcggcacaaa	9480
atcaccacgc gttaccacca cgccggccgg ccgcattggt ttgaccgtgt tcgccggcat	9540
tgccgagttc gagcgttccc taatcatcga ccgcacccgg agcgggcgcg aggccgccaa	9600
ggcccgaggc gtgaagtttg gccccgccc taccctcacc ccggcacaga tcgcgcacgc	9660
ccgcgagctg atcgaccagg aaggccgcac cgtgaaagag gcggctgcac tgcttgccgt	9720

gcacgcctcg accctgtacc gcgcacttga gcgcagcgag gaagtgaacg ccaccgaggc	9780
caggcgggcg ggtgccttcc gtgaggacgc attgaccgag gccgacgccc tggcgggccg	9840
cgagaatgaa cgccaagagg aacaagcatg aaaccgcacc aggacggcca ggacgaaccg	9900
tttttcatta ccgaagagat cgaggcgagg atgatcgcg cggggtacgt gttcgagccg	9960
cccgcgcacg tctcaaccgt gcggctgcat gaaatcctgg ccggtttgtc tgatgccaa	10020
ctggcgccct ggccggccag cttggccgct gaagaaaccg agcgccgccc tctaaaaagg	10080
tgatgtgtat ttgagtaaaa cagcttgctg catgcggctg ctgcgtatat gatgcgatga	10140
gtaataaacc aaatacgcga ggggaacgca tgaaggttat cgctgtactt aaccagaaag	10200
gcgggtcagg caagacgacc atcgcaaccc atctagcccg cgccctgcaa ctgcggggg	10260
ccgatgttct gttagtcatg tccgatcccc agggcagtg cgcgattgg gcggccgtgc	10320
gggaagatca accgctaacc gttgtcggca tcgaccgccc gacgattgac cgcgacgtga	10380
aggccatcgg ccggcgcgac ttcgtagtga tcgacggagc gccccaggcg gcggacttgg	10440
ctgtgtccgc gatcaaggca gccgacttcg tgctgattcc ggtgcagcca agcccttacg	10500
acatatgggc caccgcccgc ctggtggagc tggtaagca gcgcattgag gtcacggatg	10560
gaaggetaca agcggccctt gtctgttcgc gggcgatcaa aggcacgcgc atcgcggtg	10620
aggttgccga ggcgctggcc ggttacgagc tgcccattct tgagtcctct atcacgcgc	10680
gcgtgagcta cccaggcaact gccgcgccc gcaaacctt tcttgaatca gaaccgagg	10740
gcgacgtgc ccgcgaggtc caggcgctgg ccgctgaaat taaatcaaaa ctcatctgag	10800
ttaatgaggt aaagagaaaa tgagcaaaag cacaacacg ctaagtgcg gcgctccgag	10860
cgcacgcagc agcaagctg caacgttggc cagcctggca gacacgccag ccatgaagcg	10920
ggtcaacttt cagttgccg cgaggatca caccaagctg aagatgtacg cggtacgcca	10980
aggcaagacc attaccgagc tgctatctga atacatcgcg cagctaccag agtaaatgag	11040
caaatgaata aatgagtaga tgaattttag cggctaaagg aggcggcatg gaaaatcaag	11100
aacaaccagg caccgacgcc gtggaatgcc ccatgtgtgg aggaacgggc ggttggccag	11160
gcgtaagcgg ctgggttgtc tgccggccct gcaatggcac tggaaacccc aagcccgagg	11220
aatcgcgctg agcggctgca aaccatccg ccggtacaa atcggcgcg cgctgggtga	11280
tgacctggtg gagaagtga aggccgcgca ggccgcccag cggcaacgca tcgaggcaga	11340
agcacgcccc ggtgaatcgt ggcaagcggc cgctgatcga atccgcaaag aatcccgcca	11400
accgcggca gccggtgcgc cgtcgattag gaagccgccc aagggcgacg agcaaccaga	11460
ttttttcgtt ccgatgctct atgacgtggg caccgcgat agtcgcagca tcatggacgt	11520
ggcggtttct cgtctgtcga agcgtgaccg acgagctggc gaggtgatcc gctacgagct	11580
tccagacggg cacgtagagg ttccgcagg gccggccggc atggccagtg tgtgggatta	11640

cgacctggta	ctgatggcgg	tttcccatct	aaccgaatcc	atgaaccgat	accgggaagg	11700
gaagggagac	aagcccggcc	gcgtgttccg	tccacacgtt	gcggacgtac	tcaagttctg	11760
ccggcgagcc	gatggcggaa	agcagaaa	cgacctggta	gaaacctgca	ttcgggttaa	11820
caccacgcac	gttgccatgc	agcgtacgaa	gaaggccaag	aacggccgcc	tggtgacggt	11880
atccgaggg	gaagccttga	ttagccgcta	caagatcgta	aagagcgaaa	ccgggcggcc	11940
ggagtacatc	gagatcgagc	tagctgattg	gatgtaccgc	gagatcacag	aaggcaagaa	12000
ccgggacgtg	ctgacgggtc	accccgatta	ctttttgatc	gatcccgga	tcggccgttt	12060
tctctaccgc	ctggcacgcc	gcgccgcagg	caaggcagaa	gccagatggt	tgttcaagac	12120
gatctacgaa	cgcagtggca	gcgccggaga	gttcaagaag	ttctgtttca	ccgtgcgcaa	12180
gctgatcggg	tcaaatgacc	tgccggagta	cgatttgaag	gaggaggcgg	ggcaggctgg	12240
cccgatccta	gtcatgcgct	accgcaacct	gatcgagggc	gaagcatccg	ccggttccta	12300
atgtacggag	cagatgctag	ggcaaatg	cctagcaggg	gaaaaagg	gaaaaggctt	12360
ctttcctgtg	gatagcacgt	acattgggaa	cccaaagccg	tacattggga	accggaaccc	12420
gtacattggg	aaccctaaagc	cgtacattgg	gaaccggcca	cacatgtaag	tgactgat	12480
aaaagagaaa	aaaggcgatt	tttccgccta	aaactcttta	aaacttatta	aaactcttaa	12540
aaccgcctg	gcctgtgcat	aactgtctgg	ccagcgcaca	gccgaagagc	tgcaaaaagc	12600
gcctaccctt	cggtcgctgc	gtccctacg	ccccgcgct	tcgcgtcggc	ctatcgcggc	12660
cgctggccgc	tcaaaaatgg	ctggcctacg	gccaggcaat	ctaccagggc	gcggacaagc	12720
cgcgccgtcg	ccactcgacc	gcgggcggcc	acatcaaggc	accctgcctc	gcgcgtttcg	12780
gtgatgacgg	tgaaaacctc	tgacacatgc	agctcccgga	gacggtcaca	gcttgtctgt	12840
aaagcgatgc	cgggagcaga	caagcccgtc	agggcgcgtc	agcgggtgtt	ggcgggtgtc	12900
ggggcgagc	catgaccag	tcacgtagcg	atagcggagt	gtatactggc	ttaactatgc	12960
ggcatcagag	cagattgtac	tgagagtgc	ccatatgcgg	tgtgaaat	cgcacagatg	13020
cgtaaggaga	aaataccgca	tcaggcgctc	ttccgcttcc	tcgctcactg	actcgctgcg	13080
ctcggtcggt	cggtgcggc	gagcggatc	agctcactca	aaggcggtaa	tacggttatc	13140
cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	13200
gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	13260
tcacaaaaat	cgacgctcaa	gtcagagggt	gcgaaacccg	acaggactat	aaagatacca	13320
ggcggttccc	cctggaagct	ccctcgctgc	ctctcctgtt	cgcacctgc	cgcttaccgg	13380
atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	13440
gtatctcagt	tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccgt	13500
tcagcccgac	cgctgcgcct	tatccggtaa	ctatcgtctt	gagtccaacc	cggtaaagaca	13560

```

cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 13620
cgggtgctaca gagttcttga agtgggtggc taactacggc tacactagaa ggacagtatt 13680
tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 13740
cggcaaaaca accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg 13800
cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg 13860
gaacgaaaaa tcacgttaag ggatttttgt catgcatgat atatctccca atttgtgtag 13920
ggcttattat gcacgcttaa aaataataaa agcagacttg acctgatagt ttggctgtga 13980
gcaattatgt gcttagtgca tctaacgctt gagttaagcc gcgcgcgaa gcggcgctcg 14040
cttgaacgaa tttctagcta gacattattt gccgactacc ttggtgatct cgcctttcac 14100
gtagtggaca aattcttcca actgatctgc gcgcgaggcc aagcgatctt cttcttgtcc 14160
aagataagcc tgtctagctt caagtatgac gggctgatac tgggcccgca ggcgctccat 14220
tgcccagtcg gcagcgacat ccttcggcgc gattttgccc gttactgcgc tgtaccaa 14280
gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg gcgagttcca 14340
tagcgtttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag 14400
ttcctccgcc gctggacctt ccaaggcaac gctatgttct cttgcttttg tcagcaagat 14460
agccagatca atgtcgatcg tggtgggctc gaagatacct gcaagaatgt cattgcgctg 14520
ccattctcca aattgcagtt cgcgcttagc tggataacgc cagggaatga tgtcgtcgtg 14580
cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg aagccgaagt 14640
ttccaaaagg tcgttgatca aagctcgccg cggtgtttca tcaagcctta cggtcaccgt 14700
aaccagcaaa tcaatatcac tgtgtggctt caggccgccca tccactgcgg agccgtacaa 14760
atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta cctctgatag 14820
ttgagtcgat acttcggcga tcaccgcttc ccccatgatg tttaactttg ttttagggcg 14880
actgccttgc tgcgtaacat cggttgctgct ccataacatc aaacatcgac ccacggcgta 14940
acgcgcttgc tgcttgatg cccgaggcat agactgtacc ccaaaaaaac agtcataaca 15000
agccatgaaa accgccactg cgttccatg 15029

```

<210> 68
 <211> 38
 <212> ДНК
 <213> штучна

<220>
 <223> синтетична послідовність

<400> 68
 aataatggcg cgcctgggtg ttaaacactc tggtgagt

38

<210> 69

<211> 40
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 69
 aataatggcg cgcctttgac ctacaaaac aaagcagtea 40

 <210> 70
 <211> 36
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 70
 ttttttggcg cgccagttct ttgctttcga agttgc 36

 <210> 71
 <211> 37
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 71
 ttttttggcg cgcctactac gtactgtttt caattct 37

 <210> 72
 <211> 35
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 72
 aataaaggcg cgccgtccag aattttctcc attga 35

 <210> 73
 <211> 36
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 73
 aataaaggcg cgcctcttca ctatccaaag ctctca 36

 <210> 74
 <211> 39
 <212> ДНК
 <213> штучна

```

<220>
<223> синтетична послідовність

<400> 74
ttttatggcg cgcctagctt aatctcagat tcgaatcgt 39

<210> 75
<211> 42
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 75
ttttatggcg cgcctagtat ctacatacca atcatacaaa tg 42

<210> 76
<211> 34
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 76
ttttttggcg cgcctttcac gatttggaat ttga 34

<210> 77
<211> 38
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 77
ttttttggcg cgcctctaca acattaaaaac gaccatta 38

<210> 78
<211> 34
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 78
tatataggcg cgccagggtt tcgtttttgt ttca 34

<210> 79
<211> 37
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

```

```

<400> 79
tatataggcg cgccttatct cctgctcaaa gaaacca 37

<210> 80
<211> 35
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 80
tatataggcg cgccactggt taagcttcac tgtct 35

<210> 81
<211> 34
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 81
tatataggcg cgcctttctt ctaaagctga aagt 34

<210> 82
<211> 14495
<212> ДНК
<213> штучна

<220>
<223> векторна послідовність

<400> 82
gtgattttgt gccgagctgc cggtcgggga gctgttggt ggctggtggc aggatatatt 60
gtggtgtaaa caaattgacg cttagacaac ttaataacac attgcggacg tctttaatgt 120
actgaattta gttactgate actgattaag tactgatac ggtaccaagc ttccgcggct 180
gcagtgcagc gtgaccgggt cgtgccctc tctagagata atgagcattg catgtctaag 240
ttataaaaaa ttaccacata tttttttgt cacacttgtt tgaagtgcag tttatctatc 300
tttatacata tatttaaaact ttactctacg aataatataa tctatagtac tacaataata 360
tcagtgtttt agagaatcat ataatgaac agttagacat ggtctaaagg acaattgagt 420
atthtgacaa caggactcta cagttttatc ttttttagtg gcattgtgtc tccctttttt 480
ttgcaaatag cttcacctat ataatacttc atccatttta ttagtacatc catttagggg 540
ttaggggttaa tggtttttat agactaattt ttttagtaca tctattttat tctatttttag 600
cctctaaatt aagaaaaact aaactctatt ttagtttttt tatttaatag tttagatata 660
aaatagaata aaataaagt actaaaaatt aaacaaatac cctttaagaa attaaaaaaa 720
ctaaggaaac atttttcttg ttctgagtag ataatgcag cctgttaaac gccgtcgacg 780
agtcctaacg acaccaacca gcgaaccagc agcgtcgcgt cgggccaagc gaagcagacg 840

```

gcacggcatc tctgtcgtg cctctggacc cctctcgaga gttecgctcc accgttggac	900
ttgtctcgct gtcggcatcc agaaattgcg tggcggagcg gcagacgtga gccggcacgg	960
caggcggcct cctctctctc tcacggcacc ggcagctacg ggggattcct tteccaccgc	1020
tccttcgctt tcccttctc gcccgccgta ataatagac accccctcca caccctcttt	1080
ccccaacctc gtgttggtcg gagcgacac acacacaacc agatctcccc caaatccacc	1140
cgtcggcacc tccgcttcaa ggtacgcgc tcgtctctcc cccccccccc cctctctacc	1200
ttctctagat cggcggtccg gtccatggtt agggcccggt agttctactt ctgttcatgt	1260
ttgtgttaga tccgtgtttg tgtagatcc gtgtgctag cgttcgtaca cggatgcgac	1320
ctgtacgtca gacacgttct gattgctaac ttgccagtgt ttctcttttg ggaatcctgg	1380
gatggctcta gccgttcgc agacgggac gatttcatga ttttttttgt ttcggtgcat	1440
agggtttggg ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc	1500
atcttttcat gctttttttt gtcttggttg tgatgatgtg gtctgggttg gccgtcgttc	1560
tagatcggag tagaattctg tttcaacta cctggtggat ttattaattt tggatctgta	1620
tggtgtgccc atacatattc atagttacga attgaagatg atggatggaa atatcgatct	1680
aggataggtg tacatgttga tgcgggtttt actgatgcat atacagagat gctttttgtt	1740
cgtctggttg tgatgatgtg gtgtggttg gccgtcgttc attcgttcta gatcggagta	1800
gaatactggt tcaaaactacc tgggtgtattt attaatcttg gaactgtatg tgtgtgtcat	1860
acatcttcat agttacgagt ttaagatgga tggaaatata gatctaggat aggtatacat	1920
gttgatgtgg gttttactga tgcataata tgatggcata tgcagcatct attcatatgc	1980
tetaaccttg agtacctatc tattataata aacaagtatg ttttataatt atttcgatct	2040
tgatatactt ggatgatggc atatgcagca gctatatgtg gattttttta gccctgcctt	2100
catacgctat ttatttgctt ggtactgttt cttttgtcga tgctcaccct gttgtttggt	2160
gttacttctg cagcccgggg gatccactag ttctagaaac catggccacc gccgccgccg	2220
cgtctaccgc gtcactggc gccactaccg ctgcgccaa ggcgaggcgc cgggcgcacc	2280
tcctggccac ccgcccgcgc ctccgcgcgc ccatcaggtg ctacgaggcg tcacccgcc	2340
tgccgatggc tccccggcc accccgctcc ggcgtgggg cccaccgat ccccgcaagg	2400
gcgccgacat cctcgtcgag tccctcgagc gctgcggcgt ccgcgacgtc ttgcctacc	2460
ccggcgacac gtccatggag atccaccagg cactcaccgc ctccccgtc atcgccaacc	2520
acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgtctctcgg	2580
gccgcgtcgg cgtctgcac gccacctccg gccccggcgc caccacacct gtctccgcgc	2640
tcgccgacgc gctgctcgat tccgtcccca tggtcgccat caccgggacag gtgcgcgac	2700
gcatgattgg caccgacgcc ttccaggaga cgcccatcgt cgaggtcacc cgtccatca	2760

ccaagcacia ctacctggtc ctcgacgtcg acgacatccc ccgcgtcgtg caggaggctt	2820
tcttcctcgc ctccctctgt cgaccggggc cgggtgcttg cgacatcccc aaggacatcc	2880
agcagcagat ggcgggtgct gtctgggaca agcccatgag tctgcctggg tacattgcgc	2940
gccttcccaa gccccctgcg actgagttgc ttgagcagggt gctgcgtctt gttggtgaat	3000
cccggcgccc tgttctttat gttggcggtg gctgcgcagc atctggtgag gatttgcgac	3060
gctttgtgga gctgactgga atcccggta caactactct tatgggcctc ggcaacttcc	3120
ccagcgacga cccactgtct ctgcgcgtgc taggtatgca tggcacgggt tatgcaaatt	3180
atgcagtgga taaggccgat ctgttgcttg caettggtgt gcggtttgat gatcgtgtga	3240
cagggaagat tgaggtcttt gcaagcaggg ctaagattgt gcacgttgat attgatccgg	3300
ctgagattgg caagaacaag cagccacatg tgtccatctg tgcagatggt aagcttgctt	3360
tgcagggcat gaatgctctt cttgaaggaa gcacatcaaa gaagagcttt gactttggct	3420
catggaacga tgagttggat cagcagaaga gggaaatccc ccttgggtat aaaacatcta	3480
atgaggagat ccagccacia tatgctatcc aggttcttga tgagctgacg aaaggcgagg	3540
ccatcatcgg cacaggtgtt gggcagcacc agatgtgggc ggcacagtac tacacttaca	3600
agcgcccaag gcagtgggtg tcttcagctg gtcttggggc tatgggattt ggtttgccgg	3660
ctgctgctgg tgcttctgtg gccaaaccag gtgttactgt tgttgacatc gatggagatg	3720
gtagctttct catgaacgtt caggagctag ctatgatccg aattgagaac ctcccgtga	3780
aggtctttgt gctaaacaac cagcacctgg ggaatgggtg gcagtgggag gacaggttct	3840
ataaggccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatatc	3900
cagatttctg gacgatccgc aaagggttca acattccagc ggtccgtgtg acaaagaaga	3960
acgaagtccg cgcagcgata aagaagatgc tcgagactcc agggccgtac ctcttgata	4020
taatcgctcc acaccaggag catgtgttgc ctatgatccc taatggtggg gctttcaagg	4080
atatgatcct ggaatggatg ggcaggactg tgtactgac taaaatccag caagcaactg	4140
atctaaaatc cagcaagcac cgctccctg ctagtacaag ggtgatatgt ttttatctgt	4200
gtgatgttct cctgtattct atcttttttt gtaggcccgc agctatctgt tatggtaatc	4260
ctatgtagct tccgacctg taattgtgta gtctgttgtt ttccttctgg catgtgtcat	4320
aagagatcat ttaagtgcct ttgctacat ataaataaga taataagcac tgctatgcag	4380
tggttctgaa ttggcttctg ttgccaaatt taagtgtcca actggctcctt gcttttgttt	4440
tcgctatttt tttcttttt tagttattat tatattggta atttcaactc aacatatgat	4500
gtatggaata atgctagggc tgcaatttca aactatttta caaaccagaa tggcattttc	4560
gtggtttgag gggagtgaat aaaaatgagg catttgactg aattagttac ctgatccatt	4620
ttcgtggttt ggaatcattg aattaaatc cattctaata atagtaattt tggcatatat	4680

caattaagtt aattcgggtt tatgcaaaat atatttgtat actattatta tcaagatgtc	4740
ggagatattt atatgctaca tttttactat acaggagtga gatgaagagt gtcagtgaag	4800
ttacacagta gaaacaaatt ctattaatgc ataaaatcat ttccatcatc caccctatga	4860
atttgagata gacctatata taaactttga aaagtgggtg aatatcaaat tccaaattaa	4920
ataagttatt ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgcctc	4980
taaagctaat ttggaacctc aaactttctt agcattggag gggattgaga aaaaatatta	5040
attcattttc atctcaatca ttcaatctcc aaagagattt gagttcctta ttagtctgtt	5100
ccatgcatca aatcggtcca atgtgtcatt atttgccatg acgattgacg agttgttctg	5160
gggcctagcg ctttccacgc cgatgtgctg gggcctggtc ctggagaaga cagcttgata	5220
tttaaagcta tcaattgttt caattgatcc ccaacttcatt tttctaaatg tagaaaacgg	5280
tgacgtataa gaaaaagaat gaattaggac ttttattccg tacactaatc tagagcggcc	5340
caagcttgta cactagtacg cgtcaattga tttaaattta attaatcccg tgtcgtcaa	5400
tgtgatacta ctagcatagt actagtacca tgcatacaca cagcaggctg gccgcctgga	5460
tggatcgatg atgatactac atcatcctgt catccatcca ggcgatctag aaggggcgtg	5520
gctagctagc aaactgtgac cgggttttct acgccgataa taatactttg tcatggtaca	5580
gacgtacagt actggttata tatactctga gatttcaact gaaaagctag gatagctaga	5640
ttaatctctg agaaacacag ataaaattcg agcttggcta tagatgacaa aacggaagac	5700
gcatgcattg gacgacgtat gcaatgcgag cgcgtctcgt gtcgtcccg ccaagtctgg	5760
cgatctcagc ccacgtgctc aacagctcaa ggactgttcg tcaccagcgt taaattcatt	5820
gaagggatga gcgatttcgg catttgatcat tgcttgtagc tatatatata tatccaacag	5880
atttctctca agctttttgta tgctggaatg taaagtctag cttatacgac agcacgtgca	5940
gatatattaa cgtcattatt aggtggagag caagatgcat gatctggtag aaattgtcga	6000
aaacacaaga gagagtgaag tgcacacttc tggatatagga gtgtatacgc cgctgggtgg	6060
tgggcaatgc gcgccgcaat attggccaat gaaacctagc aacgcccact cgccacgccc	6120
catgaatggc ccccgccacg cagcgagcca gccagtgcgc gcgcgcgcc cagccggagt	6180
cggcggaacg cgccacgggg gacgaggcgc ccgagggcgc aggcagcgcg gcatggcaag	6240
caagccgaag cgggcaagcg acctgcatgc agccccctgc cctcgccctc gtcagtcgtc	6300
ccagcctccc actggaatcc acccaaccgc ccttctctct ccaaagcacg cgccccgcga	6360
ctcgccctcg cctacgtgtc ggcagcgtcc ccgcccgtcg cccacgtacc ccgccccgtt	6420
ctcccaagtg cccctccctc tgcgcgcgtc cgattggctg acccgccctt cttaagccgc	6480
gccagcctcc tgtccgggcc ccaacgcctg gctccgtcgt cgtctccgcc cccagagtga	6540
tcgagcccac tgacctggcc cccgagcctc agctcgtgag tccggcgccg ctgggtgctta	6600

aacactctgg tgagttctag tacttctgct atgatcgatc tcattaccat ttcttaaatt	6660
tctctcccta aatattccga gttcttgatt ttgataaact tcagggtttc tctttttgat	6720
aaatctgggc tttecatitt ttttttttgg tggttaattt agtttccctat gttcttcgat	6780
tgtattatgc atgatctgtg tttggattct gttagattat gtattgggtga atatgtatgt	6840
gtttttgcat gtctgggttt ggtcttaaaa atgttcaaat ctgatgattt gattgaagct	6900
tttttagtgt tggtttgatt cttctcaaaa ctactgttaa ttactatca tgttttccaa	6960
ctttgattca tgatgacact tttgttctgc ttgtttataa aattttgggt ggtttgattt	7020
tgtaattata gtgaattttt gttaggaatg aacatgtttt aatactctgt ttctgatttg	7080
tcacacattc gaattattaa tcgataattt aactgaaaat tcatggttct agatcttgtt	7140
gtcatcagat tatttgtttc gataattcat caaatatgta gtctttttgc tgatttgcca	7200
ctgtttcatt ttttctcaaa attgtttttt gttaagttaa tctaacagtt atcgttgtca	7260
aaagtctctt tcattttgca aaatctctct ttttttttgg ttgttaactt tgttttttaa	7320
gtacacatt tagtctgtaa aatagcatcg aggaacagtt gtcttagtag acttgcatgt	7380
tcttgtaact tctatttgtt tcagtttgtt gatgactgct ttgattttgt aggtcaaagg	7440
cgcgcgggat ccccggtggg tcagtcctt atgttacgtc ctgtagaaac cccaaccct	7500
gaaatcaaaa aactcgacgg cctgtgggca ttcagtctgg atcgcaaaa ctgtggaatt	7560
gatcagcgtt ggtgggaaag cgcgttacaa gaaagccggg caattgctgt gccaggcagt	7620
tttaacgac agttcgccga tgcagatatt cgtaattatg cgggcaacgt ctggatcag	7680
cgcgaagtct ttataccgaa aggttgggca ggccagcgta tcgtgctgcg ttctgatgcg	7740
gtcactcatt acggcaaagt gtgggtcaat aatcaggaaag tgatggagca tcagggcggc	7800
tatacgccat ttgaagccga tgcacgccc tatgttattg ccgggaaaag tgtacgtaag	7860
tttctgcttc tacctttgat atatataaa taattatcat taattagtag taatataata	7920
tttcaaatat ttttttcaaa ataaaagaat gtagtatata gcaattgctt ttctgtagtt	7980
tataagtgtg tatattttaa ttataactt ttctaataa tgacccaaat ttgttgatgt	8040
gcaggatatca ccgttttgtt gaacaacgaa ctgaactggc agactatccc gccgggaatg	8100
gtgattaccg acgaaaacgg caagaaaaag cagtcttact tccatgattt ctttaactat	8160
gccggaatcc atcgacgcgt aatgctctac accacgccga acacctgggt ggacgatatc	8220
accgtgggtg cgcagtgcgc gcaagactgt aaccacgcgt ctgttgactg gcaggtgggtg	8280
gccaatgggt atgtcagcgt tgaactgcgt gatgcggatc aacaggtggt tgcaactgga	8340
caaggcacta gcgggacttt gcaagtgggt aatccgcacc tctggcaacc ggggtgaagg	8400
tatctctatg aactgtgcgt cacagccaaa agccagacag agtgtgatat ctaccgcgtt	8460
cgcgtcggca tccggtcagt ggcagtgaag ggcgaacagt tccatgattaa ccacaaaccg	8520

tcttacttta ctggctttgg tctcatgaa gatgaggact tgcgtggcaa aggattcgat	8580
aacgtgctga tgggtgcacga ccacgcatta atggactgga ttggggccaa ctctaccgt	8640
acctcgcat acccttacgc tgaagagatg ctcgactggg cagatgaaca tggcatcgtg	8700
gtgattgatg aaactgctgc tgtcggcttt aacctctctt taggcattgg tttcgaagcg	8760
ggcaacaagc cgaagaact gtacagcgaa gaggcagtca acggggaac tcagcaagcg	8820
cacttacagg cgattaaaga gctgatagcg cgtgacaaa accaccaag cgtggtgatg	8880
tggagtattg ccaacgaacc ggatacccg ccgcaagggtg caggggaata tttcgcgcca	8940
ctggcggaag caacgcgtaa actcgacccg acgcggtccga tcacctgcgt caatgtaatg	9000
ttctgcgacg ctacacccga taccatcagc gatctctttg atgtgctgtg cctgaaccgt	9060
tattacggat ggtatgtcca aagcggcgat ttggaacgg cagagaaggt actggaaaa	9120
gaacttctgg cctggcagga gaaactgcat cagccgatta tcatcaccga atacggcgtg	9180
gatacgtag ccgggctgca ctcaatgtac accgacatgt ggagtgaaga gtatcagtg	9240
gcatggctgg atatgtatca ccgctcttt gatcgcgtca gcgccgtcgt cgggtaacag	9300
gtatggaatt tcgccgattt tgcgacctcg caaggcatat tgcgcgttgg cggtaacaag	9360
aaagggatct tcaactcgca ccgcaaacg aagtgcggcg cttttctgct gcaaaaacgc	9420
tggactggca tgaactcgg tgaaaaacg cagcaggag gcaacaatg aatcaaac	9480
tctctggcg caccatcgct ggctacagcc tcgggaattg ctaccgagct cctgcaggcc	9540
taggatcggt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccgggtctt	9600
gcgatgatta tcatataatt tctgttgat tacgttaagc atgtaataat taacatgtaa	9660
tgcgtacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa	9720
tacgcgatag aaaaacaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca	9780
tctatgttac tagatcgccc ggccgtttaa acttagttac taatcagtga tcagattgtc	9840
gtttcccgcc ttcaatttaa actatcagtg ttgacagga tataattggcg ggtaaacct	9900
agagaaaaga gcgtttatta gaataatcg atatttaaaa gggcgtgaaa aggtttatcc	9960
gttcgtccat ttgtatgtca atattggggg ggggggaaa ccacgttgtg tctcaaaatc	10020
tctgatgtta cattgcacaa gataaaaata tatcatcatg aacaataaaa ctgtctgctt	10080
acataaacag taatacaagg ggtgttcgcc accatgagcc atatccagcg tgaacctcg	10140
tgctcccgcc cgcgcctcaa ttccaatatg gatgcgacc tttatggcta caagtggcg	10200
cgcgacaacg tcggccagtc gggcgcgacc atttatcggc tttatggcaa acccgatgcc	10260
ccggaactgt tcctgaagca cggcaaaggc agcgtcgcaa acgatgtcac cgatgagatg	10320
gtccgcctga actggcttac cgagttcatg ccgctgcga cgattaagca tttcatccgt	10380
accccgagcg atgcctgggt cttgaccag gccattccgg gcaaaacggc ctttcaggtc	10440

cttgaagagt acccggaact cggtagaat atcgtggacg ccctcgcggt cttcctecgc	10500
cgtttgcata gcatccccgt gtgcaactgc cccttcaact cggaccgggt tttccgcctg	10560
gcacaggccc agtcgcgcgt gaataacggc ctcggtgacg cgagcgattt cgacgatgaa	10620
cggaaatggct ggccggtgga acaggtttgg aaggaaatgc acaaactgct tccgttctcg	10680
cgggattcgg tggtcacgca tgggtatttt tcctggata atctgatctt tgacgagggc	10740
aagctgatcg gctgcatcga cgtgggtcgc gtcgggtatcg ccgaccgcta tcaggacctg	10800
gcgatcttgt ggaattgcct cggcgagtgc tcgccctcgc tcagaagcg cctgttccag	10860
aagtacggca tcgacaaccc ggatatgaac aagctccagt tcacctcat gctggacgaa	10920
tttttttgaa cagaattggt taattggtt taacactggc agagcattac gctgacttga	10980
cgggacggcg gctttgttga ataaatcgaa cttttgtga gttgaaggat cgatgagttg	11040
aaggaccccg tagaaaagat caaaggatct tcttgagatc cttttttctt gcgcgtaatc	11100
tgctgcttgc aaacaaaaaa accaccgcta ccagcggtagg tttgtttgcc ggatcaagag	11160
ctaccaactc tttttccgaa ggttaactggc ttacgacagc cgacataacc aaatactgtc	11220
cttctagtgt agccgtagt aggccaccac ttcaagaact ctgtagcacc gcctacatac	11280
ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc	11340
gggttggaact caagacgata gttaccggat aaggcgagc ggtcgggctg aacgggggggt	11400
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt	11460
gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc	11520
ggcagggctg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggatatctt	11580
tatagtcttg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca	11640
ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt	11700
tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt	11760
attaccgect ttgagtgagc tgataccgct cgcgcgagcc gaacgaccga gcgcagcgag	11820
tcagttagcg aggaagcgga agagcgctg atgcggtatt ttctccttac gcatctgtgc	11880
ggtattttcac accgcatagg ccgcgatagg ccgacgcgaa gcgycggggc gtaggagcg	11940
cagcgaccga agggtaggag ctttttgag ctcttcggct gtgcgctggc cagacagtta	12000
tgcacaggcc aggggggttt taagagtttt aataagtttt aaagagtttt ayycggaaaa	12060
atgccttttt ttctctttta tatcagtcac ttacatgtgt gacgggttcc caatgtacgg	12120
ctttgggttc ccaatgtacg ggttccggtt cccaatgtac ggctttgggt tcccaatgta	12180
cgtgctatcc acaggaaaga gaccttttcg acctttttcc cctgctaggg caatttgccc	12240
tagcatctgc tccgtacatt aggaaccggc ggatgcttcg ccctcgatca ggttgcggtg	12300
gcgcataact aggatcgggc cagcctgccc cgcctctctc ttcaaatcgt actccggcag	12360

gtcatttgac ccgatcagct tgcgcacggt gaaacagaac ttcttgaact ctccggcgct	12420
gccactgcgt tcgtagatcg tcttgaacaa ccatctgggt tetgccttgc ctgcggcgcg	12480
gcgtgccagg cggtagagaa aacggccgat gccggggctg atcaaaaagt aatcgggggtg	12540
aaccgtcagc acgtccgggt tcttgcttcc tgtgatctcg cggtagatcc aatcagcaag	12600
ctcgatctcg atgtactccg gccgcccgtt ttgcctcttt acgatcttgt agcggctaata	12660
caaggcttca cctcggata ccgtcaccag gcggccgttc ttggccttct tggtagctg	12720
catggcaacg tgcgtggtgt ttaaccgaat gcaggtttct accaggctct cttcttgctt	12780
tccgccatcg gctcgcgggc agaactgag tacgtccgca acgtgtggac ggaacacgcg	12840
gccgggcttg tctcccttcc cttcccggtt tcggttcctg gattcgggta gatgggaaac	12900
cgccatcagt accaggctct aatccacac actggccatg ccggcggggc ctgcggaaac	12960
ctctacgtgc ccgtctggaa gctcgtagcg gatcacctcg ccagctcgtc ggtcacgctt	13020
cgacagacgg aaaacggcca cgtccatgat gctgcgacta tcgcgggtgc ccacgtcata	13080
gagcatcgga acgaaaaaat ctggttgctc gtcgcccttg ggcggtctcc taatcgacgg	13140
cgcacgggct gccggcggtt gccgggattc ttgctggatt cgatcagcgg ccccttgcca	13200
cgattcaccc gggcgtgctt ctgcctcgat gcgttgccgc tgggcggcct gcgcggcctt	13260
caacttctcc accaggctat caccacgcgc cgcgcgatt tgtaccgggc cggatgggtt	13320
gcgacgcctc acgcggattc ctccggcttg ggggttcag tgcattgca gggccggcag	13380
acaaccacgc cgttacgcgc tggccaaccg cccgttctc cacacatggg gcattccacg	13440
gcgtcgtgc ctggttgctt ttgatttcc atgcgcctc ctttagccgc taaaattcat	13500
ctactcattt attcatttgc tcatctactc tggtagctgc gcgatgtatt cagatagcag	13560
ctcggtaatg gtcttgctt ggcgtagcgc gtacatcttc agcttggtgt gatcctccgc	13620
cggcaactga aagttgacct gcttcattggc tggcgtgtct gccaggctgg ccaacgttgc	13680
agccttgctg ctgcgtgcgc tcggacggcc ggcacttagc gtgtttgtgc ttttgcctat	13740
tttctcttta cctcattaac tcaaatgagt ttgtatttaa ttccagcggc cagcgcctgg	13800
acctcgcggg cagcgtcgc ctccgggttct gattcaagaa cggttgtgct ggccggcgca	13860
gtgcctgggt agctcacgcg ctgcgtgata cgggactcaa gaatgggcag ctctgacctg	13920
gccagcgcct cggcaacctc accgccgatg cgcgtgcctt tgatcgcctc cgacacgaca	13980
aaggccgctt gtagccttcc atccgtgacc tcaatgcgtt gcttaaccag ctccaccagg	14040
tcggcggtgg cccaaatgtc gtaagggtt ggctgcaccg gaatcagcac gaagtcggct	14100
gccttgatcg cggacacagc caagtcgcgc gcctggggcg ctccgtcgat caactcgaag	14160
tcgcgcgggc cgatggcctt cacgtcgcgg tcaatcgtcg ggcggtcgat gccgacaacg	14220
gttagcggtt gatcttcccg cacggcgcgc caatcgcggg cactgccctg gggatcggaa	14280

```
tcgactaaca gaacatcggc cccggcgagt tgcagggcgc gggctagatg ggttgcgatg 14340
gtcgtettgc ctgacccgcc tttctgggta agtacagcga taaccttcac gcgttccct 14400
tgcgtatttg tttatttact catcgcatca tatacgcagc gaccgcatga cgcaagctgt 14460
tttactcaaa tacacatcac ctttttagat gatca 14495
```

```
<210> 83
<211> 33
<212> ДНК
<213> штучна
```

```
<220>
<223> синтетична послідовність
```

```
<400> 83
atatacgcgt ggtgcttaaa cactctggtg agt 33
```

```
<210> 84
<211> 38
<212> ДНК
<213> штучна
```

```
<220>
<223> синтетична послідовність
```

```
<400> 84
atatggcgcg cctttgacct acaaaatcaa agcagtca 38
```

```
<210> 85
<211> 32
<212> ДНК
<213> штучна
```

```
<220>
<223> синтетична послідовність
```

```
<400> 85
atatacgcgt agttctttgc tttcgaagtt gc 32
```

```
<210> 86
<211> 35
<212> ДНК
<213> штучна
```

```
<220>
<223> синтетична послідовність
```

```
<400> 86
atatggcgcg cctactacgt actgttttca attct 35
```

```
<210> 87
<211> 14065
<212> ДНК
<213> штучна
```

```
<220>
```


<223> векторна послідовність

<400> 87

```

gtgattttgt gccgagctgc cggtcgggga gctgttggt ggctgggtggc aggatatatt      60
gtggtgtaaa caaattgacg cttagacaac ttaataaac attgcggacg tctttaatgt      120
actgaattta gttactgacg actgattaag tactgataac ggtaccaagc ttccgcggt      180
gcagtgcagc gtgacccggt cgtgcccctc tctagagata atgagcattg catgtctaag      240
ttataaaaaa ttaccacata ttttttttgt cactctgtt tgaagtgcag tttatctatc      300
tttatacata tatttaaact ttactctacg aataatataa tctatagtac tacaataata      360
tcagtgtttt agagaatcat ataatgaac agttagacat ggtctaaagg acaattgagt      420
at ttgacaa caggactcta cagttttatc tttttagtgt gcagtgtgtc tctttttttt      480
ttgcaaatag cttcacctat ataactctc atccatttta ttagtacatc catttagggg      540
ttaggggttaa tggtttttat agactaattt ttttagtaca tctattttat tctatttttag      600
cctctaattt aagaaaacta aaactctatt ttagtttttt tatttaatat ttagatata      660
aaatagaata aaataaagtg actaaaaatt aaacaaatac cctttaagaa attaaaaaaa      720
ctaaggaaac atttttcttg ttccgagtag ataatgccag cctgttaaac gccgtcgacg      780
agtctaacgg acaccaacca gcgaaccagc agcgtcgcgt cgggccaagc gaagcagacg      840
gcacggcatc tctgtcgtcg cctctggacc cctctcgaga gttccgctcc accgttggac      900
ttgtcccgct gtcggcatcc agaaattcgc tggcggagcg gcagacgtga gccggcacgg      960
caggcggcct cctcctctcc tcacggcacc ggcagctacg ggggattcct tcccaccgc      1020
tctctcgctt tcccttctcc gcccgccgta ataatagac acccctcca caccctcttt      1080
ccccaaacct gtgtgtgtcg gagcgacac acacacaacc agatctcccc caaatccacc      1140
cgtcggcacc tccgcttcaa ggtacgcgc tegtctccc ccccccccc cctctctacc      1200
ttctctagat gggcgttccg gtccatggtt agggcccggt agttctactt ctgttcattg      1260
ttgtgttaga tccgtgtttg tgttagatcc gtgtgtctag cgttcgtaca cggatgcgac      1320
ctgtacgtca gacacgttct gattgctaac ttgccagtgt ttctcttttg ggaatcctgg      1380
gatggctcta gccgttccgc agacgggacg gatttcata ttttttttgt ttctgtgcac      1440
agggtttggt ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc      1500
atcttttcat gctttttttt gtcttggttg tgatgatgtg gtctgggttg gcggtcgttc      1560
tagatcggag tagaattctg tttcaacta cctgggtggat ttattaattt tggatctgta      1620
tgtgtgtgcc atacatatc atagttacga attgaagatg atggatggaa atatcgatct      1680
aggataggta tacatgttga tgcgggtttt actgatgcac atacagagat gcttttttgt      1740
cgcttggttg tgatgatgtg gtgtggttg gcggtcgttc attcgttcta gatcggagta      1800
gaatactggt tcaaaactac tgggtgtatt attaatattg gaactgtatg tgtgtgtcat      1860

```


acatcttcat agttacgagt ttaagatgga tggaaatata gatctaggat aggtatacat	1920
gttgatgtgg gttttactga tgcataata tgatggcata tgcagcatct attcatatgc	1980
tctaaccttg agtacctata tattataata aacaagtatg ttttataatt atttcgatct	2040
tgatatactt ggatgatggc atatgcagca gctatatgtg gatTTTTTTT gccctgcctt	2100
catacgtat ttatttgctt ggtactgttt cttttgtcga tgcacacct gttgtttggg	2160
gttactttctg cagccccggg gatccactag ttctagaaac catggccacc gccgccgcgc	2220
cgtctaccgc gctcactggc gccactaccg ctgcgcccaa ggcgaggcgc cgggcgcacc	2280
tcctggccac ccgccgcgc ctcgccgcgc ccacaggtg ctcagcgcg tcacccgcca	2340
tgcogatggc tccccggcc accccgctcc ggccgtgggg cccacccgat ccccgcaagg	2400
gcgcgcacat cctcgtcgag tccctcgagc gctgcggcgt ccgcgacgtc ttgcctacc	2460
ccggcgccac gtccatggag atccaccagg cactcaccgc ctcgccgcgc atcgccaacc	2520
acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgctcctcgg	2580
gccgcgtcgg cgtctgcac gccacctccg gcccgggcgc caccaacctt gtctccgcgc	2640
tcgcgcagc gctgctcgat tccgtccca tggctgccat caccggacag gtgcgcgac	2700
gcatgattgg caccgacgcc ttccaggaga cgcctatcgt cgaggtcacc cgctccatca	2760
ccaagcaca ctacctggtc ctcgacgtcg acgacatccc ccgcgtcgtg caggaggctt	2820
tcttctcgc ctcctctggt cgaccggggc cgggtcctgt cgacatcccc aaggacatcc	2880
agcagcagat ggcggtgcct gtctgggaca agcccatgag tctgcctggg tacattgcgc	2940
gccttcccaa gcccctcgc actgagttgc ttgagcaggt gctgcgtctt gttggtgaat	3000
cccggcgcc tgttctttat gttggcggtg gctgcgcagc atctggtgag gaggtgagac	3060
gctttgtgga gctgactgga atcccggta caactactct tatgggcctc ggcaacttcc	3120
ccagcgacga cccactgtct ctgcgcacgc taggtatgca tggcacggtg tatgcaaatt	3180
atgcagtgga taaggccgat ctgttgcttg cacttggtgt gcggtttgat gatcgtgtga	3240
cagggaagat tgaggctttt gcaagcagg ctaagattgt gcacgttgat attgatccgg	3300
ctgagattgg caagaacaag cagccacatg tgtccatctg tgcagatgtt aagcttgctt	3360
tgcagggcac gaatgctctt cttgaaggaa gcacatcaa gaagagcttt gactttggct	3420
catggaacga tgaatlllyal cagcagaaga gggaattccc ccttgggtat aaacatcta	3480
atgaggagat ccagccacaa tatgctatcc aggttcttga tgagctgacg aaaggcgagg	3540
ccatcatcgg cacagggtgt gggcagcacc agatgtgggc ggcacagtac tacacttaca	3600
agcgcccaag gcagtgggtg tottcagctg gtcttggggc tatgggattt ggtttgcggg	3660
ctgctgctgg tgcttctgtg gccaaaccag gtgttactgt tgttgacatc gatggagatg	3720
gtagctttct catgaacgtt caggagctag ctatgatccg aattgagaac ctcgccgtga	3780

aggtctttgt gctaaacaac cagcacctgg ggatgggtgg gcaagtgggag gacaggttct	3840
ataaggccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatatc	3900
cagatttctg gacgatcgcc aaaggggtca acattccagc ggtccgtgtg acaaagaaga	3960
acgaagtccg cgcagcgata aagaagatgc tcgagactcc agggccgtac ctcttgata	4020
taatcgcccc acaccaggag catgtgttgc ctatgatccc taatgggtgg gctttcaagg	4080
atatgatcct ggatgggtgat ggcaggactg tgtactgate taaaatccag caagcaactg	4140
atctaaaatc cagcaagcac cgctccctg ctagtacaag ggtgatatgt ttttatctgt	4200
gtgatgttct cctgtattct atcttttttt gtaggcgtc agctatctgt tatggtaate	4260
ctatgtagct tccgacctg taattgtgta gtctgttgtt ttccttctgg catgtgtcat	4320
aagagatcat ttaagtgcct ttgtctacat ataaataaga taataagcac tgctatgcag	4380
tggttctgaa ttggcttctg ttgccaatt taagtgtcca actggctcct gcttttgttt	4440
tcgctatttt tttccttttt tagttattat tatattggta atttcaactc aacatatgat	4500
gtatggaata atgctagggc tgcaatttca aactatttta caaccagaa tggcattttc	4560
gtggtttgag gggagtga aaatgagg catttgactg aattagttac ctgatccatt	4620
ttcgtggttt ggatcattgg aattaaattc cattctaata atagtaattt tggcatatat	4680
caattaagtt aattcggttt tatgcaaaat atatttgtat actattatta tcaagatgtc	4740
ggagatattt atatgctaca ttttactat acaggagtga gatgaagagt gtcatgtaag	4800
ttacacagta gaaacaaatt ctattaatgc ataaaatcat tccatcatc caccctatga	4860
at ttgagata gacctatctc taaactttga aaagtgggtg aatatcaa tccaaattaa	4920
ataagttatt ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgcctc	4980
taaagcta at ttggaactc aaactttctt agcattggag gggattgaga aaaaatatta	5040
at tcattttc atctcaatca ttcaatctcc aaagagattt gaggttccta ttagtctgtt	5100
ccatgcatca aatcggtcca atgtgtcatt atttgccatg acgattgacg agttgttctg	5160
gggcctagcg ctttccacgc cgatgtgctg gggcctgggc ctggagaaga cagcttgata	5220
tttaagcta tcaattgttt caattgattc ccacttcatt tttctaaatg tagaaacgg	5280
tgacgtataa gaaaaagaat gaattaggac ttttattccg tacactaatc tagagcggcc	5340
caagcttgta cactagtagc cgtcaattga tttaaattta attaatccc tgctcgtaa	5400
tgtgatacta ctagcatagt actagtacca tgcatacaca cagcaggtcg gccgcctgga	5460
tggtatgatg atgatactac atcatcctgt catccatcca ggcgatctag aagggcggtg	5520
gctagctagc aaactgtgac cggtttttct acgccgataa taatactttg tcatggtaca	5580
gacgtacagt actggttata tatatctgta gatttcaact gaaaagctag gatagctaga	5640
ttaattcctg agaaacacag ataaaattcg agcttggtca tagatgacaa acggaagac	5700

gcatgcattg gacgacgtat gcaatgcgag cgcgtctcgt gtcgtcccgt ccaagtctgg	5760
cgatctcacg ccacgtgctc aacagctcaa ggactgttcg tcaccagcgt taaattcatt	5820
gaagggatga cgcatttcgg catttgatcat tgcttgtagc tatatatata tatccaacag	5880
atttctctca agcttttgta tgcgtgaatg taaagtctag cttatacgac agcacgtgca	5940
gatataattaa cgtcattatt aggtggagag caagatgcat gatctggtag aaattgtcga	6000
aaacacaaga gagagtgaag tgcacacttc tggatatagga gtgtatacgc cgctgggtgg	6060
tgggcaatgc gcgcgcgaat attggccaat gaaacctagc aacgcccact cgccacgccc	6120
catgaatggc ccccgacggg cagcgagcca gccagtgcgc gcgcgcggcc cagccggagt	6180
cggcggaacg cgccacgggg gacgaggcgc ccgagggcgc aggcagcgcg gcatggcaag	6240
caagccgaag cgggcaagcg acctgcatgc agccccctgcc cctcgccctc gtcagtcgtc	6300
ccagcctccc actggaatcc acccaaccgc ccttctctct ccaaagcacg cgccccgcga	6360
ctcgccctcg cctacgtgtc ggcagcgctc ccgcgggtcg ccacgtacc ccgccccgtt	6420
ctcccaactg cccctccctc tgcgcgcgtc cgattggctg acccgccctt cttaagccgc	6480
gccagcctcc tgtccggggc ccaacgcctg gctccgtcgt cgtctccgcc cccagagtga	6540
tcgagcccac tgacctggcc ccgagcctc agctcgtgag tccggcgcgt ggtgcttaaa	6600
cactctggtg agttctagta cttctgctat gatcgatctc attaccattt cttaaatttc	6660
tctccctaaa tattccgagt tcttgatctt tgataaactc aggttttctc tttttgataa	6720
atctgggtctt tccatttttt tttttttgtg gtttaatttag tttcctatgt tcttcgattg	6780
tattatgcat gatctgtgtt tggattctgt tagattatgt attgggtgaat atgtatgtgt	6840
ttttgcatgt ctggtttttg tcttaaaaat gttcaaatct gatgatttga ttgaagcttt	6900
tttagtggtg gtttgattct tctcaaaact actgttaatt tactatcatg ttttccaaact	6960
ttgattcatg atgacacttt tgttctgctt tgttataaaa ttttggttgg tttgattttg	7020
taattatagt gtaattttgt taggaatgaa catgttttaa tactctgttt tcgatttgtc	7080
acacattcga attattaatc gataatttaa ctgaaaatc atggttctag atcttgttgt	7140
catcagatta tttgtttcga taattcatca aatatgtagt ccttttgctg atttgcgact	7200
gtttcatttt ttctcaaaat tgttttttgt taagtttato taacagttat cgttgtcaaa	7260
agttctcttc attttgcaaa atcttctctt tttttttgtt tgtaactttg ttttttaagc	7320
tacacattta gtctgtaaaa tagcatcgag gaacagttgt cttagtagac ttgcatgttc	7380
ttgtaacttc tatttgtttc agtttggtga tgactgcttt gatttttagt gtcaaggcgc	7440
cgccaccatg gaagacgcca aaaacataaa gaaaggccgc gcgccattct atccgctgga	7500
agatggaacc gctggagagc aactgcataa ggctatgaag agatacgcgc tggttcctgg	7560
aacaattgct tttacagatg cacatatcga ggtggacatc acttacgctg agtacttcga	7620

aatgtccgtt	cggttggcag	aagctatgaa	acgatatggg	ctgaatacaa	atcacagaat	7680
cgtcgtatgc	agtgaaaact	ctcttcaatt	ctttatgccg	gtgttgggcg	cgttatttat	7740
cggagttgca	gttgcgcccc	cgaacgacat	ttataatgaa	cgtgaattgc	tcaacagtat	7800
gggcatttcg	cagcctaccg	tgggtgttcgt	ttccaaaaag	gggttgcaaa	aaattttgaa	7860
cgtgcaaaaa	aagctcccaa	tcattccaaa	aattattatc	atggattcta	aaacggatta	7920
ccagggattt	cagtcgatgt	acacgttcgt	cacatctcat	ctacctcccg	gttttaatga	7980
atacgatttt	gtgccagagt	ccttcgatag	ggacaagaca	attgcactga	tcataaactc	8040
ctctggatct	actggctcgc	ctaaagggtg	cgctctgcct	catagaactg	cctgcgtgag	8100
attctcgcat	gccagagatc	ctatttttgg	caatcaaadc	attccggata	ctgcgatttt	8160
aagtgttgtt	ccattccatc	acgggttttg	aatgtttact	acactcggat	atttgatatt	8220
tggatttcga	gtcgtcttaa	tgtatagatt	tgaagaagag	ctgtttctga	ggagccttca	8280
ggattacaag	attcaaagtg	cgctgctggg	gccaacccca	ttctccttct	tcgccaaaaa	8340
cactctgatt	gacaaatcgc	atttatctaa	tttacacgaa	attgcttctg	gtggcgctcc	8400
cctctctaag	gaagtcgggg	aagcggttgc	caagagggtc	catctgccag	gtatcaggca	8460
aggatatggg	ctcactgaga	ctacatcagc	tattctgatt	acacccgagg	gggatgataa	8520
accgggctgc	gtcggtaaa	ttgttccatt	ttttgaagcg	aagggttggtg	atctggatac	8580
cgggaaaaac	ctgggcgtta	atcaaagagg	cgaactgtgt	gtgagaggtc	ctatgattat	8640
gtcgggttat	gtaacaatc	cgggaagcgc	caacgccttg	attgacaagg	atggatggct	8700
acattctgga	gacatagctt	actgggacga	agacgaacac	ttcttcatcg	ttgaccgcct	8760
gaagtctctg	attaagtaca	aaggetatca	ggtggctccc	gctgaattgg	aatccatctt	8820
gtccaacac	cccaacatct	tcgacgcagg	tgtcgcagg	cttcccgacg	atgacgccgg	8880
tgaacttccc	gccgcggttg	ttgttttgga	gcacggaaa	acgatgacgg	aaaaagagat	8940
cgtggattac	gtcggcagtc	aagtaacaac	cgcgaaaaag	ttgcgcggag	gagttgtgtt	9000
tgtggacgaa	gtaccgaaa	gtcttaccgg	aaaactcgac	gcaagaaaaa	tcagagagat	9060
cctcataaag	gccaagaagg	gcggaaagat	cgcgtgttaa	cctgcaggcc	taggatcggt	9120
caaacatttg	gcaataaagt	ttcttaagat	tgaatcctgt	tgccggtctt	gcgatgatta	9180
tcataataat	tctgttgaat	taagtttaag	atgtaataat	taacatgtaa	tgcataagct	9240
tattttatgag	atgggttttt	atgattagag	tcccgaatt	atacatttaa	tacgcgatag	9300
aaaacaaaa	atagcgcgca	aactaggata	aattatcgcg	cgcgggtgtca	tctatgttac	9360
tagatcgccc	ggccgtttta	acttagttac	taatcagtga	tcagattgtc	gtttcccgcc	9420
ttcactttta	actatcagtg	tttgacagga	tatattggcg	ggtaaaccta	agagaaaaga	9480
gcgtttatta	gaataatcgg	atatttaaaa	gggcgtgaaa	aggttttatcc	gttcgtccat	9540

ttgtatgtca atattggggg ggggggaaag ccacgttggtg tctcaaaatc tctgatgtta	9600
cattgcacaa gataaaaata tatcatcatg aacaataaaa ctgtctgctt acataaacag	9660
taatacaagg ggtgttcgcc accatgagcc atatccagcg tgaaacctcg tgcctccgcc	9720
cgcgcctcaa ttccaatatg gatgccgacc tttatggcta caagtgggcg cgcgacaacg	9780
tcggccagtc gggcgcgacc atttatcggc tttatggcaa acccgatgcc ccggaactgt	9840
tcctgaagca cggcaaaggc agcgtcgcaa acgatgtcac cgatgagatg gtccgcctga	9900
actggcttac cgagtccatg ccgctgccga cgattaagca tttcatccgt accccggacg	9960
atgcctggct cttgaccacg gccattccgg gcaaaacggc ctttcagggtc cttgaagagt	10020
acccggactc cggtgagaat atcgtggacg ccctcgcggt cttcctccgc cgtttgcata	10080
gcatccccgt gtgcaactgc cccttcaact cggaccgggt tttccgcctg gcacaggccc	10140
agtcgcgcat gaataacggc ctggttgacg cgagcgattt cgacgatgaa cggaatggct	10200
ggccggtgga acaggttttg aaggaaatgc acaaactgct tccgtttctc cgggattcgg	10260
tggtcacgca tgggtatttt tccctggata atctgatctt tgacgagggc aagctgatcg	10320
gctgcatcga cgtgggtcgc gtcgggtatc ccgaccgcta tcaggacctg gcgatcttgt	10380
ggaattgcct cggcgagttc tcgcccctgc tcagaaagcg cctgttccag aagtaaggca	10440
tcgacaaccc ggatatgaac aagctccagt tccacctcat gctggacgaa tttttttgaa	10500
cagaattggt taattgggtg taacactggc agagcattac gctgacttga cgggacggcg	10560
gctttgttga ataaatcgaa cttttgtcga gttgaaggat cgatgagttg aaggaccccg	10620
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc	10680
aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	10740
tttttccgaa ggtaactggc ttacgcagag cgcagatacc aaatactgtc cttctagtgt	10800
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc	10860
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	10920
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggggt tcgtgcacac	10980
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	11040
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg	11100
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg	11160
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga	11220
gcctatggaa aaacgccagc aacgcggcct ttttacgggt cctggccctt tgctggcctt	11280
ttgctcacat gttctttcct cgtttatccc ctgattctgt ggataaccgt attaccgcct	11340
ttgagtgagc tgataccgct cgcgcagcc gaacgaccga gcgcagcgag tcagtgagcg	11400
aggaagcggg agagcgctg atgcggtatt ttctccttac gcatctgtgc ggtatttcac	11460

accgcatagg ccgcgatagg ccgacgcgaa gcggcgsgggc gtagggagcg cagcgaccga	11520
agggtagggc ctttttgca gctctcggt gtgcgctggc cagacagtta tgcacaggcc	11580
aggcgsggtt taagagttt aataagttt aaagagttt aggcggaaaa atcgctttt	11640
ttctctttta tatcagtcac ttacatgtgt gaccggttcc caatgtacgg ctttgggttc	11700
ccaatgtacg ggttccggt cccaatgtac ggctttgggt tcccaatgta cgtgctatcc	11760
acaggaaaga gaccttttcg acctttttcc cctgctaggg caatttgccc tagcatctgc	11820
tccgtacatt aggaaccggc ggatgcttcg cctcgatca ggttgcggta gcgcatact	11880
aggatcgggc cagcctgccc cgctctctcc ttcaaatcgt actccggcag gtcatttgac	11940
ccgatcagct tgcgcacggg gaaacagaac ttcttgaact ctccggcgct gccactgcgt	12000
tcgtagatcg tcttgaacaa ccatctggct tctgccttgc ctgcggcgcg gcgtgccagg	12060
cggtagagaa aacggccgat gccggggctg atcaaaaagt aatcggsggtg aaccgtcagc	12120
acgtccgggt tcttgcttc tgtgatctcg cggtagatcc aatcagcaag ctcgatctcg	12180
atgtaactcg gcgcgccgt ttcgctcttt acgatcttgt agcggtaat caaggcttca	12240
ccctcggata ccgtcaccag gcggccgttc ttggccttct tggtagctg catggcaacg	12300
tgcgtggtgt ttaaccgaat gcaggtttct accaggtcgt ttttctgct tccgccatcg	12360
gctcgccggc agaacttgag tacgtccga acgtgtggac ggaacacgcg gccgggcttg	12420
tctcccttcc ctcccggtg tcggttcacg gattcgggta gatgggaaac cgccatcagt	12480
accaggtcgt aatcccacac actggccatg ccggcgsgggc ctgcggaaac ctctacgtgc	12540
ccgtctggaa gctcgtagcg gatcaactcg ccagctcgtc ggtcacgctt cgacagacgg	12600
aaaacggcca cgtccatgat gtcgcacta tcgcgggtgc ccacgtcata gagcatcgga	12660
acgaaaaaat ctggttgctc gtcgcccttg ggcggcttcc taatcgacgg cgcaccggct	12720
gcggcgsggt gccgggatc tttgcggatt cgatcagcgg ccccttgcca cgattcaccg	12780
gggcgtgctt ctgcctcgat gcgttgccgc tggcgsgcct gcgcggcctt caacttctcc	12840
accaggtcat caccacgcgc cgcgcgatt tgtaccgggc cggatggttt gcgaccgtc	12900
acgccgattc ctcgsgcttg ggggttccag tgccattgca gggccggcag acaaccacgc	12960
cgcttacgcc tggccaaccg cccgttctcc cacacatggg gcattccacg gcgtcggtgc	13020
ctggttggtc ttgattttcc atgcgcctc ctttagccgc taaaattcat ctactcattt	13080
attcattttg tcattttact tggtagctgc gcgatgtatt cagatagcag ctcggtaatg	13140
gtcttgctt ggcgtaccgc gtacatcttc agcttgggtg gatcctccgc cggcaactga	13200
aagttgacce gttcatggc tggcgtgtct gccaggctgg ccaacgttgc agccttgctg	13260
ctgcgtgcgc tcggacggcc ggcacttagc gtgtttgtgc ttttgcctat ttctctttta	13320
cctcattaac tcaaatgagt tttgatttaa tttcagcggc cagcgccctg acctcgcggy	13380


```

cagcgtcgcc ctcggttct gattcaagaa cggttgtgcc ggcggcgga gtgcctgggt 13440
agctcacgcg ctgcgtgata cgggactcaa gaatgggcag ctcgtaaccg gccagcgct 13500
cggcaacctc accgccgatg cgcgtgcctt tgatcgcccg cgacacgaca aaggccgctt 13560
gtagccttcc atccgtgacc tcaatgcgct gcttaaccag ctccaccagg tcggcggtgg 13620
cccaaagtgc gtaagggtt ggctgcaccg gaatcagcac gaagtcggct gccttgatcg 13680
cggacacagc caagtccgc gccctggggc ctcgctgat cactacgaag tcgcgcggc 13740
cgatggcctt cagtcgcgg tcaatcgctg ggcggctgat gccgacaacg gttagcggtt 13800
gatcttcccg cagggccgc caatcgcg gactgccctg gggatcgga tcgactaaca 13860
gaacatcgcc ccggcgagt tgcaggcgcc gggctagatg ggttgcgatg gtcgtcttgc 13920
ctgacccgcc tttctggtta agtacagcga taaccttcat gcgttcccc tgcgtatttg 13980
tttatttact catcgcatca tatacgcagc gaccgcatga cgcaagctgt ttactcaaa 14040
tacacatcac ctttttagat gatca 14065

```

ФОРМУЛА ВИНАХОДУ

- 5 1. Спосіб одержання вискоекспресійного насінно-специфічного та/або насінно-селективного рослинного промотору, що включає
 - а) забезпечення насінно-специфічного та/або насінно-селективного рослинного промотору, і
 - б) функціональне зв'язування з промотором молекули нуклеїнової кислоти, яка посилює експресію нуклеїнової кислоти (NEENA), гетерологічної зазначеному промотору, причому
- 10 NEENA має послідовність SEQ ID NO: 2.
2. Спосіб за п. 1, який **відрізняється** тим, що додатково включає стадії
 - а) включення NEENA, яка має послідовність SEQ ID NO: 2, у рослину або її частину, та
 - б) об'єднання зазначеної NEENA з геномом зазначеної рослини або її частини таким чином, щоб зазначена NEENA була функціонально зв'язаною з ендегенною насінно-специфічною
- 15 та/або насінно-селективною експресованою нуклеїновою кислотою, гетерологічною зазначеній NEENA.
3. Спосіб за п. 2, який **відрізняється** тим, що додатково включає стадію с) регенерації рослини або її частини, що включає зазначену NEENA, з зазначеної трансформованої клітини.
4. Спосіб за будь-яким з пп. 1-3, який **відрізняється** тим, що додатково включає стадії
 - а) забезпечення експресійного конструкта, який включає NEENA, яка має послідовність SEQ ID NO: 2, функціонально зв'язану з насінно-специфічним та/або насінно-селективним промотором і з однією або кількома молекулами нуклеїнової кислоти, причому остання є гетерологічною зазначеній NEENA, перебуваючи під контролем зазначеного насінно-специфічного та/або насінно-селективного промотору, та
 - б) об'єднання зазначеної експресійної послідовності, яка включає зазначену NEENA, з геномом зазначеної рослини або її частини.
- 25 5. Спосіб за п. 4, який **відрізняється** тим, що додатково включає стадію с) регенерації рослини або її частини, яка включає зазначені одну або кілька експресійних послідовностей, з зазначеної трансформованої рослини або її частини.
- 30 6. Спосіб за будь-яким з пп. 1-5, який **відрізняється** тим, що рослина є однодольною або дводольною рослиною.
7. Спосіб за п. 6, який **відрізняється** тим, що рослина є дводольною рослиною.
8. Спосіб за п. 6, який **відрізняється** тим, що рослина є однодольною рослиною.
9. Спосіб за будь-яким з пп. 1-8, який **відрізняється** тим, що зазначена NEENA є функціонально зв'язаною з насінно-специфічним та/або насінно-селективним промотором, наближеним до сайту початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти.
- 35 10. Спосіб за п. 9, який **відрізняється** тим, що зазначена NEENA є функціонально зв'язаною з насінно-специфічним та/або насінно-селективним промотором за 2500 п. о. або менше, переважно 2000 п. о. або менше, більш переважно 1500 п. о. або менше, ще більш переважно 1000 п. о. або менше, найбільш переважно 500 п. о. або менше, від сайту початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти.
- 40 11. Спосіб за будь-яким з пп. 1-10, який **відрізняється** тим, що зазначена NEENA є функціонально зв'язаною з насінно-специфічним та/або насінно-селективним промотором перед сайтом початку трансляції молекули нуклеїнової кислоти, експресія якої перебуває під

контролем зазначеного насінно-специфічного та/або насінно-селективного промотору.

12. Спосіб за будь-яким з пп. 1-10, який **відрізняється** тим, що зазначена NEENA є функціонально зв'язаною з насінно-специфічним та/або насінно-селективним промотором у межах 5'-UTR молекули нуклеїнової кислоти, експресія якої перебуває під контролем зазначеного насінно-специфічного та/або насінно-селективного промотору.

13. Рекombінантна експресійна послідовність, яка включає NEENA, яка має послідовність SEQ ID NO: 2.

14. Рекombінантна експресійна послідовність за п. 13, яка **відрізняється** тим, що включає NEENA, яка має послідовність SEQ ID NO: 2, функціонально зв'язану з одним або кількома насінно-специфічними та/або насінно-селективними промоторами, та одну або кілька експресованих молекул нуклеїнової кислоти, причому остання є гетерологічною зазначеній NEENA.

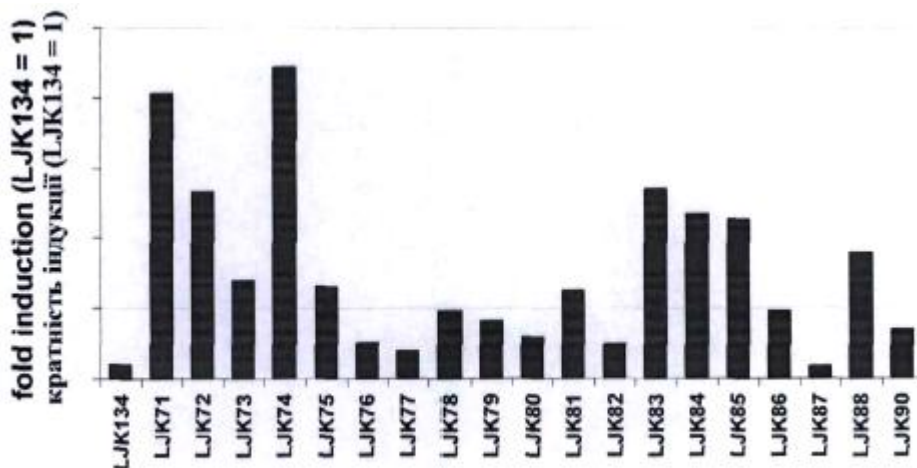
15. Трансгенна клітина або трансгенна рослина, або її частина, яка включає рекombінантний вектор експресії, який включає одну або кілька рекombінантних експресійних послідовностей за п. 13 або 14 або рекombінантну експресійну послідовність за п. 13 або 14.

16. Трансгенна клітина, яка включає рекombінантний вектор експресії, який включає одну або кілька рекombінантних експресійних послідовностей за п. 13 або 14 або рекombінантну експресійну послідовність за п. 13 або 14 причому трансгенна клітина є вибраною або походить з групи, до якої належать бактерії, гриби, дріжджі або рослини.

17. Трансгенна рослина або її частина за п. 15, яка **відрізняється** тим, що зазначена рослина або її частина є дводольною рослиною.

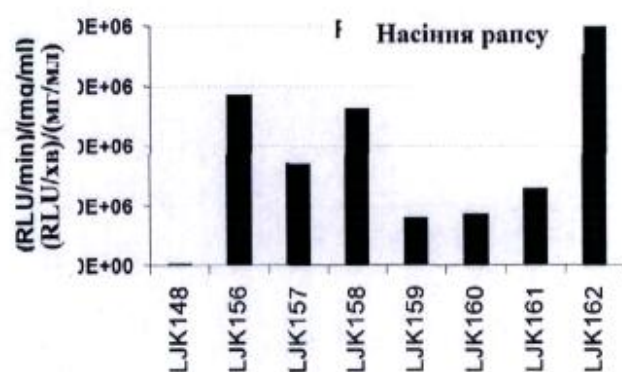
18. Трансгенна рослина або її частина за п. 15, яка **відрізняється** тим, що зазначена рослина або її частина є однодольною рослиною.

19. Застосування NEENA, яка має послідовність SEQ ID NO: 2, або рекombінантної послідовності за п. 13 або 14 для посилення експресії у рослинах або їх частинах.

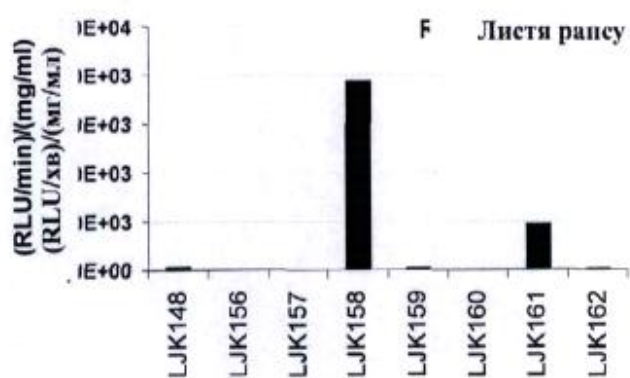


ФІГ. 1

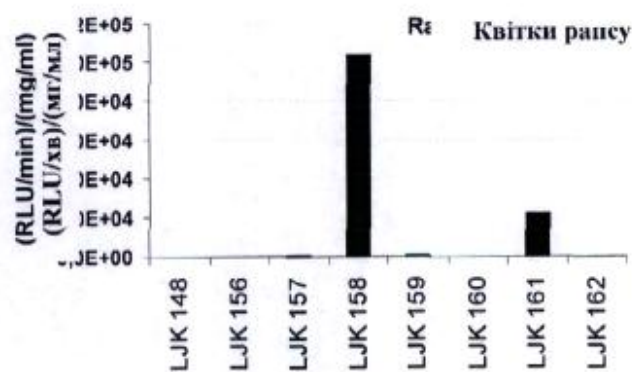
A)



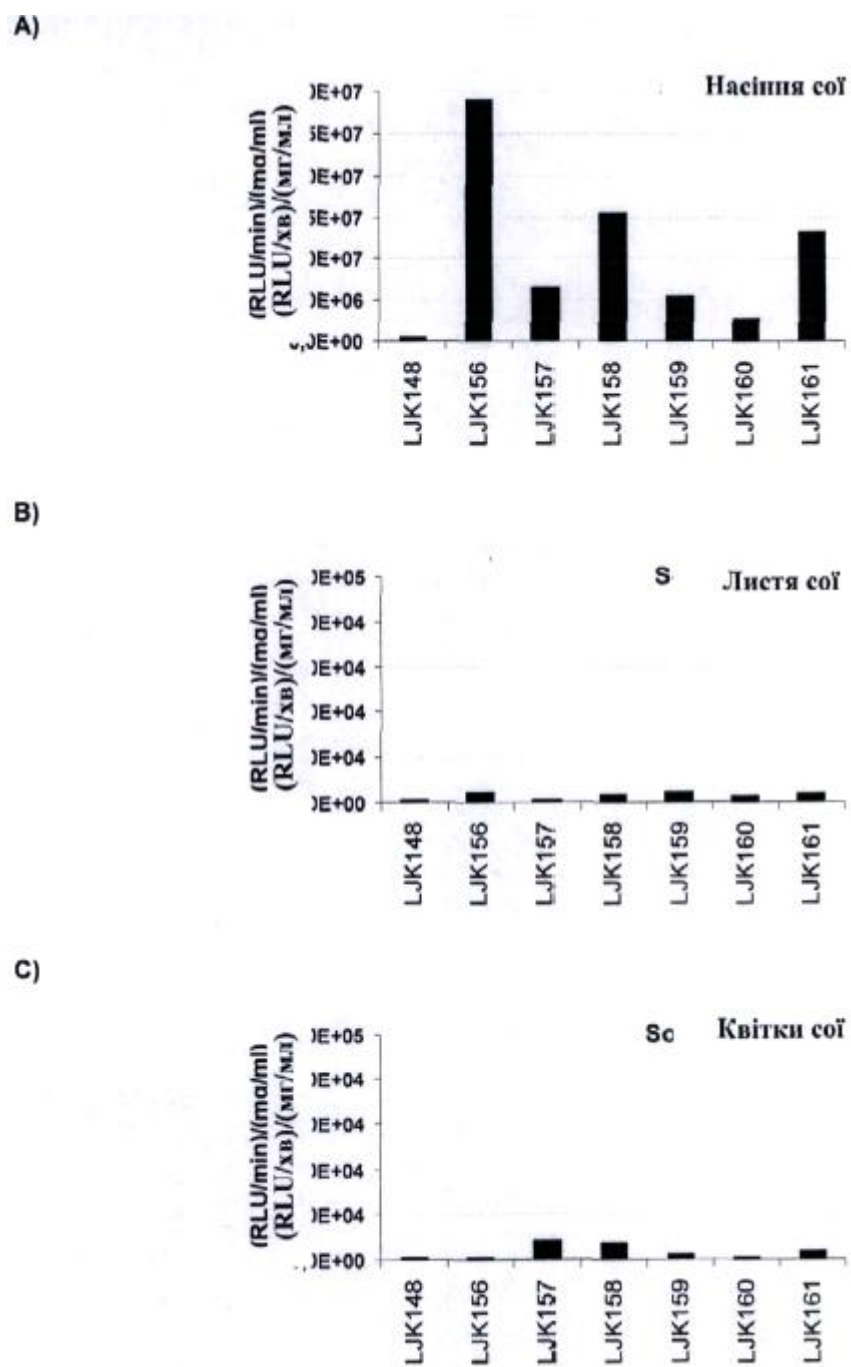
B)



C)

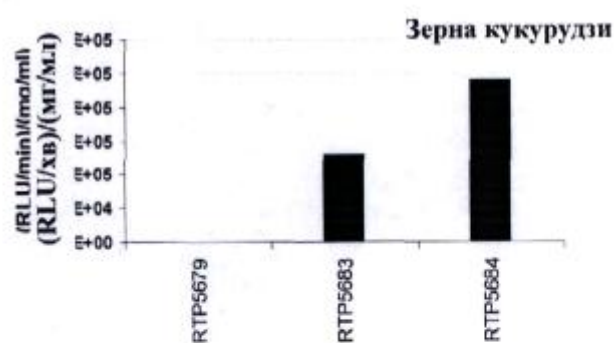


ФІГ. 2

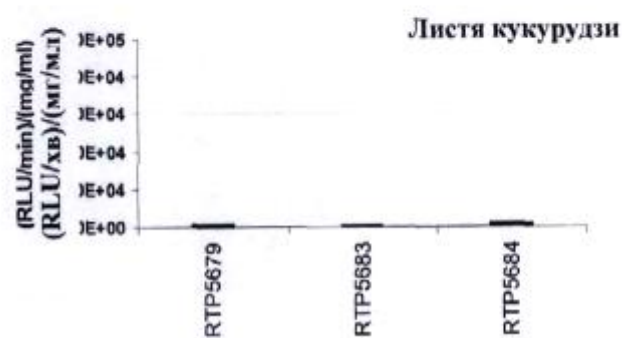


ФІГ. 3

A)

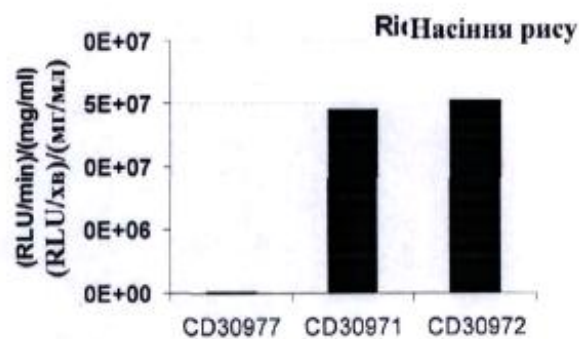


B)

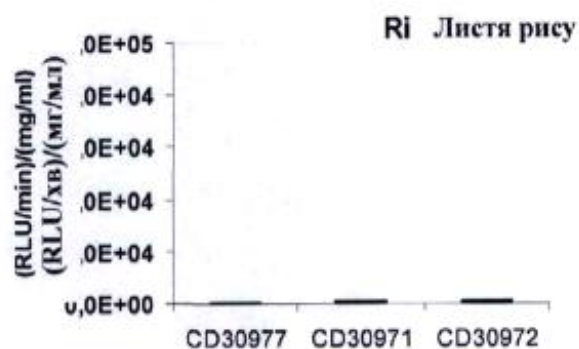


ФІГ. 4

A)



B)



ФІГ. 5

Комп'ютерна верстка С. Чулій

Міністерство розвитку економіки, торгівлі та сільського господарства України,
вул. М. Грушевського, 12/2, м. Київ, 01008, Україна

ДП "Український інститут промислової власності", вул. Глазунова, 1, м. Київ – 42, 01601